Characterization of Foregrounds and Systematics for Sensitive Radio Observations

Ph.D. Thesis

by

Aishrila Mazumder

Department of Astronomy, Astrophysics and Space Engineering

Indian Institute of Technology Indore
Khandwa Road, Simrol, Indore - 453552, India
April 2023

Characterization of Foregrounds and Systematics for Sensitive Radio Observations

A THESIS

Submitted in partial fulfillment of the requirements for the award of the degree

of

DOCTOR OF PHILOSOPHY

by

Aishrila Mazumder

Department of Astronomy, Astrophysics and Space Engineering

Indian Institute of Technology Indore
Khandwa Road, Simrol, Indore - 453552, India
April 2023

INDIAN INSTITUTE OF TECHNOLOGY INDORE

I hereby certify that the work which is being presented in the thesis entitled **CHARACTERIZATION OF FOREGROUNDS AND SYSTEMATICS FOR SENSITIVE RADIO OBSERVATIONS** in the partial fulfillment of the requirements for the award of the degree of **DOCTOR OF PHILOSOPHY** and submitted in the **DEPARTMENT OF ASTRONOMY, ASTROPHYSICS AND SPACE ENGINEERING, Indian Institute of Technology Indore**, is an authentic record of my own work carried out during the time period from July 2017 to April 2023 under the supervision of Prof. Abhirup Datta, Professor and Head, Department of Astronomy, Astrophysics and Space Engineering, Indian Institute of Technology Indore.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute.

Aishrila Mazumder

18.04.2023

AISHRILA MAZUMDER

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

18/4/2023

Prof. ABHIRUP DATTA

AISHRILA MAZUMDER has successfully given his/her Ph.D. Oral Examination held on **18.04.2023.**

18/4/2023

Prof. ABHIRUP DATTA

Dedicated

to

My Parents

"Try to make sense of what you see and wonder about what makes the universe exist "- *Stephen Hawking*

Acknowledgements

A wise person once said, *"The journey of a thousand miles begins with one step"*. The first step in my attempted journey to know how it all began is the completion of this thesis. But there are so many people without whom I never could complete this step.

First, I want to acknowledge Ma, Bapi and Didi, my rocks, my constant support. They have been the best parents and sibling one can ask for. I would not make it till here, or anywhere really, if not for them. They encouraged me when needed, stopped me when required, and were with me supporting me every step of the way. I would also like to thank Sumanjit, who so happened to be a senior and is now the dude I am married to. Beside being the one person I could rant to about literally everything, this thesis probably would not be completed on time without him, literally (his laptop came to my aid when mine started malfunctioning just weeks before the submission) and figuratively (via poking me consistently to write the thesis).

I also take this opportunity to heartily thank my supervisor, Prof. Abhirup Datta for his support and guidance. His encouragement and help has helped me an independent researcher in a very short time. I also have the fortune of having an excellent collaborator- Dr. Arnab Chakraborty. He always tolerated the numerous queries I had, and always showed me the way when I was stuck. I am also grateful to Dr. Nirupam Roy and Dr. Suman Majumdar for being there to teach or help, whenever I required, whichever I required. I also acknowledge my other collaborators - Dr. Samir Choudhuri, Dr. Yogesh Wadadekar and Dr. C.H. Iswara-Chandra for their help in my research. I would also like to thank my PSPC members, Dr. Bhargav Vaidya and Prof. Ram Bilas Pachori for their candid evaluations and valuable inputs that helped make me a better researcher. I also acknowledge my PSPC member, Dr. Siddharth Malu for his faith in my research. I would further extend my gratitude to Dr. Manoneeta Chakraborty for her help, support and showing me every step of the way that it OK to be a woman doing science. I also want to thank Desh (Dr. Avinash Deshpande) for believing I can finish what I started even on those dreadful days I wanted to give up.

I also want to thank the some excellent people who have made the famously "lonely" Ph.D. journey a little less so. My seniors and department legends - Sumanjit, Arnab, Majid, Madhurima, Ramij and Althaf, who have made me feel a part of the lab from day one, and never created the junior-senior divide. I learnt a lot of things about Ph.D., and also about life, from them. They are the best seniors anyone can ask for. I am also grateful to Deepthi, my one and only batch-mate, for being a fellow survivor of a lot. And of course, the awesome people who joined after me- Sarvesh, Akriti, Swarna, Gourab, Anshuman, Kamran, Hemapriya, Parul, Sriyasriti, Unnati, Sushmita, Harsha, Sanmoy, Arghyadeep, Sayan, Chandra Shekhar, Soumen, Chandrani. It has been a privilege knowing you people. Life would have been lonely and unbearable, being stuck inside the campus amidst a raging pandemic, if it wasn't for you guys. Thank you for all the craziness, gossip, SHASN sessions, reels, dance sessions and food ideas. I would also be lost if it wasn't for my two oldest buddies - Sreya and Shayeri. They were there, no questions asked, whenever I needed them (which was surprisingly a lot of times).

Lastly, I want to thank IIT Indore for providing me the institute fellowship to support this research, and providing the necessary computation facilities that made this thesis possible. This research made use of data observed using the Giant Metrewave Radio Telescope (GMRT). We thank the staff of GMRT for making these observations possible. GMRT is run by the National Centre for Radio Astrophysics of the Tata Institute of Fundamental Research. The work also makes use of the OSKAR simulator (https://github.com/OxfordSKA/OSKAR) & the CASA software package (https://casa.nrao.edu/). Codes used for this work have been written using the Python programming language (https://www.python.org/). I also extend my gratitude to the entire stackoverflow (https://stackoverflow.com/) community for being the one-stop solution to all my coding issues.

Abstract

Observation of the redshifted 21-cm signal from the epoch of reionization is a challenging endeavor in observational cosmology. This signal arises from a hyperfine transition in the ground state of neutral hydrogen and is detectable using radio telescopes. Since the Universe consisted principally of neutral hydrogen at the early stages, observing this signal is crucial for understanding the formation and evolution of structures in the Universe. But the target is extremely faint and prone to contamination by orders of magnitude stronger foregrounds and systematic influences. This thesis aims to understand and characterize these foregrounds and systematics that obscure the detection of the redshifted 21-cm from the Epoch of Reionization (EoR). The studies have been done using both observations and simulations. The observed data used from GMRT is of the extragalactic deep field Lockman Hole region centered at 325 MHz. The observed frequency band (Band 3) is just higher than the lowest frequency Band-2 in the (u)GMRT (120-250 MHz) which covers the most of the redshifted HI signal from CD/EoR. The advantage of Band 3 is that the systematics are comparatively lower and thus facilitates achieving tighter constraints on the frequency-dependent foreground characteristics. Most of the previous studies are either of the two categories: a) deep in sensitivity but narrow fields of view, or b) cover wide areas but shallow in sensitivity. This work, for the first time, use a relatively wider field-of-view ($6^\circ \times 6^\circ$ with angular resolution of ~ 9") with moderate sensitivity limit (RMS $50\,\mu$Jy beam^{-1}) for characterizing foregrounds. The flux, astrometry, and Euclidean normalized differential source counts derived from this catalogue are all consistent with previous studies. Three different galactic latitudes were selected within the field of view to determine the angular power spectrum (APS) of the diffuse Galactic synchrotron emission (DGSE). This was done to determine if there was any systematic variation of the synchrotron

power as a function of latitude. While no such latitudinal variation was found, the power was seen to vary between ∼1 - 100 mK^2 as a function of the angular mode for the latitudes considered. These observations of the Lockman Hole region was also used for studying source clustering properties. First, the sources were classified as SFGs (star forming galaxies) or AGNs (Active Galactice Nuclei) based on the radio luminosity function. Then, the angular and spatial clustering properties of the sources were estimated. These results from the clustering estimates (clustering lengths and bias parameter) are in agreement with the previous observations. The detailed foreground properties from Lockman Hole region emphasize the need for further deep and wide-field low-frequency observations along different lines of sight to develop better global foreground models. In addition to the observations, an end-to-end simulation pipeline was also created to characterize systematic errors in EoR observations. For the first time, two systematic errors - calibration and position errors have been studied using different telescope layouts. Mock observations are done with a theoretical model of the 21-cm signal and point source foregrounds. Their effects have been analyzed for the image plane and the power spectrum estimation. The optimal tolerance for calibration errors was found to be ∼0.1%. For position errors, an offset of ∼5" was found optimal. The pipeline developed for this work can be extended to study the effect of primary beam errors, ionosphere, etc. Thus, this thesis emphasized the need for detailed foreground characterization and lays down the platform (end-to-end pipeline) to study the same. This platform can be used for CD/EoR work as well as extended to HI Intensity mapping and other precision interferometric observations with the SKA or its precursors.

List of Publications

A. Publications Towards Thesis:

1. **Aishrila Mazumder**, Arnab Chakraborty, Abhirup Datta, Samir Choudhuri, Nirupam Roy, Yogesh Wadadekar, C H Ishwara-Chandra, *Characterizing EoR foregrounds: a study of the Lockman Hole region at 325 MHz*, Monthly Notices of the Royal Astronomical Society, Volume 495, Issue 4, July 2020 doi: 10.1093/mnras/staa1317 [arXiv: 2005.05205]

2. **Aishrila Mazumder**, Abhirup Datta, Arnab Chakraborty, Suman Majumdar, *Observing the Reionization : Effect of Calibration and Position Errors on Realistic Observation Conditions*, Monthly Notices of the Royal Astronomical Society, July 2022, doi: 10.1093/mnras/stac1994 [arXiv: 2207.06169]

3. **Aishrila Mazumder**, Arnab Chakraborty, Abhirup Datta, *A study on the Clustering Properties of Radio-Selected sources in the Lockman Hole Region at 325 MHz*, Monthly Notices of the Royal Astronomical Society, December 2022, doi: 10.1093/mnras/stac2801 [arXiv: 2208.00992]

B. Other Publications:

1. Arnab Chakraborty, Abhirup Datta, **Aishrila Mazumder**, *A Comparative Analysis To Deal With Missing Spectral Information Caused By RFI In Cosmological HI 21 cm Observations*, The Astrophysical Journal, Volume 929, Issue 1, April 2022, doi: 10.3847/1538-4357/ac5cc5 [arXiv: 2203.04994]

2. Abinash Kumar Shaw, Manoj Jagannath, **Aishrila Mazumder**, Arnab Chakraborty, Narendra Nath Patra, Rajesh Mondal, Samir Choudhuri, *Detecting galaxies in a large H*I *spectral cube*, Journal of Astrophysics and Astronomy, Volume 43, Issue 2, 2022, doi: 10.1007/s12036-022-09880-1 [arXiv: 2211.02041]

3. **Aishrila Mazumder**, Abhirup Datta, Mayuri Sathyanarayana Rao, Arnab Chakraborty, Saurabh Singh , Anshuman Tripathi, Madhurima Choudhury, *Synthetic Observations with the Square Kilometre Array (SKA) - development towards an end-to-end pipeline* Journal of Astrophysics and Astronomy, Volume 44, Issue 1, 2023, doi: 10.1007/s12036-022-09906-8 [arXiv: 2211.04302]

Table of Contents

List of Figures

List of Tables

List of abbreviations

BB	Big Bang
HI	Neutral Hydrogen
CMB	Cosmic Microwave Background
CD	Cosmic Dawn
EoR	Epoch of Reionization
DLAs	Damped Lyman α Systems
DGSE	Diffuse Galactic Synchrotron Radiation
APS	Angular Power Spectrum
EoR	Epoch of Reionization
GMRT	Giant Metrewave Radio Telescope
LOFAR	Low Frequency Array
MWA	Murchison Widefield Array
HERA	Hydrogen Epoch of ReionizationArray
SKA	Square Kilometer Array
PAPER	Donald C. Backer Precision Array to Probe the Epoch of Reionization
PS	Power Spectrum

Chapter 1

Introduction

1.1 The Cosmic Timeline

The standard cosmological model states that the Universe started with the Big Bang (BB), about 13.8 billion years ago. Starting with a hot and dense phase with infinite density and temperature, the Universe cooled and expanded through various stages. The fundamental constituents of the present Universe - protons, neutrons, electrons, neutrinos, and photons were formed. After about 100 seconds from the BB, with a temperature of 10^{10} K, the temperature was cold enough for the fusion of protons and neutrons into light elements - deuterium and Helium and their isotopes. This process is known as BB Nucleosynthesis.

At these early stages, the photons were scattered at a very high rate by electrons and other particles like protons, thereby unable to travel long distances, making the Universe opaque. The temperature began to decrease due to expansion of the Universe. Consequently, the scattering rate of photons by free electrons (and other light elements) fell below the expansion rate. Eventually, the temperature fell to

3000 K after 38000 years from the BB. Thus hydrogen atoms were formed from recombination of electrons and protons. Due to the absence of free electrons, the protons were able to stream freely, thus decoupling from baryons and turning the Universe transparent. This radiation is the Cosmic Microwave Background (CMB) [6, 7]. It is a relic from the surface of last scattering (at redshift $z \sim$1100), with the present-day temperature of $\sim$2.7 K, [6]. The expansion of the Universe caused this radiation to get "redshifted", thus shifting to the microwave wavelengths of the electromagnetic spectrum. Observation of the CMB spectrum was done by different satellite missions, with the Planck mission [1] providing the most sensitive observations to date [8].

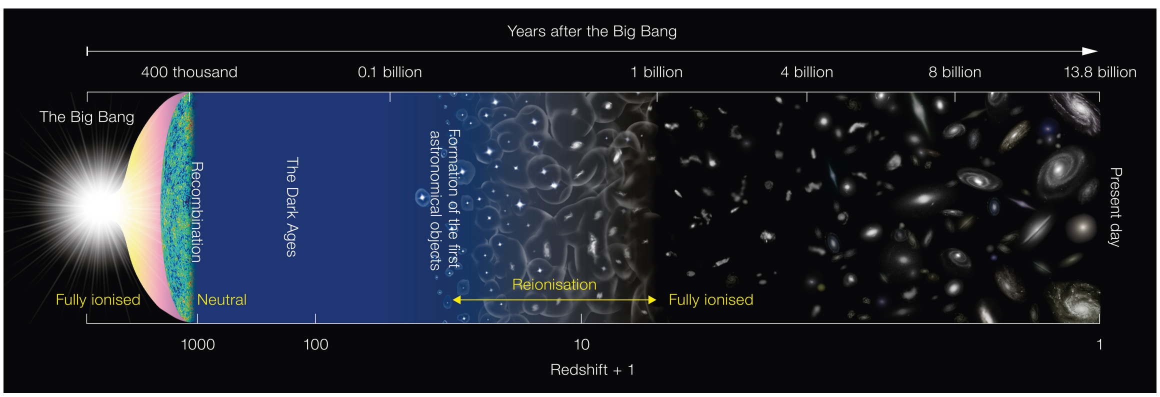

Figure 1.1: Schematic diagram for the evolution history of the Universe following the currently accepted scenario for structure formation [9].

At the end of the recombination era, the absence of significant radiating sources resulted in the Cosmic Dark Ages, which continued till $z \sim 30$ and was characterized mainly by the absence of conventional radiation sources. During this period, baryons collapsed into potential wells formed by collapsed dark matter haloes. These gave rise to the first structures i.e. the first generation of stars and galaxies - thus ending the Dark Ages and marking the beginning of the Cosmic Dawn (CD). The Cosmic Dawn approximately covers a redshift range $30 < z < 12$. The initial stellar and

[1] https://www.cosmos.esa.int/web/planck

galactic population (as well as early quasars and dark matter decay/annihilation) emitted X-Ray and ultraviolet (UV) radiation. The UV photons with energy > 13.6 eV were able to ionize the intergalactic medium (IGM), majorly neutral hydrogen (HI) [10]. This marked the Epoch of Reionization (EoR). Observational constraints from high redshift quasar absorption spectra [11–13, 15] and Thompson scattering optical depth of the CMB [25] suggest a redshift range of $12 \gtrsim z \gtrsim 6$ for EoR.

Around $z \lesssim 6$, the post-EoR era, HI was found mostly inside the galaxies, self-shielded from the ionizing radiation. Quasar absorption spectra from these redshifts indicate $\mathrm{x_{HI}} < 10^{-4}$ in the IGM [11, 15]. The regions with an abundance of HI in the post-EoR era are correlated with matter over-density. Thus, HI distribution in this period is connected with the matter distribution in the Universe [17, 18]. These HI gas clouds appear in Damped Lyα systems or DLAs in the background quasars' absorption spectrum. DLAs have very high HI column densities ($\mathrm{N}_{HI} > 10^{20.3}$ cm^{-2}), and contain the bulk of HI budget after reionization [19, 20].

1.2 Reionizing the Universe

The EoR constitutes the last phase transition in the history of the Universe where the neutral IGM was converted into one that is almost fully ionized. When the first luminous sources start appearing around z∼30, during the cosmic dawn, they emitted a considerable amount of UV photons, X-rays, and Lyα photons. These UV photons reionized the HI in the IGM and created HII bubbles called Strömgen Spheres around the radiating sources. The X-rays caused the HI in the IGM to heat up. The Lyα photons were responsible for the Wouthuysen-Field effect, i.e., coupling the kinetic temperature with the HI hyperfine transition line. The Strömgen Spheres grew gradually with time and eventually merged, creating an ionized Universe (see for example [40, 94–96]).

With improvement in the telescope sensitivity and observation techniques, as well as theoretical modeling and simulations, the understanding of this epoch has vastly improved. The CMB experiments like WMAP and Planck provided detailed probes of the epoch of recombination. At the same time, large-scale galaxy surveys like 2dF and SDSS explored the Universe at lower redshifts to understand formation and evolution of galaxies. However, the interim period, particularly the CD/EoR, remains unexplored with the same precision. Precision observations for these epochs are expected to provide information on the reionization topology, evolution and duration, nature of the first population of luminous objects, the hierarchy of structure formation, dynamics and evolution of large-scale structures, etc. [44]. A few probes provide indirect answers to the unknowns of the CD/EoR. They are discussed briefly in the following subsection.

1.3 Observations of the CD/EoR

Observations using different direct and indirect methods and their comparison with theoretical predictions are required to understand the progress of the different stages of CD/EoR. The main observables of this epoch are [22] Gunn-Peterson Effect, the redshifted 21-cm radiation, CMB anisotropy measurements, etc.

A. Observational Constraints from Indirect Probes

Gunn-Peterson Effect:

Quasar (or quasi-stellar radio source) are some of the brightest sources in the Universe. It is essentially an AGN with a supermassive black hole (mass of the order 10^6 - 10^9 M$\odot$) surrounded by an accretion disk. Quasars have reasonably smooth spectra irrespective of the line of sight they are observed along. Variations in the characteristic spectrum of quasars can be used to study the IGM between them and the observer. The resonance absorption of the Lyα photons in the spectrum of distant quasars is called the Gunn-Peterson (GP) Effect [11]. The absorption spectrum of objects beyond reionization will show a distinct trough-like feature with almost zero transmitted flux since the neutral IGM will absorb the emissions blueward of the Lyα line. Thus, observation of the GP trough can directly constrain the evolution of HI fraction in the IGM and its ionization state. GP trough was first detected in the spectrum of a quasar at z=6.28 [14]. [15] studied a sample consisting of 19 SDSS detected quasars (5.74 <z <6.42) from the SDSS survey and showed that the Universe was almost wholly ionized by z$\sim$6. Latest constraints on the neutral fraction from the high redshift quasar J1007+2115 (at z=7.5) [16] shows that the reionization process was ongoing at $z \gtrsim 7$. The volume averaged neutral fraction calculated along the sight-line of this system shows $\langle x_{HI} \rangle = 0.39^{+0.22}_{-0.13}$. The quasar also shows a weak damping wing, which provides observational evidence for patchy reionization [16]. The spectrum from J1007+2115 is shown in Figure 1.2.

CMB Anisotropies :

The other observational constraint for the reionization process and its progress comes from the CMB temperature and polarization anisotropies [23–25]. The majority of CMB photons scattered last at z$\sim$1100. However, less than 10% of such photons scattered again during the EoR. Measuring this signal can help constrain the mean redshift and boundary redshifts for EoR. Measurement of the average Thompson scattering optical depth from CMB provides a model-independent constraint on reionization. The scattering introduces two effects:

(a) The free electrons generated by the process of reionization suppress the anisotropies in the CMB and generate polarization anisotropies at low spherical harmonic multipoles.

(b) Secondary anisotropies are introduced by scattering in addition to the primary anisotropies of the background (CMB) radiation field [23, 24].

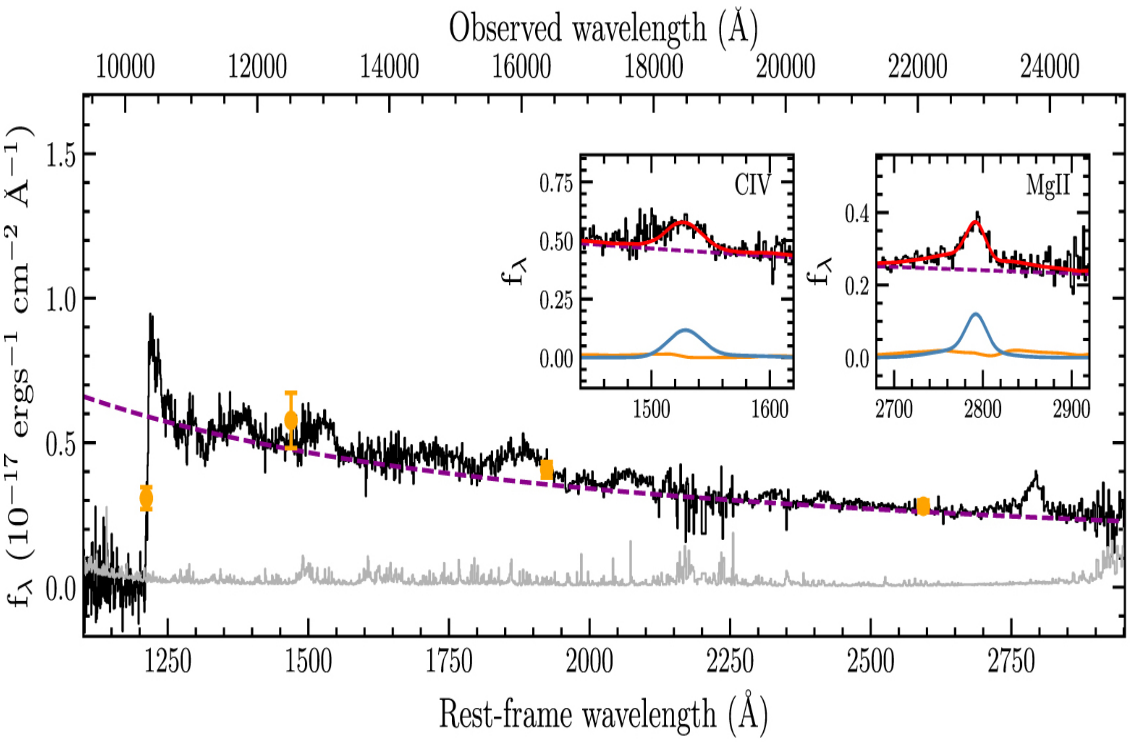

Figure 1.2: Combined spectra of J1007+2115 using data from Gemini/Gemini Near-Infrared Spectrograph, Keck/Near-Infrared Echellette Spectrometer and UKIRT Hemisphere Survey. Plots in the inset are the fits to the C IV and Mg II lines. Reproduced from [16].

The secondary anisotropies have little impact on anisotropies in CMB temperature. However, polarization angular power spectrum is significantly impacted. It creates a 'reionization bump' on large angular scales (roughly larger than the horizon at the epoch of reionization) as seen in Figure 1.3.

CD/EoR can also be probed using the kinetic Sunyaev-Zeldovich (kSZ) effect. The kSZ effect causes the CMB photons to scatter off electrons having bulk velocity with respect to the comoving frame. Unlike the thermal SZ effect, it does not cause spectral distortion of CMB but either enhances or decreases the temperature. The kSZ effect will cause secondary anisotropies in the CMB at small scales for the patchy distribution of ionized regions. These anisotropies are a probe for the reion-

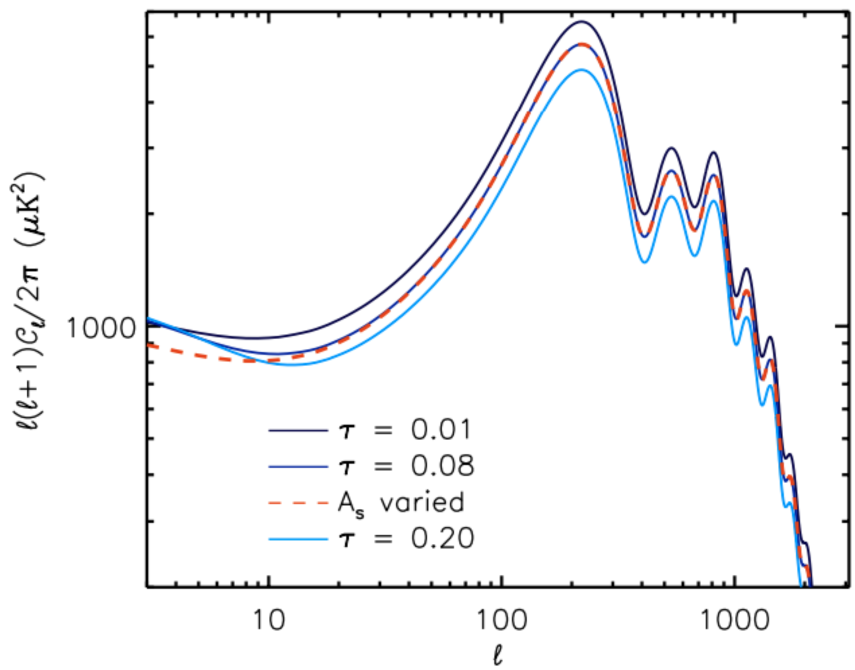

Figure 1.3: E-mode polarization power spectra of CMB. The small ℓ scale 'reionization bump' is the result of CMB photons being Thompson scattered by free electrons present in the IGM by reionization. Reproduced from [26].

ization process [27–30]. Using the latest data from the Planck Mission, the upper end of the reionization has been constrained to z_{re} <9, and a mid-point redshift of $z = 7.68 \pm 0.79$ [31].

Other Probes:

Other observables to probe the CD/EoR are- Lyα damping wing of neutral IGM, the evolution of metal absorption systems, gamma-ray bursts, luminosity function, and line profiles of Lyα galaxies, etc.

B. Redshifted 21-cm Radiation- the Direct Probe for CD/EoR:

One of the most reliable observational probes for tracing the CD/EoR is the redshifted 21-cm signal. It is a direct probe for the HI distribution in the Universe, both spatially and temporally. In the ground state, HI consists of a proton and an electron with their spins coupled and may be aligned either in parallel or anti-parallel. The transition between these two states is a hyperfine transition with either absorption or emission of a photon. However, this hyperfine transition is a forbidden transition with spontaneous emission coefficient $A_{10} = 2.85 \times 10^{-15} s^{-1}$ (or a spontaneous emission time of 11 million years for a 21-cm photon from HI) [32]. Despite the long emission time, since ∼75% of the baryons in the Universe is HI, the 21-cm signal is abundantly present. These photons, emitted at a rest-frame frequency of 1420 MHz (or wavelength ∼21-cm), get redshifted as the Universe expands. They can be observed in the frequency range of 50-200 MHz using modern radio telescopes. Due to such a redshift evolution, it is used to map the Universe in 3D. It is sensitive to the underlying astrophysics and cosmology governing the behavior of IGM and is thus an excellent probe for the underlying processes during CD/EoR. Observations of the 21-cm radiation from the early Universe can be done using both single-dish telescopes and radio interferometers. The former instruments are sensitive to the overall evolution of the signal ("global signal") as a function of time, while the latter probes the statistical nature of the signal fluctuations over space and time (power spectrum, bispectrum, etc). Interferometers can also (in principle) perform 21-cm tomography for probing the IGM in three dimensions. Since this thesis is dedicated to studying observational aspects of the redshifted 21-cm signal using radio interferometers, the physics of the 21-cm line is discussed in detail in Chapter 2.

1.3.1 Challenges of Detecting 21-cm Signal fom CD/EoR

The actual low-frequency observations of the redshifted 21-cm signal from using radio interferometers are challenging. Theoretical models predict that the differential brightness temperature is a few hundred mK at most (for instance, see [43]). The contaminants of such a weak target are numerous and vary in nature and complexity. Two significant factors that can adversely affect observations targeting detection of HI from the early Universe are instrumental systematics and astrophysical foregrounds.

Due to an interferometer's inherent complexity, certain instrumental properties are correctly determined, and many others are undetermined. Also, "optimally" dealing with known systematics can lead to a loss of sensitivity. The unknown systematics are even more harmful since the way they bias the analyses are unknown. Nevertheless, recent research into systematics has focused on the effects of instrumental gains, primary beam models, thermal noise, polarization leakage, and direction-dependent ionospheric effects. These may vary substantially depending on the instrument used, but the effects on the target signal recovery are similar. Thus, dedicated and systematic studies of their effects are required.

The astrophysical foregrounds, constituting any extraterrestrial radio emissions brighter than the target signal, are the most potent contaminant of the target signal. Foregrounds constitute DGSE, free-free emission (galactic and extragalactic), and emission from star-forming galaxies and active galactic nuclei. The foreground contamination is generally removed by considering them spectrally smooth (see, for example, [207]). Thus, foreground mitigation techniques employ the spectral smoothness assumption and use parametrized or non-parametrized strategies to model out foregrounds. The foregrounds are also "avoided" in some cases. The avoidance technique uses parts of the Fourier plane that are free from foreground contamination for detecting the 21-cm signal. While each process has pros and cons, a general caveat is that the presence of non-smooth foreground components will hamper the signal extraction. Even for smooth foregrounds, this may be caused by the instrumental feature introducing some structures. There are also chances of signal removal along with foregrounds. Since foreground is a major contaminant, handling foregrounds is one of the biggest challenges for sensitive radio observations targeting the HI 21-cm signal from CD/EoR.

The study of foregrounds was done in the extragalactic deep field in [5]. The work presented in [5] used deep observations, but the area of the sky covered was $\sim 1.5^{\circ}$. While the deeper flux density limits probed fainter sources, the smaller area resulted in a lesser number of sources for statistical studies. Robust source statistics is a prerequisite for efficient foreground modelling and requires a large number of samples in the observing volume.

1.4 Objectives of the Thesis

This thesis aims to study the challenges posed by foregrounds and systematics affecting interferometric observations of the 21-cm signal. Foreground studies use observations from the GMRT, while studies of instrumental systematics use simulations. In [5], the extragalactic deep field ELAIS-N1 was studied using the Band-3 (with a central frequency of 400 MHz) of the uGMRT. For that study, the area covered was $\sim 1.5 o^2$, but the sensitivity was 15μJy. This work uses a larger field of view ($6^{\circ} \times 6^{\circ}$) (compared to [5]) to probe a larger number of sources at 325 MHz for another well-known extragalactic deep field called the Lockman Hole. This frequency band is near band-2 relevant to CD/EoR science but has moderate systematics, and is also one of the most sensitive bands of the GMRT. Foreground studies using this particular data set are important because of both the large area covered and the moderate sensitivity of 50μJy achieved. Deep fields like Lockman Hole, ELAIS-N1, Boötes, COSMOS, XMM-LSS are important because these are some regions of the sky that have multi-frequency coverage and the observed depths and areas are such that one can study the extragalactic sky with sufficient sensitivity without compromising on the number of samples (unlike narrow area deep surveys or shallow deep field surveys). Such studies are also relevant for foreground modeling since they provide insights into their variation with frequency, flux, and line-of-sight.

The systematics themselves also constitute a major issue, and characterizing those becomes important. Systematics arising from an imperfect instrument as well as limitations of the algorithms used for data analysis cause the foregrounds present in the data to affect the 21-cm signal detection. Systematics also vary from instrument to instrument and studying their effect requires simulations using the parameters relevant to the telescope in question. Thus, simulations have been used to study different systematics and their effects on the 21-cm power spectra. An end-to-end pipeline has also been developed to generate mock observations under realistic observing conditions. The pipeline can be used to incorporate different systematic effects and study their imapct on the extraction of the target signal. Such an effort is relevant for observations and data handling of presently operating interferometers like Murchison Widefield Array (MWA) [2] and the Hydrogen Epoch of Reionization Array [3]. The upcoming Square Kilometer Array (SKA) [4], with unprecedented sensitivity, will also target such sensitive observations. Thus, the present thesis along with more such dedicated efforts would be essential in order to utilize its capacity to detect the target signal.

1.5 Contributions of the Thesis

This thesis consists of seven chapters. The main contributions of this thesis have been detailed in Chapters 4, 5, and 6. Chapters 4 & 5 deal with foregrounds, while Chapter 6 deals with instrumental systematics. Chapter 1 provides a brief motivation for the thesis, followed by a description of the physics of the 21-cm signal in Chapter 2. Chapter 3 discusses the major challenges faced in observing the 21-cm signal and the recent progress in dealing with those challenges. The conclusions derived from this research and the future scope have been discussed in

[2] https://www.mwatelescope.org/
[3] https://reionization.org/
[4] https://www.skatelescope.org/

Chapter 7.

1.5.1 Chapter-wise Organisation of the Thesis

The main findings of this thesis are:

- **Chapter 4:** In this Chapter, observations from the Giant Metrewave Radio Telescope have been used to study foregrounds present in a well-known extragalactic deep field called the Lockman Hole, studied for the first time at 325 MHz [1]. Using the 325 MHz data, a point source catalogue has been generated for a $6^\circ \times 6^\circ$ mosaiced field of view. This produced 6186 sources with excellent flux and astrometric accuracy. The Euclidean normalized source counts have been determined, which agree with previous observations and show the signature flattening at the low flux end - evidence for the domination of star-forming galaxies. The APS of the DGSE has also been calculated. Due to the large area covered, three different galactic latitudes have been selected to obtain the angular power spectrum, and a power law has been fitted to the spectrum. The synchrotron power shows amplitude between 1 to 100 mK^2, with a power-law index between 2.15-3.15 (consistent with previous estimates). While it is much lower than that near the galactic plane, it is still high enough to obscure the HI 21-cm signal from EoR. Thus, this work shows that the diffuse galactic synchrotron radiation needs to be modeled and removed carefully for CD/EoR data sets, even far away from the galactic plane.

- **Chapter 5:** In this Chapter, the 325 MHz catalogue of the Lockman Hole has been used to study the clustering of extragalactic point sources [3]. The distribution of sources across the sky follows the underlying dark matter density field and shows clustering over Poissonian distribution. Clustering provides a probe for observationally estimating the underlying cosmology of large-scale structures in the Universe. It also becomes significant for foreground modeling at angles $\gtrsim$1' for flux densities $\gtrsim$1 mJy. The two-point correlation function of the sources has been determined over both spatial and angular scales. The sources have been classified into active galactic nuclei (AGNs) and star-forming galaxies (SFGs) based on their radio luminosities. The angular clustering amplitude, spatial clustering length, and bias parameter all agree well with previous observations for both AGNs and SFGs. The clustering length for AGNs in this work is at z_{median} $\approx$1.02 is $8.30^{+0.96}_{-0.91}$ Mpc h^{-1}, while that for SFGs is $3.22^{+0.34}_{-0.32}$ Mpc h^{-1} at z_{median} $\approx$0.2. The respective bias parameters are at $3.74^{+0.39}_{-0.36}$ and $1.06^{+0.10}_{-0.10}$ for AGNs and SFGs. The clustering length and bias parameters are greater for AGNs than SFGs, which is evidence of AGNs being hosted by more massive dark matter haloes. Comparison of halo mass for SFGs from this work (as well as previous observations) with SKADS show that the latter under-predicts the masses. This indicates a requirement for more observations to better model the extragalactic sky. The data covers sources with a moderate flux density limit and a moderate survey area, bridging the gap between wide-area shallow surveys and deeper small area surveys.

- **Chapter 6:** This Chapter studies the characterization of systematic errors in EoR observations. The main aim is to develop an end-to-end simulation pipeline that can generate mock data of interferometric observations for different conditions [2]. Such a pipeline is relevant since the upcoming SKA will take thousands of hours of observations to detect the HI 21-cm signal. The data analysis algorithms being tested to process the enormous volumes of data from SKA need to account for the tolerance of different systematics. Thus, a detailed study into their effects and level of accepted tolerance needs to be done. For this work, two systematic errors - calibration and position errors have been studied using four different telescope layouts - SKA-1 Low, HERA, MWA Phase 1, and MWA Phase 2 (256 tile configuration). Mock observations are done with a theoretical model of the 21-cm signal and point source foregrounds. Their effects have been analyzed for the image plane and the power spectrum estimation. The optimal tolerance for calibration errors was found to be $\sim$0.1%. For position errors, an offset of $\sim$5" was optimal. This shows that calibration algorithms must be highly accurate to recover the 21-cm signal from EoR data sets. The position error tolerance is achievable even with the currently operational telescopes and should not be a hindrance if other sources of errors are taken into account properly. This is the first detailed study into the effects of different sources of errors across different array configurations and under a set of realistic observing scenarios. The pipeline developed for this work can be extended to study the effect of primary beam errors, ionosphere, etc.

Chapter 2

Observing the Early Universe: The Redshifted 21-cm Signal

2.1 The Redshifted 21-cm Signal: Fundamentals

The hyperfine transition frequency of the 21-cm HI emission was first calculated theoretically by Hendrik van de Hulst in 1944. Its detection was followed in 1951 by [33,34]. Theorists started studying reionization and the initial population of ionizing sources, and [38] showed how the luminous sources affected the 21-cm signal and determined the spin temperature of the IGM. With the development of telescopes with enhanced sensitivities, the high redshift Universe ($z \gtrsim 6$) became more observationally accessible. Thus began the hunt for the cosmological 21-cm signal from the early Universe.

In order to use the redshifted HI 21-cm signal as a probe of the early Universe, it is essential to understand its characteristics. The remaining part of this chapter is thus dedicated to understanding it- basic physics of the signal (Section 2.1), its time evolution (Section 2.2) and its observable signatures (Section 2.3) and advances in its observation (Section 2.4).

An electron and a proton in the $1s$ level constitutes the ground state of HI atom. The coupled spins of the electron and proton may be parallelly or anti-parallelly aligned. The parallel alignment corresponds to a higher energy state, with an energy gap of 5.87 μeV. Figure 2.1 shows an energy level diagram of a neutral hydrogen atom, with the Lyman-α and spin-flip transitions shown. This hyperfine transition is mediated by the absorption or emission of a photon with a rest wavelength of 21-cm (or a rest frequency of 1420 MHz). However, this hyperfine transition is a forbidden transition with spontaneous emission coefficient $A_{10} = 2.85 \times 10^{-15}\mathrm{s}^{-1}$ (or a spontaneous emission time of 11 million years for a 21-cm photon from HI) [32]. However, despite the long emission time, since ~75% of the baryons in the Universe is HI, the 21-cm signal is abundantly present.

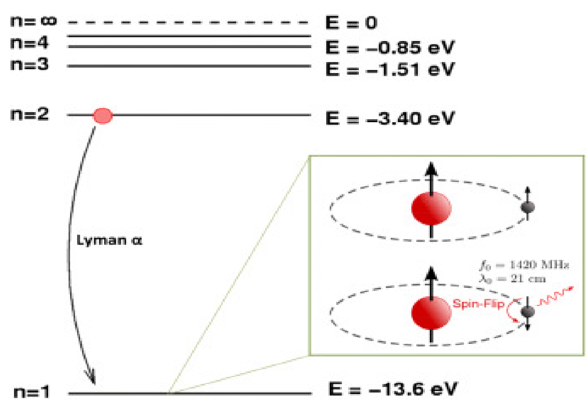

Figure 2.1: The energy level diagram of neutral hydrogen atom. The Lyα line corresponds to n = 2 to n = 1 (ground state) transition and has an energy 10.2 eV. The inset on the right shows the 21-cm spin-flip transition, corresponding to a hyperfine splitting of the n=1 state with energy 5.87 $\times 10^{-6}$eV. Reproduced from [41].

The fundamental quantity of interest is the specific intensity or brightness I_ν, the energy per unit time carried by the ray per unit area per unit frequency per unit solid angle, and expressed in $\mathrm{cm}^{-1}\mathrm{s}^{-1}\mathrm{Hz}^{-1}\mathrm{sr}^{-1}$. For a thermally emitting object at temperature T, the radiative transfer equation is given by [39]:

$$\frac{dI_\nu}{d\tau_\nu} = -I_\nu + B_\nu(\mathrm{T}) \tag{2.1}$$

where τ_ν is the optical depth of the medium for absorption and B_ν is the Planck function.

I_ν is also expressed in terms of the equivalent *brightness temperature* $\mathrm{T_b}(\nu)$, which is the temperature for an equivalent black-body radiator with spectrum B_ν, such that $I_\nu = B_\nu(\mathrm{T_b})$. Since the frequencies relevant to the 21-cm line lies on the low frequency (or long wavelength) side of the spectrum, Rayleigh–Jeans approximation gives the relation between $\mathrm{T_b}(\nu)$ and I_ν as:

$$\mathrm{T_b}(\nu) \approx \frac{I_\nu c^2}{2k_B\nu^2} \tag{2.2}$$

where, c and k_B are the speed of light and the Boltzmann constant respectively. Expressing Equation 2.1 in therms of brightness temperature (Equation 2.2), it becomes :

$$\frac{\mathrm{dT_b}}{d\tau_\nu} = -\mathrm{T_b} + \mathrm{T_{CMB}} \tag{2.3}$$

where $\mathrm{T_{CMB}}$ is the background CMB temperature. Solution to Equation 2.3 gives the temperature of the emergent radiation as :

$$\mathrm{T_b}(\nu) = T_s(1 - e^{-\tau_\nu}) + \mathrm{T_{CMB}}(\nu)e^{-\tau_\nu} \tag{2.4}$$

where T_s is the spin temperature of the intervening cloud. It is the effective excitation temperature for the hyperfine transition and dictates the properties of the 21-cm line (in terms of emission and absorption properties). Using the Boltzmann factor for the relative occupancy of the triplet with respect to the singlet state for the ground state of the hydrogen atom, T_s is given by [32]:

$$\frac{n_1}{n_0} = 3\exp\left\{\frac{-T_*}{T_s}\right\} \tag{2.5}$$

where n_0 and n_1 are the occupancy of the singlet and triplet states and $\mathrm{T_*}$=0.0682 K (from $h\nu_{21cm} = 2k_BT_* = 5.87\times10^{-6}$eV). The factor 3 is the degeneracy factor for the triplet state. T_s is the ratio between two hyperfine states, and controls the intensity of the emergent radiation from an HI cloud. It is seen from Equation 2.4 that the brightness temperature depends on the optical depth τ_ν and spin temperature T_s, and has to be determined for the optically thin regime (which is relevant to the present case). If T_s=$\mathrm{T_{CMB}}$, the brightness temperature is simply the CMB temperature, revealing nothing about the intervening HI cloud.

Due to the expansion of the Universe, the wavelength λ of a radiation coming from a cosmological redshift z expands to $\lambda(1+z)$. Hence, they are observed at a longer wavelength than the original. The optical depth for 21-cm radiation for an HI cloud with uniform T_s observed at 21(1+z) cm is given by [41]:

$$\tau(z) = \frac{3c\lambda_{21}^2 h_p A_{10} n_{HI}}{32\pi k_B T_s (1+z)(\partial v_r/\partial r)}$$

where h_p is the Planck's constant, $A_{10} = 2.85 \times 10^{-15} s^{-1}$ is the spontaneous decay rate of hyperfine transitions, $\partial v_r/\partial r$ is the radial velocity along line-of sight, where v_r is the physical radial velocity and r is the comoving distance and n_{HI} is the number density of hydrogen atoms. In completely neutral and homogeneous Universe, $n_{HI} = \bar{n}_H(z)$ and $\partial v_r/\partial r = H(z)/(1+z)$ in terms of the Hubble parameter H.

The background CMB temperature provides a contrast against which brightness temperature in either absorption or emission for 21-cm HI signal is observed. Thus, the differential brightness temperature $\delta\mathrm{T_B}(\nu)$ $(\mathrm{T_b} - \mathrm{T_{CMB}})$ is the observable quantity, defined as [32, 40]:

$$\begin{aligned}\delta\mathrm{T_b}(\nu) &= \frac{(T_s - \mathrm{T_{CMB}})}{(1+z)} e^{-\tau} \approx \frac{(T_s - \mathrm{T_{CMB}})}{(1+z)} \tau \\ &\approx 26.8\mathrm{mK} \Big(\frac{\Omega_b h}{0.0327}\Big) \Big(\frac{\Omega_m}{0.307}\Big)^{-1/2} \Big(\frac{1+z}{10}\Big)^{1/2} \Big(\frac{T_s - \mathrm{T_{CMB}}}{T_s}\Big)\end{aligned} \tag{2.6}$$

where the optical depth $\tau \ll 1$. In the Rayleigh-Jeans limit of the Planck function, the measured intensity in equivalent temperature units is the differential brightness temperature (Equation 2.2). The CMB spectrum peaks at $\lambda \sim 2$ mm making it fall in the Rayleigh-Jeans limit for 21-cm cosmology, with relevant wavelengths observed at z=0 being three orders of magnitude larger [41].

Equation 2.6 shows that the 21-cm signal can only be detected when the T_s of the gas differs from CMB temperature. This is expected since $T_s = \mathrm{T_{CMB}}$ implies thermal equilibrium between the background and the gas. Hence, there would be neither absorption nor emission above the background. Equation 2.6 also implies that T_s determines whether the signal is seen in absorption or emission against the CMB.

2.2 Evolution of the 21-cm Signal with Redshift

In Section 2.1, it was seen that T_s is critical in determining the observability of the 21-cm signal. T_s itself depends on the interplay of several astrophysical factors. Thus, it is imperative to understand how each of these factors evolves with redshift to understand their signatures in $\delta\mathrm{T_b}$. The interplay of three main processes determines the evolution of T_s - absorption and stimulated emission of CMB photons, collision of HI with other particles and the Wouthuysen-Field effect (scattering of Lyman-α (Lyα) photons through absorption and re-emission). The spin temperature of HI for the hyperfine transition can thus be written as [32]:

$$T_s = \frac{\mathrm{T_{CMB}} + y_k T_k + y_\alpha T_\alpha}{1 + y_k + y_\alpha} \tag{2.7}$$

The coupling term y_k and y_α arise due to collisional excitation and is the effect on 21-cm transition due to collisional excitation, and Wouthuysen-Field effect. y_α is for excitation due to the Wouthuysen-Field effect. It should also be mentioned for almost all cases of interest, $T_s = T_\alpha$ can be assumed [32, 40, 43].

2.3 Detection of the 21-cm Signal

The redshifted 21-cm signal from the CD/EoR can be detected through three detectable observational signatures - Global Signal, Statistical Detection, Tomography.

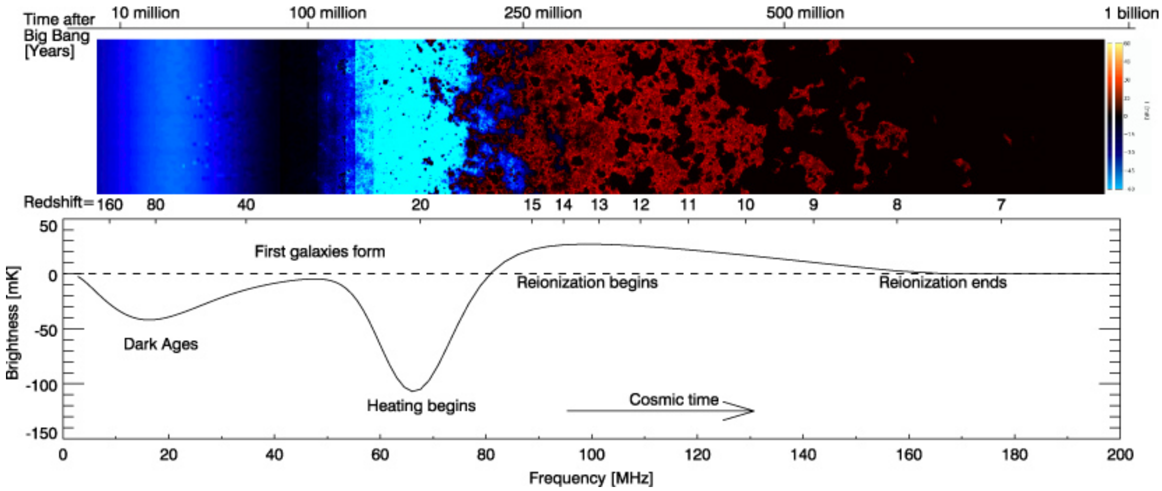

Figure 2.2: Redshift (or frequency) evolution of the 21-cm signal. **(top panel)**: slice through a simulation showing the evolution of the brightness temperature of the signal; **(bottom panel)** The 21-cm global signal as predicted by theoretical models with parameters described in [44]. The model used for describing the evolution of the signal determines its exact shape. Reproduced from [44].

2.3.1 The Global 21-cm Signature

The thermal state of the IGM as a function of redshift (or frequency) can be probed by observing the "global" 21-cm signal. The global signal is all sky averaged signal measured through the differential brightness temperature (Equation 2.6). Plugging in the evolution of T_s in Equation 2.6 gives the evolution of $\delta \mathrm{T_b}$. Figure 2.2 shows the evolution of the differential brightness temperature as a function of frequency (hence redshift). In the spectrum, the 'Dark Ages' and 'Cosmic Dawn' show two sharp absorption features. The period of reionization at $z \gtrsim 15$ gives the signal in emission. The HI in the IGM becomes completely ionized by $z \sim 6$, and the brightness temperature becomes zero. It must be emphasized here that the signal depicted in Figure 2.2 is for a particular set of parameters modeled in [44]. Its shape depends on the models' astrophysical parameters, and changing those will change the shape.

2.3.2 Statistics of the 21-cm Signal

Power Spectrum:

Another method to probe the 21-cm signal is statistically through its power spectrum (PS). This is done by looking into correlations in the signal at particular length scales. The ultimate target for 21-cm observations is 21-cm tomography- mapping the spatial distribution of HI as a function of redshift using the cosmological signal. Current theoretical estimates state the sensitivity required to be of mK order on angular scales of a few arcminutes for such observations.

Radio interferometers observe fluctuations in the 21-cm brightness temperature through its power spectrum $P(\mathbf{k}, z)$. Assuming the signal fluctuation field is statistically homogeneous and isotropic, we can write:

$$\langle \Delta\widetilde{\mathrm{T}_\mathrm{b}}(\mathbf{k}, z)\Delta\widetilde{\mathrm{T}_\mathrm{b}}^*(\mathbf{k}', z)\rangle = (2\pi)^3\delta^3(\mathbf{k}-\mathbf{k}')P(\mathbf{k}) \tag{2.8}$$

where, $\Delta\widetilde{\mathrm{T}_\mathrm{b}}(\mathbf{k}, z)$ is the Fourier transform of the differential brightness temperature $\delta\mathrm{T}_\mathrm{b}(\mathbf{r}, z)$ and $\delta^3(\mathbf{k}-\mathbf{k}')$ is the Dirac delta function. Since the 21-cm signal is isotropic, $P(\mathbf{k})$ reduces to $P(k)$.

The advantage of measuring the 21-cm PS over the global signal is that it provides information on spatial fluctuations. However, since the 21-cm signal is highly non-Gaussian, it does not encode all the available spatial information. Nevertheless, the PS is still rich in information and, observationally, is considerably easier to measure than higher-order correlations.

Decomposing the 21-cm PS provides a wealth of astrophysical information. The brightness temperature fluctuations are driven by several factors- density perturbations, ionization fraction, Lyα coupling, peculiar velocity, etc. [57]. Measurement of the PS measures the combined power from each field and their cross power. So

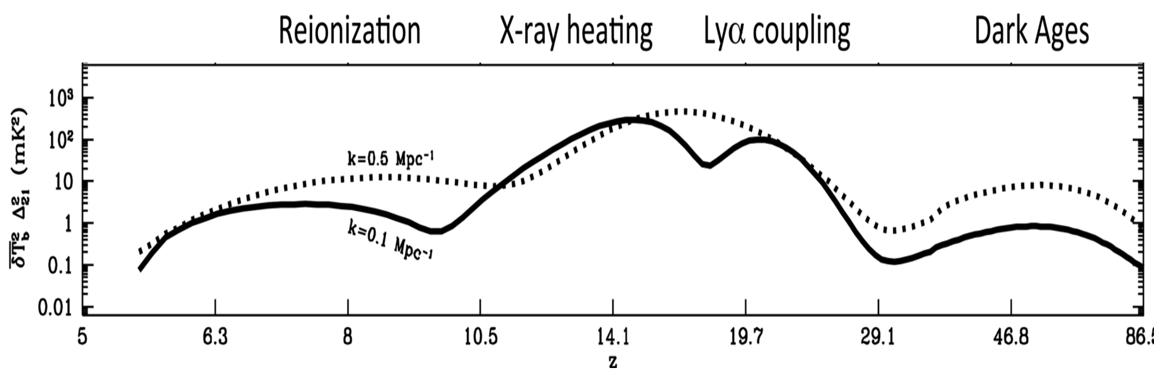

Figure 2.3: Evolution of the 21-cm power spectrum amplitude for two different Fourier modes, k =0.1 Mpc^{-1} (solid) and k = 0.5 Mpc^{-1} (dashed). Peaks in the amplitude corresponds to different cosmic milestones. Reproduced from [57].

measurements of the 21-cm PS across cosmic time will be sensitive to the dominance of each component as a function of time and spatial scales on which the signal is the strongest. This has been depicted in Figure 2.3. However, the standard method uses a range of spatial scales (rather than one Fourier mode) to recover astrophysical information from both small and large-scale physical processes.

Other Statistical Methods:

It has been also been theorized that the 21-cm signal is not a Gaussian random field, since ionized bubbles cause spatial anisotropy. In such a case, PS is not sufficient for complete statistical characterization of the signal and higher order statistics are required [44]. One point statistics like skewness and curtosis, as well three-point correlation function or the bispectrum can also be used [49–51] to study the signal statistically. However, studies into the practicalities of detecting these statistics using real interferometers is still ongoing, thus currently instruments are aiming for power spectrum detection.

2.3.3 Tomography

Tomographic imaging of the IGM using the redshifted 21-cm signal during CD/EoR is projected to be possible with SKA-1 Low [47]. Since this signal can be imaged in many different frequency bins, an cube can be made out of these images which could be used for performing a 3D analysis of the IGM at these early epochs [38]. Imaging becomes possible when the signal-to-noise for a certain spatial (or spectral) resolution becomes greater than 1. Thus, it has been projected that even instruments like LOFAR can produce tomographic images, albeit with poor resolution [48]. Tomographic imaging is essential for study of ionized regions and their connection with galaxy populations, distribution and topology of the ionized bubbles, deriving quasar properties, etc. Such information are not accessible via statistics, but their understanding important for inferring the processes at play during CD/EoR. Thus, IGM tomography with the SKA-1 Low will be interesting to explore once the telescope becomes operational.

Observations till date have mainly targeted the global signal and PS detections, with higher order statistics and tomogrphy still in the development stages. The present thesis focuses on the observational aspects of the 21-cm PS. So the remainder of this chapter will mainly discuss the different observational constraints on the 21-cm PS, with a brief mention of the progress in the global signal detection.

2.4 Observing the Redshifted 21-cm Signal

There has been a massive advancement in astronomical instrumentation over the past few decades. This has made observing the faint HI 21-cm signal from the CD/EoR an achievable target. The 21-cm signal can be studied using both the global signal and the power spectrum. Single dish telescopes (or total power radiometers) are used for global signal searches while interferometers are used for PS detections.

2.4.1 Observations of Global Signal

Global CD/EoR signal detection experiments have lead to constraints on the astrophysics of the processes involved. There are several instruments world-wide that are dedicated in searching for the global signature of the redshifted 21-cm signal. This includes instruments like Experiment to Detect Global Epoch of Reionization Signature (EDGES) [52], Broadband Instrument for Global HydrOgen ReioNisation Signal (BIGHORNS [339]), Sonda Cosmológica de las Islas para la Detección de Hidrógeno Neutro (SCI-HI [340]), Shaped Antenna measurement of the background RAdio Spectrum (SARAS [341]), Large Aperture Experiment to Detect the Dark Ages (LEDA [342]), to name a few.

The detection of an absorption trough in the global signal experiment centered at 78 MHz was reported by the EDGES team [338]. The deep and wide absorption trough reported by EDGES had lead to controversies since it indicted that the cosmological models used till date had some glaring inadequacies. This lead to theorizing a number of new models which provided explanations for the shape and depth of the EDGES spectrum [54, 55]. There were also other works that indicated some problems with the analysis done by the EDGES team [56]. However, recent results using the SARAS3 experiment ruled out the EDEGES detection with 95.3% confidence [53]. No other experiment has made any detection claims so far. The SARAS3 results indicate that the standard models in theory are still valid. Hence, the task of detecting the true global signal from CD/EoR remains.

2.4.2 Observation of PS

Radio interferometry has also seen great advancement with increased sensitivity and improved techniques. While next-generation interferometers like the SKA and HERA can detect the HI PS directly, the older instruments have been upgraded with sufficient sensitivity to place meaningful upper limits on the EoR and post-EoR PS.

The operational sensitive low frequency interferometers have placed upper limits on the HI power spectrum - the Giant Metrewave Radio Telescope (GMRT), Murchison Widefield Array (MWA), Low-Frequency Array (LOFAR), Precision Array for Probing the Epoch of Reionization (PAPER) and its successor Hydrogen Epoch of Reionization Array (HERA).

The different interferometers use different approaches to extract the information from their observations. The latest results with these interferometers are discussed briefly.

- **GMRT**: The GMRT consists of 30 dishes in a Y-shaped array spread over 25 Km, located near Pune in western India. It studied the spatial and spectral structure of the sky using visibility correlations. Upper limits on the HI signal at both low ($z \sim 1.32$) and high ($z \sim 8.6$) redshifts were obtained using the GMRT (the latter being the first direct upper limit on the EoR HI signal) [58, 59].

- **MWA**: The MWA is a 256-element interferometer in western Australia. It started operating with 128 tiles in 2013, with an upgrade to 256 tiles in 2016. It operates in two modes: an extended array with 128 tiles and long baselines and a compact array with 128 tiles but short baselines, including two 36-tile redundant sub-arrays in a hexagonal configuration. The compact array is principally used for EoR science. MWA has, till now, placed several upper limits on the 21-cm PS at different redshifts between $z \sim 6$ and $z \sim 15$ [63–67, 104].

- **LOFAR**: LOFAR is a composite low-frequency aperture array with stations of High-Band Antennas (HBA) and Low-Band Antennas (LBA) types. They operate between 120-190 MHz and 30-90 MHz, respectively, and both have been used for CD/EoR science. The LOFAR latitude allows circumpolar observations with long winter nights. Thus, a primary observing filed for LOFAR is North Celestial Pole (NCP), observable winter months for more than 12 hours. LOFAR has also placed sensitive upper limit on the HI signal in the redshift range $7 \lesssim z \lesssim 25$ [69–71].

- **PAPER and HERA**: PAPER was designed as an experimental setup to develop novel analysis techniques for 21-cm cosmology. PAPER antennas were single dipoles designed on elevated ground screens with small size, for enabling array reconfiguration. The digital correlator architecture used by the system was easily scalable with the increase in the number of antennas. The PAPER team also developed the delay spectrum a approach [72] for measuring the 21-cm PS. However, the delay spectrum has higher noise than the alternative approaches, and [73] proposed the "maximum redundancy" configuration (in which antenna arrangement produces multiple redundant baselines) as a possible solution. Despite the development of novel analysis techniques and designing a sensitive array, several errors in the analysis resulted in the PAPER upper limits being invalidated (see [74, 75] for PAPER upper limits, and [76] for details of errors).

  HERA is an interferometric array in the Karoo desert, South Africa made of fixed, zenith pointing dishes . HERA has 14m dishes packed in a hexagonal, nearly continuous core 300m across. The frequency coverage is 50–250 MHz. It is being built in a series of phases with simultaneous construction and observation. The first phase used the feeds and correlator from the PAPER experiment, but the second phase will use new feeds and analog and digital systems. The first upper limits were recently provided with data from 50 dishes [77] at $z = 7.9$ and $z = 10.4$.

Figure 2.4 shows the upper limit over the range of k-modes probed by HERA (top panel), LOFAR (bottom left), and MWA (bottom right). These three are presently the most sensitive interferometers targeting the detection of the redshifted 21-cm signal from the EoR. Table 2.1 tabulates the accepted upper limits set so far by different interferometers.

The Future of PS Observations:

With the initial construction of SKA-1 being approved, interferometric observations of the 21-cm signal is in an interesting stage. It is expected to detect the power spectrum of the redshifted 21-cm signal from CD/EoR with $\sim$1000 hours of observation and even perform 21-cm tomography [46]. In addition to SKA, other instruments are also currently undergoing commissioning or upgrade to enhance their sensitivity for providing competitive upper limits on the CD/EoR PS. The LOFAR is upgrading to LOFAR 2.0 with enhanced sensitivity and longer baselines, which is enhance its capacity to perform CD/EoR observations. Simultaneously, a new array called the Nenufar is being constructed at the Station de Radioastronomie in Nanay (France). It will consist of 96 mini-arrays in a 400 m disk which will be sensitive to large spatial scales and will be able to study the 21-cm signal from CD/EoR [78]. The HERA is also in its commissioning phase performing simultaneous observations and constructions, and is expected to provide even more sensitive upper limits in the upcoming years. Arrays like the LWA, MWA Phase-III, uGMRT etc. are also trying to achieve enhanced sensitivity which will lead to them targeting the redshifted 21-cm signal observations along with other science goals.

2.5 Estimation of Signal Parameters from Observations

The main aim of any interferometric or global signal experiment is to determine the astrophysical and cosmological parameters responsible for producing the observed signal. For the past several years, meticulous research into optimizing techniques

Table 2.1: Upper Limits on the HI 21-cm Power Spectrum from CD/EoR set by different telescopes. Reproduced from [4].

Telescope	z	k (Mpc^{-1} h)	$\Delta^2(k)$ (mK^2)	Reference
GMRT	8.6	0.65 h	$< 4.9 \times 10^3$	[59]
	8.6	0.5h	$< 6.1 \times 10^4$	[60]
	1.96	1.0	$< 3.5 \times 10^3$	[61]
	2.19	1.0	$< 3.8 \times 10^3$	
	2.62	1.0	$< 3.7 \times 10^3$	
	3.58	1.0	$< 1.1 \times 10^4$	
MWA	12.2	0.18	$< 2.5 \times 10^7$	[62]
	15.35	0.21	$< 8.3 \times 10^7$	
	17.0	0.22	$< 2.7 \times 10^8$	
	7.1	0.27	$< 2.7 \times 10^4$	[63]
	7.0	0.20	$< 3.9 \times 10^3$	[64]
	6.5	0.59	$< 3.39 \times 10^3$	[104]
	6.5	0.142	$< 1.8 \times 10^3$	[65]
	6.8	0.142	$< 3.6 \times 10^3$	
	7.1	0.142	$< 6.0 \times 10^3$	
	7.8	0.142	$< 2.4 \times 10^4$	
	8.2	0.142	$< 2.8 \times 10^4$	
	8.7	0.142	$< 6.2 \times 10^4$	
	15.2	0.14	$< 6.3 \times 10^6$	[66]
LOFAR	7.9-8.7	0.1	$< 1.1 \times 10^5$	[69]
	8.7-9.6	0.1	$< 8.7 \times 10^4$	
	9.6-10.6	0.1	$< 5.0 \times 10^4$	
	19.8-25.2	0.038	$< 2.1 \times 10^8$	[70]
	9.1	0.075	$< 5.3 \times 10^3$	[71]
HERA	7.9	0.192	$< 9.5 \times 10^2$	[77]
	10.4	0.256	$< 9.1 \times 10^3$	

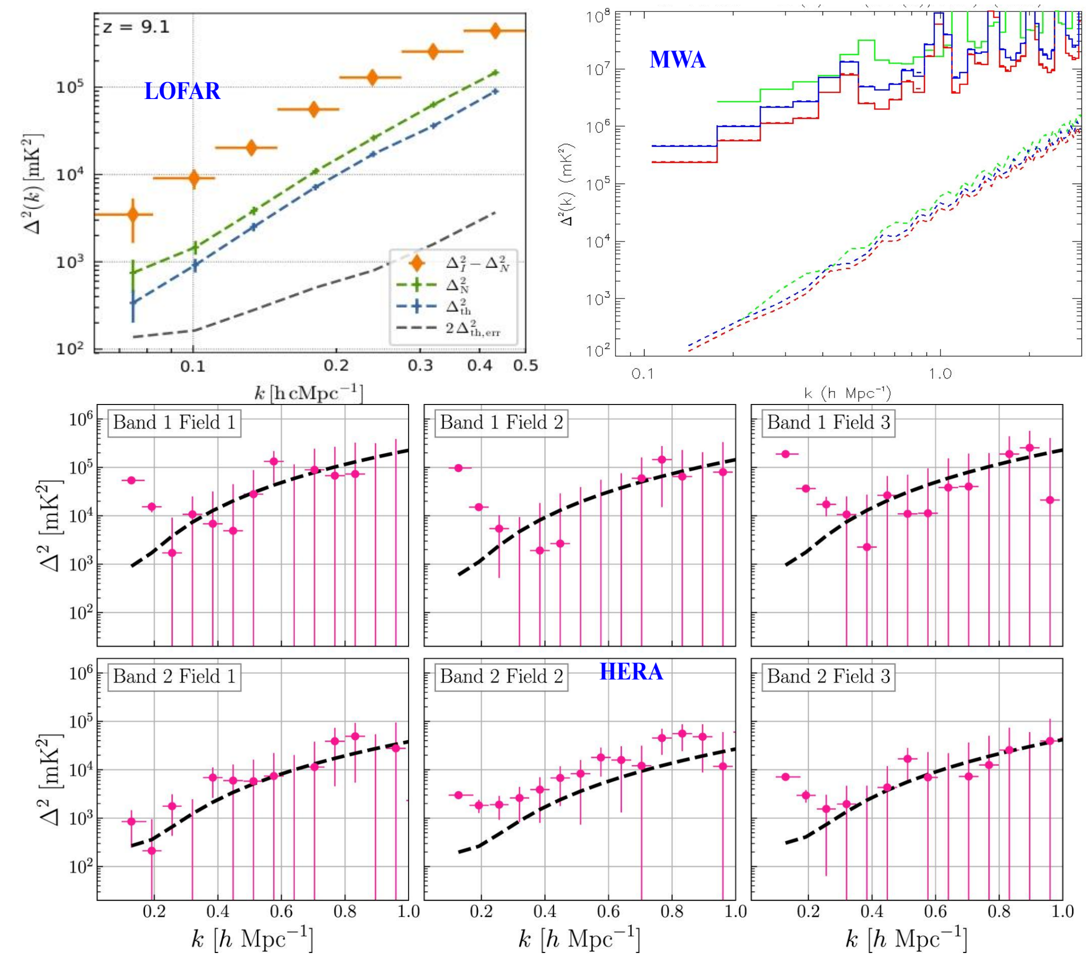

Figure 2.4: Recent upper limits on the EoR PS from HERA, LOFAR and MWA. Compiled from [77], [71] & [65].

for parameter extraction has been done. Bayesian statistics is still the most popular method for parameter estimation, but with the advent of advanced machine learning (ML) algorithms, currently these are also being explored extensively.

Using Markov Chain Monte Carlo (MCMC) methods, there have been several studies on parameter estimation from both the global signal as well as PS. Synthetic data sets have been used to extract parameters using different reionization models, which are also contaminated with foregrounds and instrument response in many cases. Some of these studies include those of [79–83], to name a few.

Bayesian inference-based methods are computation and time intensive. Thus, ML methods are now being used to efficiently extract parameters from large data sets expected from current telescopes. These are very efficient algorithms which "learn" from complex, non-linear and non-Gaussian priors. A number of different approaches using ML have been developed over the past decade. Parameter estimations using ML base emulators, as well as artificial neural networks (ANN) and convolutional neural networks (CNN) have been developed for the image domain (for example [84–87]). Similar developments in signal extraction and parameter estimation in the PS domain and global signal has been done using ANN (for example [88–91]).

However, signal extraction and parameter estimation in presence of foregrounds and systematics is a challenging task. In [91], a pair of ANNs was used for signal power spectrum extraction and parameter estimation. Their results, while encouraging also suggest the need for development of better training and sampling methods for effective signal extraction from observed data sets.

This thesis deals with the study of different factors that causes hindrance in the detection of the power spectrum from EoR and post-EoR Universe. The biases and imperfections in the 21-cm signal extraction introduced by the presence of foregrounds and an imperfect instrument (particularly radio interferometer) has been discussed in the next chapter.

Chapter 3

Challenges for the Low Frequency Observations of the 21-cm Signal

3.1 Challenges in Detecting HI from CD/EoR

The low-frequency observations of the 21-cm signal from the early Universe using radio interferometers are challenging. Theoretical models predict that the differential brightness temperature is a few hundred mK at most (for instance, see Figure 2.2 and [43]). The contaminants of such a weak target are numerous and vary in nature and complexity. The major sources that adversely affect observations targeting detection of HI from the early Universe are introduced by:

- **Astrophysical Foregrounds:** Different astrophysical sources present along the line-of-sight between the observer and the signal.
- **Instrumental Systematics:** The imperfect response of any real instrument that is present in observations.
- **Radio Frequency Interference:** The interference from the large number of terrestrial instruments that transmit and/or receive signals in radio wavelengths.

This chapter tries to provide a more detailed discussion on the nature of the challenges posed by foregrounds, systematics and RFI from the perspective of radio interferometric observations. Section 3.2 provides a brief description of the foreground emission, followed by a basic discussion on interferometric data processing in 3.3 (which is required to understand the systematics produced by the instrument). The different systematics are discussed in Section 3.4 followed by effect of RFI in Section 3.5.

3.2 Foregrounds

The astrophysical foregrounds are the most potent contaminant of the redshifted 21-cm signal from the CD/EoR. They constitute extraterrestrial radio emissions brighter than the target signal at the observing frequency. Foregrounds relevant to EoR observation frequencies are diffuse galactic synchrotron emission, free-free emission from galactic and extragalactic sources, and emission from star-forming galaxies and active galactic nuclei. These sources by themselves also provide interesting astrophysics (for example [120, 133, 134, 193–202]) and are studied actively. However, for 21-cm cosmology, they constitute contaminants that need to be removed from the data. A brief description of the foregrounds is given in the following subsections.

3.2.1 Diffuse Foregrounds

The bulk of the foreground emission at scales $\gtrsim$degree for $\nu \lesssim 150$ MHz consists of DGSE. Figure 3.1 shows the brightness temperature map of the DGSE at 408 MHz made by [193, 194]. The brightness temperature of this emission (plotted as the logarithm of temperatures in K) is brightest at the galactic plane. However, there are observations of excess synchrotron power on angular scales $\sim$ degrees at high latitudes near the North Celestial Pole and South Galactic poles [69, 203, 204].

The average brightness temperature of DGSE is approximated as a power law of the form $T(\nu) \propto \nu^{-\alpha}$. It is a function of frequency, and is thus higher at lower frequencies of interest for CD/EoR science. Beside synchrotron radiation, free-free

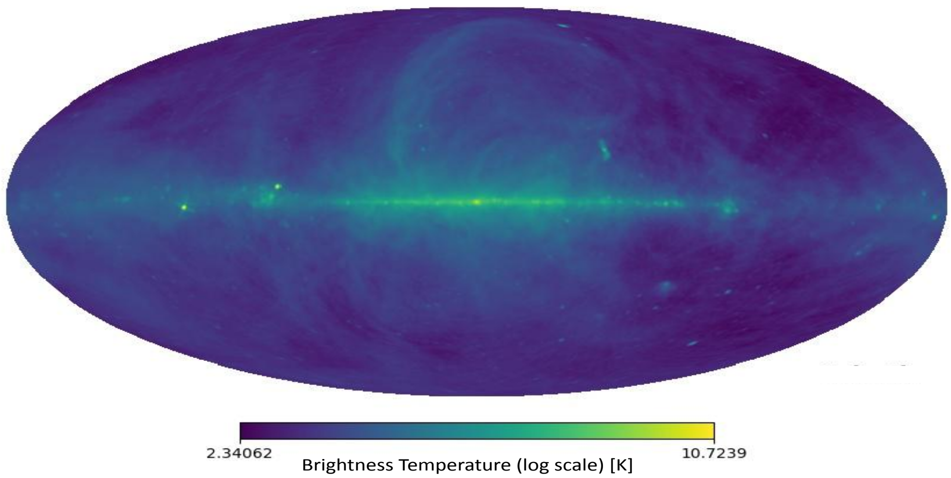

Figure 3.1: 408 MHz Haslam map [193, 194, 196] of DGSE. The emissions is clearly bright at the galactic plane, and is brighter at lower frequency. The color bar is in $\log_{10}$ scale for temperature in K and is brighter than the 21-cm emission. The plot is made using the publicly available PyGDSM package (https://github.com/telegraphic/pygdsm).

emission and a small amount of radio emission from radio haloes and relics also contribute to the diffuse foreground.

3.2.2 Point Sources

Extragalactic foregrounds at radio frequencies are compact sources consisting mostly of active galactic nuclei (AGNs) and star forming galaxies (SFGs). The low-frequency population of radio sources, especially at the faint flux density end, is not yet adequately constrained. Many surveys have tried to constrain number counts, luminosity functions, and other properties [137, 145, 205]. The current deepest 150 MHz source counts from three extragalactic deep fields - Lockman Hole, ELAIS-N1, and Boötes, obtained using the LoTSSS data [206] is shown in Figure 3.2, overlaid with counts from other surveys- Boötes at 150 MHz using LOFAR [136], ELAIS-N1 at 400 MHz using GMRT [175] and Lockman Hole at 325 MHz using GMRT [1]. In addition to source counts (i.e., flux distributions), it is also essential to understand their spatial distribution, i.e., clustering. It has been shown by [108] that spatial clustering of extragalactic sources dominates the fluctuations for angular scales $\theta \gtrsim 1'$ at 150 MHz (for flux density $\gtrsim$0.1 mJy). Thus, characterization of spatial, flux, and frequency behaviour of extragalactic sources is essential.

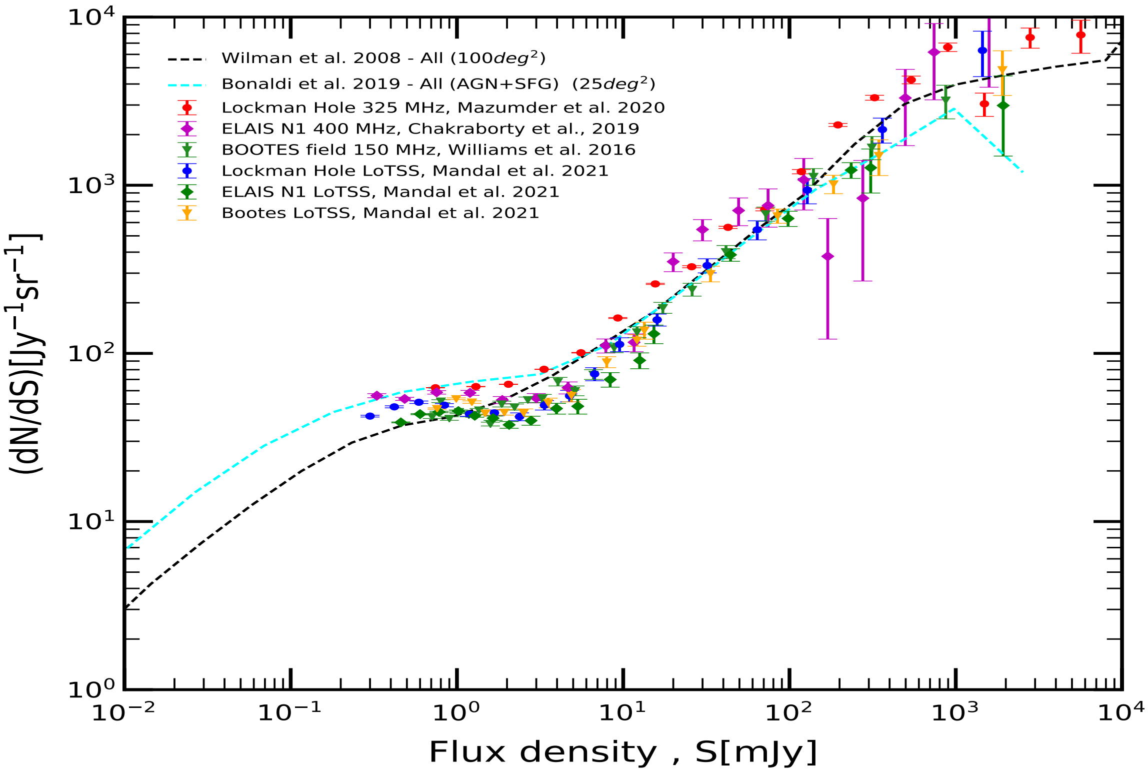

Figure 3.2: Euclidean normalised differential source counts at 150 GHz from LoTSS survey for Lockman Hole (blue circles), ELAIS-N1 (green diamonds) and Boötes (orange inverted triangle) fields from [206]. Source counts from other observations of the same field - Boötes (green inverted triangles, [136]) at 150 MHz from LOFAR, Lockman Hole (red circles, [1]) at 325 MHz from GMRT and ELAIS-N1 (magenta diamonds, [175]) at 400 MHz from from GMRT. Comparison with source counts from simulations - SKADS (black dashed curve, [121]) and T-RECS (cyan dashed curve, [122]) has also been done.

3.2.3 Challenges with Foregrounds

The relative strength of the foregrounds can be seen from Figure 3.3. It shows the angular power spectrum amplitude of different foreground components at 140 MHz

[126]. The dominant contribution is from the galactic synchrotron, followed by extragalactic point sources, with free-free emission having relatively smaller contribution, yet still above the 21-cm signal (an effect made more complicated with the introduction of systematics).

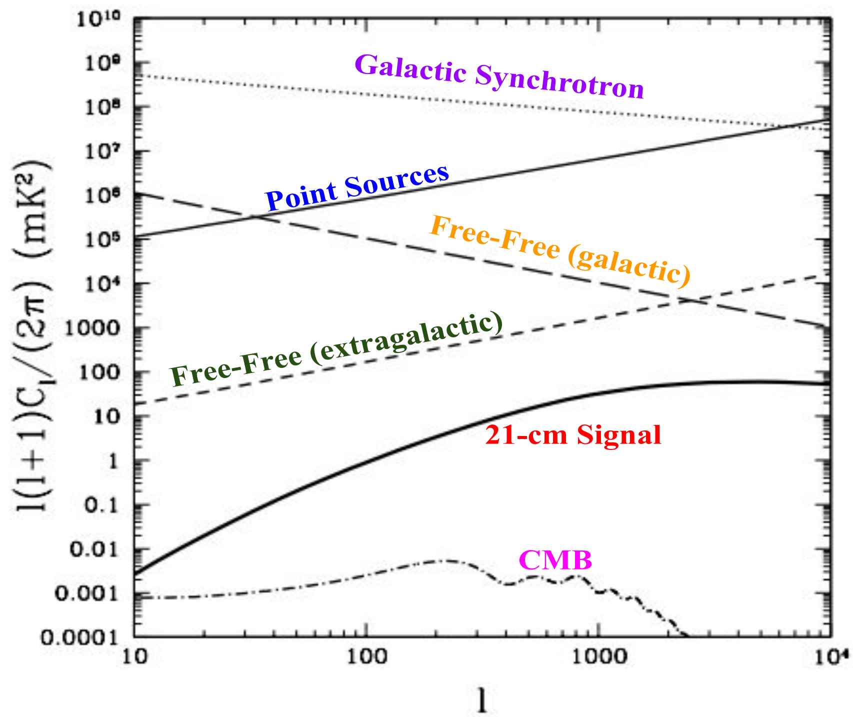

Figure 3.3: Angular Power Spectra of Different Foreground Componets along with that of the 21-cm signal and CMB. Reproduced from [126].

The foregrounds radiate synchrotron with power law behaviour with frequency resulting in spectrally smooth behaviour (see, for example, [207, 208]). Most foreground mitigation techniques employ this spectral smoothness assumption to model out foregrounds. Some standard techniques include principal component analysis, generalized morphological component analysis, independent component analysis, Gaussian process regression, etc. [209–211]. While each process has pros and cons, a general caveat is that the presence of non-smooth foreground components in the residuals will hamper the signal detection. Thus, studies over a wide frequency range as well as how spectral smoothness is affected by inclusion of instrumental features is required. Thus, simultaneous understanding of both foregrounds and systematics is required.

3.3 Understanding the Instrument: Radio Interferometer

Radio telescopes measure the incident electric field from the sky directly, rather than the energy deposited, as traditional telescopes do. A radio interferometer [1] thus measures the spatial Fourier coefficient describing the radiation pattern in the far-field approximation. The inverse Fourier transform produces a map of the intensity distribution of the radio signal in the sky (or a radio image). The basic technique followed in radio interferometric imaging is known as aperture synthesis. A finite number of spatially separated detectors or slits is used to construct an effective lens aperture, whose size is the largest separation between any two pairs of slits. This is called a synthesized aperture. It is different from the true lens aperture of the same size since it is discontinuous, being a collection of discrete and finite set of aperture points. The elements of an interferometer basically act as slits, the signals from which are points that fill the aperture. This technique is called aperture synthesis. The incident electric fields at the two aperture points are correlated to obtain the amplitude and phase of the interference patterns from each pair of slits. Thus, the incident radiation field's spatial Fourier transform is measured by a pair of slits on the aperture plane. Their spatial frequency is given in terms of physical separation between the slits in units of wavelength.

[1] The terms "radio telescope" and "radio interferometer" have been used interchangeably throughout this thesis.

3.3.1 Measurements by an Interferometer

The van-Cittert Zernike Theorem:

The basic principle for mapping distant sources using a radio interferometer is quantified using the van-Cittert-Zernike theorem of partially coherent radiation. It states that "the degree of spatial coherence of radiation from a spatially incoherent distant source is proportional to the spatial Fourier transform of the intensity distribution across the source" [92]. It is expressed mathematically as:

$$\langle E_\nu(\bar{r_1})E^*_\nu(\bar{r_2})\rangle \propto \int I_\nu(\hat{s})e^{2\pi i\nu\tau_{12}}d\Omega \equiv V(\bar{r_1}-\bar{r_2},\nu) \tag{3.1}$$

where $E_\nu(\bar{r_1})$ and $E_\nu(\bar{r_2})$ are the incident electric fields at $\bar{r_1}$ and $\bar{r_2}$ on the aperture plane, $d\Omega$ is the solid angle subtended by the instrument on the sky. The equivalent time averaged coherence function or correlation function, $V(\bar{r_1}-\bar{r_2},\nu)$ is called the visibility. It is a complex number which is a function of the separation between the antennas $(\bar{r_1}-\bar{r_2})$. A radio interferometer is a device that measures this spatial coherence function.

3.3.1.1 Coordinate System for Observations

Visibilities measured by a set of telescopes (which make up an interferometer) require a proper set of coordinates that relate the brightness distribution in the sky with the detectors' physical location and spacing.

If the antenna is pointed towards a direction $\vec{s_0}$, and $I(l,m,n) = I(\vec{s})$ where $l,\ m,\ n$ are the direction cosines along a direction $\vec{s}$ (see Figure 3.4), $\vec{s_0}$ $(l=0,\ m=0,\ n=1)$ is the reference point with respect to which the phases are measured (i.e. the phase center), and any point away from $\vec{s_0}$ is $\vec{s} = \vec{s_0} + \vec{s_\sigma}$. The "uv plane" is defined as the tangent plane at $\vec{s_0}$ onto which intensity distribution is projected and the final 2D image is formed, i.e. a plane perpendicular to the aperture plane of the array. The relation between these coordinates systems is given by:

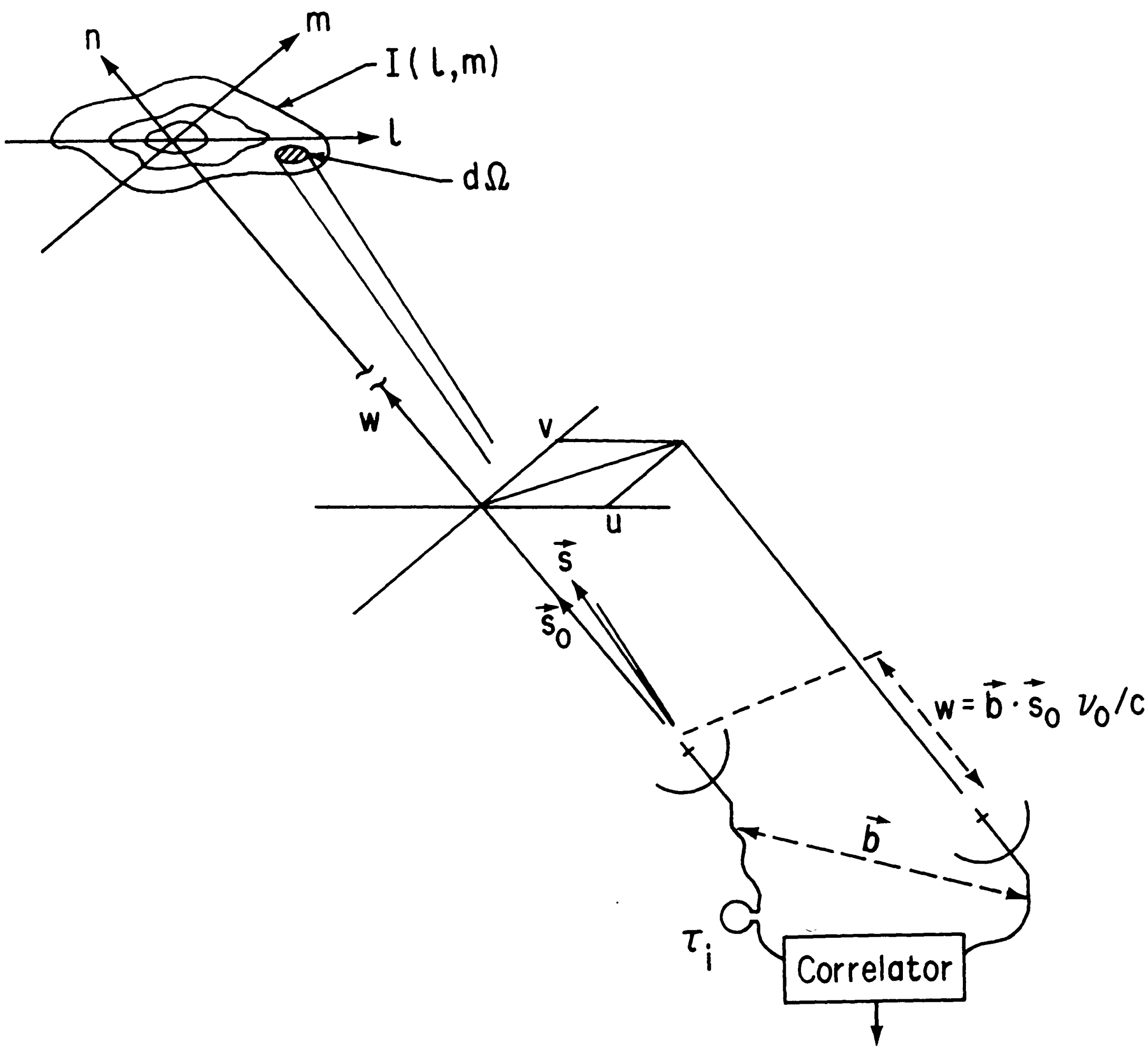

Figure 3.4: The coordinate systems used to express the source brightness distribution in the telescope aperture plane. Image adapted from [93].

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} sin(H) & cos(H) & 0 \\ -sin(\delta_0)cos(H) & sin(\delta_0)sin(H) & cos(\delta_0) \\ -cos(\delta_0)cos(H) & -cos(\delta_0)sin(H) & sin(\delta_0) \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{3.2}$$

where H and δ_0 are the hour angle and declination of the source respectively, the physical distances are represented by x,y,z and measured in meter in the XYZ coordinate system and u,v,w are the distances in the UVW coordinates in units of the wavelength λ.

The vectors $\bar{r_1}$, $\bar{r_2}$ in Equation 3.1 can be defined in the UVW coordinate system as $\bar{r_1}(u_1, v_1, w_1)$ and $\bar{r_2}(u_2, v_2, w_2)$ (in units of wavelengths). In radio interferometry, the baseline ($\vec{b}$) is the 3D vector between $\bar{r_1}$ and $\bar{r_2}$ given by $b(u, v, w) = \bar{r_1} - \bar{r_2}$, where $u = u_1 - u_2,\ v = v_1 - v_2\ and\ w = w_1 - w_2$. The vectors $\bar{r_1}$, $\bar{r_2}$ are not necessarily co-planar with the aperture plane (i.e. $\bar{r_1}.\hat{s_0} = w_1 \neq 0$ and $\bar{r_2}.\hat{s_0} = w_2 \neq 0$). Hence, any pair of detectors will not sample the same incident wavefront at a given instance, rather there will be a time delay between the detections, given by $\tau = \vec{b}.\hat{s_0}/\nu = w_2 - w_1/\nu$ (where ν is the observing frequency). This delay has to be corrected before the signals are correlated from two telescopes.

3.3.2 Aperture Synthesis

Equation 3.1 can be re-written in the UVW system as :

$$V(u, v, w) = \int\int I(l, m, n)e^{-2\pi i(ul+vm+w\sqrt{1-l^2-m^2})}\frac{dldm}{\sqrt{1-l^2-m^2}} \tag{3.3}$$

where $d\Omega = \frac{dldm}{n} = \frac{dldm}{\sqrt{1-l^2-m^2}}$, and l and n are the coordinates on the tangent plane at $\vec{s_0}$ (see Figure 3.4. Since visibilities measured in the UVW coordinates, inverse Fourier transform of Equation 3.3 should give the intensity distribution in the sky. Inversion of Equation 3.3 can be simplified to obtain $I(l, m)$ from visibility by reducing it to a two-dimensional Fourier transform if the sky is imaged close to the phase center thus ignoring the curvature of the sky (i.e. w = 0). Thus, Equation 3.3 simplifies to a 2D spatial Fourier transform relation:

$$V(u, v) = \int\int I(l, m)e^{-2\pi i(ul+vm)}dldm \tag{3.4}$$

Equation 3.4 is a 2D Fourier transform. It relates the mutual coherence function and the source brightness and is a form of the van Cittert–Zernike theorem (Equation 3.1). The baselines measure a complex visibility function at a point on the uv plane. The aperture is measured over a finite set of discrete points, and the inverse Fourier transform of Equation 3.4 is a summation over the measured points on the uv plane. However, the sampled (observed) uv plane has gaps and irregularities due to the limited number of baselines. This is mitigated by utilizing a technique called earth rotation aperture synthesis. The target is tracked as the earth rotates, increasing the number of unique projected baselines. Thus, aperture synthesis make better images of the target field by using a larger number of baselines in the uv space to sample visibilities.

3.3.3 Calibration and Imaging of Interferometric Data

As seen from Equation 3.1, data from interferometric measurements are a series of measurements of the sky in form of electrical voltages which in itself does not reveal any real astrophysical information. It requires post-processing for extracting useful information. The data is thus calibrated to account for the instrument gains, followed by imaging to deconvolve the telescope response and find the intensity distribution in the sky.

3.3.3.1 Direction Independent Calibration

The data received from the interferometer has instrumental effects combined with the actual sky intensity. The frequency-dependent complex "gains" encode all the modifications that the actual intensity goes through due to the receiver electronics. Gains modify the true visibilities by a multiplicative factor,

$$V_{ij}^{obs} = g_i g_j^* V_{ij}^{true} + n_{ij} \tag{3.5}$$

where the observed visibility V_{ij}^{obs} between an antenna pair i and j is the true visibility V_{ij}^{true} multiplied the complex gain factors g_i and g_j of the antennas and n_{ij} is the noise on this baseline. Each term in Equation 3.5 is implicitly a function of time. **Solving the gain factor of each antenna is known as calibration.**

Self Calibration:

Antenna based complex gains are calculated and applied to the observed visibilities to correct them. Some standard sources called calibrators, whose true visibilities are assumed known, are observed at regular intervals during the observation and used to calculate the gain solutions, which are interpolated to target visibilities to calibrate them. This also means that any fluctuations in the instrument gains during the target observations remain unaccounted for. Thus, self calibration is used, where a model of the target source itself is used to compute the gain solutions during the time it was observed [93]. It is an iterative combination of calibration and imaging steps.

The calibration processes described above assumes that the gain solutions for the antennas are sufficient to describe any corruptions on the across the area of interest. Stated simply, this means that g_i terms in equation 3.5 do not vary with direction, i.e. the calibration is "direction independent".

3.3.3.2 Direction Dependent Calibration

However, at the low frequencies of interest ($\sim$ few hundreds of MHz), the primary beam response changes over the wide observing field, frequency and observing time. The earth's ionosphere also introduces position and time dependent corruptions [192]. These are directional effects and are corrected using direction dependent calibration. The basic idea for direction dependent calibration is calculation of the antenna/station gains towards directions of bright sources which are located away from the phase center. This process becomes specially important at lower frequencies due to increased ionospheric activity and wider field of view.

The data calibration for this thesis uses both direction independent and dependent calibration. The details are presented in Chapter 4.

3.3.3.3 Imaging

Post calibration, the visibilities are imaged to determine the intensity distribution in the sky. The sky sampled at discrete points can be described by a sampling function $S(u,v)$. It can be represented by a collection of Kronecker delta functions as

$$S(u,v) = \sum \delta(u-u_k)\delta(v-v_k) \tag{3.6}$$

where k is the index that represents a measurement from one baseline.

An image is the direct Fourier inversion of V_{ij}^{obs} described mathematically as:

$$I^{obs}(l,m) = \int\int S(u,v)V(u,v)e^{2\pi i(ul+vm)}dldm \tag{3.7}$$

I^{obs} is called the dirty image, which is the the convolution of the true sky brightness and the instrument's point spread function (PSF). Since the observed image is a convolution of the true sky brightness with an instrumental point spread function, the true sky brightness is estimated using a deconvolution process.

3.4 Effects of Systematics

Due to the inherent complexity of an interferometer, there are certain instrumental properties correctly determined, and some others are undetermined. The known systematics can, in principle, be dealt with during data reduction. However, those may be computationally expensive. Also, "optimally" dealing with systematics can lead to sensitivity loss or other unwanted effects on the data. The unknown systematics are even more harmful since the way they bias the analyses are unknown. Nevertheless, recent research into systematics has focused on the effects of instrumental gains, primary beam models, thermal noise, polarization leakage, and direction-dependent ionospheric effects, to name a few.

Calibration Errors:

Calibration constitutes an essential part of any interferometric data processing. Sensitive observations like those targeting the redshifted 21-cm signal require very precise calibrations to avoid contamination in the final PS. Due to the non-ideal conditions during observations, individual antenna gains vary with time and frequency and might become difficult to predict. Additionally, the calibration algorithm also has inherent limitations and can introduce errors in the gain solutions. In short, for a real instrument, the terms g_i and g_j of Equation 3.5 are never the ideal value unity. This deviation from ideal value leads to "calibration error" which propagate and may cause problem in signal extraction. Two widely used calibration algorithms for EoR data sets are sky-based and redundant. The sky-based calibration uses a sky model as a prior, while redundant technique calibrates by matching visibilities from redundant baselines for highly regular arrays. Sky model incompleteness affects the former technique, while imperfect redundancy affects the latter. Thus, due to inherently incomplete sky models, imperfect redundancy, variations in the antenna response, calibration errors are always present.

Position Errors:

Since calibration processes use a sky model to solve for instrument gains. This requires the presence of high resolution source models from prior observations. Since the HI 21-cm signal decorrelates at small angles, so any sky model used for such data calibration would be derived from previous observations. However, irrespective of the observed frequency, the astrometry will be accurate to a certain extent. Thus, the astrometric inaccuracies in the model are propagated and cannot be improved by calibration. Thus, position errors in the sky model can be a dangerous source of errors that might affect the final data calibration and thus target signal detection.

Primary Beam:

The telescope's primary beam describes its spatial and frequency sensitivity. Equation 3.3 is a simplified form for an ideal instrument where the primary beam is unity. However, for a real instrument, the primary beam response deviates from the ideal case, and it is vital to quantify the primary beam pattern. The beam has a complicated spatial and polarization response which depends on frequency. The effect of these complications can be minimized by design specifications but cannot be removed entirely. There may also be variations between different feeds due to manufacturing or installation errors or positional effects. In otherwise redundant arrays, the feed-to-feed variation of beam pattern may introduce non-redundancy. It becomes a systematic error affecting redundant calibrations. Depending on the analysis technique employed, different conditions restricting properties of the primary beam are introduced. Some techniques deconvolve instrument effects using linear combinations of different visibilities, and are typically very sensitive to the knowledge of primary beams. Other techniques are not strongly dependent on such knowledge but require frequency smoothness and polarization purity. Thus, primary beams constitute a significant systematic effect that affects sensitive observations targeting the 21-cm signal from EoR.

Thermal Noise:

Thermal noise is typically considered a random error that can be minimized easily. However, there are complications generated in an interferometer that need to be understood. Close-packed arrays used in EoR observations to maximize sensitivity produces complications. It leads to many closely located antennas interacting directly, giving a source of correlated noise. This biases the visibilities in the shortest baselines,which are sensitive to the largest angular scales. An approach for such bias removal is taking advantage of the sky changing over the day, removing an appropriate average of each visibility. However, this only works for stable noise properties and thus might not be fully mitigated. Additionally, subtracting the average will also excise the average signal, which must be accounted for.

Polarization Leakage:

Radio interferometers are generally constructed from pairs of co-located antennas with orthogonal feed polarization with sensitivity to the polarization of emission in the sky. However, due to imperfections, there may be a leakage of polarized components of the emission. Polarized emission can leak in two ways- leakages due to non-orthogonal or rotated feeds and beam ellipticity (where an asymmetry in the two linear polarization of the primary telescope beam causes unpolarized signals to appear polarized, and vice versa). Dealing with foregrounds or 21-cm observations relies on their spectral smoothness. If they have a significant spectral structure within their polarized emission, the intensity estimation may be incorrect, introducing spurious spectral structure and hindering foreground removal.

Ionosphere:

Interferometric data are also affected by the earth's ionosphere. The incoming extra-terrestrial wavefronts suffer refraction from the non-uniformity of the ionospheric plasma layers. The presence of plasma gradients across the pierce points to the extragalactic sources results in offsets in source positions, while curvature can change the source flux density. At the frequencies of interest for CD/EoR science ($\lesssim$ 150 MHz), the ionospheric effects can become quite significant, affecting data calibration and signal recovery.

3.5 Radio Frequency Interference

Terrestrial radio interferometers also suffer from man-made RFI. There are numerous sources of RFI like telecommunication signals, FM broadcasts, air traffic, satellite passage through the telescope beam, etc. These signals are generally localised in time and frequency. Thus, interferometric observations in general have high time and frequency resolution which helps in localisation and excision of RFI during initial data analysis steps. However, this introduces irregularities in the otherwise uniform instrumental bandpass response, and hence in the visibilities. Since PS estimation involves Fourier transforming the visibilities from the frequency to the delay domain, the nonuniform bandpass sampling from RFI excision creates rapid fluctuation across the delay axis. Hence, there will be rapidly fluctuating components in the delay space visibilities, creating ripple across the line-of-sight k modes, thus affecting signal extraction. It has however been shown in literature (see Figure 3.5) that these fluctuations can be reduced by applying the *CLEAN* or least squares spectral analysis (LSSA) algorithms [106]. It was also shown in [106] that LSSA performs consistently across a wide range of flagging scenarios. However, more investigations into performance of other algorithms are also required for determination of the optimal one.

3.6 Radio Interferometric Observation of HI from CD/EoR

Radio interferometers are excellent instruments for determining the spatial and frequency (hence redshift) distribution of sources. Thus, the ultimate aim of sensitive 21-cm experiments is to map the spatial distribution of the signal as a function of redshift (21-cm tomography). However, owing to sensitivity limitations [2], current interferometers can only perform statistical detections (i.e. PS)[192]. Equation 3.4 shows that interferometers are inherently spatial power spectrum measuring instru-

[2]Effective 21-cm tomography requires mK sensitivity on angular scales of a few arcminutes.

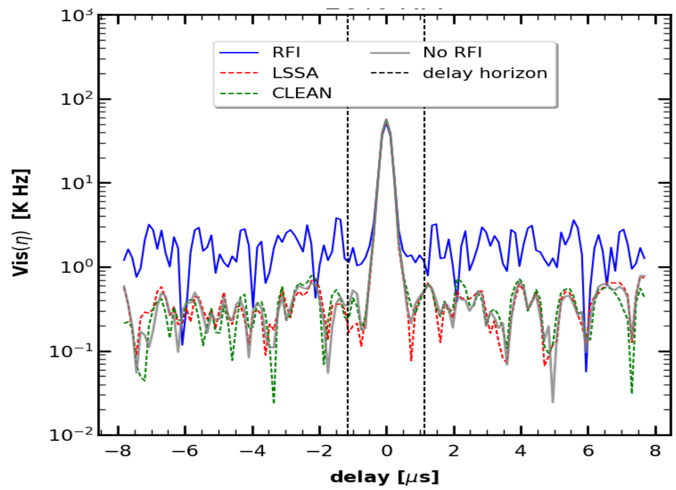

Figure 3.5: Delay spectrum for simulated visibility having 10% RFI excised. The green and red dashed curves are the delay spectra after using the *CLEAN* and the LSSA algorithms respectively. The gray curve is the delay spectrum of the baseline when there is no flagging. Reproduced from [106].

ments. The PS is computed either from a deconvolved image cube or directly from visibilities. The visibilities can be further *delay transformed* along the frequency axis [192]:

$$\tilde{V}(u, v, \tau) = \int_B V(u, v, \nu) e^{-2\pi i \nu \tau} d\nu \tag{3.8}$$

where B is the observing bandwidth and the delay τ is the Fourier conjugate of ν (i.e. frequency). From Equation 3.8 and Equation 2.8, it is seen that the delay transformed visibility is proportional to the three dimensional power spectrum where the proportionality constant transforms the visibility into power units. The baseline vector ($\vec{b}$), and delay τ are directly related to the k-modes parallel and perpendicular to the line-of-sight. Thus, it is seen that the visibilities and the 3D PS are related to each other. The exact relationship is shown schematically in Figure 3.6.

From Figure 3.6, it is evident that for wide-field observations, the line-of-sight dis-

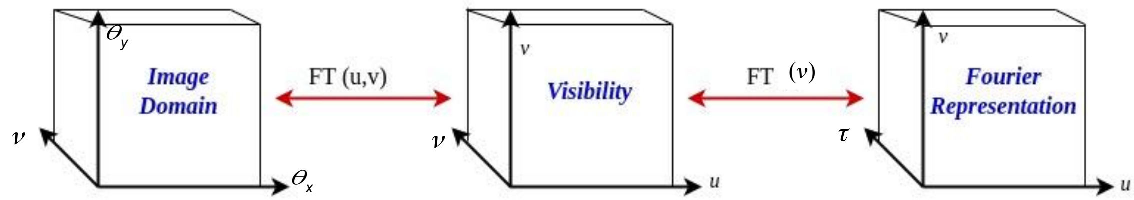

Figure 3.6: Fourier conjugate relationships between the fundamental observable of an interferometer - visibility with image (Fourier transform along the plane of sky) and the Fourier representation (Fourier transform along frequency direction). Reproduced from [4].

tance to the source maps into the observed frequency and observed angular scales map into the u,v coordinates. Since an interferometer observes the Fourier transform of the intensity distribution of the sky, the relation between the observing angles and the u,v plane gives four different representations of the signal given by Figure 3.7. The horizontal pairs are Fourier conjugates and the vertical pairs are scaling relations, related by the sky coordinates and frequency (θ and Δf) and their inverse (u, v, η) to the comoving distance $\mathbf{r}$ and its inverse $\mathbf{k}$. The exact mathematical relations amongst these quantities are given by [366]:

$$\theta_x = \frac{r_x}{D_M(z)}, \qquad \theta_y = \frac{r_y}{D_M(z)},$$
$$u = \frac{k_x D_M}{2\pi}, \qquad v = \frac{k_y D_M}{2\pi}$$
$$\Delta f \approx \frac{H_0 f_{21} E_z}{c(1+z)^2}\Delta r_z, \qquad \eta \approx \frac{c(1+z)^2}{2\pi H_0 f_{21} E_z} k_z$$

The 21-cm PS being related to the interferometric visibilities imply that imperfections in the observed visibility introduced by the instrument also propagate into the PS, making it difficult to detect. Ultimately, even the accessible k modes depend on the interferometer design. The angular wave numbers, $\mathbf{k}_\perp$, are related to the instrument layout, with the accessible modes are determined by the baseline distribution. The longest baseline limits the finest accessible angular scales (largest $k_\perp = |\mathbf{k}_\perp|$), while the shortest ones limit the largest accessible $k_\perp$ value. The lines of sight wave

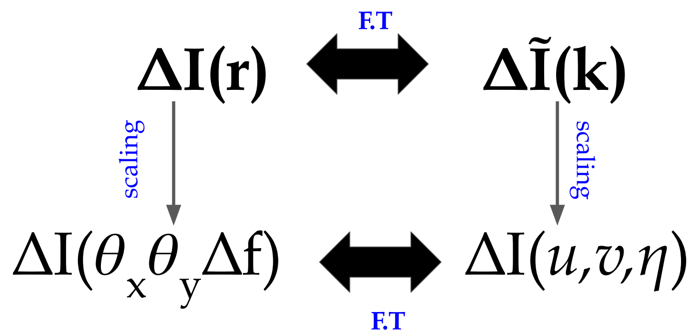

Figure 3.7: Fourier conjugates and scaling relations. Produced using the relations described in [366].

numbers, $k_{\|}$, are determined by the spectral features of the instruments. The finest modes (with highest $k_{\|}$) are spectral resolution limited, and coarsest modes are bandwidth limited. These limits are depicted qualitatively in Figure 3.8. For 21-cm experiments, the accessible region in the $\mathbf{k}_{\perp} - k_{\|}$ plane is called the "EoR window" (blue region in Figure 3.8). The EoR window is expected to be free from foreground contamination. However, this is never fully achievable. The level of contamination will depend on the accuracy of the foreground model and level of handle on the systematics. Detailed study into foregrounds have been done in Chapters 4 and 5 and into systematics have been done in Chapter 6.

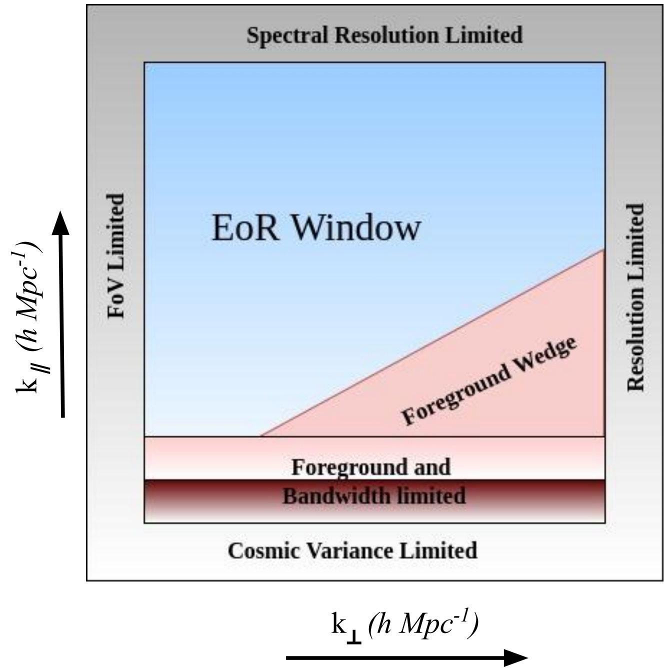

Figure 3.8: Schematic diagram for the Fourier space accessible by a radio interferometer. Reproduced from [4].

Chapter 4

Characterizing foregrounds pertinent to EoR studies

This chapter has been adapted from *Characterizing EoR foregrounds: A study of the Lockman Hole Region at 325 MHz* [1]

4.1 Introduction

The Cosmic Microwave Background (CMB), the afterglow from the Big Bang, has provided unprecedented insights into the thermal history of the Universe dating back to the surface of last scattering. However, despite the development of precision instruments for cosmology, the period after recombination (Dark Ages) of the Universe, till the neutral Universe became ionized once again remains one of the relatively unexplored realms in cosmology. It is during this time, between redshifts $\sim$ 30-12, that tiny fluctuations in matter density grew under gravitational instabilities that resulted in the generation of the first stars. This epoch is known as the Cosmic Dawn (CD) - the dawn for the luminous sources in the Universe. This first generation of stars produced a copious amount of ionizing radiation (X-rays, UV and Lyα), which converted the neutral intergalactic medium into an ionized one [94]. This era is known as the Epoch of Reionization (EoR) (see [40, 44, 95] for comprehensive reviews). Observations of quasar absorption spectra [96], as well as Thompson optical depth obtained from CMB temperature and polarization measurements [97],constrain the EoR to $6 < z < 15$.

To investigate the physical processes occurring in these early stages of the Universe, the most promising probe is the 21-cm hyperfine line of neutral hydrogen [32, 98, 99]. This signal from the early Universe is measured from its temperature contrast against the CMB temperature [38]. Various physical processes occurring in the early Universe determine whether the signal is detected in absorption or emission. Since the signal evolves with redshift (or frequency), detecting its variation with redshift is equivalent to capturing various properties of the Intergalactic Medium (IGM) during these epochs. For observational detection of the signal, mainly two approaches are taken - observation of all-sky averaged 'global' signal and observing its tiny statistical fluctuations. Several ongoing, as well as next-generation radio astronomical interferometers, are targeting the power spectrum measurements as well as tomographic imaging of the IGM during these early epochs. Instruments that are currently operational like GMRT, LOFAR, MWA have so far placed upper limits on the brightness temperature distribution of the 21-cm signal from reionization [59, 63, 70]. There are several dedicated instruments for the detection of the 21-cm fluctuations from CD and EoR. This includes the Donald C. Backer Precision Array to Probe the Epoch of reionization (PAPER, [100]), the Hydrogen Epoch of Reionization Array (HERA, [101]); also, the MWA Phase 2 upgrade [103] includes 72 new short spaced tiles for EoR science. Recent upper limits to the EoR power spectrum have been placed from PAPER [102] and MWA [104]. However, several factors have so far prevented the actual detection of the power spectrum of the 21-cm emission from CD/EoR.

The key factors responsible for obscuring the signal are instrument chromaticity, the precision of data calibration, and the presence of bright foregrounds (compact as well as diffuse). The foreground sources include diffuse emission like synchrotron emission from the galaxy [105] as well as from low redshift clusters [108], free-free emissions from both the galaxy as well as extragalactic sources [109] and emission from faint radio-loud quasars [107]. Point sources are also a significant contributor to the contaminants [110]. These foregrounds have amplitudes, which are 4-5 orders of magnitude higher than the redshifted 21-cm signal along any line of sight [111–115]. Recovering the faint signal amidst the bright sea of foregrounds requires precision instrumentation as well as sophisticated algorithms. However, any strategy used is based on the one property of the foregrounds - their spectral smoothness. The 21-cm has a spectral shape in contrast to the foregrounds, which are assumed to be spectrally smooth [44]. This property of the foregrounds can be exploited to extract the signal of interest. There are three primary strategies for handling foregrounds - foreground avoidance, foreground suppression, and foreground removal [115–117, 123]. The redshifted 21-cm signal, as opposed to the spectrally smooth foregrounds, show rapid decorrelation over a frequency separation of $\sim$ 1 MHz. Hence, the isolation of the cosmological signal from the foregrounds (see [118] and references therein) may be plausible. Thus, although the signal of interest is extremely faint, with sensitive enough telescopes and a careful investigation of the foregrounds, it is recoverable. With the inception of the Square Kilometer Array (SKA), the expected observational sensitivity would be sufficient to detect the signal statistically [47,119]. However, since there is no full-proof strategy for dealing with foregrounds (all three of the above methods come with both advantages and shortcomings), it is still a trial-and-error quest for the telescopes to determine the best strategy for signal extraction given the bright foregrounds.

Nevertheless, for the determination of a perfect strategy, it is essential to have enough observational data to produce accurate models of the foreground sources. A number of models have predicted that the frequency range of the CD and EoR observation would lie between $\sim$ 50 MHz to $\sim$ 200 MHz, since the redshift range of interest is expected to lie between $6 < z < 30$ [44, 45]. At these frequencies, the dominant foreground is the Diffuse Galactic Synchrotron Emission (DGSE). It dominates at larger angular scales (or smaller baselines) where the EoR signal is also expected to have the maximum sensitivity. The longer baselines (or smaller angular scales) have major foreground contribution from extragalactic compact source populations. The discrete source population is mainly dominated by various classes of Active Galactic Nuclei (AGNs) and Star-Forming Galaxies (SFGs). For accurate foreground modeling, the nature and population of these sources in terms of their spatial, as well as spectral characteristics, become essential. Hence, low-frequency all-sky observation to determine the discrete source distribution becomes necessary. Using available data from deep field observations as well as physical models, various state of the art simulations are being performed to simulate source population and flux distribution at low frequency for characterizing foregrounds [120–122]. However, for more accurate predictions, more observational evidence is required. In addition to spectral smoothness of the foregrounds, their angular power spectrum can also be exploited for characterizing the sky signal, which consists of the 21-cm signal together with foregrounds [117, 124, 125]. The power spectrum of DGSE is modeled in a power-law form as a function of frequency and angular scale [126, 128, 129], expressed as :

$$\mathcal{C}_\ell(\nu) = A\left(\frac{\ell}{\ell_0}\right)^{-\beta}\left(\frac{\nu}{\nu_0}\right)^{-2\alpha}, \tag{4.1}$$

with β as the power-law index of the angular power spectrum of DGSE and α as the mean spectral index. For the foreground subtraction method, a power-law fit for the DGSE is done for each pixel along the frequency direction in the data cube and hence subtracted from the data [110, 113, 130, 131]. But precise modeling of this power law is required to prevent removal of the signal along with the foreground.

Low-frequency observation of radio sky is also essential for the study of the astrophysics at play in various evolutionary stages of different galactic and extragalactic sources. Radio emission at low frequencies, together with their redshift information, can be used to infer several astrophysical properties associated with the sources. In general, the source distribution is assumed Poissonian (with the possibility of clustering following single power-law) [114, 129, 132]. Source counts are also modeled via a single power-law distribution [133–135]. However, several recent studies have shown deviation from the single power-law model [136–138]. Thus more detailed studies both in wide-field as well as deep fields are required for generating a fiducial model of sources at low frequencies. The differential source counts at these are also useful for constraining the nature of sources. At frequencies in the GHz range and upwards, the source properties are well characterized. However, at lower frequencies, there is a lack of consensus for the same. It has been seen in previous studies that at these frequencies, AGNs dominate in the flux density scales down to few 100 μJy while SFGs and radio quite AGNs become dominant below $\sim$ 100 μJy [137–145]. This is inferred from the flattening of the source counts below 1 mJy. However, such studies are very few on account of the limitations in reaching the required SNR. Thus empirical constraints at frequencies $\lesssim$ 1.4 GHz are limited. Therefore the study of the low-frequency radio sky is vital for fiducial modeling of foregrounds for 21-cm cosmology as well as constraining the physics and the astrophysics of the sources.

This work presents the low-frequency properties of an extragalactic region at high galactic latitude. The field studied is the Lockman Hole region [146], the area with the lowest HI column density in the sky, having a low infra-red background [147]. Lockman Hole is one of the most extensively studied extragalactic fields, with a large number of studies in optical, X-ray, UV, and IR bands. Many studies of the region also exists in the radio frequencies, but mostly upwards of 1.4 GHz [137, 148–156]. At lower frequencies, observations have been done at 610 MHz by [157, 158] and at

150 MHz by [159]. There is a requirement for more observational analyses of this region at the lower frequency end to complement the pre-existing analyses. The Lockman Hole field, by virtue of its high galactic latitude, is very well suited for constraining the DGSE power spectrum (since the EoR target fields selected for current and upcoming telescopes are at high galactic latitudes).

In this work, 325 MHz data of the Lockman Hole region from the legacy GMRT is analyzed. At this observing frequency, the only other observation in literature is the VLA observation of [151]. For that work, 324.5 MHz observation with a $\sim 2.3°$ FoV centered at $10^h06^m01^s$ $+34^d54^m10^s$ was considered. The RMS sensitivity reached was 70 μJy beam^{-1}. For this work, a $6° \times 6°$ map of the field reaching an RMS level of 50 μJy beam^{-1} has been produced. Due to the large field of view covered (because of multiple pointings), the data has significant coverage, which enables the study of the variation of the DGSE power spectrum with latitude. A source catalog, containing sources with fluxes down to 250 μJy containing 6186 sources has also been generated.

The catalog thus produced has been compared with previous catalogs (mostly at higher frequencies) for the same region. Euclidean normalized source count has also been determined and compared to prior studies. The Tapered Gridded Estimator (TGE, [160, 161]) has been used to determine the angular power spectra for the DGSE using three separate pointings (within the observed field of view) located at three different galactic latitudes. This has been done to study the variation in the nature DGSE with increasing distance from the galactic center. The Lockman Hole region, despite having an extensive multi-wavelength coverage, nevertheless has some gap in the lower frequency end. However, for the characterization of extragalactic sources and determination of the astrophysics at play, more multi-wavelength coverage is essential. This paper aims to fill the low-frequency observational gap for better multi-frequency source characterization of astrophysical sources.

The paper starts by describing the observation and the data reduction process in section 4.2. Section 4.3 details out the sources extracted and cataloged after reducing the data. In section 4.4, comparison with existing observations has been discussed. The flux for the sources extracted from the cataloged is corrected for various factors and binned to obtain the Euclidian normalized source counts - which is described in section 4.5. Section 4.6 describes the variation of the angular power spectrum across the field. Finally, section 4.7 concludes the paper with a discussion of the implications of this work and the further studies required.

4.2 Observation and Data Reduction

This work uses the archival data of the Lockman Hole observed by Giant Metrewave Radio Telescope (GMRT, [162]). The observations were carried out during cycle 23 (Project Code - 23_001) with the legacy GMRT (GMRT Software Backend; GSB) at 325 MHz frequency band with an instantaneous bandwidth of 32 MHz. Details of the observing parameters are outlined in Table 4.2. Data were taken over 11 nights between February 15 and March 20, 2013, and has 23 different pointing centers. The field center of entire mosaic of 23 pointings is at ($\alpha_{2000} = 10^h48^m00^s, \delta_{2000} = 58°08'0''$). Two flux calibrators- 3C48 and 3C286 are observed respectively at the beginning and end of each night's observation because the former one would have set at the time each of the observing run ended. Between each set of target observations (each set is the successive observation of the 23 targeted pointings), there is an intermittent observation of the source 1006+349, which is the phase calibrator.

The choice of the above data is due to several interesting features. It is one of the very few data having a large field of view coverage at such low frequencies. There is also a lack of study at the frequency considered here. This region, being located quite far from the galactic plane (phase center is at $l = 149.18°, b = 52.24°$), presents an intriguing field for study and characterization of extragalactic astrophysical sources.

Table 4.1: Observational details of the target field and the calibrator sources for all the observing sessions

Project code	23_001‡
Observation date	15 & 18 February 2013
	4,5,6,7,11,12,13,14,20 March 2013
(as obtained from data)	
Bandwidth	32 MHz
Frequency range	309-341 MHz
Channels	256
Correlations	RR RL LR LL
Flux calibrator	3C48 & 3C286
Phase calibrator	1006+349
Pointing centres	$01^h37^m41^s$ $+33^d09^m35^s$ (3C48)
	$10^h06^m01^s$ $+34^d54^m10^s$ (1006+349)
	$11^h03^m18^s$ $+59^d44^m50^s$ (Target)
	$10^h55^m54^s$ $+59^d44^m50^s$ (Target)
	$10^h48^m29^s$ $+59^d44^m50^s$ (Target)
	$10^h41^m05^s$ $+59^d44^m50^s$ (Target)
	$10^h33^m41^s$ $+59^d44^m50^s$ (Target)
	$10^h59^m36^s$ $+58^d53^m55^s$ (Target)
	$10^h52^m12^s$ $+58^d53^m55^s$ (Target)
	$10^h44^m47^s$ $+58^d53^m55^s$ (Target)
	$10^h37^m23^s$ $+58^d53^m55^s$ (Target)
	$11^h03^m18^s$ $+58^d03^m00^s$ (Target)
	$10^h55^m54^s$ $+58^d03^m00^s$ (Target)
	$10^h48^m29^s$ $+58^d03^m00^s$ (Target)
	$10^h41^m05^s$ $+58^d03^m00^s$ (Target)
	$10^h33^m41^s$ $+58^d03^m00^s$ (Target)
	$10^h59^m18^s$ $+57^d12^m04^s$ (Target)
	$10^h52^m12^s$ $+57^d12^m04^s$ (Target)
	$10^h44^m47^s$ $+57^d12^m04^s$ (Target)
	$10^h37^m23^s$ $+57^d12^m04^s$ (Target)
	$11^h03^m18^s$ $+56^d21^m09^s$ (Target)
	$10^h55^m54^s$ $+56^d21^m09^s$ (Target)
	$10^h48^m29^s$ $+56^d21^m09^s$ (Target)
	$10^h41^m05^s$ $+56^d21^m09^s$ (Target)
	$10^h33^m41^s$ $+56^d21^m09^s$ (Target)
	$13^h31^m08^s$ $+30^d30^m32^s$ (3C286)

‡ as given in the cover-sheet accompanying the data

4.2.1 Data Reduction

Data were reduced using the Source Peeling and Atmospheric Modelling (SPAM) pipeline [133, 163–165]. It is a fully automated software based on the Astronomical Image Processing System (AIPS, [166, 167]). The primary rationale of using SPAM is that it does a "direction-dependent" calibration, besides the traditional "direction-independent" calibration. Direction dependent approach of SPAM iteratively solves for ionospheric phases and hence corrects for ionospheric phase errors. Such phase errors are expected to be persistent, given the large field of view as well as low frequencies used for observing the current data. Therefore, modeling and correction of the ionospheric dispersive delays improve the background noise and flux-scale accuracy. The data reduction steps have been described in brief below:

(a) **Pre-processing**: Initially, SPAM computes the instrumental calibration using the best available scan of the primary calibrator and calibrates the data using these solutions. The flux densities of the primary calibrators are set following the Scafie-Heald flux density scale [168]. After the initial RFI flagging and bad data removal, both the time-varying complex gain solutions and time constant bandpass solutions are computed per antenna per polarization. In the pre-processing step, data is also averaged in time and frequency to reduce data volume for subsequent calibration steps.

(b) **Direction Independent Calibration**: This step is akin to the classical self-calibration. It uses the NVSS derived sky model for performing the self-calibration on the data. The calibrated visibilities are imaged with wide-field imaging using Briggs weighting with robust parameter -1 (which produces very well-behaved point spread functions by down-weighting the dense central UV plane coverage of GMRT).

(c) **Direction Dependent Calibration**: Gain phases and the sky model generated in the previous step are used to initiate the direction-dependent calibration. SPAM determines the gain phases by peeling the apparently bright sources in the field. The phase correction factors generated by this process is a measure of the ionospheric phase delay. These solutions are applied to the data calibrated in the previous step. It creates of visibilities with correction for ionospheric phases. At the end of the direction-dependent step, the primary beam corrected image is obtained.

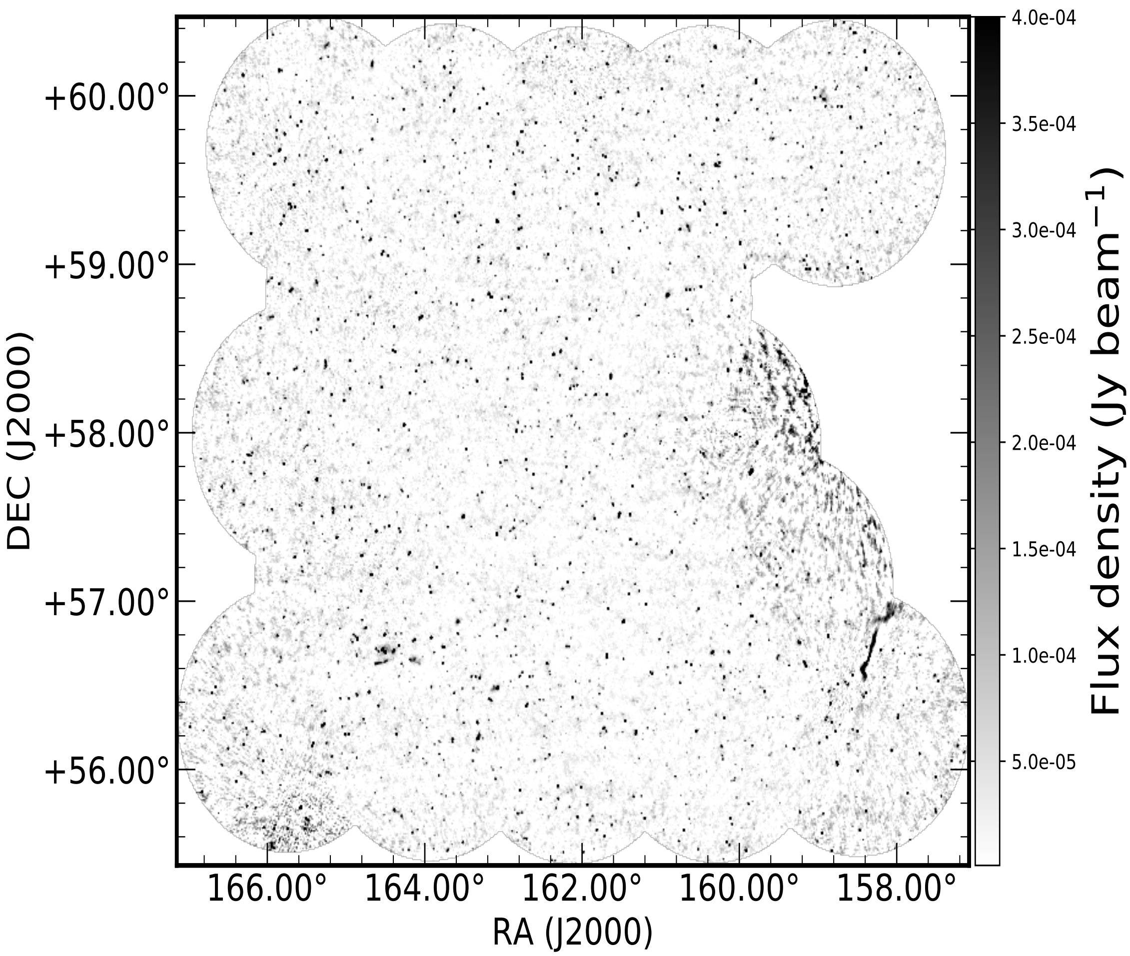

Figure 4.1: Primary beam corrected mosaic of the Lockman Hole region at 325MHz. The off source RMS at the center is $\sim 50\ \mu$Jy beam^{-1} and beam size is $9.0'' \times 9.0''$

For each observing night, the archival data is downloaded in LTA format, together with a FLAGS file containing all the information regarding the telescope system during the observation. SPAM was used to pre-calibrate the data for each night's observation. The individual pointings were processed separately, post concatenation of the data for all nights. Running the pipeline, as mentioned above, gives calibrated visibilities, which are both time and channel averaged. The visibility files thus produced consist of calibrated uvfits files (containing calibrated UV data) and residual uvfits files (containing residual visibilities). The averaged data has 42 channels, with width 781.25 KHz, i.e., the effective bandwidth remains 32 MHz. Each of such pointings was individually imaged, to produce primary beam corrected image. The primary beam corrected images of each pointing were then mosaiced to create the final image. The image is $6° \times 6°$, as shown in Figure 4.1. It is to be mentioned that two pointings (located off the center) had to be removed due to the presence of substantial noise in the final PB-corrected image. In Figure 4.2, the zoomed-in image of the phase center is shown. The off-source noise is $\sim 50\mu$Jy beam^{-1}, covering almost $6° \times 6°$ with a beam size of $9'' \times 9''$.

4.3 Source Catalog

A source catalog was produced using Python Blob Detection and Source Finder [1](PyBDSF, [169]) for characterization of the sources present in the field. PyBDSF also produced a residual map, which is the image of the field after subtracting all the modeled point sources. The source catalog was generated using the primary beam corrected mosaiced image of the field. PyBDSF uses a sliding box window to calculate the RMS variation across the field. For this work, a box of size 120 pixels every 30 pixels was used, i.e., rms_box = (120,30). Using a threshold of $100\sigma_{\rm RMS}$, bright regions with peak amplitude higher than this value were first isolated (where$\sigma_{\rm RMS}$ is the clipped RMS across the entire map). A smaller box of size rms_box_bright

[1] https://www.astron.nl/citt/pybdsf/

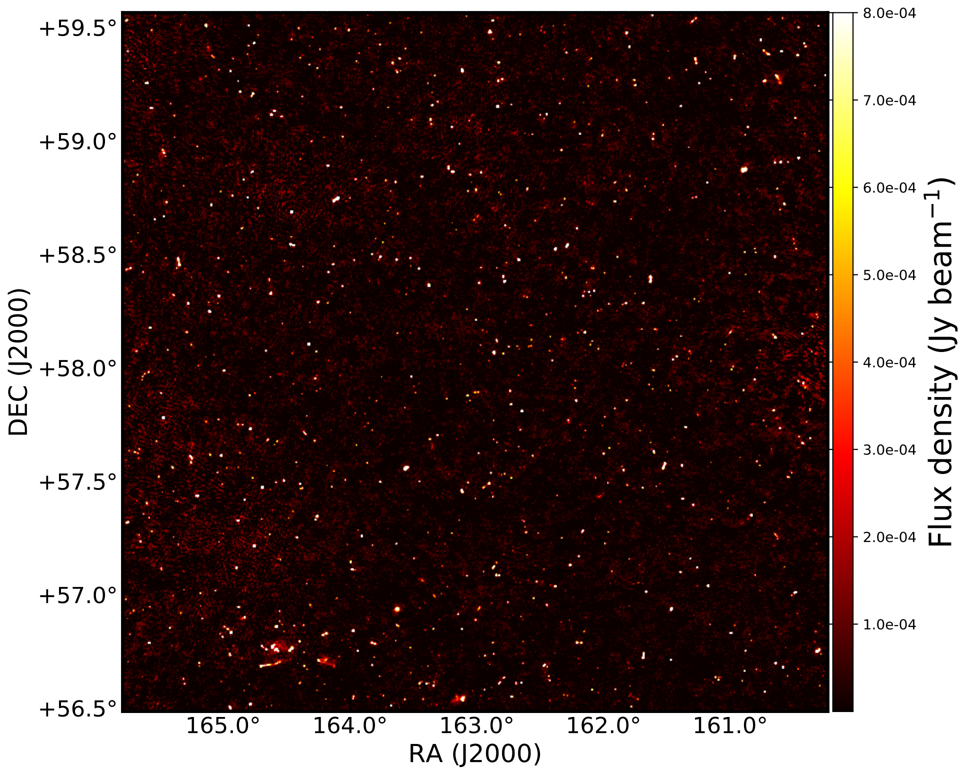

Figure 4.2: Zoomed-in Stokes I image of the central region of Lockman Hole at 325MHz, covering an area of ~ 5 deg^2. The Central off-source noise is ~ 50 μJy beam^{-1}

= (30,10) was used around the bright regions to avoid considering artifacts around these regions as real sources. It identifies islands that have contiguous emission and then fits multiple Gaussians to each island. The selection threshold for islands is $3\sigma_{\rm RMS}$ and the same for source detection is $6\sigma_{\rm RMS}$.

The sources are identified by PyBDSF by grouping neighboring Gaussians into sources. The total flux of a source is the sum of all fluxes in a group of Gaussians. The flux uncertainty is the quadrature sum of individual uncertainties of the Gaussians. The source position is the centroid, while the spatial size is determined using moment analysis using restoring beam size.

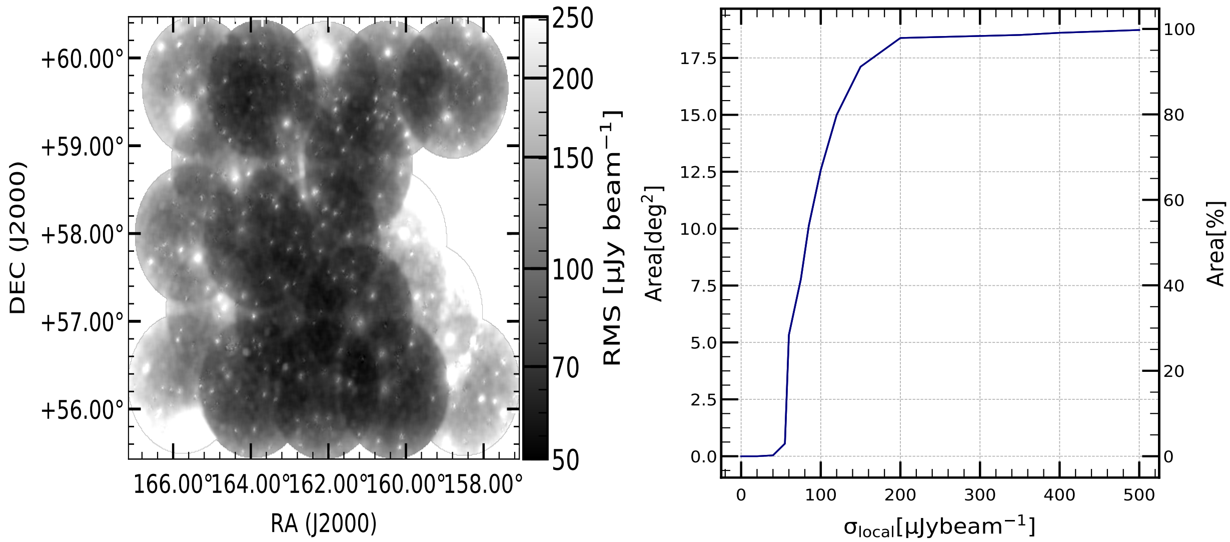

Figure 4.3: (**Left**) Local RMS noise in the final mosaic; (**Right**) Cumulative area of the final mosaic as a function of the RMS noise

The ionospheric fluctuation over the large field of view, given the low observing frequency, would considerably vary the PSF. Thus the actual PSF may be slightly different than the restoring beam at different parts of the image. Setting the option psf_vary_do = True in PyBDSF takes care of the PSF variation across the field of view (readers are referred to the PyBDSF documentation available from the link in footnote 1 for more details). An RMS map was also produced that depicted variation of the background noise across the field. As can be seen in Figure 4.3 (left), the background RMS is quite high near the bright sources and particularly high towards the edges of the field.

Every source catalog generated, besides being resolution and thermal noise limited, is also limited by confusion noise. Confusion noise is defined as the background sky brightness fluctuation due to several faint sources in one telescope beam.

Using the formula described in [153],

$$\sigma_c = 1.2\mu\text{Jybeam}^{-1}(\frac{\nu}{3.02GHz})^{-0.7}(\frac{\theta}{8''})^{10/3} \tag{4.2}$$

(where ν is observing frequency, and θ is the FWHM of the telescope beam[2]) the confusion noise limit is 8.46 μJy beam^{-1}.

The catalog assembled in this work comprises of 6186 sources above the $5\sigma_{RMS}$ threshold. A sample of the catalog is shown in Table 4.2 (full catalogue is available with the electronic version of the paper).

4.3.1 Classification

It is not very straightforward to classify the sources detected into resolved and unresolved (point) categories by just using the derived source properties. Time and bandwidth smearing might extend the sources artificially in the image plane. Also, calibration errors, if any, as well as variable noise, may scatter the ratio of integrated flux density ($S_{\rm int}$) to peak flux density ($S_{\rm peak}$). Consequently, sources cannot be classified into point and resolved by mere use of $\left(S_{\rm int}/S_{\rm peak}\right) > 1$. In Figure 4.4, $\left(S_{\rm int}/S_{\rm peak}\right)$ as a function of $\left(S_{\rm peak}/\sigma_L\right)$ is plotted, where σ_L is the local RMS.

For point sources, assuming that $\sigma_{\rm S_{peak}}$ and $\sigma_{\rm S}$ are independent (where $\sigma_{\rm S_{peak}}$ and $\sigma_{\rm S}$ are the uncertainties in the peak flux density and integrated flux density respectively), $ln\left(S/S_{peak}\right)$ follows a zero mean Gaussian distribution, with RMS given by

$$\sigma_{\rm R} = \sqrt{\left(\frac{\sigma_{\rm S}}{S_{\rm int}}\right)^2 + \left(\frac{\sigma_{S_{\rm peak}}}{S_{\rm peak}}\right)^2} \tag{4.3}$$

An extended source is detected at the 3σ level iff R $>3\sigma_R$ [170]. Using this criterion, a total of 1825 sources are resolved (magenta dots), while the remaining 4460 sources are point sources (turquoise dots) as shown in Fig 4.4.

[2] https://www.cv.nrao.edu/course/astr534/Radiometers.html

Table 4.2: Sample of 325 MHz source catalog of Lockman Hole region generated

ID ()	RA (deg)	E_RA (arcsec)	DEC (deg)	E_DEC (arcsec)	Total_flux (mJy)	Peak_flux (mJy beam^{-1})	RMS (mJy beam^{-1})	Maj (arcmin)	Min (arcmin)	PA (deg)
0.	167.22	0.09	56.52	0.13	11.99	9.11	0.23	0.18	0.14	115.13
1.	167.15	0.31	56.28	0.26	4.39	3.55	0.23	0.19	0.15	63.28
2.	167.14	0.38	56.39	0.43	3.85	2.67	0.22	0.2	0.17	36.15
3 .	167.08	0.08	56.62	0.06	42.18	23.27	0.26	0.23	0.18	25.85
4.	167.08	0.26	56.27	0.2483	5.37	3.96	0.21	0.19	0.16	133.3
5.	167.33	1.15	58.00	0.65	1.81	1.09	0.18	0.23	0.16	108.93
6.	167.31	0.18	58.65	0.10	12.32	6.94	0.17	0.24	0.17	87.42
7.	167.07	0.09	56.65	0.09	14.09	11.2	0.23	0.17	0.16	61.47
8.	167.07	0.19	58.16	0.25	12.15	6.15	0.23	0.23	0.19	24.84
9.	167.28	0.15	56.05	0.26	9.51	4.64	0.16	0.25	0.18	2.54

Notes: The columns of the final catalog (fits format) include source ids, positions, error in positions, flux densities, peak flux densities, local RMS noise, sizes, and position angle respectively.

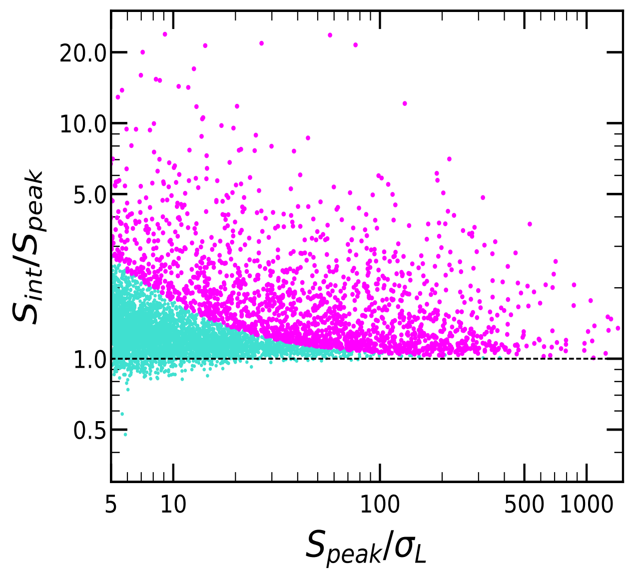

Figure 4.4: Ratio of of integrated-to-peak flux as a function of the source SNR. The teal colored dots are unresolved sources and the magenta filled dots represent resolved sources

4.4 Comparison with other radio catalogs

This section describes the cross-validation of this work with other radio catalogs covering the same region. [157] have studied the same area of the sky using 610 MHz observation of the legacy GMRT. There are also several other studies covering various parts of the Lockman Hole region at radio frequencies. Amongst those, the studies by [159] using 150 MHz LOFAR data and by [137] using 1.4 GHz Westerbork Synthesis Radio Telescope (WSRT) data have been considered for comparison. The large area VLA Faint Images of the Radio Sky at Twenty-Centimeters survey (FIRST survey) [171] has also been used for validation.

The cross-validation of the source catalog with previous findings is necessary since the ionospheric fluctuations are significant at low-frequency. This will distort the source position and smear the sources residing at a large distance from the phase center. Thus comparison with existing literature allows us for quantification of

Table 4.3: Flux limits of the catalogues considered

catalog	Frequency (MHz)	Resolution (arcsec)	$S^{\dagger}_{\mathrm{limit}}$ (mJy)	$S_{\mathrm{cut,325MHz}}$ (mJy)
GMRT (this work)	325	9.0”	0.25	0.25
FIRST [171]	1400	5.4”	1.0	2.779
LOFAR [159]	150	14.7”	2.000	1.164
GMRT [157]	610	6.0”	0.556	0.864
WSRT [137]	1400	11.0”	0.070	0.194

† S_{limit} is the flux density limit of the corresponding catalog.

any systematic offsets in flux densities as well as in source positions. Using a 5” search radius in the other catalogs, counterparts of the sources presented in this catalog have been identified. Each catalog has a flux density limit depending on the observation sensitivity and completeness. The flux limit of a given catalog was scaled to 325 MHz using $S_{\nu} \propto \nu^{\alpha}$, where α = -0.8 and denotes that limit at 325 MHz as $S_{\mathrm{cut,325MHz}}$. Only those sources with flux densities higher than this flux cut-off were chosen. Table 4.3 enlists the resolution of each of the chosen catalogs, their corresponding flux limits, and the equivalent 325 MHz cut-off.

4.4.1 Flux Density Offset

Different catalogs set the flux scales following different flux density scales. Here, the Scaife-Heald flux scale [168] has been used to set the flux of sources. This scale has also been used in LOFAR 150 MHz data of [159]. For WSRT 1.4 GHz data of [137], the FIRST catalog [171] and GMRT 610 MHz observation of [157], flux standard of

[172] was used. Due to uncertainties in the flux scale used, as well as in modeling of the primary beam, systematic offsets may arise in the flux density of the sources. The flux densities of cataloged sources have been compared with the catalogues mentioned above to check for such systematic offsets. The source selection criterion is based on [136], where only high SNR sources $\left(S_{\text{peak}} > 10\sigma\right)$ have been selected for comparison. Further, the sources are selected to be "compact." This condition implies that the sources have sizes below the resolution at a higher frequency. There was another additional constraint on the flux limit. The catalogues used have a flux limit; for instance in [157] it is 556μJy which corresponds to 920μJy at 325 MHz (assuming α = -0.8). Only sources above such limits (tabulated in Table 4.3 for all the catalogues considered) were selected for analysis.

The flux density ratios, after proper scaling of the fluxes, are calculated $\Big(S_{\text{325MHz}}/S_{\text{others}}$, where others mean the previous catalogues used$\Big)$. The median of the ratio comes out to be $0.99^{+0.3}_{-0.5}$ (errors from the 16th and the 84th percentiles) with the 610 MHz data. With FIRST, 150 MHz and 1.4 GHz data, the median values of the flux ratios are $0.97^{+0.5}_{-0.5}$, $0.82^{+0.3}_{-0.6}$, $0.90^{+0.3}_{-0.6}$ respectively. This is shown in the left panel of Figure 4.5, where it is observed that the ratio is nearly 1 for most cases, showing the reliability of the fluxes obtained.

For further validation of the flux scale reliability, sources common to the 610 MHz GMRT catalogue [157] and the 150 MHz LOFAR catalogue [159] and the catalogue obtained in this work have been selected (satisfying the same criterion described previously). The spectral indices are calculated using the flux densities obtained from the other two catalogues. The mean value for the spectral index is 0.83. Using this value of the spectral index, the fluxes are calculated for the 325 MHz catalogue. Comparison with the actual flux values obtained in the catalogue with the predictions obtained using the value of spectral index gives flux ratios with a median value 1.04 and a standard deviation of 0.43. This has been illustrated in the

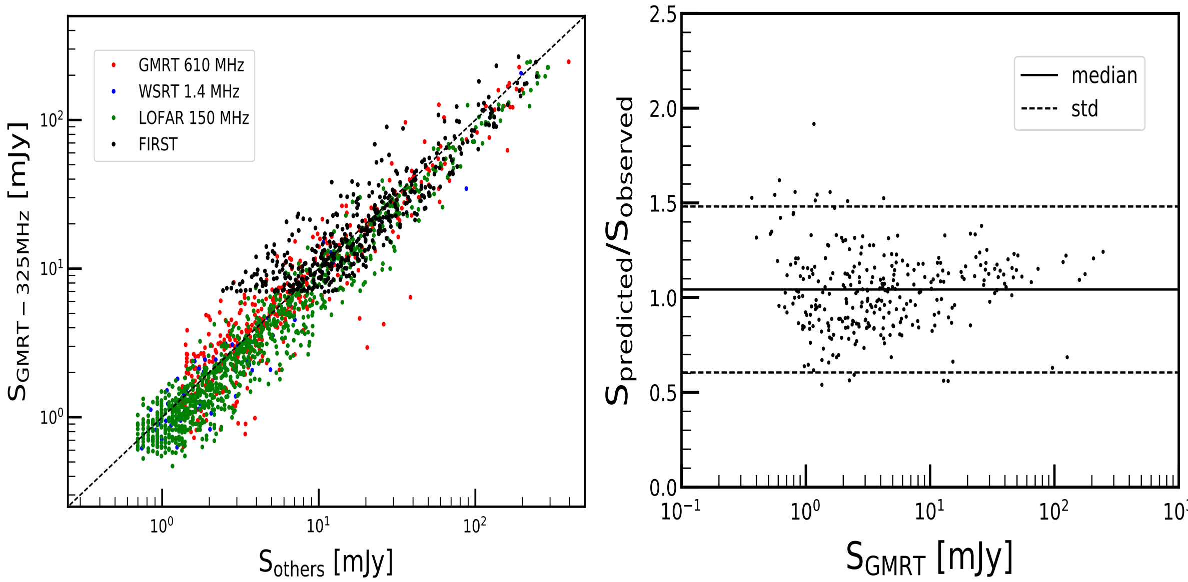

Figure 4.5: **(left)** Comparison of integrated flux densities of compact sources measured with GMRT at 325MHz with other catalogues: 610MHz GMRT(red), 1.4GHz WSRT(blue), 150MHz LOFAR(green). Fluxes have been scaled using α=l-0.8 The black dashed line corresponds to $S_{\rm GMRT}/\ S_{\rm others} = 1$. **(right)** Comparison of flux densities obtained from the 325 MHz data in this work with flux densities predicted using spectral densities obtained from [157] (610 MHz) and [159] (150 MHz). The solid black line indicates the mean ratio of the predicted and observed flux and the standard deviation of the ratios is indicated by the dashed black lines.

right panel of Figure 4.5. As it is clearly seen in the figure, the flux scale used for obtaining source fluxes for this work is reasonably accurate.

4.4.2 Positional Accuracy

The astrometric accuracy of the source positions obtained is verified by comparing the source positions of selected sample to the FIRST, LOFAR 150 MHz catalog and GMRT 610 MHz catalog. The source selection criteria also remain the same as previously mentioned in 4.4.1. FIRST, being at higher frequency (1.4 GHz), has a resolution 5.4” and faces lesser ionospheric fluctuation. All such reasons combined make the position accuracy of FIRST better than 1” [171]. The offsets have been

Table 4.4: Median values of the deviation (along with the 16th and 85th percentile errors) of RA and DEC of GMRT 325MHz source catalogue from other catalogues

catalog	Frequency MHz	$\delta_{\rm RA,median}$ (arcsec)	$\delta_{\rm DEC,median}$ (arcsec)
FIRST	1400	$-0.244^{+1.2}_{-1.0}$	$0.600^{+0.8}_{-0.9}$
GMRT	610	$-0.063^{+0.9}_{-0.8}$	$0.562^{+0.8}_{-0.6}$

set as (following [136]) :

$$\begin{aligned} \delta_{\rm RA} &= {\rm RA}_{{\rm GMRT}_{325}} - {\rm RA}_{\rm FIRST} \\ \delta_{\rm DEC} &= {\rm DEC}_{{\rm GMRT}_{325}} - {\rm DEC}_{\rm FIRST} \end{aligned} \tag{4.4}$$

The value of the median offsets with respective errors have been shown in Table 4.4.

Figure 4.6 shows the histogram of offset values from the other observations considered. No systematic deviation across the field of view has been observed. However, the offsets being lesser than the 2" (i.e., the cell size of the image), there is no significant astrometric error that may be caused by it. The final catalog source positions have been corrected by a constant value derived from the offset from FIRST (the median offset value).

4.4.3 Spectral Index Distribution

The present data covers a wide area in the Lockman Hole region with a large number as well as a variety of sources. The spectral properties of the sources are characterized by comparing the fluxes derived for this work with other available data at higher frequencies. The data used are 610 MHz data of [157] and 1.4 GHz FIRST catalog. The source selection procedure is the same as that in section 4.4.1. The number of sources used to obtain the spectral indices is 511 for GMRT data and 1696 for FIRST. Assuming a synchrotron like power-law distribution, $S_\nu \propto \nu^\alpha$, where α is the spectral index, the flux densities of the matched sources were used to

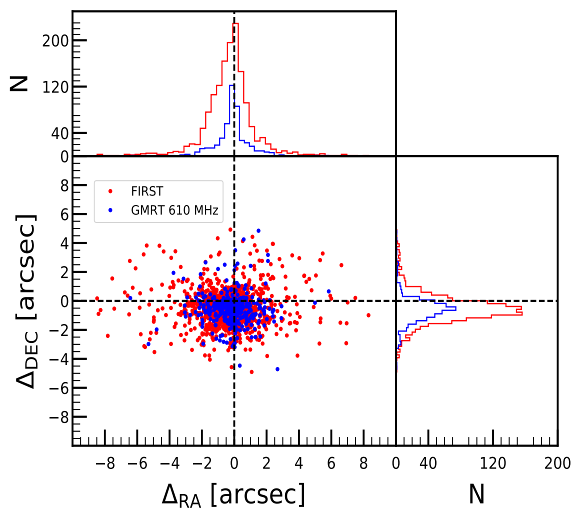

Figure 4.6: Offset of the source RA and Dec for the 325 MHz catalog from FIRST catalog (red) and GMRT 610 MHz catalog (blue).

estimate the value of α. The distribution of spectral indices of matched sources is depicted in Figure 4.7. For this work, as is seen in figure 4.7, the normalized counts are highest between the range $\sim$ -1.3 to $\sim$ -0.4. The median values of the spectral indices with errors from the 16th and 84th percentile are $-0.860^{+0.621}_{-0.578}$, $-0.708^{+0.344}_{-0.506}$ for GMRT 610 MHz and FIRST catalogues respectively. Hence, the median value for the spectral index of this catalog after matching can be taken as $\sim$ -0.8. Several low frequency radio observations for the Lockman Hole region has spectral index distribution with $|\ \alpha\ | \leq 2.0$ (for instance [152, 157, 159]). Low frequency studies for other deep fields like COSMOS [173], Boötes [136], ELAIS-N1 [175, 176] also mostly follow this distribution (except a few cases seen in [173] (Figure 14) & [176] (Figure 14)). But it can be seen in Figure 4.7 that the distribution of α is wider than what is usually observed. The exact reason for this is unknown. However, since a detailed study of spectral index requires multi-wavelength as well as wider bandwidth data, it is deferred to future works.

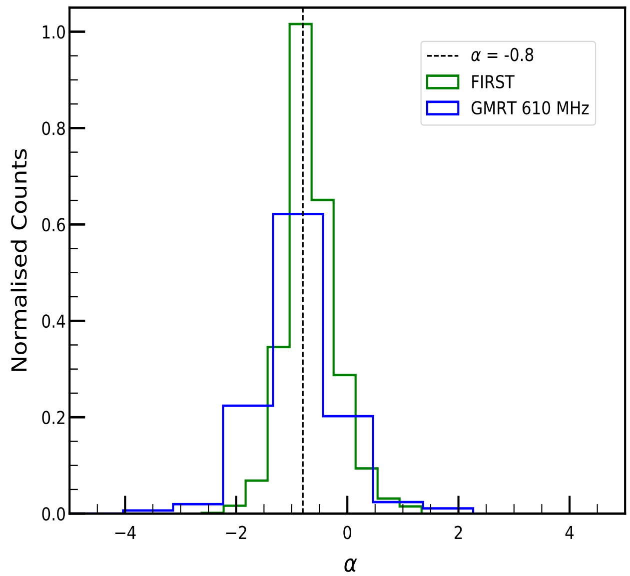

Figure 4.7: Normalised counts of the measured spectral indices in the field, obtained after matching using a 5" match radius with FIRST(green) and GMRT 610 MHz data (blue). Black dashed line corresponds to spectral index -0.8

4.5 Source counts

This section presents the differential source counts based on the flux densities arising in the PyBDSF generated outputs. At low frequencies, the most dominant source population at the faint flux end is that of SFGs and Radio-quiet Quasars (RQQs). This has been well established from various observations as well as simulations ([121, 122], and references therein). Nevertheless, observational constraints on sub-mJy source populations are very few. However, for telescopes like MWA, SKA, or LOFAR, which aim to detect the faint cosmological HI 21-cm signal, the sources from sub-mJy flux levels down toμJy level act as foregrounds which can potentially obscure the signals. Hence, characterizing the spatial and spectral nature of sources with flux densities from sub-mJy to μJy, especially at low frequencies, is essential for foreground characterization. The determination of source counts, i.e., the source distribution at the flux ranges of interest, is one of the significant steps for such characterization.

The differential source counts at 325 MHz have been measured with sources having flux densities down to 0.2mJy ($\approx 4\sigma$). However, this does not give the correct distribution, since the PyBDSF output has errors arising out of catalog incompleteness, resolution bias, false detection, Eddington bias. The problem is especially prevalent at low frequencies and for the faint end of the flux bins. The following subsections describe the corrections made to the source count distribution in detail.

4.5.1 False Detection Rate

False Detection Rate (FDR) is simply the number of spurious detection by the source finder. This occurs either due to noise spikes or due to the presence of bright artifacts in the image. If noise distribution is symmetric about the mean, i.e., positive noise spikes have equivalent negative peaks in the image, the number of spurious detection would be equal to the number of negative sources in the inverted image.

To quantify this false detection rate, PyBDSF was run on the inverted image with parameters identical to that used in the original image. This yielded a total of 71 sources with negative peaks below -5σ. Now, for FDR correction to the flux density bins, the negative sources are binned in the same manner as the actual sources and compared to the sources detected in the original image. The number of real sources in each bin (quantified as a fraction) is [138]:

$$f_{\mathrm{real},i} = \frac{N_{\mathrm{catalog},i} - N_{\mathrm{inv},i}}{N_{\mathrm{catalog},i}}, \tag{4.5}$$

where $N_{\mathrm{inv},i}$ and $N_{\mathrm{catalog},i}$ are the number of detected sources in i^{th} flux density bin for inverted and original image respectively. The errors are quantified as Poisson errors; the correction factor is multiplied with the number of sources in each flux density bin of the original catalogue. The correction factors obtained for each bin is shown in Figure 4.8.

4.5.2 Completeness

The incompleteness of a catalog is the inability to detect sources lying above the flux limit of the catalog. This is mainly due to noise variation in the image. The source catalog constructed using the source finder algorithm (PyBDSF) is completeness limited. Incompleteness can cause both underestimation as well as overestimation of the source counts derived from the image plane. Eddington bias [177] causes scattering of higher count bins into lower count ones more than the reverse, subsequently overestimating the latter. Resolution bias is another factor limiting catalog completeness. It arises because the detection probability is lesser for resolved sources than unresolved ones, resulting in an underestimation of source counts.

Simulations were performed on the image plane to quantify and correct for the above biases. 3000 sources were artificially injected in the residual RMS map (same as [136]) using the open-source software Aegean[3] [180,181]. Of these, 1000 are extended sources (major and minor axes greater than 9") while the rest are unresolved point sources. This number distribution is done following the actual source classification in the catalog, where ∼30% are resolved. The flux density is generated using power-law distribution of the form dN/dS $\propto S^{-1.6}$ [178, 179]. The flux values are chosen randomly and constrained between 200μJy to 4100μJy. The source positions were also randomly chosen, spanning the entire right ascension and declination range of the catalog. Following [175], 100 different realizations of the said simulations were done. The simulations take into account the visibility area effects and source confusion limitations [135, 136, 138].

The sources are extracted separately from each image using PyBDSF, setting the same parameters as described in Section 4.3. The recovered sources are binned into the same number of bins as the actual catalog. The correction factor is calculated as:

[3] https://github.com/PaulHancock/Aegean

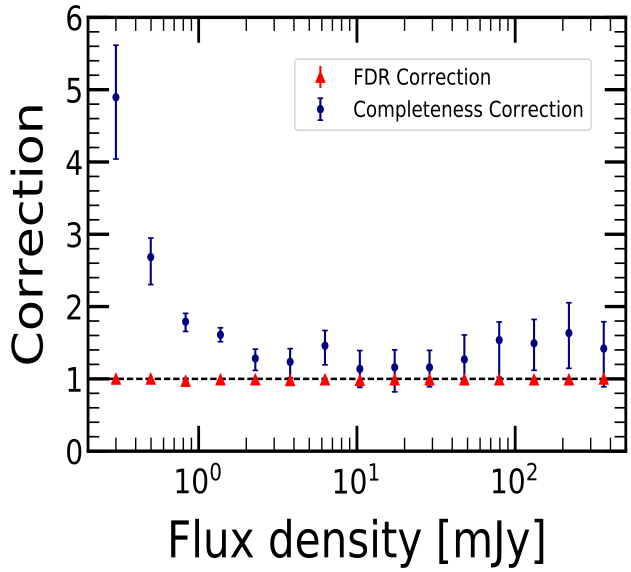

Figure 4.8: Correction factors for false detection (red triangles) and completeness (blue circles).

$$\text{Correction}_{,i} = \frac{N_{\text{injected},i}}{N_{\text{recovered},i}} \tag{4.6}$$

where, $\text{Correction}_{,i}$ is the completeness correction factor in the i^{th} flux density bin $N_{\text{injected},i}$ is the number of injected sources and $N_{\text{recovered},i}$ is the number of sources recovered after subtracting original pre-simulation sources in the i^{th} bin [138].

Figure 4.8 show the FDR and completeness correction factors for each of the flux bins. The median value of each flux bin from the 100 simulations is taken as the correction factor, while the errors are those associated with the 16th and 85th percentiles.

4.5.3 Differential Source Count

The Euclidean-normalised differential source counts have been estimated from the PyBDSF generated source list, after correcting for incompleteness and FDR. The correction factor for each bin is multiplied to the respective uncorrected source counts. Since noise is variable across the image (see Figure 4.3), correction is also done for effective area per bin (i.e., area over which a source can be detected). To determine the effective area per bin, the fraction(f) of the area over which a source having a specific flux density is detectable has been determined. The source count in each bin is then weighted by f^{-1} [182]. The fluxes have been binned into 18 logarithmic bins down to 0.30mJy (6σ), and Poisson errors on the source counts have been evaluated. The source counts and associated errors are given in Table 4.5. Post incorporation of the required corrections, the normalized differential source counts have been plotted in Figure 4.9. Comparison has also been made with state of the art simulations as well as other observed counts.

4.6 Power Spectrum of Diffuse Emission

Diffuse Galactic Synchrotron Emission (DGSE) is one of the major contributing factors to the foregrounds, which obscure the cosmological HI signal from reionization and cosmic dawn. It remains dominant even after modeling and removal of the point sources from the data-sets. However, the spectrally smooth nature of the emission may be exploited for isolating it from the data to extract the faint cosmological target signal. Therefore, detailed knowledge of the spectral as well as spatial characteristics is essential for reasonable removal of foregrounds (i.e., removing it without removing the signal).

The spectrally smooth nature of the DGSE is captured by modeling it as a power law in both frequency and angular scales (Equation 4.1). Observational measurements of the angular power spectrum of DGSE for various fields have constrained its power-law index in the range [1.5 to 3.0] [129, 183–186]. Another measurement of the angular power spectrum of galactic synchrotron was done by [188] using the 408 MHz Haslam map [187] and 1.4 GHz map of [189]. The spectral index obtained from the aforementioned sources (which fall in the range [2.9 to 3.2] at different

Table 4.5: Euclidian-normalized differential source counts for the Lockman Hole field.

S (mJy)	S_c (mJy)	N	$S^{2.5}$dN/dS ($\mathrm{Jy}^{1.5}\mathrm{sr}^{-1}$)	FDR	Completeness	$S^{2.5}$dN/dS ($\mathrm{Jy}^{1.5}\mathrm{sr}^{-1}$)
0.3 - 0.5	0.4	260	4.9 ± 0.1	1.0 ± 0.0	$4.9^{+0.7}_{-0.9}$	24.7 ± 0.05
0.5-0.8	0.7	1315	9.3 ± 0.2	1.0 ± 0.0	$2.7^{+0.3}_{-0.4}$	25.1 ± 0.06
0.8 - 1.4	1.1	1316	14.9 ± 0.4	0.97 ± 0.0	$1.8^{+0.1}_{-0.1}$	25.9 ± 0.05
1.4 - 2.3	1.8	906	19.9 ± 0.7	0.99 ± 0.0	$1.6^{+0.1}_{-0.1}$	31.8 ± 0.06
2.3 - 3.8	3.0	680	31.4 ± 1.2	0.99 ± 0.0	$1.3^{+0.1}_{-0.2}$	39.9 ± 0.15
3.8 - 6.3	5.0	536	53.0 ± 2.3	0.98 ± 0.0	$1.2^{+0.2}_{-0.2}$	64.2 ± 0.42
6.3 - 10.4	8.4	337	70.9 ± 3.9	0.99 ± 0.0	$1.5^{+0.2}_{-0.3}$	102.4 ± 0.81
10.4 - 17.3	13.9	258	116.1 ± 7.2	0.98 ± 0.0	$1.1^{+0.3}_{-0.3}$	129.5 ± 1.84
17.3 - 28.8	23.0	201	193.6 ± 13.7	0.99 ± 0.0	$1.2^{+0.2}_{-0.3}$	222.1 ± 3.3
28.8 - 47.7	38.3	120	247.2 ± 22.6	0.99 ± 0.0	$1.2^{+0.2}_{-0.3}$	283.7 ± 5.3
47.7 - 79.3	63.5	86	379.0 ± 40.9	0.99 ± 0.0	$1.3^{+0.3}_{-0.3}$	476.2 ± 13.9
79.3 - 131.6	105.4	63	593.9 ± 74.8	0.99 ± 0.0	$1.5^{+0.3}_{-0.5}$	903.2 ± 18.8
131.6 - 218.5	175.1	44	887.4 ± 133.8	0.99 ± 0.0	$0.8^{+0.2}_{-0.2}$	1310.8 ± 44.2
218.5 - 362.8	290.6	23	1035.5 ± 211.4	0.99 ± 0.0	$1.5^{+0.3}_{-0.4}$	1674.2 ± 88.8
362.8 - 602.3	482.5	20	1846.1 ± 412.8	1.00 ± 0.0	$1.4^{+0.4}_{-0.5}$	2619.6 ± 153.2
602.3 - 1000.0	801.2	4	789.9 ± 394.9	0.99 ± 0.0	$1.5^{+0.5}_{-0.4}$	1205.8 ± 190.8
1000.0 - 2018.5	1509.3	5	1877.9 ± 839.8	0.97 ± 0.0	$1.6^{+0.5}_{-0.9}$	2992.7 ± 410.7
2018.5 - 4074.3	3046.4	2	2154.1 ± 1523.2	0.99 ± 0.0	$1.5^{+0.5}_{-0.9}$	3098.6 ± 689.9

Notes: This table includes the flux density bins, central of flux density bin, the raw counts, normalized source counts, False Detection Rate (FDR), completeness and corrected normalized source counts.

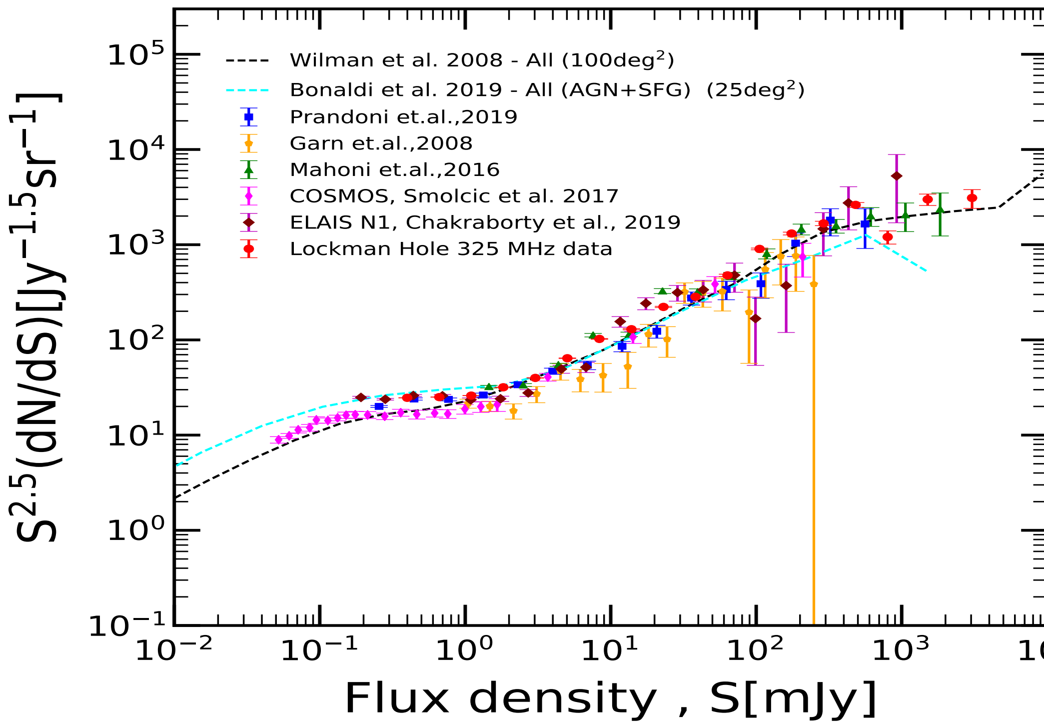

Figure 4.9: Euclidean normalised source counts for the 325 MHz data (shown in red circles) after correcting for false detection and incompleteness . It has been compared against source counts modelled by simulations - S^3 [121] plotted as black dashed curve and T-RECS [122] plotted as cyan dashed curve. Comparison has also been shown with other observation covering the same region of the sky - WSRT 1.4 GHz [137] (blue squares), GMRT 610 MHz [157] (orange pentagons) & LOFAR 150 MHz [159] (green triangles). Comparison has also been done with source counts for VLA COSMOS 3 GHz survey [173] (magenta diamonds) & GMRT 400 MHz observation of ELAIS N1 [175](maroon diamonds)

galactic latitudes), were used to extrapolate up-to 23 GHz using a single power-law fit. However, in [175], using wideband GMRT data, they have shown that the spectral nature of the DGSE has a hint of a broken power law.

The angular power spectrum for the field was quantified using TGE, [160, 161]. It estimates the angular power spectrum using the correlation between grided visibilities. Bright sources near the edge of the FoV can make the spectrally smooth diffuse foreground oscillate, thereby making it challenging to remove. Also, if side lobes of these sources are near the nulls of the primary beam of the telescope, complications arise in extracting the signal. TGE handles this by tapering the primary beam much before the first null is reached. It is also computationally efficient since it grids the visibilities before computing the power spectrum.

The estimator ($\hat{\mathrm{E}}_{\mathrm{g}}$) is defined as :

$$\hat{\mathrm{E}}_{\mathrm{g}} = M_g^{-1}\Big(|\mathcal{V}_{cg}|^2 - \Sigma_i|\tilde{\omega}_g(\mathbf{U_g} - \mathbf{U_i})|^2|\mathcal{V}_i|^2\Big) \tag{4.7}$$

where $\mathcal{V}_{cg}$ is the convolved visibility at every grid point g, $\mathcal{V}_i$ is the measured visibility, $\tilde{\omega}_g$ is the Fourier transform of the window function used for tapering the sky response $\mathbf{U_g}$ refers to baseline corresponding to the grid point g and M_g^{-1} is a normalization factor (refer to [161] for further details). The tapering function used is a Gaussian window function of the form $\mathcal{W}(\theta) = \exp(-\frac{\theta^2}{\theta_{\mathrm{w}}})$, where $\theta_w = \mathrm{f}\theta_0$ ($\theta_0 = 0.6 \times \theta_{FWHM}$ of the telescope primary beam).

The residuals obtained from visibilities calibrated by SPAM should contain only diffuse emission (considering perfect modeling and removal of discrete sources). However, since there are imperfections in the modelling, there will be residual(unsubtracted) point sources with the diffuse emission. To determine the angular power spectrum of the DGSE, TGE was applied to the residual visibilities (obtained after subtracting the point source contribution from the calibrated visibilities). The 23 pointings are spread over a galactic latitude coverage of $\sim 3°$, from which three pointings were chosen at three different galactic latitudes to characterize the variation of DGSE. The values used for determination of $\mathcal{W}(\theta)$ are f=1.0 and θ_0=44' for all the three pointings. Figure 4.10 shows the angular power spectrum ($\mathcal{C}_\ell$) as a function of the angular mode ℓ plotted for galactic latitudes $b \sim 50°$, $52°$ & $55°$. Power-law of the form $A\ell^{-\beta}$ has been fitted to the residual data choosing an ℓ range where there is a plunge in the amplitude of power spectrum. The best fit values for A and corresponding power-law index β is shown in the plots and also indicated in Table 4.6.

The power-law indices obtained from the best fit values are $\beta = 2.15$, 2.27 & 3.15 at $b = 49.8^\circ$, 52.4° & 54.9° respectively. The fall of the $\mathcal{C}_\ell$ values as a function of ℓ is consistent with the nature expected for DGSE. The ℓ range for fit, along with other details of the pointing, is shown in Table 4.6.

The value of U_{min} (corresponding to shortest baseline) for the three pointngs are 50λ, 80λ and 40λ. These values translate to ℓ_{min} 314, 502, and 251, respectively. But these values are ideal; estimating values of ℓ_{min} without any bias require consideration for the convolution effects of the primary beam, tapering window, actual uv-coverage (see equation 6 of [161]). The value of ℓ_{min} for each field was chosen conservatively from a careful inspection of the obtained power spectrum. In [183], it was shown that at larger angular scales, since convolution effects become important, it may lead to the drop-off in the power seen at ℓ below ℓ_{min}.

DGSE dominates the power at small angular scales up to a specific ℓ_{max}. This value of ℓ_{max} was chosen by visual inspection at the ℓ, where the power-law behavior shows clear breakage. Beyond this point, residual point sources, bright artifacts, and calibration effects may contribute to the obtained power. Beyond the ℓ ranges considered, the power spectrum behavior is the same as seen in earlier studies [183].

Unsubtracted point sources present in the residual can cause Poisson fluctuations in the values of $\mathcal{C}_\ell$. The green dashed curve in Figure 4.10 shows the value of $\mathcal{C}_\ell$ due to presence of residual point sources below a threshold flux (which corresponds to the maximum flux in the residual data) of $\mathrm{S_c} = 4.6$, 1.9 & 1.7 mJy respectively from lowest to highest galactic latitudes considered. This is predicted using the formulation described in [129]. As seen in Figure 4.10, the values of $\mathcal{C}_\ell$ where the power-law fitting is obtained lie above the theoretical threshold for residual point source contamination. It is to be noted that the threshold (indicated by green dashed curves) are the lower limits of the expected residual point source contribution under ideal conditions with an assumed dN/dS distribution. However, for the case of

Table 4.6: Details of the pointings used for determining the Angular Power Spectrum and the best fit values for A & β

(RA , Dec) (h:m:s , d.m.s)	(l, b) (deg,deg)	ℓ_{min}	ℓ_{max}	A	β	$\chi^2_{reduced}$
(10:33:41, 59.44.51)	(149.12, 49.79)	1200	2800	6.26±0.58	2.15±0.31	0.89
(10:48:30,58.03.00)	(149.26, 52.37)	1300	3280	3.84±0.22	2.27±0.20	0.30
(10:55:54, 56.21.09)	(146.14, 54.97)	1040	3020	6.18±0.59	3.15±0.29	3.60

real data, considering the presence of residuals, non-Gaussian noise, and calibration errors, a flat floor is not obtained (as seen in Figure 4.10. Previous studies by [183] for TGSS fields have also found similar results.

The $\mathcal{C}_\ell$ range obtained varies between ∼1 mK^2 to ≤100 mK^2 for all three pointings across the entire range of angular modes. Despite being located quite far away from the galactic plane, these values may be sufficiently high to obscure the 21-cm signal (for which analytical calculations show a value ∼0.1 mK^2). The range of amplitude of the angular power spectrum for the residual data is slightly smaller than that obtained by [174]. They have done the analysis for $b = 44.89°$, and the obtained values of $\mathcal{C}_\ell$ in the residuals vary between $\sim 10-100\,\mathrm{mK}^2$ for $1115 \leq \ell \leq 5083$ using direction-dependent approach. Previous works by [184], [118], [185], [183] were done for lower galactic latitudes (within ± 20°), with the obtained values of β varying between ∼1.8 to ∼2.9. This work demonstrates for the first time that the power-law index for diffuse emission is also in the same range of values for locations far away from the galactic plane.

4.7 Conclusion

Characterizing foregrounds is one of the most challenging steps in recovering the redshifted 21-cm signal targeted for studying CD and EoR. The foregrounds vary spatially across the entire sky, and hence more observation at low frequencies are

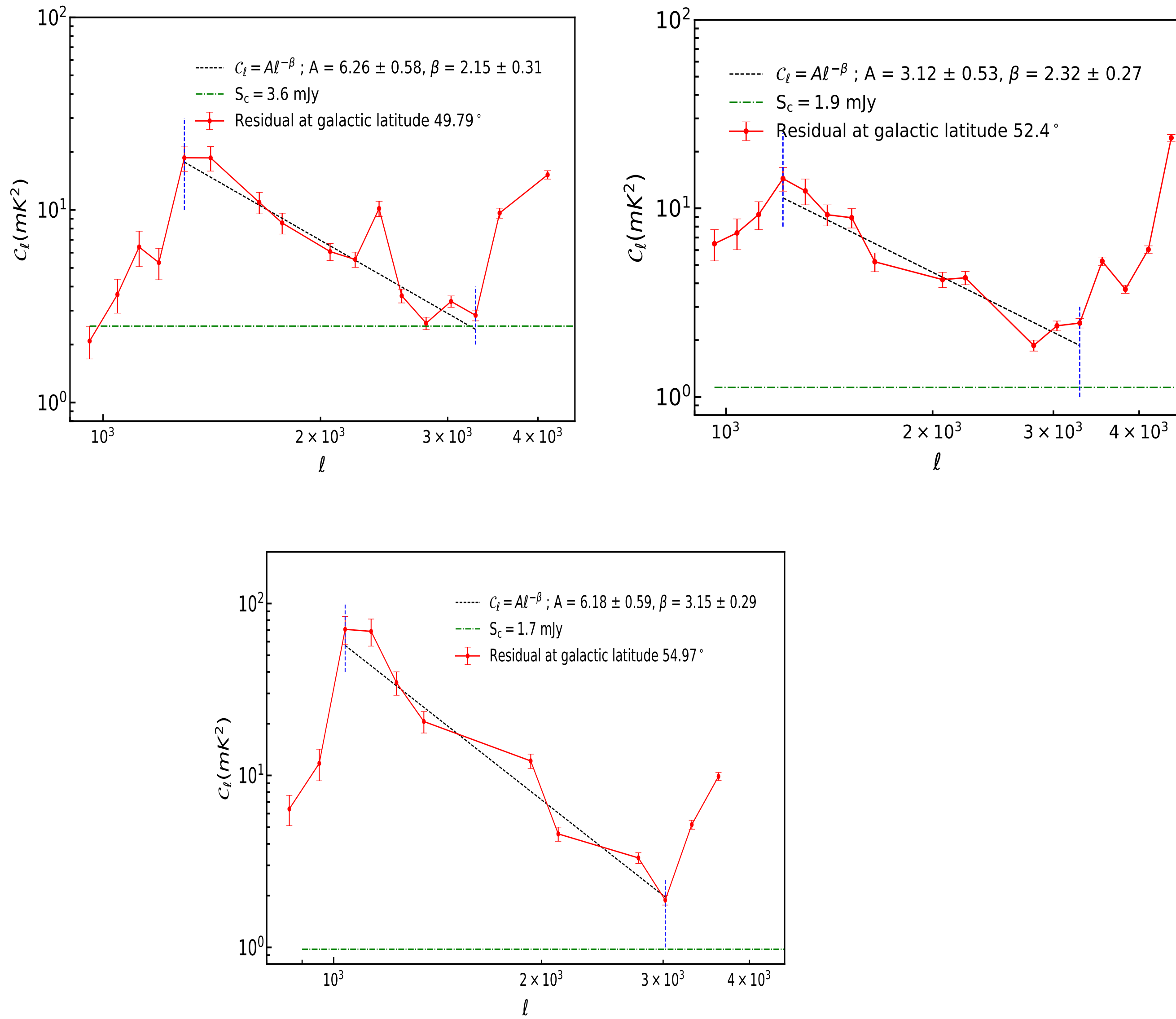

Figure 4.10: Estimated Angular Power Spectrum, $\mathcal{C}_\ell$, with $1-\sigma$ error bars for 3 different galactic latitudes plotted as a function of angular mode, ℓ. Red circles represent the observed values and black dashed curve is the best fit power law curve of the form $A\ell^{-\beta}$. The blue dashed lines is the ℓ range where the power law could be fit. The green dot-dashed line is the contribution from the unmodelled point sources in the residual data.

essential for generating fiducial foreground models. This paper presents the results of such a low-frequency observation of the Lockman Hole field located at very high galactic latitude. The data obtained from the GMRT archive is at 325 MHz covering almost 6 square degrees. A mosaic image, $6^\circ \times 6^\circ$ across, has been produced after direction-dependent calibration of the data. The RMS level reached is ~ 50 μJy beam^{-1}. The sources recovered above $5\sigma_{\rm RMS}$ have been considered to produce a catalog of the field. The total sources recovered are 6186. Comparison has also been made with previous observation covering a part or whole of the same field to check for flux and position accuracy. The recovered fluxes and positions are found to be consistent with previous observations. Euclidean normalized source counts have been determined for the cataloged sources, after correcting for different errors and biases. The final counts are consistent with the previous observation for the same field as well as other parts of the sky.

The paper additionally probes variation in the power spectrum of DGSE at these frequencies for locations far off from the galactic plane. The angular power spectrum $\mathcal{C}_\ell$ as a function of ℓ has been determined for three different galactic locations of this field using TGE. The values of $\mathcal{C}_\ell$ lies between $\sim$1 mK^2 to $\sim$100 mK^2 for all the pointings. Despite being far-off from the galactic center, this is a very high value for $\mathcal{C}_\ell$ that can render recovery retrieval of the cosmological 21-cm difficult. Power law of the form $A\ell^{-\beta}$ has also been fitted in the angular power spectrum obtained. The fitted values for power-law index lie between 2.15 and 3.15 and also varies with varying the galactic latitude, thereby showing the necessity for more low-frequency observations for characterizing foregrounds.

Chapter 5

Clustering Properties of sources in the Lockman Hole Region

This chapter has been adapted from *A study on the Clustering Properties of Radio-Selected sources in the Lockman Hole Region at 325 MHz* [3]

5.1 Introduction

Observations of the extragalactic sky at radio frequencies are essential for the study of both large-scale structures (LSS) and different populations of sources present in the Universe. The initial research on LSS using clustering was performed with the reporting of slight clustering signals from nearby sources [212, 213]. With the advent of large-area surveys like FIRST (Faint Images of the Radio Sky at Twenty-Centimeters, [171]) and NVSS (NRAO VLA Sky Survey, [214]), the studies became more precise due to the large number of sources detected in these surveys.

The extragalactic sky at radio frequencies is dominated by sources below mJy flux densities (at frequencies from MHz to a few GHz, see for example [137, 139–142]). The source population can be divided into Active Galactic Nuclei (AGNs) and Star-Forming Galaxies (SFGs) [139, 145, 156, 215–217]. The dominant sources at these fluxes are SFGs, AGNs of Fanaroff-Riley type I (FR I, [218]), and radio-quiet quasars

[219]. Emission mechanism dominating populations at low frequencies ($\lesssim$10 GHz) is synchrotron emission, modeled as a power law of the form $S_\nu \propto \nu^{-\alpha}$, where α is the spectral index. Study of the extragalactic population using synchrotron emission can help trace the evolution of the LSS in the Universe. It also helps to map their dependence on various astrophysical and cosmological parameters [220–222]. Radio continuum surveys, both wide and deep, help constrain the overall behavior of cosmological parameters and study their evolution and relation to the environment [223–227]. The clustering pattern of radio sources (AGNs and SFGs) can be studied to analyze the evolution of matter density distribution. Clustering measurements for these sources also provide a tool for tracing the underlying dark matter distribution [21, 228–231]. The distribution of radio sources derived from clustering is related to the matter power spectrum and thus provides insights for constraining cosmological parameters that define the Universe. The relationship of the various galaxy populations with the underlying dark matter distribution also helps assess the influence of the environment on their evolution. Clustering studies are also required for extragalactic foreground characterization for EoR and post-EoR science. Spatial clustering of extragalactic sources with flux density greater than the sub-mJy range (around$\sim$150 MHz) dominate fluctuations at angular scales of arcminute range. Thus, their modeling and removal allow one to detect fluctuation of the 21-cm signal on the relevant angular scales.

The definition of clustering is the probability excess above a certain random distribution (taken to be Poisson for astrophysical sources) of finding a galaxy within a certain scale of a randomly selected galaxy. This is known as the two-point correlation function [232]. The angular two-point correlation function has been studied in optical surveys like the 2dF Galaxy Redshift Survey [233–235], Sloan Digital Sky Survey [236–240] and the Dark Energy Survey [241]. Optical surveys provide redshift information for sources either through photometry or spectroscopy. This information can be used to obtain the spatial correlation function and the bias parameter [242–244]. But for optical surveys, observations of a large fraction of the sky is expensive in terms of cost and time. Additionally, optical surveys suffer the limitation of being dust-obscured for high redshift sources. However, at radio wavelengths, the incoming radiation from these sources do not suffer dust attenuation and thus can be used as a mean to probe such high z sources [245–249]. The highly sensitive radio telescopes like GMRT [162], ASKAP [250], LOFAR [251] are also able to survey larger areas of the sky significantly faster. They are thus efficient for conducting large-area surveys in lesser time than the old systems while detecting lower flux densities. Therefore, radio surveys provide an efficient method for investigation of the clustering for the different AGN populations. Additionally, at low-frequencies ($\lesssim$ 1.4 GHz), synchrotron radiation from SFGs provide insight into their star-formation rates [252–256]. These insights have lead to clustering studies of SFGs as well at low frequencies [257, 258]. Through clustering studies of radio sources, deep radio surveys help trace how the underlying dark matter distribution is traced by luminous matter distribution. In addition to this, the two-point correlation functions can also provide other information relevant for cosmology by fitting parameterized models to the data to obtain acceptable ranges of parameters. These include the bias parameter, dark energy equations of state, and Ω_m (total density of matter), to name a few [232, 259–262].

Extensive observations at multiple frequencies can help understand the relationship of the various source populations with their host haloes and individual structures (stars) present. It has been inferred from clustering observations that AGNs are primarily hosted in more massive haloes than SFGs and are also more strongly clustered [222, 257, 263, 327]. While AGNs are more clustered than SFGs, for the latter, the clustering appears to be dependent on the rate of star formation. SFGs with higher star formation rates are more clustered than the ones with a lower rate (since star formation rate is correlated to stellar mass, which in turn is strongly correlated to the mass of the host halo, see [264–266] and references therein). Studying the large-scale distribution of dark matter by studying the clustering pattern of luminous baryonic matter is vital for understanding structure formation. From linear perturbation theory, galaxies are "biased" tracers of the underlying matter density field since they are mostly formed at the peak of the matter distribution [232]. Bias parameter (b) traces the relationship between overdensity of a tracer δ and the underlying dark matter overdensity (δ_{DM}), given by $\delta = \mathrm{b}\delta_{DM}$. The linear bias parameter is the ratio between the dark matter correlation function and the galaxy correlation function ([232, 267, 268], also see [269] for a recent review). Measurement of the bias parameter from radio surveys will allow measurements which probe the underlying cosmology governing the LSS, and probe dark energy, modified gravity, and non-Gaussianity of the primordial density fluctuations [270–274].

Analysis of the clustering pattern for extragalactic sources is also important for observations targeting the 21-cm signal of neutral hydrogen (HI) from the early Universe. These weak signals from high redshifts have their observations hindered by many orders of magnitude brighter foregrounds - namely diffuse galactic synchrotron emission [105], free-free emission from both within the Galaxy as well as extragalactic sources [109], faint radio-loud quasars [107] and extragalactic point sources [108]. [108] showed that spatial clustering of extragalactic sources with flux density $\gtrsim$0.1 mJy at 150 MHz (the equivalent flux density at 325MHz is $\sim$0.05 mJy) dominate fluctuations at angular scale $\theta \gtrsim$1'. Thus, their modeling and removal allow one to detect fluctuation of the 21-cm signal on relevant angular scales. So their statistical modeling is necessary to understand and quantify the effects of bright foregrounds. Many studies have modeled the extragalactic source counts as single power-law or smooth polynomial [133, 135] and the spatial distribution of sources as Poissonian [129] or having a simple power-law clustering. However, a Poisson distribution of foreground sources is very simplistic and may affect signal recovery for sensitive observations like those targeting the EoR signal [129, 132]. Thus more observations are required for low-frequency estimates of the clustering pattern of compact sources.

A number of studies have been done in recent years for observational determination of the clustering of radio selected sources (for instance [138, 221, 222, 257, 258, 275–278]). However, more such studies are required for modeling the influence of different processes on the formation and evolution of LSS in the Universe. The sample used for such analyses should not be limited to small deep fields, since the limited number of samples makes clustering studies of different populations (AGNs/ SFGs) sample variance limited. Studies on the statistics of the source distribution are also essential for understanding the matter distribution across space. Thus, observations using sensitive instruments are required to conduct more detailed studies. At 1.4 GHz and above, many clustering studies are present (for instance [221, 222, 257, 275, 276, 279]); however there extensive studies at low frequencies (and wider areas) are still required. The TIFR GMRT Sky Survey (TGSS) [133] is a wide-area survey of the northern sky at 150 MHz. But the available catalog from the TGSS- Alternate Data Release (TGSS-ADR) suggests that the data is systematics limited. Thus it is unsuitable for large-scale clustering measurements [280]. The ongoing LOFAR Two-metre Sky Survey (LoTSS [201]) at a central frequency of 144 MHz is expected to have very high sensitivity and cover a very wide area and thus provide excellent data for studying source distribution statistics at low frequencies [278]. However, to constrain cosmological parameters, consensus for the overall behavior of sources along different lines of sight and across frequencies is also required, and there data sets like the one analyzed here become important [270, 271, 281]. Radio data has the advantage that even flux-limited samples contain high-z sources [282]. Thus, using the entire radio band provides insights into physical processes driving the evolution of different galaxy populations and helps create a coherent picture of the matter distribution in the Universe. Therefore, studies at radio frequencies would help constrain the cosmology underlying structure formation and evolution.

The recent study of the clustering of the ELAIS-N1 field centered at 400 MHz using uGMRT by [258] was extremely sensitive, with an RMS (σ_{400})[1] of 15 μJy beam^{-1}. But the area covered was significantly smaller (∼1.8 deg ^{2}) than this work. This smaller field of view makes measurement of clustering properties on large angular scales impossible. Smaller areas also lead to smaller sample sizes for statistics, resulting in studies limited by cosmic variance. Another study of the HETDEX spring field at 144 MHz (using the data release 1 of LOFAR Two meter Sky Survey) by [278] has a sky coverage of ∼350 square degree, but the mean σ_{150} is ∼91 μJy beam^{-1}. However, despite the sensitivity achieved in the survey, the analysis by [278] is limited to flux densities above 2 mJy. Motivated by the requirement for a study in the intermediate range (in terms of flux density, area covered, and frequency), this work aims to quantify the clustering of the sources detected in the Lockman Hole field. The data analyzed here fall in the intermediate category, with a survey area ∼6 deg^2 with $\sigma_{325} \sim 50$ μJy beam $^{-1}$. It is thus ideal for clustering studies with a sizeable area of the sky covered (thus large angular scales can be probed) and moderately deep flux threshold (catalogue will have fluxes reliable to a lower value). Additionally, the Lockman Hole region has excellent optical coverage through surveys like SDSS and SWIRE; thus, associated redshift information is available to study spatial clustering and bias parameters. This frequency also has the additional advantage of having lesser systematics than the 150 MHz band while still being sensitive to the low-frequency characteristics of sources. New data releases for the LoTSS surveys promise greater sensitivity and source characterization over various deep fields targeted by these observations [283, 284]; all these observational data at multiple frequencies will put more precise constraints on the various parameters

[1] Unless otherwise stated, $\sigma_{frequency}$ is the RMS sensitivity at the quoted frequency throughout the text.

governing the structure formation and evolution.

This work uses archival GMRT data at 325 MHz covering a field of view of $6° \times 6°$ through multiple pointings. In [1], data reduction procedure is described in detail. This work used the source catalogue obtained there for clustering analyses. However, the entire dataset could not be used due to limiting residual systematics at large angular scales. The clustering pattern and linear bias parameter are determined for the whole population and sub-populations, i.e., AGNs and SFGs, separately. The previous work by [1] had determined the flux distribution of sources (i.e., differential source count) and characterized the spatial property and the angular power spectrum of the diffuse galactic synchrotron emission using the same data.

This paper is arranged in the following manner: In section 5.2, a brief outline of the radio data as well as various optical data used is discussed; the classification into source sub-populations is also using radio luminosity of sources is also shown. The following section, i.e., Section 5.3 shows the clustering quantification - both in spatial and angular scales and calculation of linear bias for all the detected sources. Section 5.4 discusses the clustering property and bias for classified population, with a brief discussion on the choice of the field of view for this analysis discussed in Section 5.5. Finally, the paper is concluded in Section 5.6.

For this work, the best fitting cosmological parameters obtained from the Planck 2018 data [97] has been used. The values are $\Omega_{\rm M} = 0.31$, $\Omega_{\Lambda} = 0.68$, $\sigma_8 = 0.811$, & $H_0 = 67.36$ km s^{-1} Mpc $^{-1}$. The spectral index used for scaling the flux densities between frequencies is taken as α=0.8.

5.2 Observations and Source Catalogues

This work uses 325 MHz GMRT archival data of the Lockman Hole region. The details of the data reduction procedure have been described in [1], here it is discussed very briefly. The data were reduced using the SPAM pipeline [133, 163, 165], which performs direction-independent as well as direction-dependent calibration techniques. The observation had 23 separate pointings , centered at (α_{2000} = $10^h48^m00^s, \delta_{2000} = 58°08'00"$), each of which was reduced separately. The final image is a 6° × 6° mosaic having off-source RMS of 50μJybeam^{-1} at the central frequency. Figure 4.1 shows the primary beam corrected final mosaic image of the observed region. This image was used to extract a source catalogue using Python Blob Detection and Source Finder [2](PyBDSF, [169]) above a minimum flux density S_{325}^{cut} 0.3mJy (i.e., above $6\sigma_{325}$). A total of 6186 sources were detected and cataloged. The readers are referred to [1] for details on catalogue creation and subsequent comparison with previous observations.

The redshift information for the sources are derived by matching with optical data from the Sloan Digital Sky Survey (SDSS)[3] and the Herschel Extragalactic Legacy Project (HELP) [4,5] [285]. The SDSS [286, 287] has been mapping the northern sky in the form of optical images as well as optical and near-infrared spectroscopy since 1998. The latest data release (DR16) is from the fourth phase of the survey (SDSS-IV, [288]). It includes the results for various survey components like the extended Baryon Oscillation Spectroscopic Survey eBOSS, SPectroscopic identification of ERosita Sources SPIDERS, Apache Point Observatory Galaxy Evolution Experiment 2 APOGEE-2, etc. The surveys have measured redshifts of a few million galaxies and have also obtained the highest precision value of the Hubble parameter $H(z)$ to date [289]. An SQL query was run in the CasJobs [6] server to obtain the optical data corresponding to the radio catalogue, and the catalogue thus obtained was used for further analysis.

HELP has produced optical to near-infrared astronomical catalogs from 23 extra-

[2]https://www.astron.nl/citt/pybdsf/
[3]https://www.sdss.org/
[4]http://herschel.sussex.ac.uk/
[5]https://github.com/H-E-L-P
[6]https://skyserver.sdss.org/casjobs/

galactic fields, including the Lockman Hole field. The final catalogue consists of ∼170 million objects obtained from the positional cross-match with 51 surveys [285]. The performance of various templates and methods used for getting the photometric redshift is described in [290, 291]. Each of the individual fields is provided separate database in the *Herschel Database in Marseille* site [7] where various products, field-wise and category wise are made available via "data management unit (DMU)". For the Lockman Hole field, the total area covered by various surveys is 22.41 square degrees with 1377139 photometric redshift objects. The Lockman Hole field is covered well in the Spitzer Wide-area InfraRed Extragalactic Legacy Survey (SWIRE) with photometric redshifts obtained as discussed in [292, 293]. However, additional data from other survey catalogues like Isaac Newton Telescope - Wide Field Camera (INT-WFC, [294]), Red Cluster Sequence Lensing Survey (RCSLenS, [295]) catalogues, Panoramic Survey Telescope and Rapid Response System 3pi Steradian Survey (PanSTARRS-3SS, [296]), Spitzer Adaptation of the Red-sequence Cluster Survey (SpARCS, [297]), UKIRT Infrared Deep Sky Survey - Deep Extragalactic Survey (UKIDSS-DXS, [298]), Spitzer Extragalactic Representative Volume Survey (SERVS, [299]) and UKIRT Hemisphere Survey (UHS, [300]) resulted in more sources being detected and better photometric determination. The publicly available photometric catalogue for the Lockman Hole region was used to determine the redshift information for matched sources. The source catalogue derived from the 325 MHz observation is pre-processed, matched to add redshift information, and then further analysis is done. The following subsections describe these steps in detail.

5.2.1 Merging multi-component sources

The final map produced has a resolution of 9". The source finder might resolve an extended source into multiple components for such high-resolution maps. Such sources are predominantly radio galaxies that have a core at the center and hotspots

[7] https://hedam.lam.fr/HELP/

that extend along the direction of the jet(s) or at their ends; these structures may be classified as separate sources [137, 301–303]. Using the NVSS catalogue, it has been shown in [304] that large radio sources with unmarked components can significantly alter clustering measurements. Thus, for unbiased estimation of source clustering, such sources need to be identified and merged properly. A strong correlation between the angular extent of radio sources and their fluxes has been discovered by [305]. The angular extent (θ) of a source is related to its flux density (S) by the θ-S relation, $\theta \propto \sqrt{\mathrm{S}}$. This relation was used to identify resolved components of multi-component sources in surveys like the FIRST survey [301].

Identification of multi-component sources in the Lockman Hole catalogue resolved as separate sources are made using two criteria. The maximum separation between pairs of sources (using the θ-S relation) is given by $\theta_{\mathrm{max}} = 20\sqrt{\mathrm{S}_{\mathrm{total}}}$, where $\mathrm{S}_{\mathrm{total}}$ is the summed flux of the source pairs [137, 258, 306]. Sources identified by the above criteria have been considered as the same source if their flux densities differ by less than a factor of 4 [306].

Figure 5.2.1 shows the separation between the nearest neighbour pairs from the 325 MHz catalogue as a function of the separation between them. Above the black dotted line, the sources have separation less than θ_{max} as mentioned above. Blue triangles are sources that have flux density differences less than a factor of 4. The two criteria mentioned gave a sample of 683 sources (out of 6186 total) to have two or more components. After merging multi-component sources and filtering out random associations, 5489 sources are obtained in the revised catalogue. The position of the merged sources are the flux weighted mean position for their components.

5.2.2 Adding Redshift Information

As already mentioned, optical cross-identification have for the sources detected has been done using the HELP and SDSS catalogues. A positional cross-match with 9”

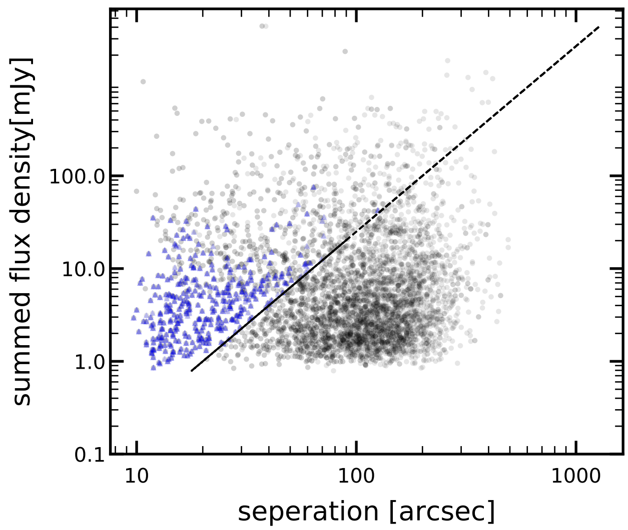

Figure 5.1: Sum of nearest neighbour flux densities from the 325 MHz catalogue as a function of separation between nearest neighbours in the catalogue are shown by light blue circles. Sources above the black doted line (having an angular separation less than $\theta_{\max}$) shown by blue triangles have their flux densities differing by a less than a factor of 4.

matching radius (which is the resolution for this observation) was used for optical cross-matching. Since the positional accuracy of the catalogue is better than 1" [1], a nearest neighbour search algorithm was used to cross-match sources with the optical catalogue with a search radius r_s. The rate of contamination expected due to proximity to optical sources is given by [221]:

$$P_c = \pi r_s^2 \sigma_{opt}$$

where σ_{opt} is the surface density of the optical catalogue. For surface density of 1.4×10^4 deg^{-2}, a matching radius r_s = 9" gives a contamination of <10%. This radius was thus used to ensure valid optical identification of a large number of radio sources.

Large FoVs are helpful for observational studies of like the one used here is useful for studying LSS, since the presence of a large number of sources provides statistically

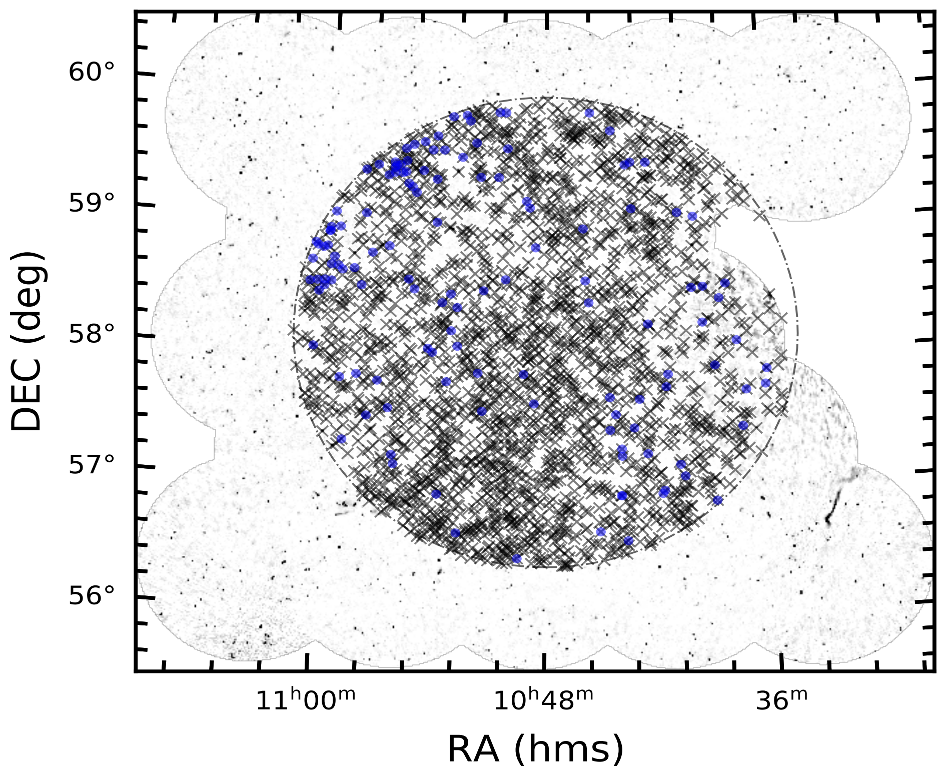

Figure 5.2: Sources with optical cross matches within 1.8° radius of phase center for the full mosaic. The area considered is represented by the black dot-dashed circle, and the sources in this region are represented by "x" marks. The blue circles represent the sources without any redshift matches.

robust results and also reduces the effect of cosmic variance. Accordingly, the data used for this work had an Fov of 6° × 6°. However, cross-matching with optical catalogue produced matches with only 70% total sources over entire FoV, and 30% sources remained unclassified. Further investigation also revealed the presence of unknown systematics, which resulted in excess correlation and deviation from power-law behavior at large angles. The most probable cause for such a deviation seems to be either the presence of many sources with no redshifts at the field edge or the presence of artifacts at large distances from the phase center. Analysis done by reducing the area of the field increased percentage of optical matches and reduced the observed deviation from the power-law nature. Thus, the cause of such a deviation has been attributed to the former one.

Hence, the clustering properties of sources at large angular scales are not reliable for this observation. Thus, the analysis was restricted to a smaller area of the Lockman Hole region around the phase center; large-scale clustering properties could not be estimated. Taking a cut-off with 1.8° radius around the phase center resulted in ∼95% sources having an optical counter-part. Hence, it is expected that the unclassified sources present would not affect the signal significantly (a detailed discussion on the choice of the FoV cut-off is discussed in Section 5.5). This FoV cut-off yielded 2555 sources in the radio catalogue, out of which 2424 sources have optical matches within the aforementioned match radius. This is shown in Figure 5.2.2, where the area considered is represented with black dot-dashed circle, and the "x" marks denote the sources in radio catalogue; the blue circles represent the sources without any optical cross-matches in either photometry or spectroscopy. A total of 2415 photometric and 664 spectroscopic matches were obtained after the cross-match with optical catalogues. Out of these, 650 sources had both photometric and spectroscopic detection. For such cases, the spectroscopic identifications were taken. Combined photometry and spectroscopic identifications were obtained for a total of 2424 sources, of which 27 sources were discarded from this analysis since they were nearby objects with 0 or negative redshifts [221]. The final sample thus had 2397 sources, which is ∼94% of the total catalogued radio sources within 1.8° radius of the phase center. The redshift matching information for both the full and restricted catalogues have been summarised in Table 5.1. The redshift information from the optical catalogues was incorporated for these sources and was used for further analysis. Figure 5.3 shows the distribution of redshifts for the sources detected in both HELP and SDSS. In the left panel, the photometric redshifts are plotted as a function of the spectroscopic redshifts. As can be seen, the two values are

Table 5.1: Summary of number of sources with redshift information

Area	Number of sources	Redshift matches	Percentage of matches	AGNs	SFGs
$6° \times 6°$	5489	3628	66	2149	1479
3.6° radius	2555	2397	95	1821	576

in reasonable agreement with each other for most cases. Additionally to check for the reliability of obtained photometric redshifts, following [291], the outlier fraction defined by $\frac{|z_{phot}-z_{spec}|}{1+z_{spec}}$ >0.2, is plotted as a function of the spectroscopic value (right panel of Figure 5.3). For this work, the drastic outliers are the points with values >0.5. The fraction of outliers with drastically different values between photometric and spectroscopic redshifts is ∼10%. While a detailed investigation is beyond the scope of this work, the outliers may be present due to the combination of uncertainties in the different surveys used in the HELP catalogue. As can be seen, the outlier fraction is not very drastic except for some cases; however, the reason for deviations in these sources is unknown. The median redshift for all the sources with redshift information comes out to be 0.78. The top panel of Figure 5.4 shows the distribution N(z) as a function of source redshift, with the black dashed line indicating median redshift.

5.2.3 Classification using Radio Luminosity Function

The catalogued sources with optical counterparts were divided into AGNs and SFGs using their respective radio luminosities. Assuming pure luminosity evolution, the luminosity function evolves approximately as $(1+\mathrm{z})^{2.5}$ and $(1+\mathrm{z})^{1.2}$ for SFGs and AGNs respectively [309]. The value for AGNs differ slightly from those of [310] and [311] for the COSMOS field at 3 GHz and ELAIS N1 field at 610 MHz respectively. However, they are consistent with those of [312] for the GAMA fields. The values for redshift evolution of SFGs also agree broadly for the GAMA fields [312] and the ELAIS N1 field [313].

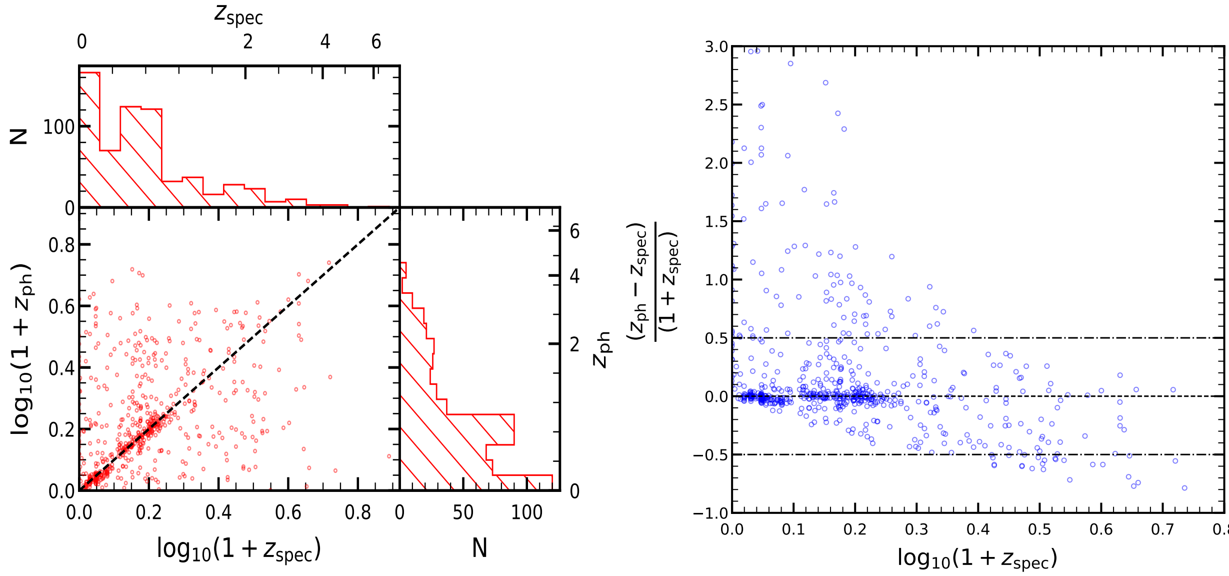

Figure 5.3: Comparison of photometric and spectroscopic redshifts for the sources with matches for both. **(left panel)** Photometric redshift plotted as a function of the spectroscopic redshift, with the black dashed line indicating equal values for both. The redshift distribution is shown via the histograms for photometric (right subplot) and spectroscopic (top subplot). **(right panel)** Distribution of $\frac{\mathrm{z_{phot}}-\mathrm{z_{spec}}}{1+\mathrm{z_{spec}}}$ plotted as a function of $\mathrm{z_{spec}}$. The black dashed line represents 0 deviation between the photometric and spectroscopic redshifts.

It has been shown in [257, 307, 308] that radio selected galaxies powered by AGNs dominate for radio powers beyond a radio power $\mathrm{P_{cross}}(z)$ which is related with the redshift z as:

$$\log_{10}\mathrm{P_{cross}} = \log_{10}\mathrm{P_{0,cross}} + z \tag{5.1}$$

upto z ~1.8, with P (at 1.4 GHz) in W $\mathrm{Hz^{-1}sr^{-1}}$. In the local Universe, the value of $\mathrm{P_{cross}}$ is $10^{21.7}$ (W $\mathrm{Hz^{-1}sr^{-1}}$), coinciding with the observed break in the radio luminosity functions of SFGs [316], beyond which their luminosity functions decrease rapidly and the numbers are also reduced greatly. Thus contamination possibility between the two population of radio sources is very low using the radio luminosity based selection criterion [257, 307].

The radio luminosity has been calculated for the sources from their flux as [307]:

$$P_{1.4GHZ} = 4\pi S_{1.4GHz} D^2 (1+z)^{3+\alpha} \tag{5.2}$$

where D is the angular diameter distance, and α is the spectral index of the sources in the catalogue. The individual spectral index for the sources was not used (since all sources do not have the measured values). The median value of 0.8 for α was derived by matching with high-frequency catalogues in [1]. Since the probability of finding a large number of bright, flat-spectrum sources is very low [257], the median value of 0.8 was used to determine the luminosity functions of the sources in the Lockman Hole field detected here.

Besides the radio luminosity criterion described above, there are several other methods to classify sources into AGNs and SFGs. X-ray luminosity can also be used to identify AGNs since it can directly probe their high energy emissions [314]. Color-color diagnostics from optical data (like IRAC) can also be used for identifying AGNs [315]. Classification can also be done using the q_{24} parameter, which is the ratio of 24 μm flux density to the effective 1.4 GHz flux density [217]. Based on the resuts of [309], it was shown by [307] that the radio luminosity function for SFGs fall of in a much steeper manner than AGNs for all redshifts, and this reduces the chances of contamination in the two samples. Additionally, these different multi-wavelength methods are not always consistent with each other, and a detailed investigation into any such discrepancy is beyond the scope of this work. Hence, only the radio luminosity criterion has been used for classification.

The sources with redshifts up to 1.8 were classified into AGNs and SFGs according to whether their luminosity is greater than or less than the threshold in Equation 5.1 (with P_{cross} determined using Equation 5.2). At higher redshifts (i.e. >1.8), $P_{0,cross}$ is fixed to $10^{23.5}$ [$WHz^{-1}sr^{-1}$] [309]. Of the 2397 sources, 1821 were classified as AGNs and 576 as SFGs using the radio luminosity criteria. The median redshifts for AGNs and SFGs are 1.02 and 0.2 respectively.

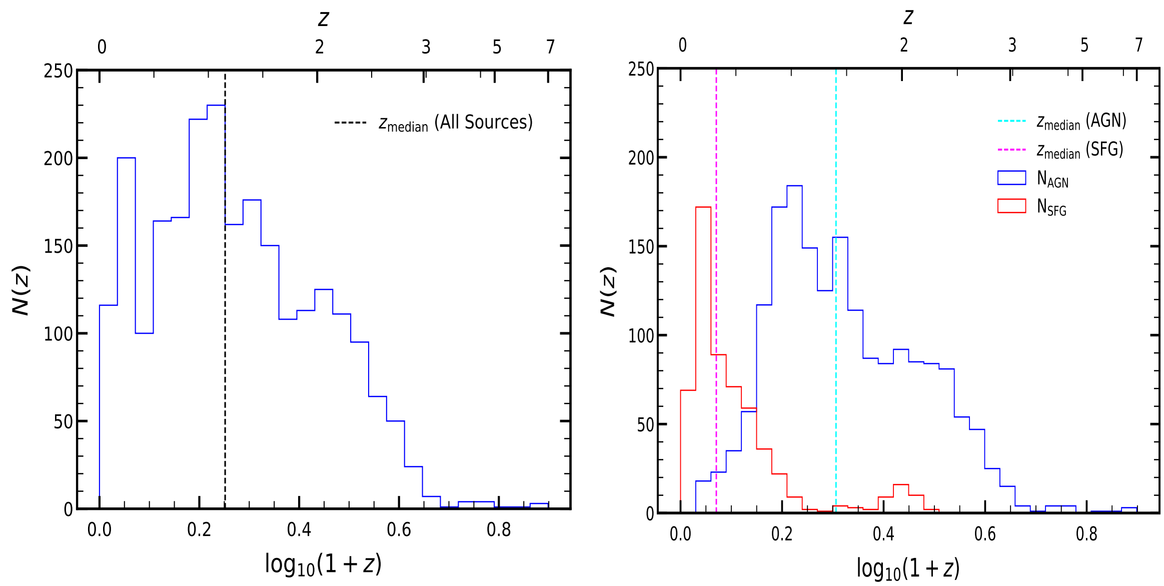

Figure 5.4: Redshift distribution (N(z)) for sources with either photometric (HELP) or spectroscopic (SDSS) redshifts. The left panel shows N(z) distribution for all sources with the black dashed line indicating the median redshift for all sources. The right panel shows the same for sources classified into AGNs (blue curve) and SFGs (red curve) using the radio luminosity criterion discussed in the text, with the cyan and magenta dashed lines indicating the respective median redshifts.

5.3 Estimation of Correlation Function: Combined Sources

5.3.1 The Angular Correlation Function

The angular two-point correlation function $w(\theta)$ is used to quantify clustering in the sky on angular scales. While several estimators have been proposed in literature (for a comparison of the different types of estimators see [317] and Appendix B. of [278]), this work uses the LS estimator proposed by [318]. It is defined as :

$$w(\theta) = \frac{\mathrm{DD}(\theta) - 2\mathrm{DR}(\theta) + \mathrm{RR}(\theta)}{\mathrm{RR}(\theta)} \tag{5.3}$$

Here DD(θ) and RR(θ) are the normalised average pair count for objects at separation θ in the original and random catalogues, respectively. Catalogue realizations generated by randomly distributing sources in the same field of view as the real observations have been used to calculate RR(θ). The LS estimator also includes the normalized cross-pair separation counts DR(θ) between original and random catalogue, which has the advantage of effectively reducing the large-scale uncertainty in the source density [276, 304, 318, 319]. The uncertainty in the determination of $w(\theta)$ is calculated using the bootstrap resampling method [320], where 100 bootstrap samples are generated to quote the 16th and 84th percentile errors in determination of $w(\theta)$.

5.3.1.1 Random Catalogue

The random catalogues generated should be such that any bias due to noise does not affect the obtained values of the correlation function. The noise across the entire $6° \times 6°$ mosaic of the field is not uniform (see Figure 3 of [1]). This can introduce a bias in estimating the angular two-point correlation function since the non-uniform noise leads to the non-detection of fainter sources in the regions with higher noise.

PyBDSF was used for obtaining the noise map of the image. Assuming the sources follow a flux distribution of the form dN/dS $\propto S^{-1.6}$ [178, 179], random samples of 3000 sources were generated in the given flux range (with lower limit corresponding to 2 times the background RMS of the image) and assigned random positions to distribute them in the entire FoV. The sources constitute a mixture of 70% unresolved sources and 30% extended sources, which is roughly in the same ratio as the actual source catalogue [1]. These were injected into the residual map, and using the same parameters in PyBDSF as the ones used in the extraction of the original sources (see [1]), the random catalogues were extracted. 100 such statistically independent realizations were used to reduce the associated statistical uncertainty.

For clustering analysis of AGNs and SFGs, two sets of random catalogues were generated uing the publicly available catalogues for these source types from the T-RECS simulation [122]. These catalogues have source flux densities provided at different frequencies between 150 MHz to 20 GHz. The flux densities at 300 MHz were considered for the randoms. They were scaled to 325 MHz using $\alpha = 0.8$, and 2000 sources were randomly chosen within flux density limit for the radio catalogue of AGNs and SFGs. They were assigned random positions within the RA, Dec limits of the original catalogues and injected into the residual maps. Then using the same parameters for PyBDSF as the original catalogue, the sources were recovered. 100 such realizations were done for AGNs and SFGs separately. The recovered random catalogues were used for further clustering analysis of the classified populations. It should also be mentioned here that the lower cut-off of flux density for the random catalogues was $\sim$0.1 mJy, which is 2 times the background RMS. As already seen from [1], even a flux limit of 0.2 mJy (4 times the background RMS) takes care of effects like the Eddington bias [177]. Thus 0.1 mJy is taken as the limiting flux for both the combined and the the classified random catalogues. The final random samples for AGNs and SFGs consisted of a total of $\sim$120000 sources each, while for the combined sample, it was $\sim$200000. This is much higher than the number of sources in the radio catalogue. Thus, it does not dominate the errors. As has already been stated, the point and extended sources in the random catalogues (generated for the whole sample and the classified sources) are taken in the same ratio as that of the original radio catalogue. The drawback of this assumption is that there is a chance of underestimating extended sources in the random catalogue, which may lead to spurious clustering signals at smaller angular scales. However, since no evidence of any spurious signal is seen, taking point and extended sources in the same ratio as the original catalogue seem reasonable.

5.3.2 Angular Clustering Pattern at 325 MHz

The angular correlation function of the sources detected in this observation is calculated using the publicly available code `TreeCorr` [8][321]. The 325 MHz catalogue

[8]https://github.com/rmjarvis/TreeCorr

was divided into 15 equispaced logarithmic bins between $\theta \sim 36"$ to $2°$. The lower limit corresponds to the four times the PSF at 325 MHz, and the upper limit is the half-power beamwidth at this frequency. Figure 5.3.2 shows the angular correlation function of the 325 MHz in red circles; the error bars are estimated using the bootstrap method as discussed earlier. A power law of the form $w(\theta) = A\theta^{1-\gamma}$ is also fitted. The power law index, γ is kept fixed at the theoretical value of 1.8. The parameter estimation for this fit is done using Markov chain Monte Carlo (MCMC) simulation by generating 10^6 data points by applying the Metropolis-Hastings algorithm in the A parameter space. The first 10^2 samples have been removed from the generated chains to avoid the burn-in phase. From the sampled parameter space, χ^2 is used to estimate the most likely values of the parameters. The best fit parameters are $\log(\mathrm{A}) = -2.73^{+0.11}_{-0.15}$, with the error bars being the 1-σ error bars from the 16th and 84th percentiles of the chain points.

5.3.2.1 Comparison with previous Observations

The best fit values obtained for parameters A and γ of the 325 MHz catalogue have been compared with those for other observations at radio frequencies. The parameters obtained for different radio surveys, namely from [138, 221, 222, 258, 277–279] have been summarised in Table 5.2. The scaled flux limit at 325 MHz for the [279] catalogue at (originally 1.4 GHz, scaled using a spectral index of 0.8) is $\sim$0.4 mJy, very close to the flux limit for this work. However, their estimates are higher than all previous estimates (they particularly compare with [257]), which they assign partly to the presence of sample variance. While the area probed by [279] is also included within the region this work probes, the area covered are different, the one covered here being larger. This might be the reason for differences between the estimates in this work and [279], despite both having similar flux density cut-offs. The clustering amplitude for this work is similar to [222] at almost all the angular scales. One possible reason is that the flux limit for the study at 3 GHz was 5.5

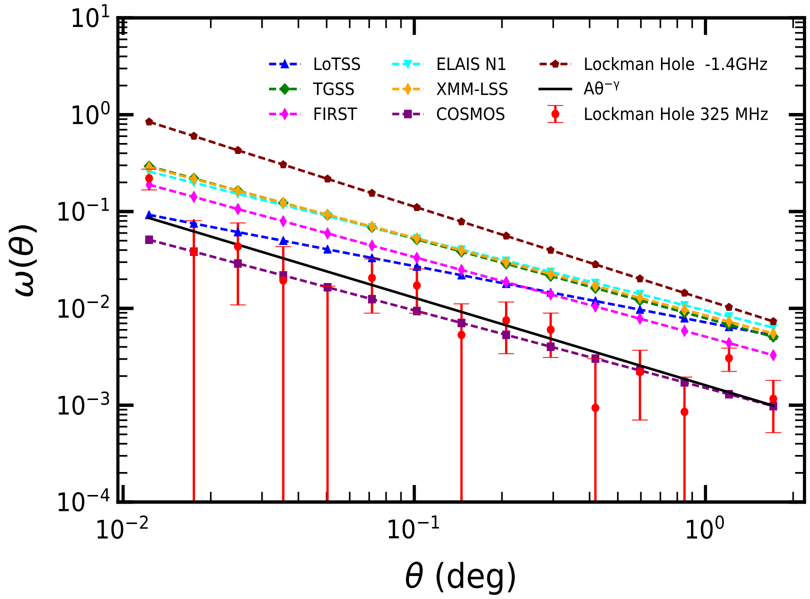

Figure 5.5: Angular correlation function for all sources in the Lockman Hole region at 325 MHz. The dashed black line is the best-fitting power law to $w(\theta)$. Comparison has been done with best fitted power laws obtained from previous studies of LoTSS (blue triangle, [278]), TGSS-ADR (green diamonds, [277]), FIRST (magenta thin diamonds, [221]), ELAIS-N1 (cyan inverted triangles, [258]), XMM-LSS (orange thin diamonds, [138]), COSMOS (purple squares, [222]), & Lockman Hole 1.4 GHz observation (maroon pentagons, [279]). Different studies mentioned here have different flux density limits (see Table 5.2).

times the 2.3 μJy beam^{-1} limit corresponding to a flux of ∼0.1 mJy at 325 MHz, which is near the flux cut-off for this work (0.3 mJy), and thus can trace similar halo masses and hence clustering amplitudes.

The clustering properties of the radio sources in the VLA-FIRST survey [171] has been reported in [221], where log(A) is -2.30$^{+0.70}_{-0.90}$. [222] and [138] have reported log(A) value of -2.83 and -2.08 for the COSMOS and XMM-LSS fields, respectively, by fixing γ at the theoretical value of 1.80. The clustering amplitude of the 150 MHz TGSS-ADR [133] has been shown by [277] for a large fraction of the sky and

Table 5.2: Clustering Parameters for Observed Data. The columns indicate the name of the survey (Observation), observing frequency in MHz (Frequency), the flux density cut-off at the observing frequency ($S_{cut,\nu}$), the equivalent 325 MHz flux-density ($S_{cut,325}$), best fit clustering amplitude ($\log_{10}(A)$) and best fit power-law index (γ) respectively.

Observation	Frequency (MHz)	$S^{\dagger}_{cut,\nu}$ (mJy)	$S^{*}_{cut,325}$	$\log_{10}(A)$	γ	Reference
FIRST	1400	1.00	3.21	$-2.30^{+0.70}_{-0.90}$	1.82±.02	[221]
COSMOS	3000	0.013	0.08	$-2.83^{+0.10}_{-0.10}$	1.80	[222]
XMM-LSS	144	1.40	0.73	$-2.08^{+0.05}_{-0.04}$	1.80	[138]
TGSS-ADR	150	50	26.9	$-2.11^{+0.30}_{-0.30}$	1.82±.07	[277]
ELAIS-N1	400	0.10	0.12	$-2.03^{+0.10}_{-0.08}$	1.75±0.06	[258]
Lockman Hole	1400	0.12	0.39	$-1.95^{+0.005}_{-0.005}$	1.96±.15	[279]
LoTSS	144	2.00	1.04	$-2.29^{+0.6}_{-0.6}$	1.74±.16	[278]
Lockman Hole	325	0.30	0.30	$-2.73^{+0.11}_{-0.15}$	1.80	This work

† $S_{cut,\nu}$ is the flux density limit at the respective observing frequencies; * $S_{cut,325}$ is the scaled flux density (α=0.8) limit at 325 MHz

at different flux density cut-offs. In the recent deep surveys of the ELAIS-N1 field at 400 MHz [258], log(A) and the best fit power law index have the values $-2.03^{+0.10}_{-0.08}$ and 1.75±0.06 respectively.

Comparison has also been made with the wide-area survey of LoTSS data release 1 [278]. This study (with data obtained at a central frequency of 144 MHz) employed various masks on the data to obtain the angular clustering values. The survey covers a wider area, but the flux cut-off threshold is above 1 mJy for all of the masks due to systematic uncertainties. A wide range of angles, $0.1° \leq \theta \leq 32°$ was fixed to determine the angular clustering. Taking three different flux density limits- at 1, 2 and 4 mJy and different masks, the values of log(A) and power-law index were obtained(the fitting for the power-law form was done for $0.2° \leq \theta \leq 2°$). [278] have applied various flux density cuts and masks to their sample for obtaining the angular clustering parameters. They have concluded that the flux density cut-off of 2 mJy provides the best estimate for the angular clustering parameters, and the same has been used here for comparison. Comparison of the present work with LoTSS 2 mJy flux cut shows that the values of log(A) agree well. The best fit power-law index is also consistent within 1σ error bars. Hence, it is seen that the angular correlation function obtained in the present work gives values for the parameters log(A) and γ consistent with those reported in previous surveys. Additionally, since this survey has both wider coverage than the recent EN1 data and a lower flux density threshold than the LoTSS data used by [278], it provides an intermediate data set along a different line of sight to probe cosmology.

5.3.3 The Spatial Correlation Function at 325 MHz

For known angular clustering $w(\theta)$, the spatial clustering of sources is quantified by the two-point correlation function $\xi(r)$. Using the Limber inversion [322], $\xi(r)$ can be estimated for known redshift distribution. Gravitational clustering causes the spatial clustering to vary with redshift, and thus a redshift dependent power-law spatial correlation function can be defined as [276, 322]:

$$\xi(\mathrm{r}, \mathrm{z}) = (\mathrm{r_0}/\mathrm{r})^{\gamma}(1+\mathrm{z})^{\gamma-(3+\epsilon)} \tag{5.4}$$

where the clustering length r is in comoving units, ϵ specifies clustering models [276] and $\mathrm{r_0}$ is the clustering length at z=0. For this work, comoving clustering model, in which the correlation function is unchanged in the comoving coordinate system and with $\epsilon = \gamma$-3, is used. The comoving cluster size is constant. The correlation length is calculated using [232]:

$$\mathrm{A} = \mathrm{r}_0^{\gamma}\mathrm{H}_{\gamma}(\mathrm{H_0}/\mathrm{c})\frac{\int_0^{\infty}\mathrm{N}^2(\mathrm{z})(1+\mathrm{z})^{\gamma-(3+\epsilon)}\chi^{1-\gamma}(\mathrm{z})\mathrm{E}(\mathrm{z})\mathrm{dz}}{[\int_0^{\infty}\mathrm{N}(\mathrm{z})\mathrm{dz}]^2} \tag{5.5}$$

where $\mathrm{H}_{\gamma} = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{\gamma-1}{2})}{\Gamma(\frac{\gamma}{2})}$,

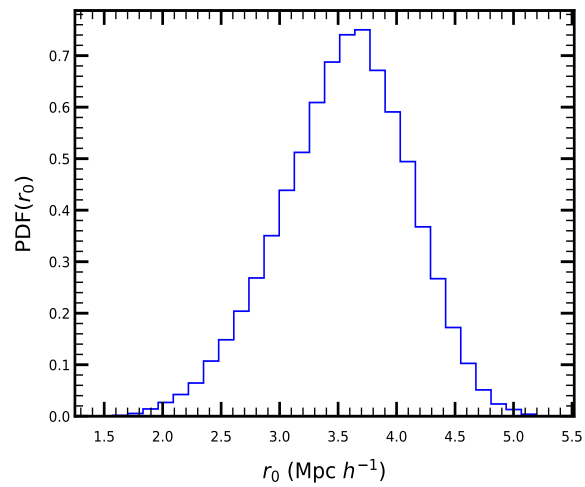

Figure 5.6: Probability distribution function of spatial clustering length (r_0) for the entire sample at 325 MHz.

$\mathrm{E(z)} = \sqrt{\Omega_{\mathrm{m},0}(1+\mathrm{z})^3 + \Omega_{\mathrm{k},0}(1+\mathrm{z})^2 + \Omega_{\Lambda,0}}$ is the cosmological factor, N(z) is the redshift distribution of the sources and χ(z) is the line of sight comoving distance. Equation 5.5 can be used to estimate $\mathrm{r_0}$ using the angular clustering amplitude A and the redshift distribution shown in Figure 5.4.

The theoretical value of 1.8 for γ, as predicted by [232] is consistent with the values across various surveys, as well as within 2 σ of the current analysis (tabulated in Table 5.2). Thus the theoretical value of γ, the distribution of A obtained from the MCMC distribution discussed in 5.3.2 and the combined redshift distribution distribution discussed in Section 5.2.2 are used to estimate the value of $\mathrm{r_0}$. Figure 5.6 shows the probability distribution function (PDF) of the spatial clustering length. As already mentioned, the median redshift of the samples is $\sim$0.78, and at this redshift, the median value of $\mathrm{r_0}$ is $3.50^{+0.50}_{-0.50}$ Mpc $\mathrm{h^{-1}}$, where the errors are the 16th and 84th percentile errors.

5.3.4 The Bias Parameter

The bias parameter is used to quantify the relation between the clustering property of luminous sources and the underlying dark matter distribution. The ratio of the galaxy to the dark matter spatial correlation function is known as the scale-independent linear bias parameter $b(z)$ [267, 268, 323]. For cosmological model with dark matter governed only by gravity, following [221, 222, 258], $b(z)$ is calculated as :

$$b(z) = \left(\frac{r_0(z)}{8}\right)^{\gamma/2} \frac{J_2^{1/2}}{\sigma_8 D(z)/D(0)} \tag{5.6}$$

where $J_2 = 72/[(3\text{-}\gamma)(4\text{-}\gamma)(6\text{-}\gamma)2^{\gamma}]$, $D(z)$ is the linear growth factor, calculated from CMB and galaxy redshift information [324], and σ_8^2 is the amplitude of the linear power spectrum on a comoving scale of 8 Mpc h^{-1}.

For this work, the bias parameter has been calculated using the median redshift value of the r_0 distribution with the 16th and 84th percentile errors. The value of the bias parameter $b(z)$ at z=0.78 is $2.22^{+0.33}_{-0.36}$.

5.4 Estimation of Correlation Function: AGNs and SFGs

This section discusses the angular and spatial correlation scales and the bias parameter obtained for the two separate populations of sources (i.e., AGNs and SFGs). The obtained values are also compared with previously reported values using radio and other bands data. Following a similar procedure as that done for the entire population, initially, the angular clustering was calculated, and a power law of the form $A\theta^{1-\gamma}$ was fitted. The best value of clustering amplitude A is determined, once again keeping γ fixed at the theoretical value of 1.8 for both AGNs and SFGs populations. Figure 5.7 shows the angular correlation function of AGNs (left panel) and SFGs(right panel). Using the MCMC simulations as discussed previously, the clustering amplitudes, log(A) have values $-2.18^{+0.20}_{-0.20}$ and $-1.69^{+0.10}_{-0.10}$ respectively for

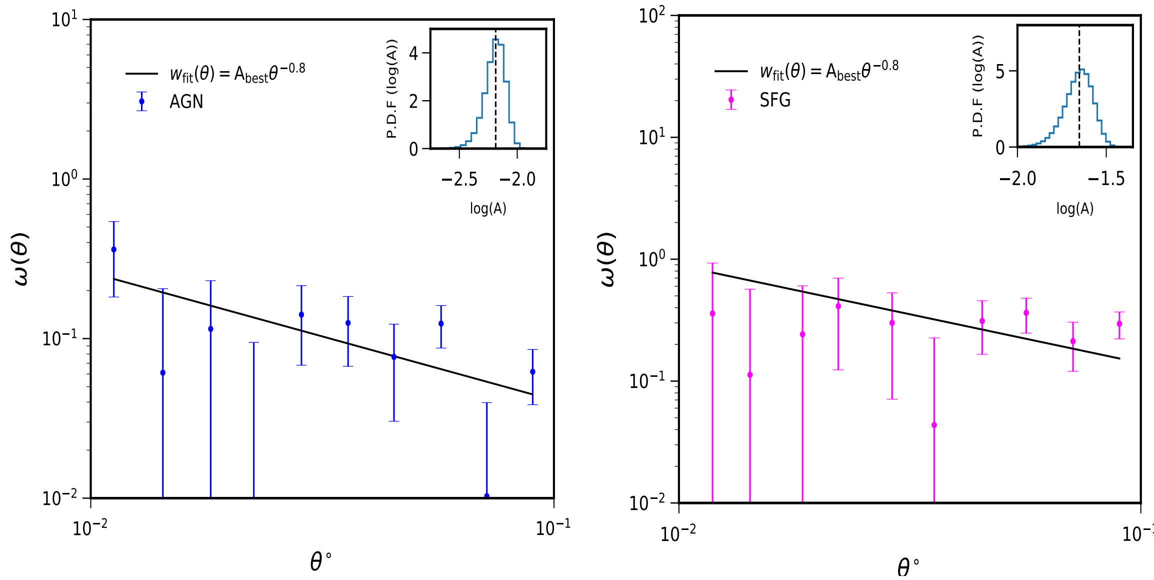

Figure 5.7: Angular correlation function for sources classified as AGNs **(left panel)** and SFGs **(right panel)**. The slope of the best fit function γ is fixed at 1.8 and the distribution of correlation amplitude (A) is shown in the inset of each panel.

AGNs and SFGs. The results of the fit and the subsequent values of clustering length and bias parameter obtained here and results from previous surveys in radio wavelengths are also tabulated in Table 5.3.

The spatial clustering length and bias parameter b_z for the AGNs with z_{median}=1.02 are $8.30^{+0.96}_{-0.91}$ Mpc h^{-1} and $3.74^{+0.39}_{-0.36}$. For SFGs with z_{median} ≈0.20, the values are r_0= $3.22^{+0.34}_{-0.32}$ Mpc h^{-1} and b_z=$1.06^{+0.1}_{-0.1}$. It is seen that the spatial clustering length and consequently the bias factor for AGNs is more than SFGs, which implies that the latter are hosted by less massive haloes, in agreement with previous observations [222, 257, 258, 327–330].

5.4.1 Comparison with previous Observations

Figure 5.8 shows the observationally determined values from surveys at various wavebands for r_0 as a function of their redshift, while Figure 5.9 shows the same for the bias parameter. The left and right panels of Figure 5.8 are for AGNs and SFGs, respectively. Table 5.3 summarizes the values obtained in radio surveys only, while Figures 5.8, 5.9 show the observed values for surveys at radio as well as other wavebands, e.g. IR and X-Ray.

The clustering length for AGNs in this work is at z_{median} $\approx$1.02 is $8.30^{+0.96}_{-0.91}$ Mpc h^{-1}. Using X-Ray selected AGNs in the COSMOS field, [327] obtained clustering lengths at redshifts upto $\sim$3.0. They divided their sample into a number of bins, to obtain r_0 at different median redshifts. For their entire sample, taking slope of the angular correlation function as 1.80, r_0 was $8.39^{+0.41}_{-0.39}$ Mpc h^{-1}, for a median redshift of 0.98. It is consistent with the value obtained here at a similar redshift. The clustering length with this work is also consistent within error bars for AGNs at 400 MHz and 610 MHz of [258]. For their work, they obtain an r_0 value of $7.30^{+1.14}_{-1.12}$ Mpc h^{-1} at $z \approx 0.91$, $6.00^{+1.5}_{-1.3}$ Mpc h^{-1} at $z_{median} = 0.84$. The clustering length estimates for radio selected AGNs in the COSMOS field at 1.4 GHz [257] and 3 GHz [222] observed with the VLA also agree within error bars with the estimates obtained here. [257] have found a clustering length of $7.84^{+1.75}_{-2.31}$ Mpc h^{-1} at z $\approx$1.25 while [222] obtained $6.90^{+0.60}_{-0.70}$ Mpc h^{-1}, $9.60^{+0.70}_{-0.70}$ Mpc h^{-1} and $7.30^{+0.90}_{-0.90}$ Mpc h^{-1} at $z \approx$ 0.70, 1.24, 1.77 respectively. Using X-ray selected AGNs in the CDFS field, [326] obtained a value of $10.30^{+1.7}_{-1.7}$ Mpc h^{-1} at z $\approx$ 0.84. This value though higher than the values for radio selected AGNs, is still consistent within error bars.

For the SFGs population (right panel of Figure 5.8), the median redshift is 0.20. At this redshift, the clustering length is $3.22^{+0.34}_{-0.32}$ Mpc h^{-1}. This estimate is at a redshift lower than previous observations. An extensive study at mid-IR frequency has been done by [329] for SFGs. The lowest redshift probed in their study is 0.31, where r_0 is $3.41^{+0.18}_{-0.18}$ Mpc h^{-1}. Thus, the value is consistent with that obtained here at a nearby redshift. [331] studied the clustering of SFGs using the Herschel PACS Evolutionary Probe observations of the COSMOS and Extended Groth Strip fields.

They found clustering lengths for SFGs out to $z \approx 2$. For the ELAIS-N1 field at 400 MHz and 610 MHz, [258] reported clustering length of $4.62^{+0.39}_{-0.40}$ Mpc h^{-1} and $4.16^{+0.70}_{-0.80}$ Mpc h^{-1} at redshifts 0.64 and 0.87 respectively. The 3 GHz COSMOS field studies of [222] gave clustering lengths $5.00^{+0.50}_{-0.60}$ Mpc h^{-1} and $6.1^{+0.60}_{-0.70}$ Mpc h^{-1} respectively at z$\approx$ 0.62 and 1.07. The mid-IR selected samples for Lockman Hole give r_0 values $4.98^{+0.28}_{-0.28}$ Mpc and $8.04^{+0.69}_{-0.69}$ Mpc h^{-1} at $z \approx 0.7$ and 1.7 respectively. Similarly, the mid-IR sample for [328] has clustering lengths r_0 is $4.25^{+0.12}_{-0.12}$ Mpc h^{-1} and $3.81^{+0.10}_{-0.10}$ Mpc h^{-1} for $z \approx 0.67$ and 0.73.

The results have also been compared with the assumed bias models of the semi-empirical simulated catalogue of the extragalactic sky, the Square Kilometer Array Design Studies (referred to as SKADS henceforth, [121]). This simulation models the large-scale cosmological distribution of radio sources to aid the design of next-generation radio interferometers. It covers a sky area of $20° \times 20°$, with sources out to a cosmological redshift of $z \sim$20 and a minimum flux 10 nJy at 151, 610 MHz & 1.4, 4.86 and 18 GHz. The simulated sources are drawn from observed and, in some cases, extrapolated luminosity functions on an underlying dark matter density field with biases to reflect the measured large-scale clustering. It uses a numerical Press–Schechter [228] style filtering on the density field to identify clusters of galaxies. The SKADS catalogue has been used here for statistical inference of the spatial and angular clustering variations of the sources with redshift. It should be mentioned here that the T-RECS catalogue [122] incorporates more updated results from the recent observations. However, the evolution of the bias parameter and clustering length with redshift is not available for the same; hence SKADS has been used.

The bias parameter for AGNs and SFGs at z_{median} 1.02 and 0.20 is $3.74^{+0.39}_{-0.36}$ and $1.06^{+0.10}_{-0.10}$ respectively. Although the value for AGNs is slightly higher than those obtained by [222] and [258], it is still with reasonable agreement with the SKADS for FR-I galaxies. Comparison with the population distribution of the SKADS simulation of [121], in terms of both clustering length and the bias parameter (solid magenta pentagons in Figure 5.9) show that the AGN population is dominated by FR-I type galaxies hosted in massive haloes with $\sim M_h = 5\times10^{13}\mathrm{h}^{-1}\mathrm{M}_\odot$. It can also seen from Figure 5.9, that the mass of the haloes hosting the SFG samples of the current sample is $\sim M_h = 3\times10^{12}\mathrm{h}^{-1}\mathrm{M}_\odot$. Thus, it is seen that the SFGs have a lower range of halo masses compared to AGNs, which implies that the latter inhabits more massive haloes and are more biased tracers of the dark matter density field.

5.5 Discussion

The analysis of clustering properties of radio selected sources in the Lockman Hole region presented in this work is one of the first results reported at 325 MHz. Similar studies were previously done at 400 MHz for the ELAIS-N1 field [258], however for a much smaller area. Beside analysing at a mostly unexplored frequency, this work also presents a comparatively large area with a significant number of sources. Previous clustering study of the same area using 1.4 GHz data from WSRT by [279] used 1173 sources with a flux density cut-off of 0.12 mJy at 1.4 GHz (or 0.4 mJy at 325 MHz). Their obtained clustering amplitude is slightly higher than previous surveys, as acknowledged by the authors. However, further investigation is required to ascertain the reason for the deviation. Clustering analysis was also done using the recent LoTSS observation of the HETDEX spring field [278]. The clustering analyses were produced with many flux density cut-offs and masks, and the most reliable estimate was for a flux density limit of 2 mJy at 150 MHz (or ∼1.0 mJy at 325 MHz). The clustering amplitude estimate at 2 mJy limit for LoTSS is consistent with that obtained here within error bars. As seen from Figure 5.3.2, the clustering amplitude obtained here agrees with previous observations. The slightly higher values for the bias parameter of the AGNs for this work, compared to that of [222] and [258], may be attributed to the different flux limits of the studies. This implies

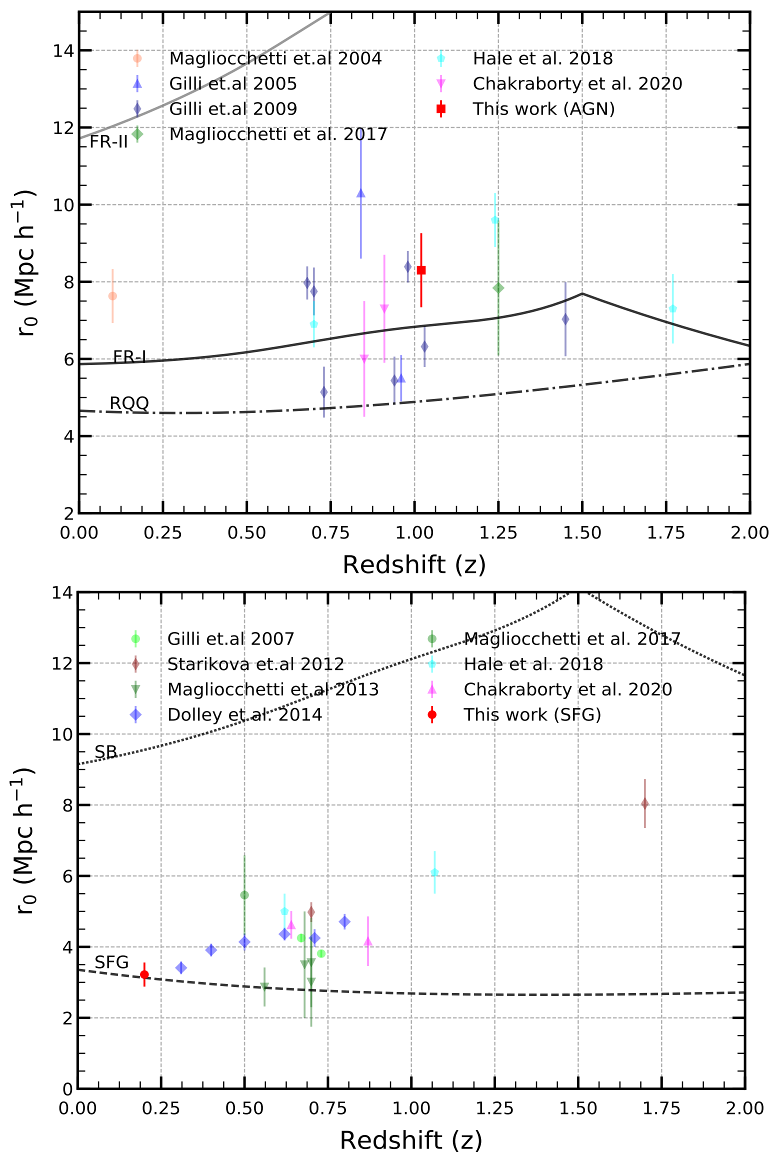

Figure 5.8: **(top panel)** Spatial correlation length (r_0) for AGNs as a function of redshift (red square). For comparison, the values for the same obtained by [325], [326], [327], [257], [222] & [258] have been plotted. Predictions from SKADS [121] have also been plotted for Radio Quiet Quasars (RQQ), FR-I and FR-II radio galaxies. **(bottom panel)** Spatial correlation length (r_0) SFGs as a function of redshift (red circle). For comparison, values obtained by [328], [330], [331], [329], [257], [222], & [258] have been plotted. Predictions from SKADS [121] have also been plotted for SFGs and Starburst galaxies (SB).

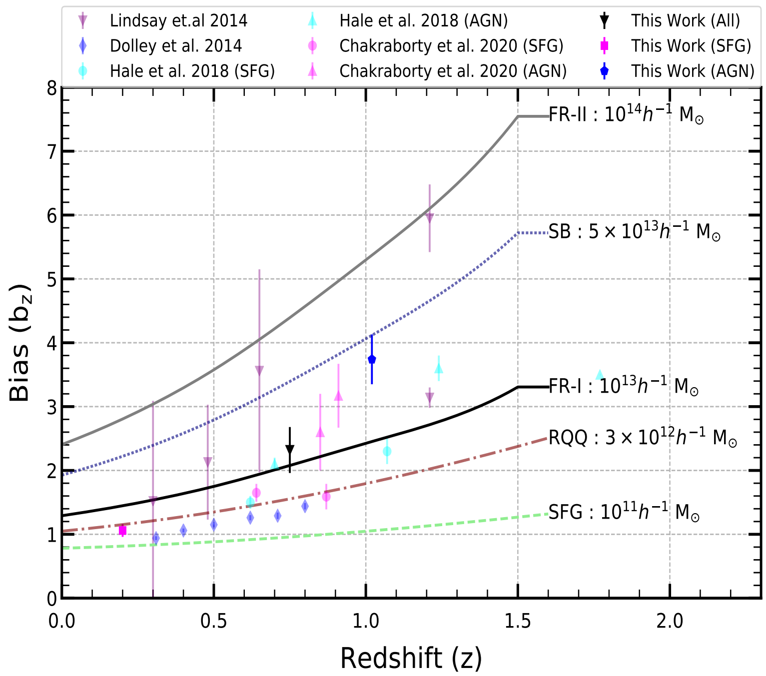

Figure 5.9: Bias parameter b_z for whole population (black inverted triangle) as well AGNs (blue pentagons) and SFGs (magenta squares). Bias parameters from previous observations by [221] (purple inverted traingles), [329] (blue diamonds), [222] (cyan triangles for AGNs and cyan circles SFGs) & [258] (magenta triangles and circles for AGNs and SFGs respectively) are shown. Predictions from SKADS is also shown by the continuous curves.

that each of these observations are probing slightly different populations of sources (with slightly different luminosities as discussed later). Nevertheless, as seen from Figure 5.9, the values are broadly consistent with each other.

The angular clustering amplitude for this work as shown in Figure 5.7 agree with previous observations, as seen in Table 5.3. The angular clustering of sources for this work are calculated with `TreeCorr` using the default values for most parameters. It has been shown in [278] that using default value of 1 for the parameter `bin_slop` gives less accurate results than for values ≤ 1. [278] obtained the most accurate values for `bin_slop`=0. They also showed that angular clustering amplitudes deviate largely from precise values (calculated from a separate brute-force algorithm, see [278] for details) at angular scales $\gtrsim 1°$. However, the computation times also significantly increased for `bin_slop`=0. The use of default parameters in this work might be the cause slight oscillation of the correlation function seen around the best fit curves. Nevertheless, since results obtained here are in reasonable agreement with the previous observations and owing to constraints in the available computing power, the default values have been used.

The clustering lengths and bias parameters obtained here also agree with previous studies, as evident from Table 5.3 and Figures 5.8, 5.9. Comparison of bias parameter with SKADS simulation [121] shows that the expected mass for dark matter haloes hosting AGNs is orders of magnitude higher than that for SFGs. This trend is consistent with previous observations (see for example [222, 257, 258]). Studies on the luminosity of the AGNs and SFGs suggest that it is correlated with source clustering, and hence with the bias parameter and host halo mass [222]. Using observations of high and low luminosity AGNs in the COSMOS field, [222] showed that the luminosity and clustering are correlated, with the higher luminosity AGNs residing in more massive galaxies [332]. Thus, they are hosted by more massive haloes. Again, this points to AGNs (which are in general more luminous than SFGs) being hosted by more massive haloes. However, it should also be mentioned that the two population of AGNs in [222] studied are at different redshifts, with the low excitation population studied at $z \lesssim 0.65$. So it is possible that this population may evolve into higher mass haloes at higher redshifts. Moreover, some works (for example [333]) do not find any relation between clustering and luminosity. Nevertheless, following [222], studies using a larger population of samples covering a large range of luminosities is required for probing the relationship with clustering.

It is also observed in Figure 5.9 that the bias parameter for SFGs (solid blue square) is higher than the SKADS predicted values for this population. This trend is consistent with that observed in previous studies of [222] (cyan circles) and [258] (light magenta triangles). The trend observed in [329] (light blue diamonds) is almost similar as well. There may be two reasons that cause this variation- contamination of the SFG sample by star-burst (SB) galaxies or underestimation of halo mass for SKADS. If there exists a few SB samples in the SFG population, comparison with SKADS (blue dotted curve in Figure 5.9) shows that the overall value for bias (as well as r_0) will be higher than an uncontaminated sample. However, the most likely reason remains the second one, the halo mass used in the SKADS is not a correct representation, which has also been hinted at by comparing the values obtained from previous observations (for instance [222, 258, 329]) in Figure 5.9. However, it is also seen from Figure 5.8 spatial clustering length for SFGs agrees with SKADS. The exact reason for agreement of r_0 and disagreement of $b(z)$ for SFGs in this work with SKADS is unclear, and will be investigated in detail in later works.

Analyses like the one presented here are important for fully understanding how bias scales with redshift as well as with source properties like luminosity. This is important for cosmology, since the bias relates to dark matter distribution, and is thus essential for understanding the underlying cosmological parameters that define the Universe.

It should be mentioned here that the current study also has certain limitations. Due to unknown systematics at large scales, and lack of optical matches, the entire observed field could not be utilised for this study. Additionally, the source classification is done based solely on the radio luminosity of the sources. As has already been mentioned previously, [307] showed that the chances of contamination between the populations is very less using this criteria. Nonetheless, there are several other methods that can also be used to classify AGNs and SFGs in a sample (detailed in Section 5.2.3). Future works will present a more detailed analysis using the different multi-wavelength classification schemes available.Such multi-frequency studies, combined with the present work and similar studies with other fields will enhance the knowledge of the extragalactic sources and provide more insights into the processes governing their formation and evolution.

5.6 Conclusion

This work investigates the higher-order source statistics, namely the angular and spatial clustering of the sources detected in the Lockman Hole field. The data was observed by the legacy GMRT at 325 MHz. The details of data analysis and catalogue extraction are discussed in [1]. The initial step involved merging the multi-component sources present in the raw catalogue. The resultant catalogue was cross-matched with SDSS and HELP catalogues to identify sources with either spectroscopic or photometric redshift information. A region of radius 1.8° around the phase center was selected for optical identifications, yielding ~95% matches. All the sources with redshift distribution were separated into AGN and SFG populations using the criterion for radio luminosity. The angular correlation function was determined for the combined population for separation between 36" to 2°. A power law fitted to this function, keeping a fixed power law index of $\gamma = 1.80$, as estimated theoretically [232]. This gave the value of clustering amplitude $\log_{10}(\mathrm{A}) = -2.73^{0.11}_{-0.15}$.

The source population was further divided into AGNs and SFGs based on their radio luminosity, and clustering analyses were done for these populations as well. Using the redshift information and the clustering amplitude, spatial correlation length was determined using Limber inversion for the AGNs and SFGs. The correlation length and bias parameters have been obtained for the full sample, as well as the classified AGN and SFG population. For the full sample at $\mathrm{z}_{median} \approx 0.78$, $\mathrm{r}_0 = 3.50^{+0.50}_{-0.50}$ Mpc h^{-1} and $b(z) = 2.22^{+0.33}_{-0.36}$. For AGNs, the values are $\mathrm{r}_0 = 8.30^{+0.96}_{-0.91}$ Mpc h^{-1} and $b(z) = 3.74^{+0.39}_{-0.36}$ at $\mathrm{z}_{median} \approx 1.02$. At $\mathrm{z}_{median} \approx 0.20$, SFGs have values $\mathrm{r}_0 = 3.22^{+0.34}_{-0.32}$ Mpc h^{-1} and $b(z) = 1.06^{+0.10}_{-0.10}$. The clustering length for AGNs is reasonably consistent with [327] as well as [258]. For SFGs, the values are consistent with [329].

The obtained values have also been compared with the SKADS simulation of [121]. The comparative analysis suggests that the AGNs are dominated by FR-I galaxies, with host dark matter halo masses of M_h=5-6 $\times$ $10^{13}h^{-1}M_{\odot}$. For SFGs, the estimated halo mass obtained from SKADS is lower compared to the value $\sim M_h$=3 $\times 10^{12}h^{-1}M_{\odot}$ obtained here. The halo mass obtained here are in agreement with previous literature [222,258,329]. It is worthwhile to mention that while the current classifications are based on radio luminosity for each source alone, the results are in agreement with clustering properties for populations of AGNs and SFGs in X-Ray and mid-IR surveys as well. However, there are some deviations from the predictions of the SKADS simulation. This deviation, seen in other observations as well, emphasizes the need for wider and deeper low-frequency observations. These will be able to constrain the host properties better, leading to better models formation and evolution of the different sources, and better understanding of the distribution of these populations over space and time (redshift).

The study done in this work aims to characterize the clustering and bias of an observed population of radio-selected sources, both as a combined population and as distinct classes of sources (namely AGNs and SFGs). This work is the first to report the clustering properties of radio-selected sources at 325 MHz. Thus, current data, being at a frequency with little previous clustering study, bridges the gap between low frequency and high-frequency studies. It also has the advantage of covering a wider area than many recent studies with moderately deep RMS values, thus probing a larger number of sources with fluxes at the sub-mJy level ($\sim$0.3 mJy). More such studies using observational data are required for constraining cosmology and probing how different source populations are influenced by their parent halos and how they evolve with time (redshift). Additionally, for sensitive observations of CD/EoR and post-EoR science, accurate and realistic models for compact source populations that comprise a significant fraction of foregrounds are required. Studies on the effect of imperfections in foreground modeling in the power spectrum estimates will require detailed observational studies of source position and flux distributions. For realistic estimates, second-order statistics like clustering also cannot be ignored. Thus many more analyses like the one done in this paper will be required for better understanding the effects of the interplay between the various cosmological parameters on the different populations of sources and putting constraints on said parameters. Sensitive large area surveys like the MIGHTEE [248] with the MeerKat telescope, LoTSS [201,284,334] with the LOFAR, EMU [335,336] with the ASKAP, as well as those to be done with the upcoming SKA-mid telescope would provide wider as well as deeper data for doing cosmology. The current work demonstrates that even instruments like the legacy GMRT can provide reasonable data depth and coverage for cosmological observations.

Table 5.3: Spatial Clustering Length and Bias Parameter from Different Observations. The columns are respectively name of the survey field, observing frequency in MHz, type of radio source (AGNs/SFGs), median redshift, angular clustering amplitude, spatial clustering length in Mpc h^{-1} & bias parameter value.

Observation	Frequency (MHz)	Source type	z_{median}	$\log_{10}(A)$	r_0 (Mpc h^{-1})	$b_{z_{median}}$	Reference
COSMOS	3000	AGNs	0.70	$-2.30^{+0.1}_{-0.1}$	$6.9^{+0.60}_{-0.70}$	$2.1^{+0.2}_{-0.2}$	[222]
		AGNs	1.24	$-2.60^{+0.1}_{-0.1}$	$9.6^{+0.70}_{-0.70}$	$3.6^{+0.2}_{-0.2}$	
		AGNs	1.77	$-2.60^{+0.1}_{-0.1}$	$7.3^{+0.90}_{-0.90}$	$3.5^{+0.4}_{-0.4}$	
		SFG	0.62	$-2.60^{+0.1}_{-0.1}$	$5.0^{+0.50}_{-0.60}$	$1.5^{+0.1}_{-0.2}$	
		SFG	1.07	$-2.90^{+0.1}_{-0.1}$	$6.1^{+0.60}_{-0.70}$	$2.3^{+0.2}_{-0.2}$	
VLA-COSMOS	1400	AGNs	1.25	$-2.79^{+0.1}_{-0.1}$	$7.84^{+1.75}_{-2.31}$	-	[257]
		SFG	0.50	$-2.36^{+0.3}_{-0.3}$	$5.46^{+1.12}_{-2.10}$	-	
ELAIS N1	400	AGNs	0.91	$-2.22^{+0.16}_{-0.16}$	$7.30^{+1.4}_{-1.2}$	$3.17^{+0.5}_{-0.5}$	[258]
		SFG	0.64	$-2.16^{+0.05}_{-0.06}$	$4.62^{+0.39}_{-0.40}$	$1.65^{+0.14}_{-0.14}$	
ELAIS N1	612	AGNs	0.85	$-2.30^{+0.02}_{-0.03}$	$6.0^{+1.5}_{-1.3}$	$2.6^{+0.6}_{-0.5}$	[258]
		SFG	0.87	$-2.19^{+0.01}_{-0.02}$	$4.16^{+0.7}_{-0.8}$	$1.59^{+0.2}_{-0.2}$	
Lockman Hole	325	AGNs	1.02	$-2.18^{+0.20}_{-0.20}$	$8.30^{+0.96}_{-0.91}$	$3.74^{+0.39}_{-0.36}$	This work
		SFG	0.20	$-1.65^{+0.1}_{-0.1}$	$3.22^{+0.34}_{-0.32}$	$1.06^{+0.10}_{-0.10}$	

Chapter 6

Simulating Effects of Systematic Errors on EoR Observations

This chapter has been adapted from *Observing the Reionization : Effect of Calibration and Position Errors on Realistic Observation Conditions* [2]

6.1 Introduction

The thermal history of the evolution of the Universe can be traced back to the surface of the last scattering with the help of the Cosmic Microwave Background (CMB). The evolutionary history of the Universe between the post-recombination era until the period after the neutral intergalactic medium (IGM) became ionized once again remains largely unexplored with observational data. Theoretical models suggest that the Universe's first structures- the earliest stars and galaxies formed during this period (viz. redshifts between ~ 30 and ~ 12). The first stars, generated at Cosmic Dawn (CD), formed due to the small-scale fluctuations in matter density arising from the gravitational instabilities. Ionizing radiations produced from these objects ionized the IGM causing the last phase transition in the evolution of the universe [94]. This phase transition period is called Epoch of Reionization (EoR) (for comprehensive reviews see [40, 41, 44, 45, 95, 337]). Using quasar absorption spectrum [96] and Thompson scattering optical depth [97], the extended Epoch of Reionization (EoR) is constrained between the redshift interval of $6 \lesssim z \lesssim 15$.

Understanding the conditions prevailing during these early stages of the Universe requires a reliable probe that can trace thermal evolution history. The hyperfine transition line of neutral hydrogen, with a rest wavelength of 21-cm, is the most promising probe into the IGM of these early cosmic epochs [32, 98, 99]. The contrast of the brightness temperature corresponding the 21cm line transition against the CMB gives a differential brightness temperature which is the observable through the all sky-averaged "global signal" (using a single total power radio telescope) or as fluctuations in the spatial correlations (using radio interferometers) [38]. Radio interferometric observations are necessary for estimation of astrophysical parameters and constraining the nature of reionizing sources. Detection of an absorption trough in the global signal experiment centered at 78 MHz was reported by the Experiment to Detect Global Epoch of Reionization Signature (EDGES) team [338]. Besides EDGES, single antenna experiments like BIGHORNS [339], SCI-HI [340], SARAS [341], LEDA [342] are also aiming to detect the global signal but are yet to report any detection.

The most sensitive operational interferometers like the GMRT, MWA, LOFAR, HERA have all set upper limits on the power spectrum (PS) amplitude of the signal [59, 63, 65, 70, 71, 77] but and no confirmed detection of the cosmological HI 21cm signal. The cosmological signal is prone to contamination due to bright astrophysical foregrounds [1, 112–115, 175, 183, 343]. Other sources of contamination are Earth's ionosphere and several instrumental systematics. Substantial amount of focused research has taken place over the past decade to understand and quantify each components of these contamination, namely instrument model [344–347], calibration [69, 348–354], foreground avoidance or removal [117, 123, 211, 355–359], PS estimation [100, 160, 161, 362–365] and Earth's ionosphere [360, 361]. Using these advancements, highly sensitive next generation interferometers like the Hydrogen Epoch of Reionization Array (HERA, [101]) and the Square Kilometer Array (SKA1-Low, [119]) are expected to detect the 21-cm signal and characterize the multi-redshift PS, leading to tighter constrains on the astrophysical parameters in the early Universe. The upcoming SKA1-Low is being designed to be sensitive enough to detect the PS precisely. Additionally, it is also projected to be able to create tomographic images of the HII regions [47]. The EoR signal is inherently isotropic in spatial wavenumber (k) space. Additionally, it also shows spectral structure. Thus, despite the spectrally smooth foreground contamination can be distinguished from the foreground contaminants [364, 366–369, 371, 372]. The spectrally smooth nature of the foregrounds, along with the inherent instrumental chromaticity, keeps the contamination confined to the "wedge" in the cylindrical Fourier space (i.e., in the 2D PS). The region outside the wedge where the foregrounds are subdominant compared to the signal is called the"EoR window" [370]. However, the interaction of astrophysical foregrounds and the instrument causes the wedge power to leak into the clean modes of the window- an effect called "mode mixing" [370]. If not properly mitigated, mode mixing is likely to confuse the detection.

One of the major limiting systematic that causes the problem in cosmological signal detection is improper calibration. In general, the most common calibration approach for radio astronomy is "sky-based" calibration. Since Cosmic Dawn/Epoch of Reionization (EoR) observations are limited to shorter baselines, these observations suffer from poor angular resolutions and higher confusion noise. This may lead to inaccurate sky models for calibration. Errors or inaccuracies in this process propagate as residual errors and ultimately hinder the target cosmological signal detection [110,117,349,351]. The other calibration approach being explored by various interferometers is the redundant calibration approach . Telescopes like the HERA use the redundant calibration approach. In this approach, repeated simultaneous measurements of the interferometers' redundant baselines can simultaneously solve for the incoming sky signal and instrumental parameters. The redundant method calibrates for real visibility by considering the prior that sky visibilities are equal for redundant baselines, having the advantage of not requiring a sky model for performing gain solutions for antennas. Nevertheless, this works only for highly regular arrays and identical antenna beam patterns [373,374].

It must be pointed out that irrespective of the accuracy of the calibration approach employed, the ultimate limitation is set by the signal-to-noise ratio (SNR). It has been shown in [110], that calibration errors ∼0.1% causes a drastic lowering of the dynamic range, thereby obscuring the signal[1]. Thus, studies are necessary for estimating the extent of the tolerance of the imperfections, which potentially contribute to the noise that obscure the detection of the faint cosmological signal. This paper aims to quantify this level of tolerance (or the accuracy of the calibration algorithm) required for successful detection of the redshifted 21-cm signal from the cosmic reionization using various sensitive telescopes.

Astrophysical foregrounds like diffuse galactic synchrotron emission, extra-galactic free-free emission, and compact sources substantially affect the interferometric data sets used for EoR signal recovery. Currently, most studies focus on determining the best strategy- for avoidance, suppression, or subtraction for handling the foregrounds. While diffuse emission is a significant contaminant at large angular scales, EoR data sets also require the careful handling of the extragalactic point sources. Inaccurate removal of bright, compact sources can lead to large residuals in the data and obscures the detection of the cosmological signal. Besides ionosphere, incomplete UV-plane coverage gives rise to large sidelobes that contaminate the data. All these corruption effects lead to constructing an imperfect sky model, which causes

[1]Dynamic range is the ratio of the peak flux in the image and the RMS noise in a source-free region. [110] showed that with calibration error of ∼0.1%, the dynamic range lowers to 10^5 compared to the required 10^8

residuals that adversely affects the cosmological signal recovery. The study on the impact of incomplete sky model has been done by [110, 117, 123, 349, 373]. However, there is still a lack of detailed study on how the effects compare for different telescope arrays targeting the cosmological 21-cm signal from reionization.

Our current work presents a proof of concept on developing an end-to-end pipeline to simulate radio interferometric observations using a realistic sky model and telescope observing parameters. The studies have been carried out taking four array configurations - SKA1-Low core (stations around $\sim$ 2000 m of the central station), HERA 350 dish configuration, MWA Phase-1 (MWA1) 128 tile configuration, and MWA 256 tile configuration (MWA-256). This paper is organized in the following manner: in Section 6.2, the detailed simulation methodology is described. Section 6.3 discusses briefly the observational effects considered in this work; Section 6.4 discusses the analyses formalism which is followed by the Section 6.5. The implications for the assumptions considered in this work is presented in Section 6.6; finally, the paper is concluded Section 6.7.

6.2 Synthetic Observations

This section describes the steps involved in the synthetic observation and analysis pipeline developed. Figure 6.1 shows the blocks that are used for generating the observations. The various simulation parameters and their respective values are described in Table 6.1. In the following subsection, each of the parameters in the input block is briefly described.

The simulations have been performed using the OSKAR software [375][2] package for SKA1-Low configuration and Common Astronomy Software Application (CASA, [376])[3] for HERA, MWA-1 & 2. The observations track the sky with phase center at α=15h00m00s and δ=-30°00'00" for 4 hours ($\pm$ 2 HA). The observing bandwidth

[2] https://github.com/OxfordSKA/OSKAR/releases

[3] https://casaguides.nrao.edu/index.php?title=Main_Page

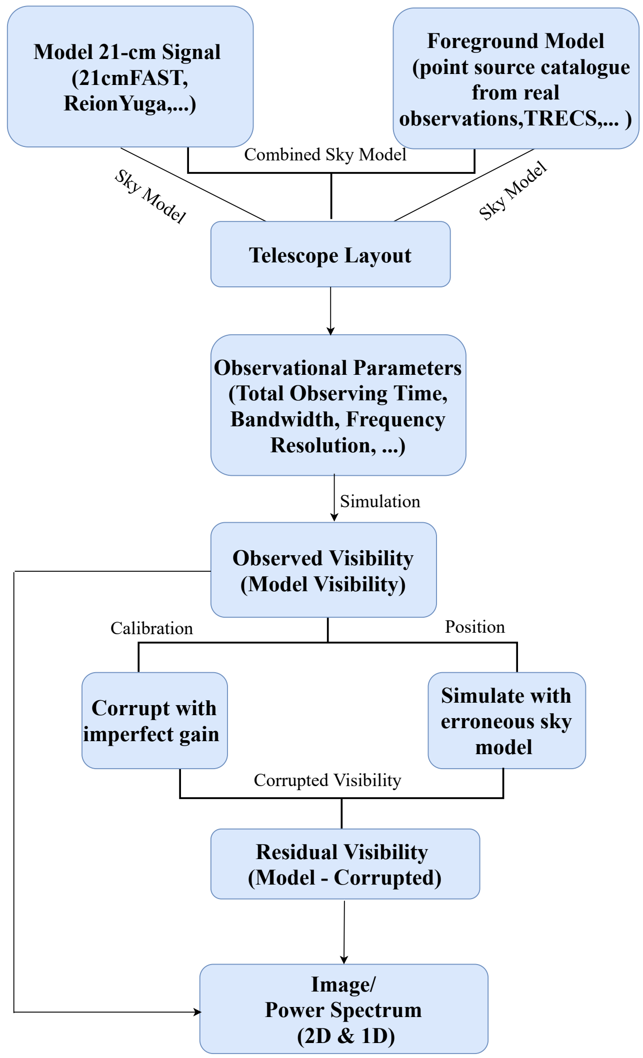

Figure 6.1: Schematic diagram of the observation pipeline.

is 8 MHz, with a channel separation of 125 kHz. For simplicity, the simulations are noise-free. It should also be mentioned that throughout the paper, best fitted cosmological parameters from the Planck 2018 results [97] are used: $\Omega_{\rm M} = 0.31$, $\Omega_\Lambda = 0.68$, $\sigma_8 = 0.811$, $H_0 = 67.36$ km s^{-1} Mpc^{-1}.

Table 6.1: Parameter values for the simulations

Parameter	Value
Central Frequency	142 MHz (z$\sim$9)
Bandwidth	8 MHz
Number of frequency channels	64
Field of view	4°
Number of array elements:	
SKA1-Low	296
HERA	350
MWA Phase-1	128
MWA-256	256
Maximum baseline (m):	
SKA1-Low	$\sim$2000
HERA	$\sim$880
MWA Phase-1	$\sim$2800
MWA-256	$\sim$5300
Synthesised beam (arcmin):	
SKA1-Low	$\sim$2.5
HERA	$\sim$8.1
MWA Phase-1	$\sim$2.6
MWA-256	$\sim$1.3
Thermal Noise (mJy beam^{-1}):	
SKA1-Low	$\sim$0.07
HERA	$\sim$0.52
MWA Phase-1	$\sim$17.59
MWA-256	$\sim$8.79

6.2.1 Telescope Model

Though several interferometers are targeting the detection of the redshifted 21-cm signal, this work undertakes a comparative analysis under identical error conditions for four different arrays- SKA1-Low, HERA, and MWA1 & MWA2. The telescope layouts are shown in Figure 6.2.

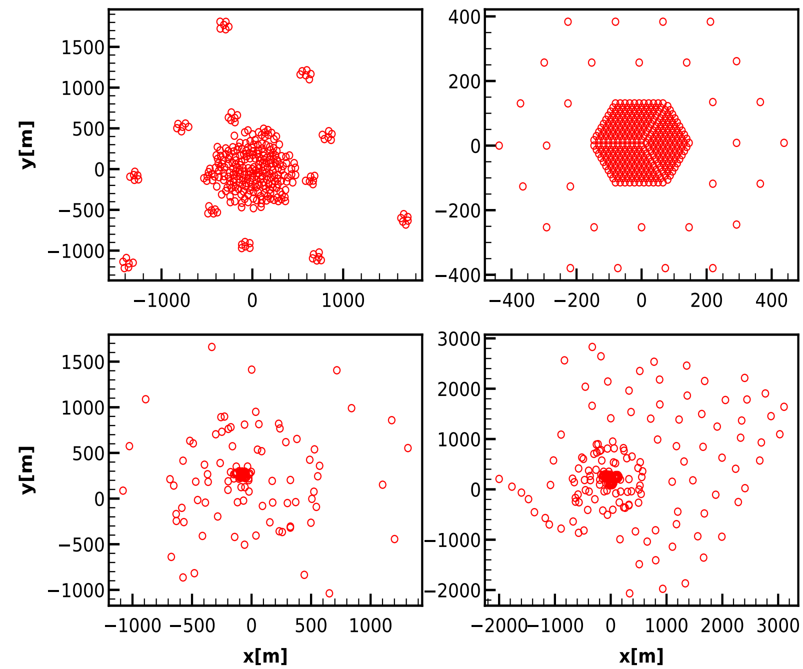

Figure 6.2: Telescope layouts used in the simulation. Clockwise from top left are SKA1-Low (Low), HERA, MWA-256 & MWA-I respectively.

The SKA[4], one of the most sensitive upcoming telescopes, is expected to detect the signal statistically [119] and also have just enough signal-to-noise for performing tomography [47]. SKA will be built at two sites as two separate arrays, a low-frequency array in Australia (frequency coverage between 50-350 MHz) and a mid-frequency array in South Africa (frequency coverage between 350 MHz to 14 GHz). The redshifted 21-cm signal from the EoR is present at frequencies of up to a few hundred MHz (∼200 MHz at z ∼6) and thus will be targeted by the former array (SKA1-Low). While the available documentation states that the SKA1-Low will have 512 "stations" of ∼ 40m diameter spread around the core or central station (having a maximum baseline length of ∼65 km, [377]), this work does not explore the entire layout. Since compact baselines will be sensitive to the target signal,

[4] https://www.skatelescope.org/

all stations within 2 km of the core (i.e., a maximum baseline of 2000m around the central station) are used [378]. This results in 296 stations around the central station, with a synthesized beam size of ∼2.5' at 142 MHz.

HERA[5] is an interferometer designed to search for the fluctuations in the 21-cm signal during the cosmic dawn and reionization [101]. Located in the South African Karoo Radio Astronomy Reserve, the complete array would comprise 350 parabolic dishes of 14m diameter each, spread in a highly redundant configuration (320 in a dense core and 30 outriggers). The instrument is planned to be operational between 100-200 MHz. This work uses the entire 350 tile configuration resulting in a maximum baseline of ∼880 m. At 142 MHz, this results in an ∼8' synthesized beam.

MWA[6], an SKA precursor facility, is located at the Murchison Radio Observatory in Australia (adjacent to the SKA1-Low site). The array has an operating frequency range of 70-300 MHz. In the initial deployment stage, the array had 128 square-shaped "tiles" (a 4×4 array of dipoles). In Phase 2, the 128 tile configuration was extended to include additional 128 tiles [379]. The upgraded array has a compact configuration with 56 tiles from earlier and 72 new tiles arranged in two compact hexagons, which enhance EoR PS capacity [379]. For this work, two MWA configurations are used- entire Phase 1 with 128 tiles and entire Phase 2 with 256 tiles [7]. The synthesized beam-width at 142 MHz are ∼2.5' ∼1.3' respectively for Phases 1 & 2 respectively.

6.2.2 The HI 21-cm Maps

We use two semi-numerical approaches, 21cmFAST [383, 384] and ReionYuga (formerly known as Sem-Num) [380–382] to generate the redshifted 21-cm signal maps

[5] https://reionization.org/

[6] https://www.mwatelescope.org/

[7] The MWA coordinates are available at https://www.mwatelescope.org/telescope/configurations/phase-ii

for our simulated observations. Both of these methods are developed independently, and they use different approaches of excursion set formalism to simulate the 21-cm maps from the EoR. Both of these methods have been tested against full radiative transfer simulations and found to be capable of simulating the signal power spectrum at large length scales with a reasonable accuracy [381]. They, however, show significant differences at small length scale power spectrum. Our motivation behind using two separate signal simulations is to test the robustness of the outcome of our data analysis pipeline. It is going to use both extracted image and power spectrum statistics for its evaluation.

21cmFAST generates the 21-cm maps from a matter density field obtained via Zeldovich approximation. It uses a guided excursion set formalism to convert the matter density field at a given redshift into an ionization field which is finally translated into 21-cm brightness temperature fluctuations. The excursion set approach allows it to generate multiple realizations of the 21-cm maps at a very low computing cost compared to a full radiative transfer approach.

The ReionYuga or Sem-Num (a serial version of ReionYuga), on the other hand, uses a N-body simulation to generate the matter distribution at a given redshift and further uses a Friends-of-Friends algorithm to identify collapsed halos in this matter distribution. These halos are then considered the hosts of the first sources of lights (e.g., galaxies, quasars, etc.), which produces the ionizing photons responsible for reionizing the Universe. The density of the neutral hydrogen field (assuming that the hydrogen follows the dark matter distribution) and the ionizing photon field are then estimated at a much coarser grid compared to the original N-body simulation resolution, and using a guided excursion set formalism, an ionization field is created. This ionization field is then converted into a 21-cm brightness temperature field.

For our analysis in this paper, we use the output of an already existing run of Sem-Num, precisely the fiducial light-cone map of [385]. This light-cone volume has sky plane comoving extent of 500×500 $(h^{-1}\mathrm{Mpc})^2$ and the line-of-sight or redshift extent of $7.0 \le z \le 12.0$. The 21-cm map in this light-cone volume is saved on a grid of size $232 \times 232 \times 562$. We extract a cuboid of grid dimension $232 \times 232 \times 64$ from this light-cone volume and project it to a World Coordinate System (WCS) of the same dimension. This extracted cuboid is used as the signal model for our analysis. Similarly, we generate a 21-cm light-cone cuboid using 21cmFast of the same volume and treat it as our other signal model.

Figure 6.3 shows a slice of the signal at the central frequency (142 MHz, z$\sim$9) for 21cmFAST (top panel) and Sem-Num (bottom panel). It is evident from this figure that the reionization histories followed by these two models are not the same; thus, at the same redshift, they arrive at a significantly different level of IGM ionization. Therefore, the levels of resulting 21-cm fluctuations are also considerably different. Figure 6.4 shows the power spectrum from these two signal models as observed by four different telescope configurations discussed earlier. Here we assume that the observational data contains the signal alone. We follow that the resulting power spectra for the two signal models are not drastically different at the length scales of our interest. As 21cmFAST is a faster simulation to implement, the subsequent analysis has been done using this signal model alone.

6.2.3 Foreground Models

The foreground used for this work consists of compact sources only. Once again, two separate catalogs have been used - the 400 MHz catalog obtained using uGMRT of the field ELAIS N1 [175] and the Tiered Radio Extragalactic Continuum Simulation (T-RECS, [122]) to demonstrate the ability of the pipeline. The sources were redistributed around a declination of -30°. The fluxes were converted to 142 MHz values using $S_\nu \propto \nu^{-\alpha}$, with α value of -0.8. The minimum flux value present at 400 MHz was 100 μJy, and the highest flux is 0.6 Jy. T-RECS is a simulation

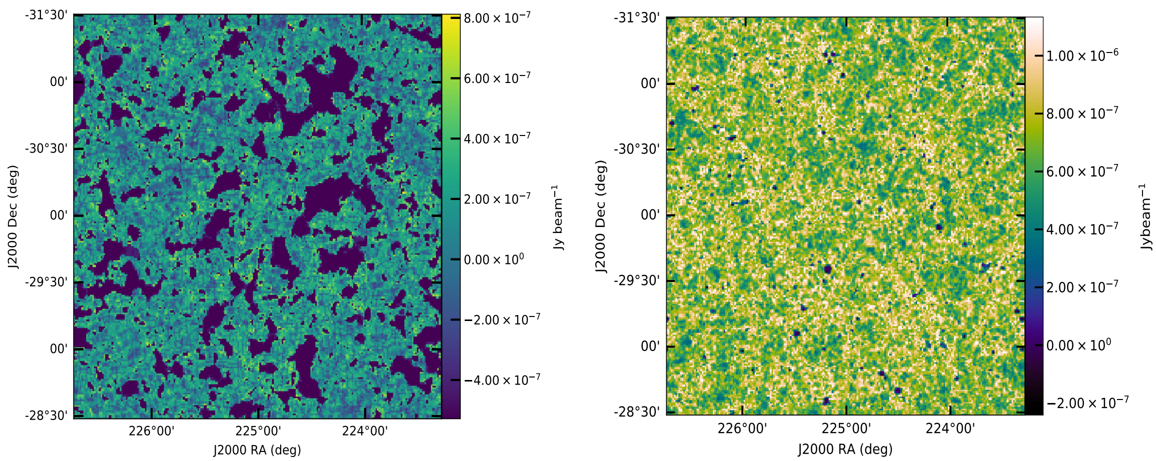

Figure 6.3: Slice through the HI 21-cm light cone at 142MHz, projected on a WCS using 21cmFAST (top) and SemNum (bottom). The light cones produced have different reionization histories and are thus not identical.

modeling the continuum radio sky in the range 150MHz to 20GHz. It models Active Galactic Nuclei (AGNs) and Star-Forming Galaxies (SFGs), comprising most radio galaxy populations. These simulations consider real data from recent observations to model the sources. This gives realistic cosmological evolution of the luminosity functions, the total intensity number counts, polarized intensity, and clustering properties. The source catalogs are available publicly for 1 deg^2, 25 deg^2, and 100 deg^2. For this work, from the 25 deg^2 catalog, sources within $\sim$4 deg^2 having flux cut-off the same as EN1 catalog is chosen (the brightest source in the constructed catalogue has a flux of 1.7 Jy at 150 MHz). Fluxes at 150 MHz were scaled to 142 MHz using $\alpha \approx$ -0.8. The catalog comprises 2522 sources in the chosen field of view (FoV) and flux cut-off.

It is to be noted that no diffuse emission model was used in the foreground model to keep the present analysis simple. Future works will incorporate models of the diffuse emission for investing their effect.

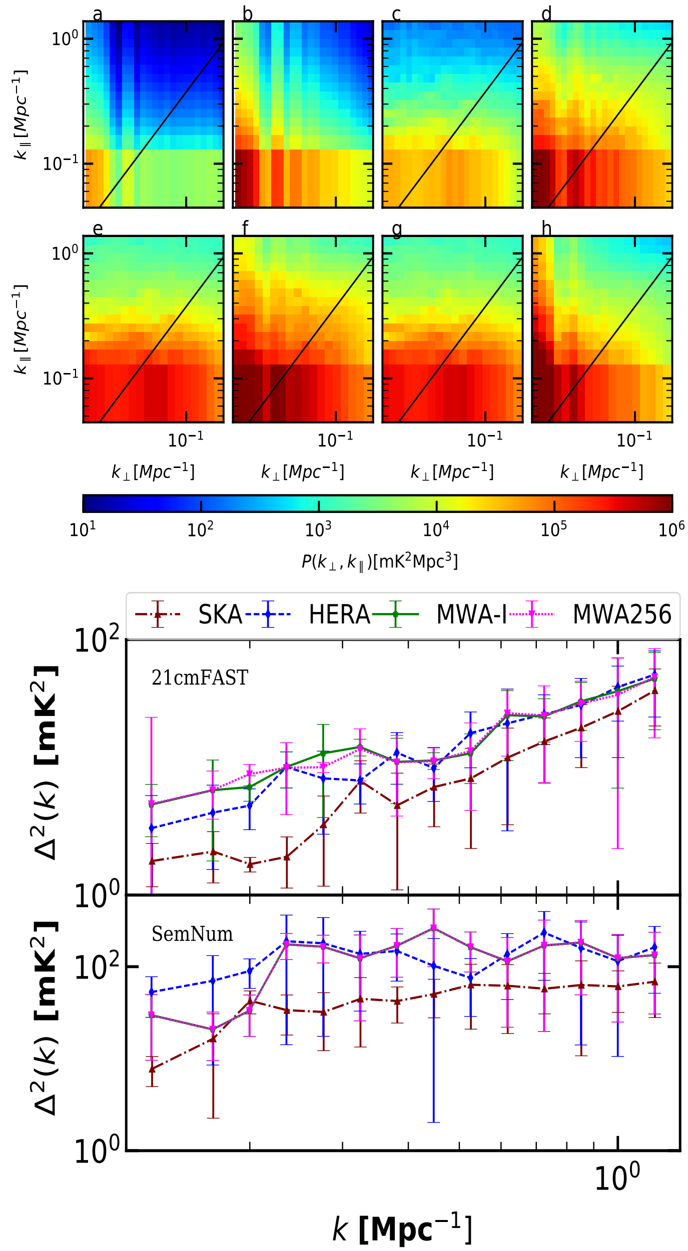

Figure 6.4: **(top)** Power spectrum obtained for the two signal models used when observed with the four telescope configurations. The top panel shows the cylindrical PS with SKA1-Low (panel a-21cmFAST, b-Sem-Num), HERA (panel c-21cmFAST, d-Sem-Num), MWA-I (panel e-21cmFAST, f-Sem-Num), MWA-256 (panel g-21cmFAST, h-Sem-Num). **(bottom)** The spherical PS with 21cmFAST **(upper subplot)** and SemNum **(lower subplot)**. For both subplots, the layouts are SKA (maroon dot-dash line with triangle markers), HERA (blue dashed line with diamond markers), MWA-I (green solid line with pentagon markers) and MWA-256 (magenta dotted lines with inverted triangle markers).

6.3 Non-ideal observing conditions

Observational effects on the cosmological 21-cm signal are numerous. However, this work focuses on calibration errors and position errors.

Radio interferometers observe the Fourier transform of the sky intensity distribution. The observed visibilities (i.e., spatial coherence function of sources) are calculated by cross-correlating the detected electric field between antenna pairs using digital correlators. The individual antenna (or tile) measures the incident electric field modified by a complex gain factor (arising due to the electronics). Calibration aims to solve for these complex gains. For observations that target the cosmological 21-cm signal, either a "traditional" sky-based technique or a "redundant" calibration technique is used. In the former method, a sky model is used to calibrate instruments [93]. The latter is useful for arrays that have a large number of redundant baselines [386]. However, irrespective of the technique used, the basic equation is the measurement equation relating the measured sky visibility to true sky visibility. The simplest form of the equation is :

$$V_{ij}^{m} = g_i(t)g_j^{*}(t)V_{ij}^{t} \tag{6.1}$$

where V_{ij}^{m} is the measured visibility from the antenna pairs i and j, V_{ij}^{t} is the true sky visibility and $g_i(t)$ and $g_j^{*}(t)$ are the complex gains of the respective antennae. While gain calibrations aim to solve for these gains, there is a limit to the accuracy of these. The uncalibrated gains (or residual gains) that remain are propagated in the subsequent steps and leads to errors in extracting the target signal. The complex gain can be modeled by

$$g_i = (a_i + \delta a_i)exp(-i(\phi_i + \delta\phi_i)) \tag{6.2}$$

where the phase, ϕ_i is in radian, and the amplitude, a_i is dimensionless. The terms δa_i and $\delta\phi_i$ are the errors in the gain solutions. For the ideal case, a_i is 1 and ϕ_i is 0, giving gain of 1. However, due to the presence of residual gain errors, the error terms will give the resultant gain as

$$g_i = (1 + \delta a_i)exp(-\delta\phi_i) \tag{6.3}$$

An "efficient" calibration algorithm will be one that can minimize δa_i and $\delta\phi_i$. The target signal itself is feeble, thus attaining an SNR >1 becomes a challenge. The thermal noise (one of the most fundamental limits for any instrument) can be lowered significantly with increasing observing time. However, other noise sources like faint unmodelled sources [351], ionospheric variation, instrument response, etc., cannot be reduced by increasing integration time. Thus accurate calibration can increase SNR, but up to a certain limit as allowed by observing conditions. This work quantifies the actual level of calibration accuracy required for upcoming radio interferometers to obtain "good" SNR levels.

The bright radio sources that contaminate the data sets are expected to be removed before extracting the cosmological 21-cm signal. While the predominant source of the bright foreground is the diffuse galactic emission, the point source contamination also poses a problem. The problem arises mainly because modeling and removal of each source is challenging. Any calibration imperfections give rise to artifacts in the residual that may affect the signal extraction adversely [110, 117]. Moreover, these observations have limited resolution due to the signal being coherent at short baselines only. Thus, high-resolution observations are required for accurate modeling and removal of discrete foregrounds. Hence, there are very stringent accuracy requirements for the sky model. Further calibration cannot minimize any error inherent to such a model, and the same will result in contaminated residuals. Thus position errors pose a challenge for the detection of the cosmological signal. Moreover, due to phase errors from directional effects, source positions may also shift. This will result in incorrect point source removal resulting in artifacts affecting signal extraction. Most sources have fluxes $\sim$ mJy or lower at the target frequencies for real observations. Since the cosmological signal is feeble, compact sources with even such "low" fluxes remaining in the residuals pose a problem. It has been shown that while redundant calibration is potentially better for arrays with a high degree of redundancy like HERA and the MWA Hexagonal configuration, offsets in source positions systematically affect the phase solutions [346]. A global sky model (having only compact sources) as observed in actual observation is considered for this work. The effect of errors in the source position through a systematic error in the right ascension of the sources is investigated here.

6.4 Formalism for Analyses

The visibilities obtained from the synthetic observations described in Section 6.2 are used for further analysis. Figure 6.5 shows the simulated image of the sky with T-RECS as foreground model for SKA1-Low, HERA, MWA-I & II respectively (top to bottom respectively). Figure 6.6 shows the point spread function (PSF) at the central frequency of observation for each of the telescopes used. It can be seen from the figure that the best PSF response comes from SKA1-Low (magenta curve), both in terms of width and amplitude of side lobes. PSF for HERA (shown by green curve) has the widest response, which is, because it has the shortest maximum baseline. However, due to the layout and configuration of the array, it has a small side-lobe amplitude. The MWA configurations (Phase-1 in cyan and Phase-2 in blue) show high amplitude oscillatory features in the sidelobes.

For determination of the effect of calibration errors, the ideal visibilities are corrupted by complex gains (as described in Equation 6.1) . The errors in gain phases and amplitudes (see Equation 6.3) were drawn from a 0 mean Gaussian distribution with a standard deviation equal to the percentage of error introduced. For this work, 0.001%, 0.01%, 0.1%, 1.0% & 10.0% gain errors are introduced in both the gain phases and amplitudes. The model visibilities are subtracted from the

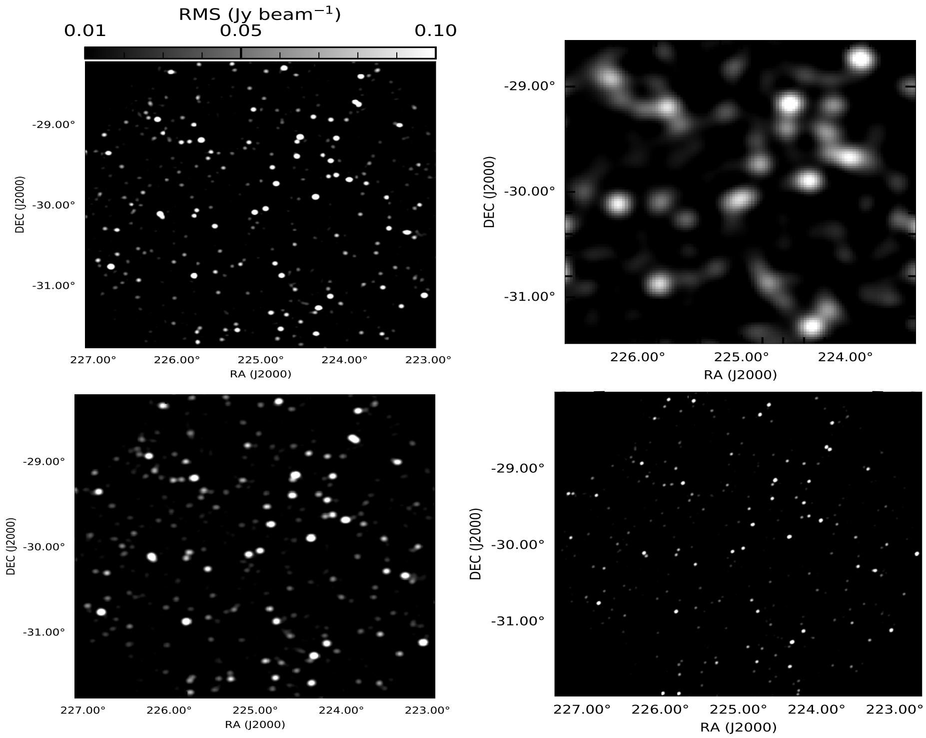

Figure 6.5: Simulated sky with T-RECS as foreground model for (top to bottom) SKA1-Low, HERA, MWA-I & MWA-256.

corrupted ones to produce residual visibilities, i.e.,

$$V_{ij}^{residual} = V_{ij}^{observed} - V_{ij}^{model} \tag{6.4}$$

Here, $V_{ij}^{observed}$ is the corrupted visibility (the label 'observed' implies actual observations the instrument records where the visibilities come with various imperfections).

The effect of position errors (i.e., in the global sky model) is determined by simulating observations with erroneous sky models. These inaccurate sky models are generated, taking the PSF size of HERA at the observing frequency (enlisted in Table 6.1) as a metric to quantify the offsets. A zero mean Gaussian distribution with standard different deviations- 0.01%, 0.1%, 1.0%, 10.0% & 100.0% of 8.1' (486") [8] are used to generate different sets of erroneous sky models. The position errors thus

[8]This means that the maximum displacement allowed are $\sim$ 0.05", 0.5", 5.0", 50" and 500" respectively for the percentages of error mentioned.

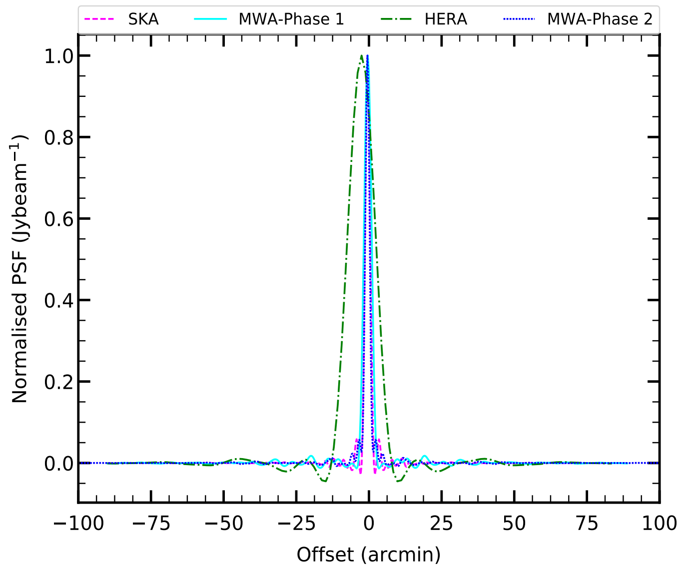

Figure 6.6: Normalized Point Spread Function using robust parameter 0 at the central frequency of observation for SKA1-Low(magenta), HERA(green), MWA-1(cyan) and MWA-2(blue).

generated were added to the RA of the sources to create a new catalog with inherent errors in the source positions. The residuals are obtained similarly, as discussed for Equation 6.4. Residuals obtained by subtracting the corrupted sky from the true one have been used to determine the effect in both the image plane and the PS. The formalism used to quantify the image plane effects and PS is described briefly in the following subsections. It should be mentioned that the position error in only the RA has been introduced for simplicity. Such errors can be present in either one of RA or DEC or both in a point source model. Such errors will obscure the EoR signal beyond a certain offset.

6.4.1 Image Plane

While most EoR detection experiments do not target imaging due to low signal-to-noise ratio (SNR), the SKA1-Low is projected to produce tomographic images with

SNR$\gtrapprox$1 [47]. There are various techniques, including machine learning approaches and the traditional *CLEAN* [388] algorithm that is being explored for maximum information extraction in the image domain (for instance, see [190] for a comprehensive summary). However, the efficacy and robustness of these methods are not determined for a wide variety of actual observations. Thus it is worthwhile to investigate how the errors that have been introduced here affect the image plane. The analysis is done by constructing a simple dirty map in CASA and using the RMS in the image plane as the quantifying metric (note that a comparison of various techniques for image recovery has not been made since it is beyond the scope of the current work). The pure HI signal used in the input has a peak flux of ~ 0.8 μJybeam^{-1} at 142 MHz, hence to get a significant detection, the residual RMS should lie below this level. Thus, performance in terms of the RMS has been compared. Additionally, image plane performance is also quantifiable using its dynamic range (DR) - the ratio of the peak flux in the image and the residual RMS. In the case of the presence of imperfect calibration, the theoretical DR is defined as [93] :

$$DR = \sqrt{\frac{N(N-1)}{2(a^2+\phi^2)}} \tag{6.5}$$

with a and ϕ as amplitude and phase, respectively, and N is the number of interferometer elements. The lowest RMS obtained from the residual image near the phase center is used to quantify the effect of the errors mentioned earlier. Since two different foreground models are used for this work, they are represented via the dashed (EN1) and dot-dashed (T-RECS) curves.

Figure 6.7 summarises the results obtained from the gain residuals. For the dynamic range, the general trend is consistent with theoretical expectations. The desirable calibration accuracy obtained from these simulations translates to a dynamic range requirement of $\sim 10^5$ or higher, which in turn implies the necessity of a highly accurate gain solution for observations with the SKA1-Low. Figure 6.8 shows the variation of RMS in the residuals with position error as a function of angular size.

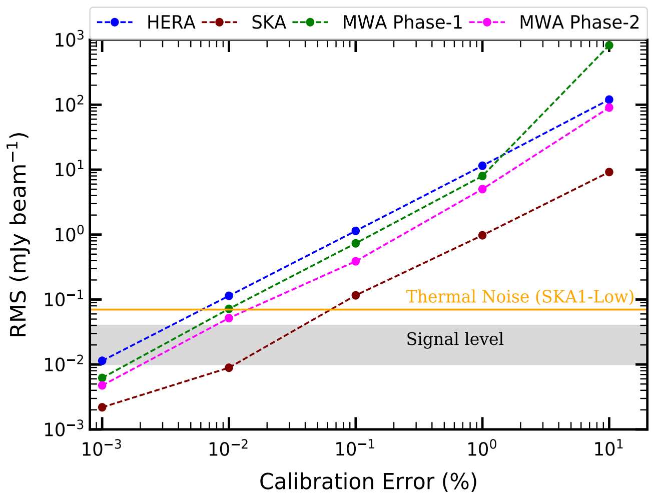

Figure 6.7: Variation of RMS in the residual image as a percentage of calibration error for T-RECS catalog as foreground model. The grey region shows the amplitude of range of observed signal. The solid orange line is the thermal noise limit for a 4-hour observation with SKA1-Low.

In both Figures 6.7 and 6.8, the gray band represents the maximum and minimum signal amplitudes in the image plane.

In the best case scenario, the sensitivity of the instrument will be determined by the thermal noise in the system. The thermal noise limit is given by [96]:

$$\sigma_T = \left(\frac{1.9}{\sqrt{\Delta\nu_{kHz} t_{hr}}}\right)\left(\frac{T_{sys}}{A_{eff} N_{ant}}\right) \tag{6.6}$$

where σ_T is the thermal noise in the system, $\Delta\nu_{kHz}$ is the channel width in kHz, t_{hr} is the integration time in hours, T_{sys} is the system temperature, A_{eff} is the antenna collecting area and N_{ant} is the number of antennas. At the frequencies of interest, T_{sys} is equivalent to the sky temperature. The sky temperature, T_{sky} is calculated using $T_{sky} \sim 180\left(\frac{\nu}{180MHz}\right)^{-2.6}$ K [40]. The values of thermal noise is tabulated in Table 6.1. As can be seen, the least thermal noise level for a 4 hour observation is obtained for SKA1-Low, which is shown via the solid orange line in Figures 6.7 & 6.8. In the case of SKA1-Low, [387] states that the spectral line sensitivity at

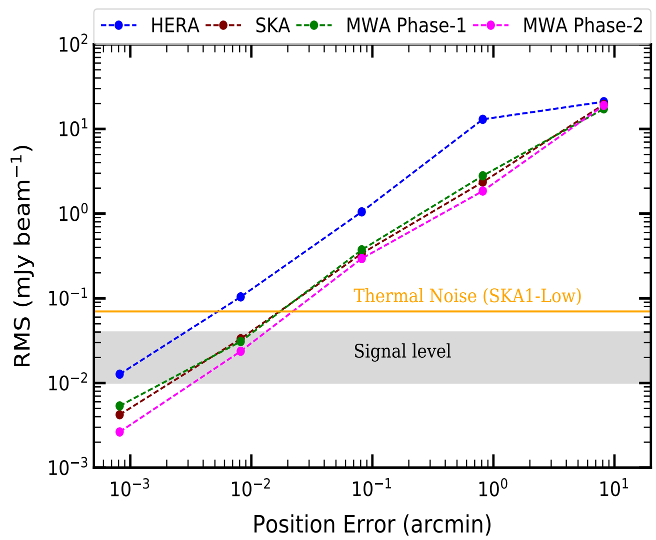

Figure 6.8: Variation of RMS in residual image with position error using T-RECS as foreground model for the four telescope configurations considered. The RMS is plotted as a function of the maximum displacement of the sources (in arcmin) from the true position. The solid orange line is the thermal noise limit for a 4-hour observation with SKA1-Low.

the frequency band of interest is 1258 μJy beam^{-1} for a fractional bandwidth of 10^{-4}, and integration time $\Delta\tau$ of 1 hour; for the specifications used in this work, the sensitivity comes down to $\sim$220 μJy beam^{-1}.

6.4.2 Power Spectrum

Statistical detection of the redshifted cosmological 21-cm signal is the primary target of most low-frequency interferometric arrays. For EoR experiments, there is a formalism to estimate the PS from both visibility (for instance, see [65]) as well as from the image plane (for example [71]); this work uses visibility for its determination. The interferometric visibility is the correlation between the signals received between two antenna pairs, which is given by [93] :

$$V(\mathbf{U},\nu) = \iint A(\hat{\mathbf{s}},\nu)B(\nu)I(\hat{\mathbf{s}},\nu)e^{-i2\pi\nu\mathbf{U}.\hat{\mathbf{s}}}d\Omega, \tag{6.7}$$

where, $\mathbf{U}$ is the baseline vector, $I(\hat{\mathbf{s}}, \nu)$ and $A(\hat{\mathbf{s}}, \nu)$ are the specific intensity and antenna beam pattern as a function of frequency (ν) and $B(\nu)$ is the instrumental bandpass response; the unit vector as, $\hat{\mathbf{s}} \equiv (l, m, n)$, where l, m, n are the direction cosines towards east, north and zenith respectively with $n = \sqrt{1 - l^2 - m^2}$ and $d\Omega = \frac{dldm}{\sqrt{1-l^2-m^2}}$. For this work, $A(\hat{\mathbf{s}}, \nu)$ is taken to be 1, i.e. the effect of primary beam is not considered.

Inverse Fourier transform of $V(\mathbf{U}, \nu)$ along the the frequency axis gives the visibility in the delay domain (henceforth represented as τ), $V(\mathbf{U}, \tau)$. Using this formalism, the cylindrical PS (as discussed in [366]) is given by:

$$P(\mathbf{k}_\perp, k_\parallel) = \left(\frac{\lambda^2}{2k_B}\right)^2 \left(\frac{X^2 Y}{\Omega B}\right) |V(\mathbf{U}, \tau)|^2, \tag{6.8}$$

with unit $K^2(Mpc/h)^3$. In the above equation, λ is the wavelength corresponding to central frequency (2.1 m for these simulations), k_B is the Boltzmann constant, Ω is the primary beam response, B is the bandwidth (8 MHz for this work), X and Y are the conversion factors from angle and frequency to transverse co-moving distance (D(z)) and the co-moving depth along the line of sight, respectively [366]. The quantity $\mathbf{k}_\perp$ represents the Fourier modes perpendicular & $k_\parallel$ represent the k modes along the line of sight, given by

$$\mathbf{k}_\perp = \frac{2\pi|\mathbf{U}|}{D(z)} \qquad \& \qquad k_\parallel = \frac{2\pi\tau\nu_{21} H_0 E(z)}{c(1+z)^2}$$

where ν_{21} is the rest-frame frequency of the 21 cm spin flip transition of HI z is the redshift corresponding to the observing frequency, H_0 is the Hubble parameter and $E(z) \equiv [\Omega_\mathrm{M}(1+z)^3 + \Omega_\Lambda]^{1/2}$. Ω_M and Ω_Λ are matter and dark energy densities, respectively [389].

From the 2D power spectrum, the 1D PS (obtained by spherical averaging of $P(\mathbf{k}_\perp, k_\parallel)$)is calculated as :

$$\Delta^2(k) = \frac{K^3}{2\pi^2} \langle P(\mathbf{k}) \rangle_k \tag{6.9}$$

where, $k = \sqrt{k_\perp^2 + k_\parallel^2}$. The uncertainty in PS has been calculated using a modified version of 21cmSense [390] for each array. The 3σ error bars have been put in the spherically averaged PS to show the uncertainty in power obtained.

In the case of SKA1-Low, it is still not decided whether signal PS estimation will be done using foreground removal, avoidance, or some other technique. This work uses the generic method of averaging over all k-modes to determine the 1D power. The same approach is also used for both MWA configurations and HERA to maintain consistency (although in [65] and [77], the upper limits have been placed using avoidance methodology).

The results obtained for each array configuration for imaging and PS estimation are described in the following section. It should be mentioned that both the foreground models give consistent results for all the telescope configurations considered. The power spectrum performance plots shown in the following subsections use T-RECS as the foreground model. A representative case using EN1 as foreground model has been shown in the appendix for all the configurations to emphasize that the pipeline performs consistently irrespective of the sky model used.

6.5 Results

This section describes the results obtained from the simulations and discusses their implications for actual observations. The results for each of the configuration considered is described in separate subsections below.

6.5.1 SKA1-Low

The synthetic observation pipeline is primarily being developed to simulate the effects of non-ideal conditions that are present in actual observations. The primary motivation is to determine how these errors will affect the performance of SKA1-Low for the sensitive low-frequency observation done for CD/EoR science. The following subsections describe the effect of calibration and position errors for SKA1-Low.

6.5.1.1 Calibration Error

The maroon curve shows the effect of calibration error on SKA1-Low in the image plane in Figure 6.7. It is seen that for calibration inaccuracy $\sim 10^{-2}\%$, the residual RMS is well below the signal level. 0.01% represents the borderline case where the RMS and thermal noise are nearly similar. However, the residual RMS level exceeds the signal level by almost an order of magnitude, indicating that the limiting factor for these observations would be systematics (instead of thermal noise) for sufficiently large integration times using the SKA1-Low. For even higher errors, the image RMS is high enough to confuse or even obscure the faint cosmological signal detection.

In Figure 6.9, the left panel shows the 2D power spectrum of the residual visibilities for SKA1-Low. It is seen that up to 0.01% calibration error, the wedge stays $k_\perp$ modes below $\sim$ 0.1 and below $k_\parallel \sim 0.2$, and the EoR window region have lower contamination. For cases exceeding 0.1% error, the EoR window has a very high power amplitude, implying the residual foregrounds dominate above inaccuracies $\sim$ 0.1% in gain amplitude.

The right panel of Figure 6.9 shows the 1-dimensional spherically averaged power spectrum for gain error residuals. The zoomed-in part shows the region near the signal power at k $\sim$0.2 and above. The green curve shows the power for the observed sky (i.e., the sky with both signal and foregrounds). It is seen that at 0.001% error (blue curve), the residual power overlaps with the signal power (red curve). At 0.01%, the power spectrum starts deviation from the signal level at k=0.5 Mpc^{-1} and above, but are still consistent within error-bars. But for 1% and above errors in calibration, the foreground contribution in the residuals overrides the signal power completely. Thus, if all other sources of error are well understood and mitigated, the residual calibration error can render the signal obscured if the algorithm accuracy

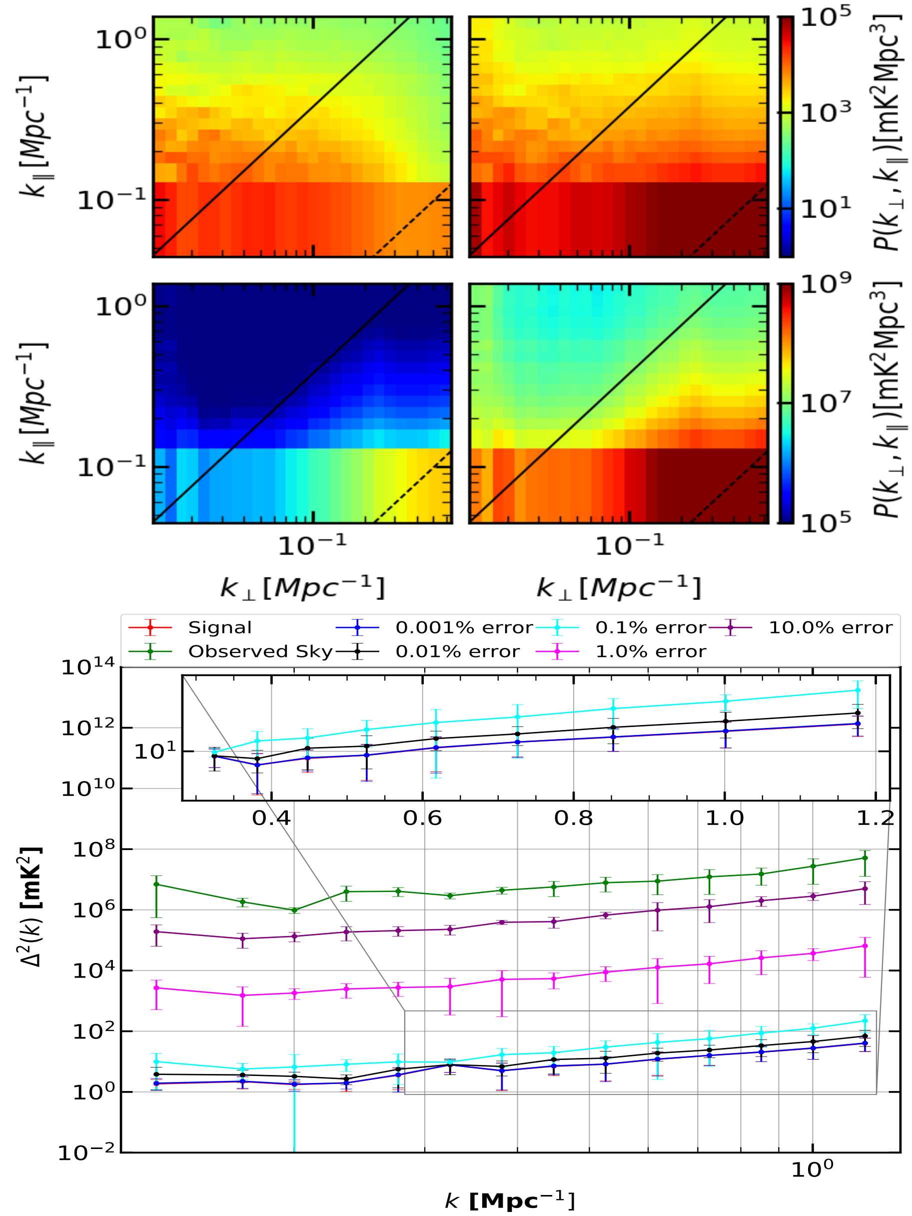

Figure 6.9: Residual power spectrum for gain errors with SKA-1 Low. **(top)** 2D PS of the residual visibilities for calibration errors for 0.001% (top-left), 0.01% (top-right), 0.1% (bottom-left) & 1.0 % (bottom-right). The solid black line represents the horizon line, while the dashed black line is the FoV limit. **(bottom)** Spherically averaged PS with calibration residuals compared with the signal power and foreground power. The error bars are 3σ uncertainties for the k-bins including sample variance and thermal noise.

is ∼1% or worse.

6.5.1.2 Position Error

The maroon curves show the effect of position error in the image plane in Figure 6.8. The top panel shows the RMS as a function of maximum displacement from the actual position, and the bottom curve represents the same as a function of the relative percentage of error. For errors up to $\sim$ 0.048" (0.1% of the PSF size), the residual RMS stays well below the signal level, thereby making its detection feasible. At 1% error (or 0.48" displacement), the RMS residual RMS starts getting dominated by the residual foregrounds, which again can potentially confuse detection of the redshifted signal. It is also seen that up to a maximum displacement of $\sim$ 0.048", the error remains below the thermal noise level for 4-hour observation (orange line), again emphasizing for a sufficiently large integration time, observations would be systematics limited rather than noise limited.

In Figure 6.10, the residual power for position errors with SKA1-Low is shown. The left panel shows the 2D PS, while the right panel 1D spherically averaged PS. It is seen for the residual 2D power spectrum that with offsets $\geq$ 10% of PSF size (i.e., 40" and above), the residual foregrounds cause leakage into the k-modes outside the wedge, giving a significantly high amplitude of power. Subsequently, in the case of the 1D PS, the residual power overrides the signal at all k modes by orders of magnitude. This is also reflected in the 1D power spectrum. For a position error of ∼0.1% (i.e., 4"), the signal and the residual foreground powers are at par with each other. Thus, it can be seen from both the figures that the position accuracy required is ∼0.01' to render the signal detectable. Hence in the case of PS estimation, it can be seen that for achievable positional accuracy at the observing frequency, the statistical detection of the cosmological signal is possible.

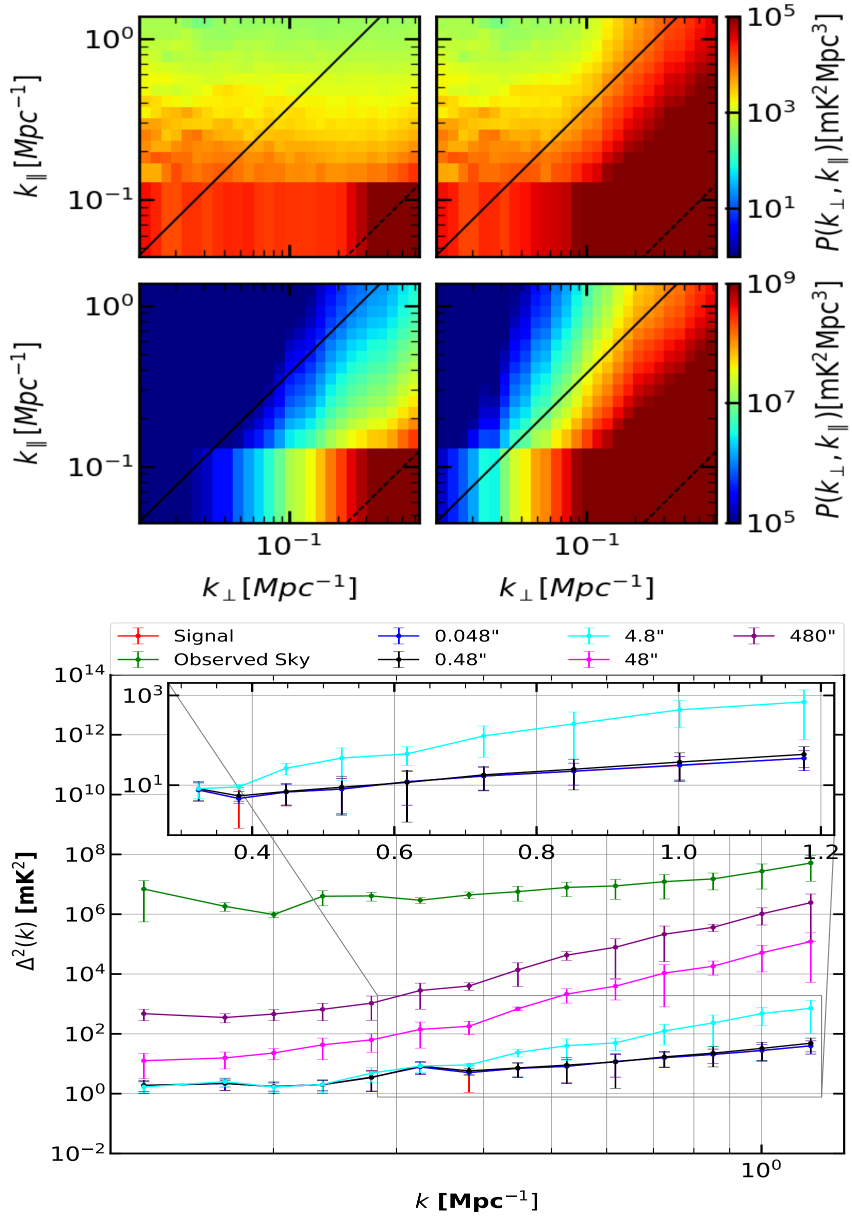

Figure 6.10: Residual power spectrum for position errors with SKA1-Low. **(top)** 2D cylindrical averaged power spectra of the residual visibilities for position errors with 0.1% (top-left), 1% (top-right), 10% (bottom-left) & 100 % (bottom-right). The solid black line represents the horizon line, while the dashed black line is the FoV limit. **(bottom)** Spherically averaged PS with position residuals compared with the signal power and foreground power.The error bars are 3σ uncertainties for the k-bins including sample variance and thermal noise.

6.5.2 HERA

HERA is one of the key EoR instruments that has already begun its observation and has recently provided its first upper limit using the data from the initial observing phase using 52 antennas [77]. For this work, the entire 350 antenna configuration (see Figure 6.2, top right panel) has been used. The effect of errors seen in both image and PS domain is described below.

6.5.2.1 Calibration Errors

The blue curves in Figure 6.7 show the effect of calibration residuals in the image plane for HERA. For both the sky models, it is seen that the residual RMS for gain solutions that have as small as 0.01% and above errors exceeds the signal power level. Thus, under identical observing conditions and in the presence of similar errors, SKA1-Low outperforms HERA in terms of calibration error tolerance by an order of magnitude in the image plane.

The PS for HERA for gain error residuals is shown in Figure 6.11. The left panel is the 2D PS, while the right panel is the 1D PS. In the $k_\parallel$- $k_\perp$ plane, the wedge can be observed in the modes within $k_\perp \sim 0.01$ and 0.3 and $k_\parallel \geq 0.3$ are free from very high amplitude of foregrounds.

The residual 2D PS shows that for calibration errors above ∼0.1%, the foregrounds power becomes high and starts to dominate at all k modes. This is also evident from the 1D power spectrum of the calibration residuals (Figure 6.11 right panel), where it is observed that upto 0.1%, the residual power very slightly deviates from the signal power. However, for inaccuracies ≥0.1%, the residual power exceeds the signal by a significant factor, overruling the signal and obscuring it.

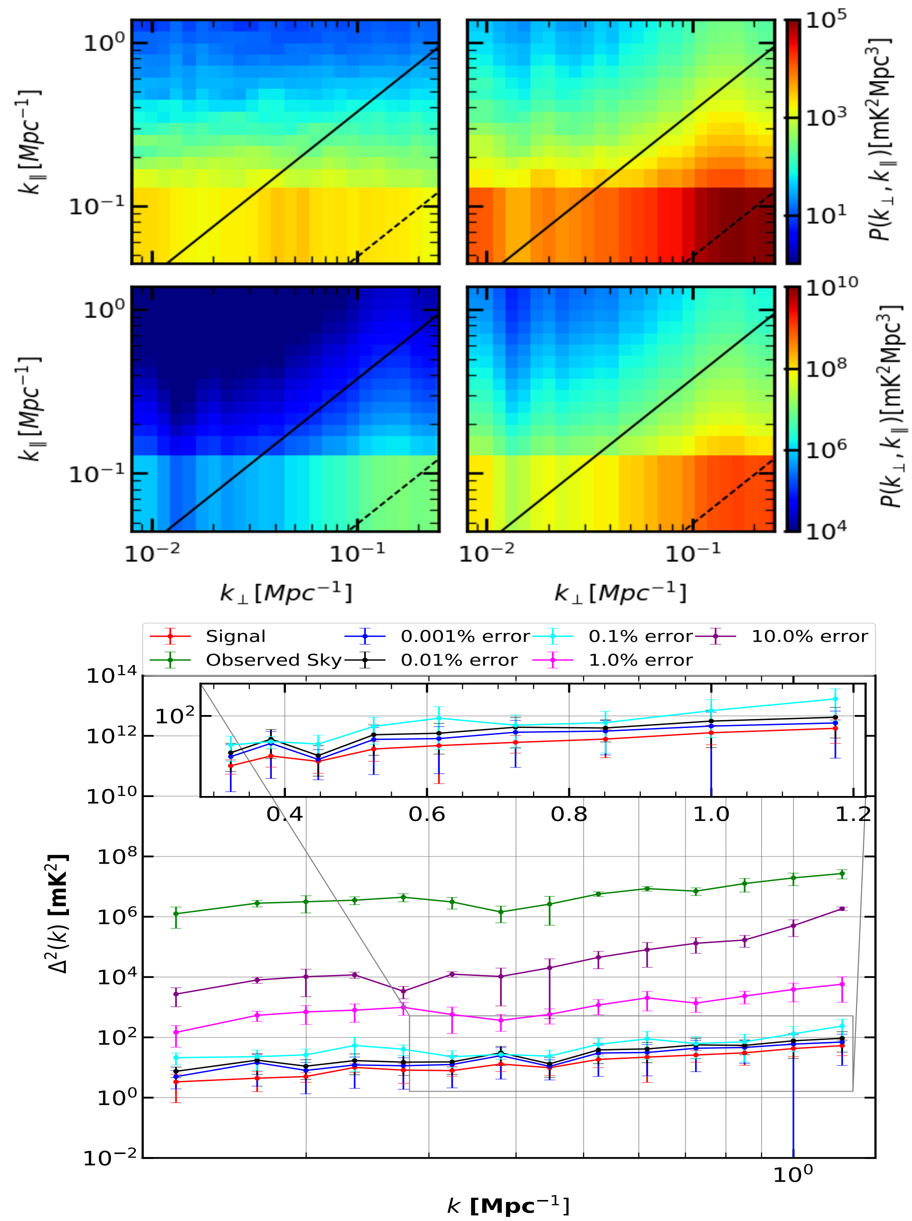

Figure 6.11: Residual power spectrum for gain errors with HERA. **(top)** 2D PS of the residual visibilities for calibration errors for 0.001% (top-left), 0.01% (top-right), 0.1% (bottom-left) & 1.0 % (bottom-right). The solid black line represents the horizon line, while the dashed black line is the FoV limit. **(bottom)** Spherically averaged PS with calibration residuals compared with the signal power and foreground power.The error bars are 3σ uncertainties for the k-bins including sample variance and thermal noise.

6.5.2.2 Position Errors

In Figure 6.8, the blue curves represent the residuals of calibration error for synthetic observations with HERA. In this case, an offset of 0.4" makes the residual RMS low enough to detect the signal. Hence, for HERA, offsets less than 0.4" are required for the image RMS to lie below the signal level.

The residual PS for position error residuals for HERA is shown in Figure 6.12. It is observed from the left panel that for deviations of >1% of the "actual" position, the power from the residual foregrounds contaminates most of the k-modes, causing the signal to be obscured. A similar trend is seen in the right panel, where the signal power spectrum (red curve) is being followed (with slight deviations of a factor of $\sim$1.3) for position errors up to $\sim$0.01" (0.1% of PSF size). At 1% of the PSF size ($\sim$0.1"), the residual power deviates from the signal power at k modes above 0.5 Mpc $^{-1}$. Above 1% error, the deviations are orders of magnitude higher, which results in the foreground residuals obscuring the signal.

6.5.3 MWA

The MWA is one of the precursor facilities to the SKA1-Low. In the initial phase of operation, it consisted of 128 square tiles extended up to $\sim$3 km. The second phase of the MWA, after its upgrade, consists of 256 tiles, which spreads out to $\sim$5 km. Both the old and the upgraded configurations with all the tiles have been used for this work. The results for the introduction of the errors as mentioned earlier are discussed in the following subsections.

6.5.3.1 Calibration Error

Green and magenta curves in Figure 6.7 show the performance of MWA-I and MWA-256 respectively in the image domain in the presence of residual calibration errors. For both the cases, RMS corresponding to residuals of 0.01% is below the signal level,

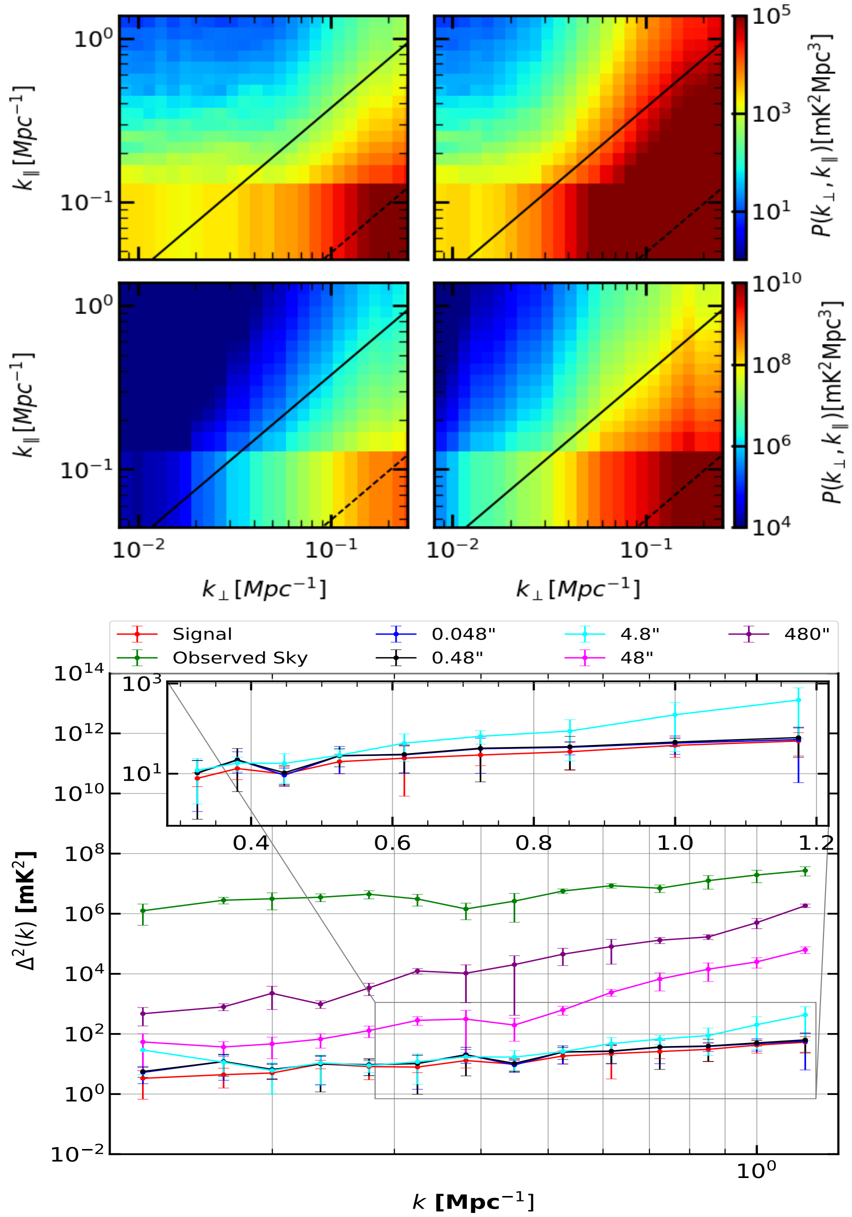

Figure 6.12: Residual power spectrum for position errors with HERA. **(top)** 2D cylindrical averaged power spectra of the residual visibilities for position errors with 0.1% (top-left), 1% (top-right), 10% (bottom-left) & 100 % (bottom-right). The solid black line represents the horizon line, while the dashed black line is the FoV limit. **(bottom)** Spherically averaged PS with position residuals compared with the signal power and foreground power.The error bars are 3σ uncertainties for the k-bins including sample variance and thermal noise.

with values $\sim 10^{-1}$ mJy beam^{-1}. Any algorithm producing inaccuracies greater than this can cause the residual RMS to override the signal by order of magnitude, thereby obscuring it. This performance is consistently seen for both configurations.

The PS recovery performance for both MWA configurations is similar, which may be attributed to the fact that baselines within a maximum of 2 km of the core are used to determine the power spectrum (baselines of longer lengths are not of interest for EoR statistics). Figures 6.13 & 6.14 show the PS recovered from the residual visibility of with calibration error using MWA-I & MWA-256. For both cases, it is observed from the left panel that for the 2D PS, gain errors beyond 0.1%, the residual foregrounds have power from the wedge leaking into the EoR window. Thus all the $k_{\perp}$ and $k_{\parallel}$ modes have higher power compared to the signal power, which implies the signal being obscured. The right hand panel for both Figures 6.13 & 6.14 show similar trends for the 1D power spectrum. For calibration inaccuracies $\leq 0.1\%$, the residual power follows the signal power spectrum within error bars. This is also evident from the zoomed-in parts of both figures, where it can be seen that the residuals follow the signal power up to $\sim$0.01%. Above 0.1% , the residual foregrounds start to dominate over the signal exceeding the signal in orders of magnitude, thus obscuring it.

6.5.3.2 Position Error

The green and magenta curves in Figure 6.8 show the residual RMS as a function of the angular displacement from the actual source position (top panel) and the relative percentage of error introduced (bottom) for MWA-I & MWA-256, respectively. As seen from the green curves, in the case of MWA-I, an offset of 1% (or $\sim$1") or more with respect to the actual position results in the residual RMS level exceeding the signal level, thereby obscuring it. Similarly, for the magenta curve, i.e., MWA-256, the residuals lie within the signal level up to an offset of $\sim$ 0.1", beyond which it

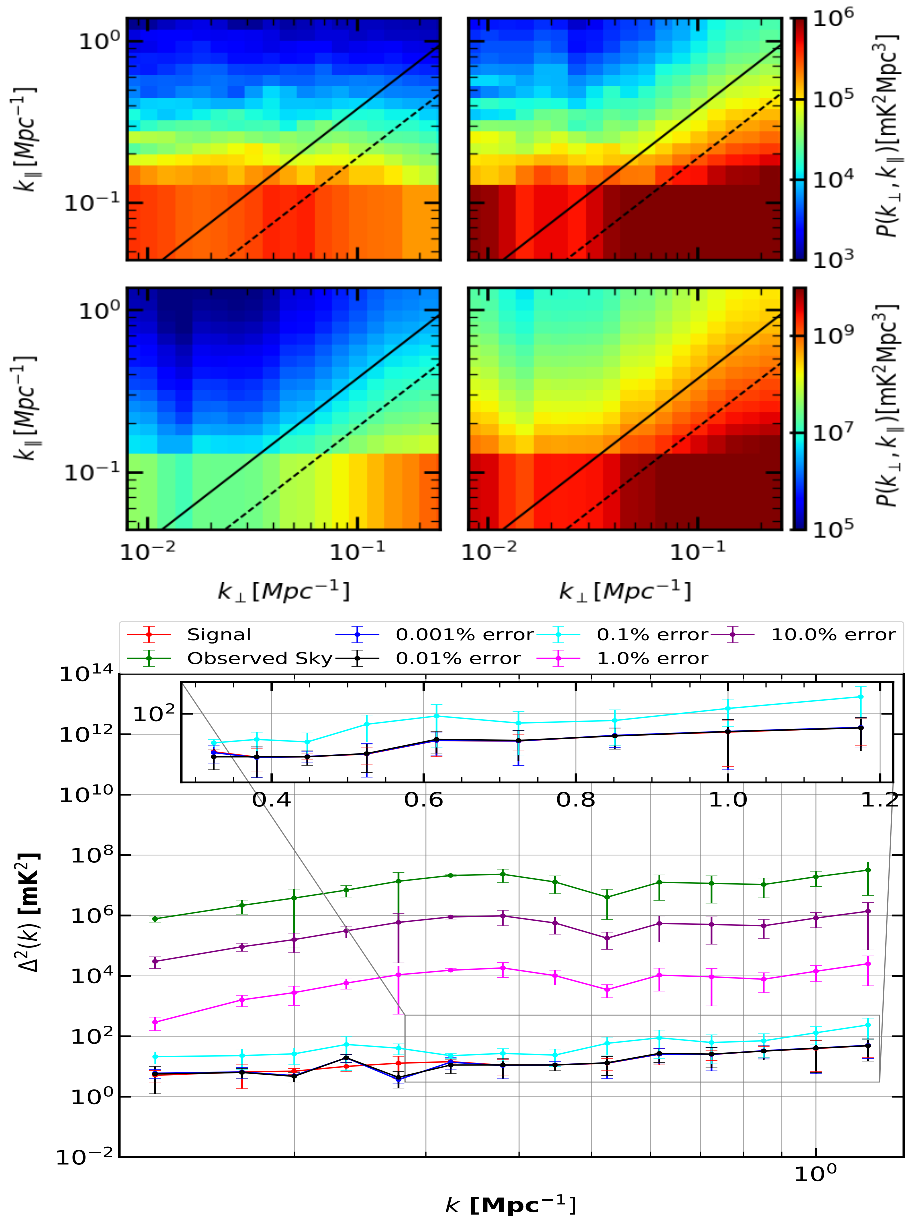

Figure 6.13: Residual power spectrum for gain errors with MWA Phase I. **(top)** 2D PS of the residual visibilities for calibration errors for 0.001% (top-left), 0.01% (top-right), 0.1% (bottom-left) & 1.0 % (bottom-right). The solid black line represents the horizon line, while the dashed black line is the FoV limit. **(bottom)** Spherically averaged PS with calibration residuals compared with the signal power and foreground power.The error bars are 3σ uncertainties for the k-bins including sample variance and thermal noise.

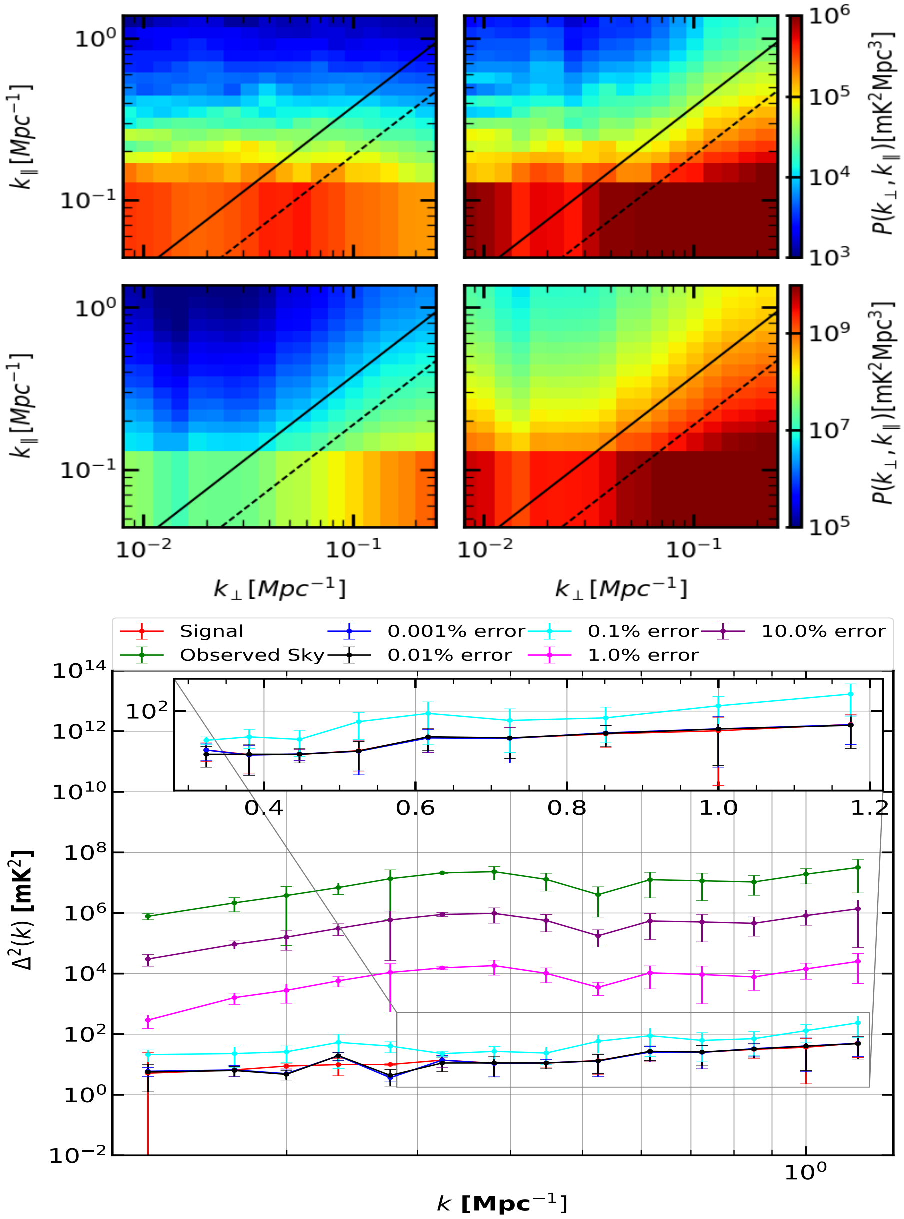

Figure 6.14: Residual power spectrum for gain errors with MWA-256. **(top)** 2D PS of the residual visibilities for calibration errors for 0.001% (top-left), 0.01% (top-right), 0.1% (bottom-left) & 1.0 % (bottom-right). The solid black line represents the horizon line, while the dashed black line is the FoV limit. **(bottom)** Spherically averaged PS with calibration residuals compared with the signal power and foreground power. The error bars are 3σ uncertainties for the k-bins including sample variance and thermal noise.

exceeds the signal level and the thermal noise level for SKA1-Low, thereby rendering the signal undetectable.

The residual PS for MWA configurations in the presence of position error are shown in Figures 6.15 & 6.16 - the left panel is the cylindrical PS, and the right panel is the spherical PS. As in the case of calibration errors, there is no significant difference between the two MWA configurations. From both the figures, it can be seen that in the entire $k_{\parallel}$-$k_{\perp}$ plane, the residual power has a higher magnitude than signal power at position error $\geq$1% (i.e., displacement $\gtrapprox$0.1'). From the 1D PS (right panel of both figures), it is can likewise be observed that at 1% error (cyan curve), the power spectrum of residual visibility starts deviating from the signal level (the residuals below this value follows the signal exactly). This implies that position accuracy better than 0.1" is required to detect the redshifted 21-cm signal from the EoR.

6.5.4 Discussion

After comparing the performances of the arrays considered here in terms of both imaging and statistics (power spectrum), the following has been noted:

- In the presence of residual calibration error, it is seen that for SKA1-Low, inaccuracies of 0.001%, there is sufficient DR in the image ($\sim 10^6$) to detect the signal in the image domain. A careful inspection also reveals that for SKA1-Low, up to an error of $\sim$0.07%, the residual RMS, though above the signal level, is below the daily 4-hour thermal noise level (orange line). The observations targeting EoR signal are expected to be systematics limited (rather than thermal noise limited). Since the systematics are expected to be unrelated from day to day, reducing them becomes more critical than thermal noise. From this consideration, these simulations suggest that the image plane performance of SKA-1 Low is marginally better than the other arrays with identical observing strategy. However, it is worth mentioning that these sim-

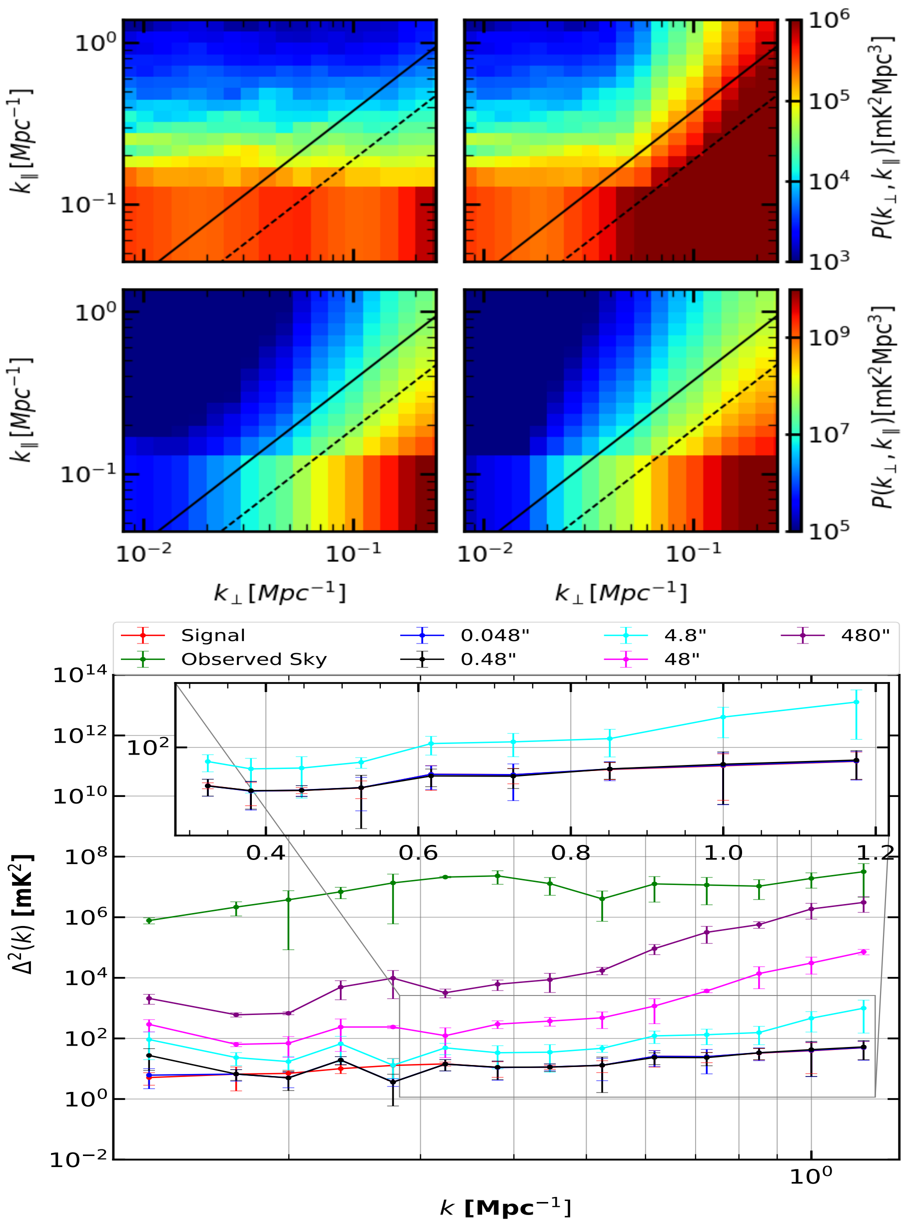

Figure 6.15: Residual power spectrum for position errors with MWA Phase I. **(top)** 2D PS of the residual visibilities for position errors for 0.01% (top-left), 0.1% (top-right), 1.0% (bottom-left) & 10.0 % (bottom-right). The solid black line represents the horizon line, while the dashed black line is the FoV limit. **(bottom)** Spherically averaged PS with position error residuals compared with the signal power and foreground power. The error bars are 3σ uncertainties for the k-bins including sample variance and thermal noise.

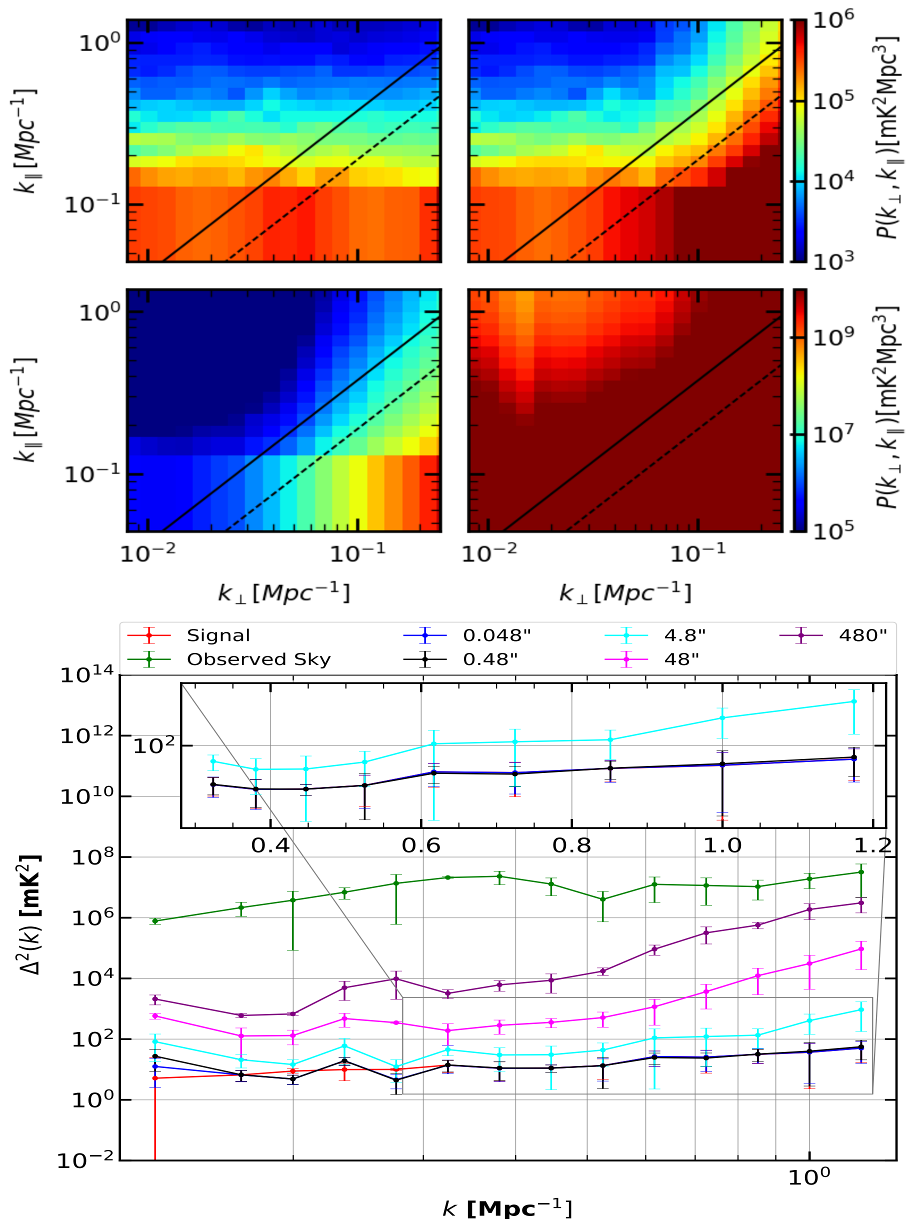

Figure 6.16: Residual power spectrum for position errors with MWA-256. **(top)** 2D PS of the residual visibilities for position errors for 0.01% (top-left), 0.1% (top-right), 1.0% (bottom-left) & 10.0 % (bottom-right). The solid black line represents the horizon line, while the dashed black line is the FoV limit. **(bottom)** Spherically averaged PS with position error residuals compared with the signal power and foreground power.The error bars are 3σ uncertainties for the k-bins including sample variance and thermal noise.

ulations do not consider thermal noise. The noise for MWA is theoretically expected to be much higher compared to the other configurations under identical observation conditions. Nevertheless, detailed image plane analysis in the presence of noise will be pursued in future works.

- In the presence of position error, the accuracy of the order of $\gtrapprox$0.5" is required. At displacement 5" or more, the residual RMS lies above both the signal and 4-hour thermal noise level (orange line). This is consistently observed for all four array configurations. This level of astrometric precision is in principle achievable for observations using the SKA-1. Recent results from LOFAR and MWA show an astrometric accuracy of ∼1" (for example, the LoTSS survey between 120–168 MHz [201] and the GLEAM survey between 70-231 MHz [134, 391]). Thus it is expected that the SKA-1, with improved sensitivity, should be able to produce sufficiently accurate sky models.

- For PS estimation, it is seen that residual calibration error causes the EoR window to get contaminated significantly for inaccuracies above 0.1%. This trend is almost consistent in all the arrays considered and reflected in the spherically averaged PS. It is observed for the 1D PS that the power spectrum follows the signal reasonably well for errors below 0.1%. After that, the residual amplitude in the $k_\perp$ - $k_\parallel$ plane starts exceeding the signal level. In the 1D PS, this manifests as deviation from the signal power at 0.1% and above errors.

- From the PS estimates, it is seen that the residual calibration error causes significant contamination of the EoR window with calibration errors of ∼0.1% and above. Below this value, the residual PS follow the "observed" signal power, and hence can be considered as the tolerance threshold for calibration errors.

- Residual PS in the presence of position error shows that for reasonably accurate astrometry, i.e., for a maximum position offset $\sim$0.5", the residual power follows the signal power. Position offset of 5" appears to be the threshold for tolerance, with some k modes lying within the signal uncertainty limits. Thus, it can be concluded that 5" is the optimal tolerance for position error. Here, it should be reemphasized that the sky model used for data calibration is expected to be derived from higher-resolution observations. If they are not sufficiently accurate, the errors propagating into subsequent analyses would potentially confuse signal detection.

- For the signal detection in PS domain, these simulations do not distinctly favour any of the four array configurations. However, it should be pointed out that these analyses are done under some very simplistic assumptions. In order to compare performances under even more realistic conditions, future works will explore various unavoidable systematics that can hamper observations targeting the redshifted 21-cm signal from CD/EoR.

6.6 Influence of Array Specific Parameters on Calibration and Position Error

This work considers the effects of residual calibration error and position error on the extraction of the 21-cm signal for four distinct array configurations. However, a detailed study of the influence of different array-specific parameters on the systematics considered here is beyond the scope of the present study. Nevertheless, this section briefly discusses a few parameters that vary across arrays depending on their location and specifications.

With diameters of 40m, 14m, and 4m for SKA-1 Low, HERA, and MWA, the FoV is $\sim 3^\circ$, 8° and 30° respectively at 142 MHz. A larger FoV implies observation of a greater number of sources. Thus, an observation incorporating a large FoV results in more robust foreground statistics like source counts, clustering, etc. Such robust statistics are required for precision foreground modelling and removal. However, a larger FoV can result in the inclusion of bright sources in the telescope primary beam. Presence of extremely bright objects like the A-Team sources [134] in the southern sky and Cassiopeia A and Cygnus A in the northern sky hinder data calibration. Thus they need to be removed very carefully from the data [65]. A larger FoV may also result in inclusion nearby bright sources in the side lobes of the telescope beam. Their presence increase the already difficult problem of modelling and removal of foregrounds from EoR data. Additionally, any residual bright artifact generated by such sources can also obscure the 21-cm signal. Thus, the different FoV of the different arrays considered here provide both advantages and disadvantages.

Different FoV are also affected differently by ionospheric errors. The earth's ionosphere introduces phase corruptions to extragalactic signals at low frequencies ($\lesssim$1 GHz). These effects are "directional", meaning they vary across different directions, causing an apparent shift of sources from their original positions by introducing additional phases. [392] provided a description of the different regimes for performing ionospheric phase calibration, based on the observing FoV along with baseline lengths and irregularity size[9]. Variable ionospheric conditions can affect the arrays considered here differently [163, 392, 393]. The compact array layout combined with the large beam size makes MWA fall on a calibration regime where there it is susceptible to having additional phases in the residuals due to ionosphere. This increases the probability of higher residual errors, even after self-calibration. The ionosphere-induced calibration inaccuracy in MWA EoR data was reported in [361] and [65] used ionospheric metric to find usable data for providing upper limits on the EoR PS. SKA-1 Low will be situated near the present MWA site and will face similar ionospheric conditions. However, each SKA station is projected to have a diameter of ∼40 m, making its FoV smaller than the MWA at a similar frequency. Thus, it

[9] [392] designated the different regimes of calibration as regimes 1, 2, 3 and 4. The interested readers are directed to Figure 1 of the same.

will fall in a different calibration regime of [392], where the smaller primary beam (i.e. FoV) is expected to introduce lesser ionospheric phase errors. For HERA, both the ionospheric conditions and FoV are different from SKA-1 Low and MWA, owing to its completely different location and dish diameter. However, a detailed study for HERA has not been done, and the analysis by [77] also does not consider any directional effects. Sensitive observations like those targeting the detection of the EoR need to understand the local ionosphere and calibrate its effects. But a detailed study into these effects is beyond the scope of the present work and is deferred till future works.

The baseline distribution as well minimum and maximum lengths of baselines also vary across arrays. The EoR signal is most sensitive at large angular scales and thus shorter baselines. The smallest $k_\perp$ mode accessible by an array is also determined by the minimum baseline length present. The arrays considered here have different baseline lengths and distributions. [351] showed that the chromaticity increases with baseline lengths, causing longer baselines to have chromatic gain errors. So applying gain solutions to the shorter baselines mixes the contamination from the longer baselines to the shorter ones, which contaminates the EoR window. Thus, they suggest up-weighting the shorter baselines compared to the longer ones. [351] also suggest that sky-based EoR experiments use short baselines to calibrate data used for PS estimation to prevent signal loss due to directional effects. Here, "short baselines" is a relative term depending on the array layout and the exact length (for example, $\sim$42 m for SKA and $\sim$7 m for MWA) should not matter. Longer baselines are essential for producing point source models for sky-based calibration. The recent MWA Long Baseline Epoch of Reionisation Survey (LoBES) [202] produced catalogues with better angular resolution than previous ones, which successfully removed foreground power from smaller angular scales. So it is evident that for successful detection of EoR PS, both long and short baselines are required for the best possible data processing.

The observing strategy also varies across arrays and is governed to a large extent by the actual physical properties of the dishes/tiles that make up the array. Most interferometers that are currently operational use tracking scans, i.e., tracks the phase center throughout the observation. Conversely, HERA is a drift scan array, meaning it does not track the sources but instead allows the sky to drift across its beam [101]. The feasibility of doing drift scan observations for EoR PS estimation with the MWA has also been explored [394, 395]. The biggest advantage of drift scans is that the beam is not steered to track the source, thus providing better instrumental stability. Stable instruments are extremely important for sensitive observations targeting EoR signal. [394] used MWA EoR experiment specifications to show that drift scans produce slightly lower uncertainties in the signal power and slightly higher SNR for the 1D PS compared to tracked observations. Simulations from [394] also show that drift scans perform better in terms of reducing cosmic variance, thus providing better sensitivity at low k modes. But higher k modes for these observations are thermal noise dominated causing greater uncertainties due to the reduced coherence and subsequent increase in thermal noise.

A major challenge for drift scan observations is finding suitable calibrator sources to perform sky-based calibrations [396]. Thus, HERA has targeted to replace sky-based techniques with redundant calibration methods. MWA Phase 2 layout also has many redundant baselines for EoR observations. Still, both HERA and MWA use sky-based calibration at some point. The HERA team uses a final step of absolute calibration for breaking degeneracies of redundant calibration [77, 397]. They build their models using bright sources from the GLEAM catalogue that coincides with the HERA track [77, 396]. If such sources are not available, then calibration for drift scan observation constitutes a problem. It was also shown in [373, 396, 397] that the redundant calibration technique works only for very high redundancy and identical primary beam response. This was also seen by the HERA team while calibrating the observed data (see Section 3.2 of [77]). Since it is impossible to have "perfect" redundancy, non-redundancies in a real array also lead to calibration errors. Thus, calibrating a drift scan instrument can be challenging due to the absence of "good" calibrator sources and imperfect redundancy. This may increase the calibration error and leak foregrounds into the EoR window. The observation and calibration strategy for SKA-1 Low is yet to be decided. Thus, it is worthwhile to simulate the merits and disadvantages of each method for SKA. But it should be reiterated that such detailed simulations are beyond the scope of the present work and are deferred till future works.

6.7 Summary and Conclusion

Detection of the redshifted HI 21-cm signal from the CD to EoR transitions is the most challenging undertaking for next-generation radio interferometers. The cosmological signal is weak, and prone to contamination by astrophysical foregrounds and instrumental systematics. The actual observations would require thousands of hours of data, which may be contaminated by improper calibration or position inaccuracy. Hence,it is extremely important to understand and quantify the effects of these systematic errors in order to be able to detect the cosmological signal. Hence, development of an end-to-end pipeline to check the effects of such systematic through simulations is essential. This work presents an end-to-end pipeline dealing with synthetic radio interferometric observations of the radio sky at low radio frequencies. The sky model includes the redshifted 21-cm signal and astrophysical foregrounds. Through various simulations the effects of these systematic errors in the extraction of the redshifted 21cm power spectrum has been shown. A comparison of results obtained for different array configurations- SKA1-Low, MWA-I, MWA-256, and HERA has also been demonstrated. The effect of the errors in the image plane detection of the cosmological signal has also been studied.

The simulations performed show that the optimal error in the case of calibration errors is 0.1%, beyond which foreground domination in the residuals overrides the signal. This translates to a DR $\sim 10^5$ or higher, which would require unprecedented precision in the algorithm employed. For the case of PS recovery, the signal is overridden by residual foregrounds at errors >0.1%. Since the determination of the signal power spectrum is the primary aim for most interferometers, it can be said that the target precision for the calibration algorithm used should be of the order of 0.1% or better.

It is also found that position errors have a significant impact only if the displacement is $\geq$ 5". PS derived from residual visibilities show that the residual power follows the signal power amplitude if the displacements are below $\sim$5", valid for all the cases considered. Thus it is concluded that for the next generation interferometers, position errors alone should not be a major systematic limitation upto a reasonable level. However, it is also important to note that these errors are not removable by calibration (being inherent in the sky model used in calibration), unless they are treated carefully, they may become a significant contaminant.

The simulations are done using simplistic assumptions for foregrounds and instruments. For such a scenario, they do not point to a preferred configuration among the four considered. The optimal tolerance for both calibration and position error are at a similar level for all four. However, it is seen that in presence of calibration error, the image plane performance of SKA and the PS performance of MWA-256 is marginally better compared to others. This points to the requirement of more detailed simulations with effects like primary beam chromaticity to determine the preferred configuration. It should also be mentioned that for simplicity, the simulations are noise-free. Hence, the limitations set by thermal noise have not been considered in the analysis. In our future works, we will incorporate thermal noise which will result in a more realistic case, to explore their impact on signal recovery.

Appendix:

The power spectra with the 4 telescope configurations with EN1 as sky model for 0.1% calibration and position errors. The trends observed with actual observed sky model is consistent with that obtained with a simulated sky model (i.e. T-RECS). The synthetic visibilities from both sky models show that the in the limit that only calibration or only position error is present, the optimum error should ideally be below 0.1% for calibration errors and ∼1" for position.

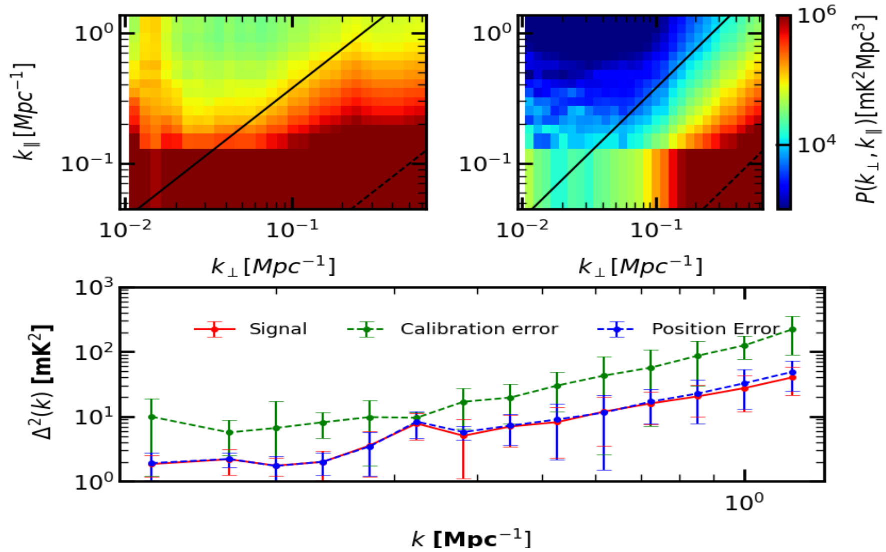

Figure 6.17: Residual power spectrum with SKA-1 Low using EN1 as foreground model. **top-left** 2D PS with 0.1% calibration error and **top-right** 2D PS with 0.1% position error. **bottom** Spherical power spectrum for 0.1% calibration error (green) and 0.1% position error (blue).

To demonstrate the capability of the pipeline to incorporate any signal model, Figure 6.19 shows the spherically averaged power spectrum for observation using the Sem-Num model in four cases - observation of pure HI only (red curve), the observed sky with signal and foreground (green curve), and residuals with 0.1% calibration errors (blue curve) and 0.1% position error (black curve). Comparison with the case for

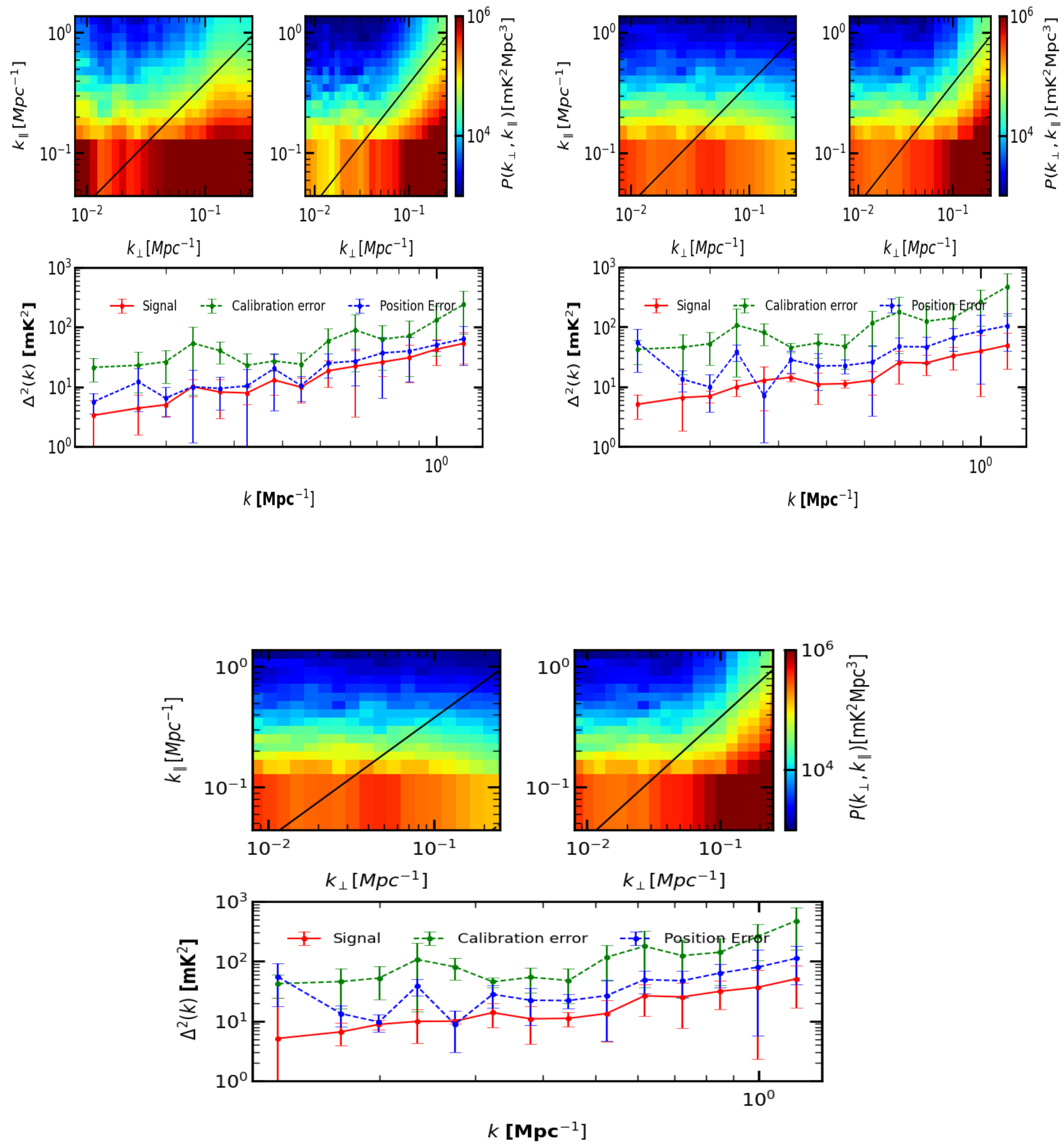

Figure 6.18: Same as Figure 6.17 using HERA (top left), MWA-I (top right), MWA-256(bottom)

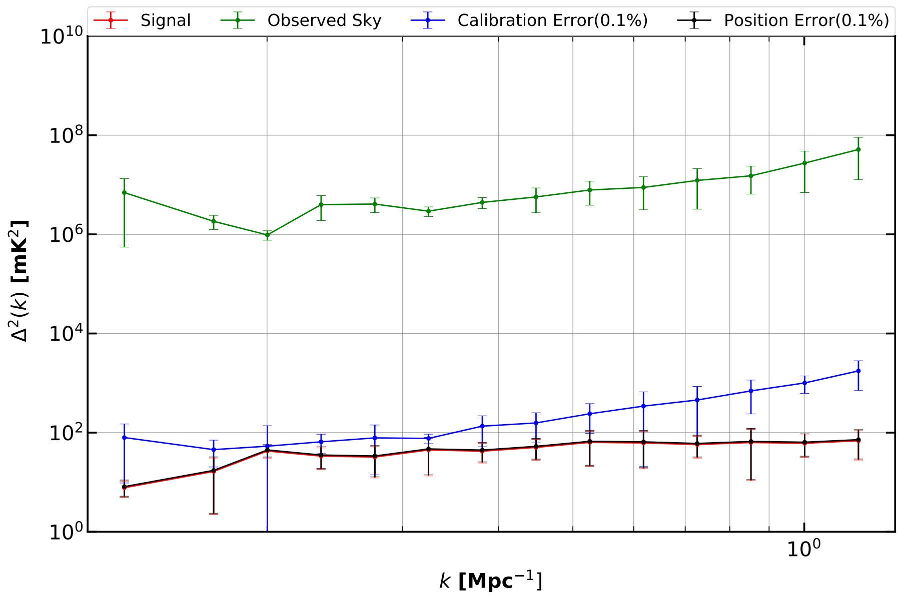

Figure 6.19: Spherically averaged power spectrum as observed with SKA1-Low for signal only (red curve), calibration error residual (blue curve), position error residual (black curve) and the sky i.e. signal+foreground (green curve).

SKA1-Low (Figures 6.9 and 6.10) shows that the results are consistent with those obtained for 21cmFAST signal model. This shows, once again the flexibility of the developed observational simulation pipeline to work with realistic scenarios.

Chapter 7

Summary and Future Scope

The redshifted HI 21-cm signal is one of the most reliable observational probes of the Epoch of Reionization but affected by orders of magnitude brighter astrophysical foregrounds and instrumental systematics. Both these sources of contamination need careful treatment for extracting the target from the observed data sets.

7.1 Summary

This thesis uses radio interferometers to characterize foreground emissions and quantify systematic error tolerance for EoR observations. Foregrounds and systematic errors have been studied using actual observations and state-of-the-art simulations.

Observed data from the GMRT is used to characterize foregrounds in the extragalactic deep field called the Lockman Hole observed at 325 MHz. Point sources up to sub-mJy fluxes have been catalogued, and the obtained source counts agree with previous observations. The analysis also suggests that the diffuse galactic synchrotron emission must be dealt with very carefully, even for locations far away from the galactic plane (where synchrotron is the brightest).

Clustering of extragalactic sources has also been studied for aiding the development of precise foreground models. Radio sources were divided into AGNs and SFGs populations based on radio luminosity. The bias parameter of different source populations shows a discrepancy for the dark matter halo mass for SFGs with respect to the SKADS simulation. This trend for SKADS is also seen when compared to other observations. It shows that existing point source models must incorporate more observational data to provide a realistic model for their evolution.

Systematic errors left in residual data after foreground removal are also potential contaminants of the EoR signal. These can vary depending on the observing conditions, but the underlying effect remains the same. Thus, simulations have been done to study the effects of two systematics - calibration and position errors. It is seen that EoR data sets require extremely accurate calibration algorithms (with error tolerance of $\sim 0.01\%$) and foreground estimates (position offset tolerance of $\lessapprox 5$") for extracting the target signal. This work has also led to developing an end-to-end pipeline for simulating interferometric observations under realistic conditions. This pipeline will help study different systematics and is extremely useful for predicting the performance of the SKA.

7.2 Future Scope

The interferometric detection of the HI 21-cm signal from EoR is a frontier of modern observational cosmology. Beyond the the statistics of the foregrounds and tolerance towards systematics done for this thesis, there are still prospects for further development. Some of the possible extensions of this thesis are as follows:

- The observations used here, though widefield, do not probe the very faint flux density sources. Foreground modeling requires spatial and spectral behavior of the foreground emissions to be characterized at the sub-mJy level (at a frequency of a few hundreds of MHz) -achievable with deeper observations over a wider bandwidths.

- Foreground studies with similar data sets as the one used for this thesis (and also using deeper observations) needs to be performed for other deep fields like Boötes, COSMOS, XMM-LSS to obtain insights into variation of foregrounds along different lines of sight.

- Observations using the frequency band used here will also be useful for characterizing the ionospheric contributions to the overall systematics for any sensitive radio observations. At these low frequencies, the first order effects of ionospheric perturbations cause shifts in the apparent source positions in the image. For example, for the 30KM baseline of GMRT, ionospheric distortions caused by the a 0.05 TECU variations in the ionospheric electron content shifts the apparent source position by $\sim$1" at 325 MHz (or as high as $\sim$6" at 150 MHz). Additionally, if there is significant higher order contributions, sources can even become distorted or disappear completely. For CD/EoR observations, propagation of even the first order errors can lead to the original signals being obscured.

- The clustering study done here has a fixed power law index of 1.8, (which is the average value estimated in most studies) due to the sample size. However, with larger number of samples, the power law index can be also be left as a free parameter, and comparison can be done to test the robustness of the slope assumed. Thus, such studies with deeper observations will provide more insights into clustering behaviour of extragalactic sources and if the power law slopes are consistent everywhere.

- The simulation pipeline developed here from scratch, can be extended to handle diffuse galactic emissions. Additionally, the simulations consider no primary beam effects and the field of is restricted to 4°, which is smaller than the primary beam of both HERA and MWA. Thus, a study on the effect of the different primary beam characteristics of different telescopes on the 21-cm signal recovery can also be done. Particularly, since the extent of the wedge is determined by sources at large angular separations from the phase center, effect of larger beam-sizes on the signal recovery can be also be studied.

- In its current form, the pipeline can also be used to study the effects of ionospheric perturbations, i.e. direction-dependent effects on sensitive low-frequency observations.

- The influence of thermal noise has not been considered in the study done for this thesis. Figure 7.1 shows the relative thermal noise levels for a 4 hour track using SKA-1 Low, HERA, MWA Phase-1, MWA Phase-2 (calculated using Equation 6.6). It can be seen that even for a constant system temperature, the interplay of the number of dishes/stations and the effective collecting area causes drastic variations in the thermal noise levels. Consequently, the effect on signal recovery across different configurations is important to consider.

The above are just a few possible extensions of the current thesis. This thesis, in conjunction with [5] has laid the ground work for foreground studies at the 150 MHz band. Data from the uGMRT Band-2 (150 MHz) is now being analysed to extend the study of foregrounds. Sensitive observations using imperfect instruments require careful mitigation of the effects influencing the actual science. The simulations done here can thus be extended to incorporate not only more realistic systematics, but signal and foreground models for more robust predictions. Hence, there is a lot of scope beyond the cases mentioned above for extending the work presented here, especially with the initial phase of the Square Kilometer Array being on track to start observing soon!

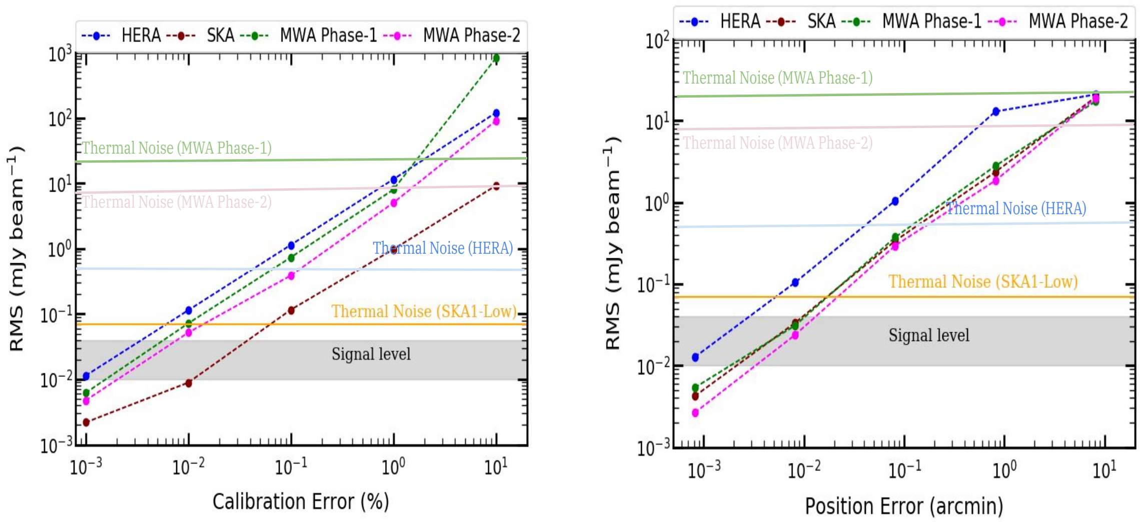

Figure 7.1: Thermal noise levels for a 4 hour track of the telescopes over-plotted with the residual image RMS for calibration (left) and position (right) errors. The expected 21-cm signal level and the used are plotted for reference.

References

[1] Mazumder, A., Chakraborty, A., Datta, A., Choudhuri, S., Roy, N., Wadadekar, Y., Ishwara-Chandra, C. H. (2020), Characterizing EoR foregrounds: a study of the Lockman Hole region at 325 MHz, MNRAS 495 (4), 4071 (DOI 10.1093/mnras/staa1317)

[2] Mazumder, A., Datta, A., Chakraborty, A., Majumdar, S. (2022), Observing the Reionization : Effect of Calibration and Position Errors on Realistic Observation Conditions, MNRAS (DOI: 10.1093/mnras/stac1994)

[3] Mazumder, A., Chakraborty, A., Datta, A. (2022), A study on the Clustering Properties of Radio-Selected sources in the Lockman Hole Region at 325 MHz, MNRAS (DOI: 10.1093/mnras/stac2801)

[4] Mazumder, A., Datta, A., Sathyanarayana Rao, M., Chakraborty, A., Singh, S., Tripathi, A., Choudhury, M. (2022), Synthetic Observations with the Square Kilometre Array (SKA) - development towards an end-to-end pipeline, *under review in JAA*

[5] Chakraborty, A. (2021), Searching for diffuse Hi 21 cm signal using upgraded Giant Metrewave Radio Telescope, Ph.D. Thesis (URL: http://dspace.iiti.ac.in:8080/jspui/handle/123456789/3115)

[6] Penzias, A. A., Wilson, R.W. (1965), A Measurement of Excess Antenna Temperature at 4080 Mc/s, ApJ 142, 419 (DOI: 10.1086/148307)

[7] Dicke, R. H., Peebles, P. J. E., Roll, P. G., Wilkinson, D. T. (1965), Cosmic Black-Body Radiation, ApJ 142, 414 (DOI: 10.1086/148306)

[8] Planck Collaboration (2020), Planck 2018 results - I. Overview, the cosmological legacy of Planck, A&A 641, A1 (DOI: 10.1051/0004-6361/201833880)

[9] Image obtained from https://www.eso.org/public/images/eso1620a/

[10] Dayal, P. and Ferrara, A. (2018), Early galaxy formation and its large-scale effects, Phys. Repts. 780, 1 (DOI: 10.1016/j.physrep.2018.10.002)

[11] Gunn, J. E. and Peterson, B. A. (1965), On the Density of Neutral Hydrogen in Intergalactic Space, ApJ 142, 1633 (DOI: 10.1086/148444)

[12] Storrie-Lombardi, L. J. and Wolfe, A. M. (2000), Surveys for $z > 3$ Damped Lyα Absorption Systems: The Evolution of Neutral Gas, ApJ 543(2), 552 (DOI: 10.1086/317138)

[13] Prochaska, J. X., Herbert-Fort, S., Wolfe, A. M. (2005), The SDSS Damped Lyα Survey: Data Release 3, ApJ 635(1), 123 (DOI: 10.1086/497287)

[14] Becker,R.H., Fan, X., White, R.L., Strauss, M.A., Narayanan, V.K., Lupton, R.H., Gunn, J.E., Annis, J., Bahcall, N.A., Brinkmann, J., Connolly, A. J., Csabai, I., Czarapata, P.C., Doi, M., Heckman, T.M., Hennessy, G. S., Ivezić, Zeljko, Knapp, G. R., Lamb, D.Q., McKay, T.A., Munn, J.A., Nash, T., Nichol, R., Pier, J.R., Richards, G.T., Schneider, D.P., Stoughton, C., Szalay, A.S., Thakar, A.R., York, D. G. (2006), Evidence for Reionization at z~6: Detection of a Gunn-Peterson Trough in a z=6.28 Quasar, AJ 122, 2850 (DOI: 10.1086/324231)

[15] Fan, X., Strauss, M. A., Becker, R. H., White, R. L., Gunn, J. E., Knapp, G. R., Richards, G. T., Schneider, D. P., Brinkmann, J., Fukugita, M. (2006),

Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with $z \sim 6$ Quasars. II. A Sample of 19 Quasars, AJ 132(1), 117 (DOI: 10.1086/504836)

[16] Yang, J. , Wang, F., Fan, X., Hennawi, J.F., Davies, F. B., Yue, M., Banados, E., Wu, X., Venemans, B., Barth, A.J., Bian, F., Boutsia, K., Decarli, R., Farina, E.P., Green, R., Jiang, L., Li, J., Mazzucchelli, C., Walter, F. (2020), Pōniuā'ena: A Luminous z = 7.5 Quasar Hosting a 1.5 Billion Solar Mass Black Hole, ApJ 897, L14 (DOI:10.3847/2041-8213/ab9c26)

[17] Bharadwaj S., Nath B.B., Sethi S.K. (2001), Using HI to Probe Large Scale Structures at $z \sim 3$, JoAA 22(1), 21 (DOI: 10.1007/BF02933588)

[18] Bull P., Ferreira P.G., Patel P., Santos M.G. (2015), Late-time Cosmology with 21 cm Intensity Mapping Experiments, ApJ, 803(1), 21 (DOI: 10.1088/0004-637X/803/1/21)

[19] Wolfe, A. M., Turnshek, D. A., Smith, H. E., Cohen, R. D. (1986), Damped Lyman-Alpha Absorption by Disk Galaxies with Large Redshifts. I. The Lick Survey, ApJS 61, 249 (DOI: 10.1086/191114)

[20] Gardner, J. P., Katz, N., Weinberg, D. H., Hernquist, L. (1997), Testing Cosmological Models against the Abundance of Damped Lyman-Alpha Absorbers, ApJ 486(1), 42 (DOI: 10.1086/304526)(1997)

[21] Mo, H., van den Bosch, F. C., White, S. (2010), Galaxy Formation and Evolution, Cambridge University Press

[22] McQuinn, M. (2016), The Evolution of the Intergalactic Medium, ARA&A 54(1), 313 (DOI: 10.1146/annurev-astro-082214-122355)

[23] Rees, M. J (1968), Polarization and Spectrum of the Primeval Radiation in an Anisotropic Universe, ApJ 153, L1 (DOI: 10.1086/180208)

[24] Hu, W., White, M. (1997), A CMB polarization primer, NewA 2, 323 (DOI: 10.1016/S1384-1076(97)00022-5)

[25] Planck Collaboration (2016), Planck intermediate results - XLVII. Planck constraints on reionization history, A&A 596, A108 (DOI: 10.1051/0004-6361/201628897)

[26] Reichardt, C. L. (2016), Observing the Epoch of Reionization with the Cosmic Microwave Background, Understanding the Epoch of Cosmic Reionization: Challenges and Progress 423, 227 (DOI: 10.1007/978-3-319-21957-8_8)

[27] Sunyaev, R. A., Zeldovich, Y. B. (1972), The Observations of Relic Radiation as a Test of the Nature of X-Ray Radiation from the Clusters of Galaxies, Comments on Astrophysics and Space Physics 4, 173 (URL: https://ui.adsabs.harvard.edu/abs/1972CoASP...4..173S/abstract)

[28] Ostriker, J. P., Vishniac, E. T. (1986), Generation of Microwave Background Fluctuations from Nonlinear Perturbations at the ERA of Galaxy Formation, ApJ 306, L51 (DOI: 10.1086/184704)

[29] Carlstrom, J. E., Holder, G. P., Reese, E. D. (2002), Cosmology with the Sunyaev-Zel'dovich Effect, AnRevA&A 40, 643 (DOI: 10.1146/annurev.astro.40.060401.093803)

[30] McQuinn, M., Furlanetto, S. R., Hernquist, L., Zahn, O., Zaldarriaga, M. (2006), The kinetic Sunyaev Zel'dovich effect from reionization, NewAR 50, 84 (DOI: 10.1016/j.newar.2005.11.025)

[31] Planck Collaboration (2020), Planck 2018 results. VI. Cosmological parameters, A&A 641, A6 (DOI: 10.1051/0004-6361/201833910)

[32] Field, G. B. (1958), Excitation of the Hydrogen 21-CM Line, Proc. of IRE 46, 240 (DOI: 10.1109/JRPROC.1958.286741)

[33] Ewen, H.I. and Purcell, E.M. (1951), Observation of a Line in the Galactic Radio Spectrum: Radiation from Galactic Hydrogen at 1,420 Mc./sec, Nature 168, 356 (DOI: 10.1038/168356a0)

[34] Muller, C. and Oort, J. (1951), Observation of a Line in the Galactic Radio Spectrum: The Interstellar Hydrogen Line at 1,420 Mc./sec., and an Estimate of Galactic Rotation, Nature 168, 357–358 (DOI: 10.1038/168357a0)

[35] van de Hulst, H.C., Muller, C.A., Oort, J.H (1954), The spiral structure of the outer part of the Galactic System derived from the hydrogen emission at 21 cm wavelength, Bulletin of the Astronomical Institutes of the Netherlands 12, 117 (URL: https://ui.adsabs.harvard.edu/abs/1954BAN....12..117V)

[36] van de Hulst, H.C., Raimond, E., van Woerden, H. (1957), Rotation and density distribution of the Andromeda nebula derived from observations of the 21-cm line, Bulletin of the Astronomical Institutes of the Netherlands 14, 1 (URL: https://ui.adsabs.harvard.edu/abs/1957BAN....14....1V)

[37] Wild, J.P. (1952), The Radio-Frequency Line Spectrum of Atomic Hydrogen and its Applications in Astronomy, ApJ 115, 206 (DOI: 10.1086/145533)

[38] Madau, P., Meiksin, A., Rees, M.J. (1997), 21 Centimeter Tomography of the Intergalactic Medium at High Redshift, ApJ 475(2), 429 (DOI: 10.1086/303549)

[39] Rybicki, G.B. and Lightman, Alan P.(1986), Radiative Processes in Astrophysics, Wiley-VCH

[40] Furlanetto, S.R., Oh, S.P., Briggs, F.H. (2006), Cosmology at low frequencies: The 21 cm transition and the high-redshift Universe, Phys. Repts. 433(4-6), 181 (DOI: 10.1016/j.physrep.2006.08.002)

[41] Barkna, R. (2016), The rise of the first stars: Supersonic streaming, radiative feedback, and 21-cm cosmology, Phys. Repts. 645, 1 (DOI: 10.1016/j.physrep.2016.06.006)

[42] Wouthuysen, S. A. (1952), On the excitation mechanism of the 21-cm (radio-frequency) interstellar hydrogen emission line, AJ 57, 31 (DOI: 10.1086/106661)

[43] Zaroubi S. (2013), The Epoch of Reionization, Astrophysics and Space Science Library (396) Springer, Berlin, Heidelberg (DOI:10.1007/978-3-642-32362-1_2)

[44] Pritchard, J.R., Loeb A. (2012), 21 cm cosmology in the 21st century, Reports on the Progress in Physics 75(8), 086901 (DOI: 10.1088/0034-4885/75/8/086901)

[45] Loeb, A.,Furlanetto S.R. (2013), The First Galaxies in the Universe, Princeton Series in Astrophysics, Princeton University Press

[46] Koopmans,L., Pritchard, J., Mellema, G., Aguirre,J., Ahn,K., Barkana,R., van Bemmel, I., Bernardi, G., Bonaldi, A., Briggs, F., de Bruyn, A.G., Chang, T.C., Chapman,E., Chen, X., Courty, B., Dayal, P., Ferrara, A., Fialkov, A., Fiore, F., Ichiki, K., Illiev, I.T., Inoue, S., Jelic, V., Jones, M., Lazio, J., Maio, U., Majumdar, S., Mack, K.J., Mesinger, A., Morales, M.F., Parsons, A., Pen, U.L., Santos, M., Schneider, R., Semelin, B., de Souza, R.S., Subrahmanyan, R., Takeuchi, T., Vedantham,H., Wagg, J., Webster,R., Wyithe, S., Datta, K.K., Trott, C. (2015), The Cosmic Dawn and Epoch of Reionisation with SKA, Advancing Astrophysics with the Square Kilometre Array - PoS (214), 001 (DOI: 10.22323/1.215.0001)

[47] Mellema, G., Koopmans, L., Shukla, H., Datta, K.K., Mesinger, A., Majumdar, S. (2015), HI tomographic imaging of the Cosmic Dawn and Epoch of Reionization with SKA, Advancing Astrophysics with the Square Kilometre Array - PoS (214), 010 (DOI: 10.22323/1.215.0010)

[48] Zaroubi, S., de Bruyn, A. G., Harker, G., Thomas, R. M., Labropolous, P., Jelić, V., Koopmans, L. V. E., Brentjens, M. A., Bernardi, G., Ciardi, B., Daiboo, S., Kazemi, S., Martinez-Rubi, O., Mellema, G., Offringa, A. R., Pandey, V. N., Schaye, J., Veligatla, V., Vedantham, H., Yatawatta, S. (2012), Imaging neutral hydrogen on large scales during the Epoch of Reionization with LOFAR, MNRAS 425, 2964 (DOI: 10.1111/j.1365-2966.2012.21500.x)

[49] Majumdar, S., Kamran, M., Pritchard, J.R., Mondal, R., Mazumdar, A., Bharadwaj, S., Mellema, G. (2020), Redshifted 21-cm bispectrum - I. Impact of the redshift space distortions on the signal from the Epoch of Reionization, MNRAS 499, 5090 (DOI: 10.1093/mnras/staa3168)

[50] Harker, G. J. A., Zaroubi, S., Thomas, R.M., Jelić, V., Labropoulos, P., Mellema, G., Iliev, I.T., Bernardi, G., Brentjens, M.A., de Bruyn, A. G., Ciardi, B., Koopmans, L.V. E., Pandey, V. N., Pawlik, A.H., Schaye, J., Yatawatta, S. (2009), Detection and extraction of signals from the epoch of reionization using higher-order one-point statistics, MNRAS 393, 1449 (DOI: 10.1111/j.1365-2966.2008.14209.x)

[51] Watkinson, C. A., Pritchard, J. R. (2015), The impact of spin-temperature fluctuations on the 21-cm moments, MNRAS 454, 1416 (DOI: 10.1093/mnras/stv2010)

[52] Bowman, J.D., Rogers, A.E.E. Hewitt, J.N. (2008), Toward empirical constraints on the global redshifted 21 cm brightness temperature during the Epoch of Reionization, ApJ, 676(1), 1 (DOI: 10.1086/528675)

[53] Singh, S., Jishnu, N.T., Subrahmanyan, R., Udaya Shankar, N., Girish, B. S., Raghunathan, A., Somashekar, R., Srivani, K. S., Sathyanarayana Rao, M. (2022), On the detection of a cosmic dawn signal in the radio background, NatAst, 6, 607, (DOI: 10.1038/s41550-022-01610-5)

[54] Barkana, R. (2018), Possible interaction between baryons and dark-matter particles revealed by the first stars, Nature 555, 71 (DOI: 10.1038/nature25791)

[55] Fialkov, A., Barkana, R. (2019), Signature of excess radio background in the 21-cm global signal and power spectrum, MNRAS 486, 1763 (DOI: 10.1093/mnras/stz873)

[56] Hills, R., Kulkarni, G., Meerburg, P.D., Puchwein, E. (2018) Concerns about modelling of the EDGES data, Nature 564, E32 (DOI: 10.1038/s41586-018-0796-5)

[57] Greig, B. (2019), Inference from the 21 cm Signal, The Cosmic 21-cm Revolution 2514, (4-1 - 4-40), IOP Publishing (DOI: 10.1088/2514-3433/ab4a73ch4)

[58] Ghosh, A., Bharadwaj, S., Ali, S.S., Chengalur, J. N. (2011), GMRT observation towards detecting the post-reionization 21-cm signal, MNRAS 411(4), 2426–2438 (DOI: 10.1111/j.1365-2966.2010.17853.x)

[59] Paciga, G., Chang, T., Gupta, Y., Nityanada, R., Odegova, J., Pen, U., Peterson, J.B., Roy, J., Sigurdson, K. (2011), The GMRT Epoch of Reionization experiment: a new upper limit on the neutral hydrogen power spectrum at $z \approx 8.6$, MNRAS 413(2), 1174–1183 (DOI: 10.1111/j.1365-2966.2011.18208.x)

[60] Paciga, G., Albert, J.G., Bandura, K., Chang, T., Gupta, Y., Hirata, C., Odegova, J., Pen, U., Peterson, J.B., Roy, J., Shaw, J.R., Sigurdson, K., Voytek, T. (2013), A simulation-calibrated limit on the H I power spectrum from the

GMRT Epoch of Reionization experiment, MNRAS 433,639 (DOI: 10.1093/mnras/stt753)

[61] Chakraborty, A., Datta, A., Roy, N., Bharadwaj, S., Choudhury, T.R., Datta, K.K., Pal, S., Choudhury, M., Choudhuri, S., Dutta, P., Sarkar, D. (2021) First Multi-redshift Limits on Post-Epoch of Reionization 21 cm Signal from z = 1.96-3.58 Using uGMRT, ApJL 907, L7 (DOI: 10.3847/2041-8213/abd17a)

[62] Ewall-Wice, A. et al. (2016), First limits on the 21 cm power spectrum during the Epoch of X-ray heating, MNRAS 460, 4320 (DOI: 10.1093/mnras/stw1022)

[63] Beardsley, A. P., Hazelton, B. J., Sullivan, I. S., Carroll, P., Barry, N., Rahimi, M., Pindor, B., Trott, C. M., Line, J., Jacobs, Daniel C., Morales, M. F., Pober, J. C., Bernardi, G., Bowman, Judd D., Busch, M. P., Briggs, F., Cappallo, R. J., Corey, B. E., de Oliveira-Costa, A., Dillon, Joshua S. Emrich, D., Ewall-Wice, A., Feng, L., Gaensler, B. M., Goeke, R., Greenhill, L. J., Hewitt, J. N., Hurley-Walker, N., Johnston-Hollitt, M., Kaplan, D. L., Kasper, J. C., Kim, H. S., Kratzenberg, E., Lenc, E., Loeb, A., Lonsdale, C. J., Lynch, M. J., McKinley, B., McWhirter, S. R., Mitchell, D. A., Morgan, E., Neben, A. R., Thyagarajan, Nithyanandan, Oberoi, D., Offringa, A. R., Ord, S. M., Paul, S., Prabu, T., Procopio, P., Riding, J., Rogers, A. E. E., Roshi, A., Udaya Shankar, N., Sethi, Shiv K., Srivani, K. S., Subrahmanyan, R., Tegmark, M., Tingay, S. J., Waterson, M., Wayth, R. B., Webster, R. L., Whitney, A. R., Williams, A., Williams, C. L., Wu, C., Wyithe, J. S. B (2016), First Season MWA EOR Power Spectrum Results at Redshift 7, ApJ 833(1), 102 (DOI: 10.3847/1538-4357/833/1/102)

[64] Barry, N., Wilensky, M., Trott, C. M., Pindor, B., Beardsley, A. P., Hazelton, B. J., Sullivan, I. S., Morales, M. F., Pober, J. C., Line, J., Greig, B., Byrne, R., Lanman, A., Li, W., Jordan, C. H., Joseph, R. C., McKinley, B., Rahimi,

M., Yoshiura, S., Bowman, J. D., Gaensler, B. M., Hewitt, J. N., Jacobs, D. C., Mitchell, D. A., Udaya Shankar, N., Sethi, S. K., Subrahmanyan, R., Tingay, S. J., Webster, R. L., Wyithe, J. S. B. (2019), Improving the Epoch of Reionization Power Spectrum Results from Murchison Widefield Array Season 1 Observations, ApJ 884(1), 1 (DOI: 10.3847/1538-4357/ab40a8)

[65] Trott, C.M., Jordan, C. H., Midgley, S., Barry, N., Greig, B., Pindor, B., Cook, J. H., Sleap, G., Tingay, S. J., Ung, D., Hancock, P., Williams, A., Bowman, J., Byrne, R., Chokshi, A., Hazelton, B. J., Hasegawa, K., Jacobs, D., Joseph, R. C., Li, W., Line, J. L. B., Lynch, C., McKinley, B., Mitchell, D. A., Morales, M. F., Ouchi, M., Pober, J. C., Rahimi, M., Takahashi, K., Wayth, R. B., Webster, R. L., Wilensky, M., Wyithe, J. S. B., Yoshiura, S., Zhang, Z., Zheng, Q. (2020), Deep multiredshift limits on Epoch of Reionization 21 cm power spectra from four seasons of Murchison Widefield Array observations, MNRAS 493(4), 4711 (DOI: 10.1093/mnras/staa414)

[66] Yoshiura, S., Pindor, B., Line, J. L. B., Barry, N., Trott, C. M., Beardsley, A., Bowman, J., Byrne, R., Chokshi, A., Hazelton, B. J., Hasegawa, K., Howard, E., Greig, B., Jacobs, D., Jordan, C. H., Joseph, R., Kolopanis, M., Lynch, C., McKinley, B., Mitchell, D. A., Morales, M. F., Murray, S. G., Pober, J. C., Rahimi, M., Takahashi, K., Tingay, S. J., Wayth, R. B., Webster, R. L., Wilensky, M., Wyithe, J. S. B., Zhang, Z., Zheng, Q. (2021), A new MWA limit on the 21 cm power spectrum at redshifts 13, MNRAS 505(4), 4775 (DOI: 10.1093/mnras/stab1560)

[67] Rahimi, M., Pindor, B., Line, J. L. B., Barry, N., Trott, C. M., Webster, R. L., Jordan, C. H., Wilensky, M., Yoshiura, S., Beardsley, A., Bowman, J., Byrne, R., Chokshi, A., Hazelton, B. J., Hasegawa, K., Howard, E., Greig, B., Jacobs, D., Joseph, R., Kolopanis, M., Lynch, C., McKinley, B., Mitchell, D.

A., Murray, S., Morales, M. F., Pober, J. C., Takahashi, K., Tingay, S. J., Wayth, R. B., Wyithe, J. S. B., Zheng, Q. (2021), Epoch of reionization power spectrum limits from Murchison Widefield Array data targeted at EoR1 field, MNRAS 508(4), 5954 (DOI: 10.1093/mnras/stab2918)

[68] Kazemi, S., Yatawatta, S., Zaroubi, S., Lampropoulos, P., de Bruyn, A.G., Koopmans, L. V. E., Noordam, J. (2011), Radio interferometric calibration using the SAGE algorithm, MNRAS 414(2), 1656–1666 (DOI: 10.1111/j.1365-2966.2011.18506.x)

[69] Patil, A. H., Yatawatta, S., Koopmans, L. V. E., de Bruyn, A. G., Brentjens, M. A., Zaroubi, S., Asad, K. M. B., Hatef, M., Jelić, V., Mevius, M., Offringa, A. R., Pandey, V. N., Vedantham, H., Abdalla, F. B., Brouw, W. N., Chapman, E., Ciardi, B., Gehlot, B. K., Ghosh, A., Harker, G., Iliev, I. T., Kakiichi, K., Majumdar, S., Mellema, G., Silva, M. B., Schaye, J., Vrbanec, D., Wijnholds, S. J. (2017), Upper Limits on the 21 cm Epoch of Reionization Power Spectrum from One Night with LOFAR, ApJ 838(1), 65 (DOI: 10.3847/1538-4357/aa63e7)

[70] Gehlot, B. K., Mertens, F. G., Koopmans, L. V. E., Brentjens, M. A., Zaroubi, S., Ciardi, B., Ghosh, A., Hatef, M., Iliev, I. T., Jelić, V., Kooistra, R., Krause, F., Mellema, G., Mevius, M., Mitra, M., Offringa, A. R., Pandey, V. N., Sardarabadi, A. M., Schaye, J., Silva, M. B., Vedantham, H. K., Yatawatta, S. (019), The first power spectrum limit on the 21-cm signal of neutral hydrogen during the Cosmic Dawn at z = 20-25 from LOFAR, MNRAS 488(3), 4271 (DOI: 10.1093/mnras/stz1937)

[71] Mertens, F. G., Mevius, M., Koopmans, L. V. E., Offringa, A. R., Mellema, G., Zaroubi, S., Brentjens, M. A., Gan, H., Gehlot, B. K., Pandey, V. N., Sardarabadi, A. M., Vedantham, H. K., Yatawatta, S., Asad, K. M. B., Ciardi,

B., Chapman, E., Gazagnes, S., Ghara, R., Ghosh, A., Giri, S. K., Iliev, I. T., Jelić, V., Kooistra, R., Mondal, R., Schaye, J., Silva, M. B. (2020), Improved upper limits on the 21 cm signal power spectrum of neutral hydrogen at z $\approx$ 9.1 from LOFAR MNRAS 493(2), 1662 (DOI: 10.1093/mnras/staa327)

[72] Parsons, A., Pober, J., McQuinn, M., Jacobs, D., Aguirre, J. (2011), A Sensitivity and Array-configuration Study for Measuring the Power Spectrum of 21 cm Emission from Reionization, ApJ 753(1), 81 (DOI: 10.1088/0004-637X/753/1/81)

[73] Parsons, A.R., Pober, J.C., Aguirre, J.E., Carilli, C.L., Jacobs, D.C., Moore, D.F. (2012), A Per-baseline, Delay-spectrum Technique for Accessing the 21 cm Cosmic Reionization Signature, ApJ 756(2), 165 (DOI: 10.1088/0004-637X/756/2/165)

[74] Ali, Z.S., Parsons, A.R., Zheng, H., Pober, J.C., Liu, A., Aguirre, J.E., Bradley, R.F., Bernardi, G., Carilli, C.L., Cheng, C., DeBoer, D.R., Dexter, M.R., Grobbelaar, J., Horrell, J., Jacobs, D.C., Klima, P., MacMahon, D.H.E., Maree, M., Moore, D.F., Razavi, N., Stefan, I.I., Walbrugh, W.P., Walker, A. (2015), PAPER-64 Constraints on Reionization: The 21 cm Power Spectrum at z = 8.4, ApJ 809(1), 61 (DOI: 10.1088/0004-637X/809/1/61)

[75] Jacobs, D.C., Pober, J.C., Parsons, A.R. ; Aguirre, J.E., Ali, Z,S., Bowman, J., Bradley, R.F., Carilli, C.L., DeBoer, D.R., Dexter, M.R., Gugliucci, N.E., Klima, P., Liu, A., MacMahon, D.H.E., Manley, J.R., Moore, D.F., Stefan, I.I., Walbrugh, W.P. (2015), Multiredshift Limits on the 21 cm Power Spectrum from PAPER, ApJ 801(1), 51 (DOI: 10.1088/0004-637X/801/1/51)

[76] Cheng, C., Parsons, A.R., Kolopanis, M., Jacobs, D.C., Liu, A., Kohn, S.A., Aguirre, J.E., Pober, J.C., Ali, Z.S., Bernardi, G., Bradley, R.F., Carilli, C.L.,

DeBoer, D.R., Dexter, M.R., Dillon, J.S., Klima, P., MacMahon, D.H.E., Moore, D.F., Nunhokee, C.D., Walbrugh, W.P., Walker, A. (2018), Characterizing Signal Loss in the 21 cm Reionization Power Spectrum: A Revised Study of PAPER-64, ApJ 868(1), 26 (DOI: 10.3847/1538-4357/aae833)

[77] The HERA Collaboration (2021), First Results from HERA Phase I: Upper Limits on the Epoch of Reionization 21 cm Power Spectrum, arXiv-eprint (2108.02263)

[78] Mertens, F. G., Semelin, B., Koopmans, L. V. E. (2021), Exploring the Cosmic Dawn with NenuFAR, SF2A-2021: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, 211 (URL: http://sf2a.eu/proceedings/2021/2021sf2a.conf.211M.pdf)

[79] Greig, B., Mesinger, A. (2015), 21CMMC: an MCMC analysis tool enabling astrophysical parameter studies of the cosmic 21 cm signal, MNRAS 449, 4246 (DOI: 10.1093/mnras/stv571)

[80] Mirocha, J., Harker, G. J. A., Burns, J. O. (2015), Interpreting the Global 21-cm Signal from High Redshifts. II. Parameter Estimation for Models of Galaxy Formation, ApJ 813, 11 (DOI: 10.1088/0004-637X/813/1/11)

[81] Bernardi, G., Zwart, J. T. L., rice, D., Greenhill, L. J., Mesinger, A., Dowell, J., Eftekhari, T., Ellingson, S. W., Kocz, J., Schinzel, F. (2016), Bayesian constraints on the global 21-cm signal from the Cosmic Dawn, MNRAS 461, 2847 (DOI: 10.1093/mnras/stw1499)

[82] Mirocha, J., Muñoz, J. B., Furlanetto, S. R., Liu, A., Mesinger, A. (2022), A galaxy-free phenomenological model for the 21-cm power spectrum during reionization, MNRAS 514, 2010 (DOI: 10.1093/mnras/stac1479)

[83] Maity, B., Choudhury, T. R. (2022), Constraining the reionization and thermal history of the Universe using a seminumerical photon-conserving code SCRIPT, MNRAS 515, 617, (DOI: 10.1093/mnras/stac1847)

[84] Schmit, C. J., Pritchard, J. R. (2018), Emulation of reionization simulations for Bayesian inference of astrophysics parameters using neural networks, MNRAS 475, 1213 (DOI: 10.1093/mnras/stx3292)

[85] Chardin, J. , Uhlrich, G., Aubert, D., Deparis, N., Gillet, N., Ocvirk, P., Lewis, J., (2019), A deep learning model to emulate simulations of cosmic reionization, MNRAS 490, 1055 (DOI: 10.1093/mnras/stz2605)

[86] Cohen, A., Fialkov, A., Barkana, R., Monsalve, R. A. (2020), Emulating the global 21-cm signal from Cosmic Dawn and Reionization, MNRAS 495, 4845 (DOI: 10.1093/mnras/staa1530)

[87] Zhao, X., Mao, Y., Cheng, C., Wandelt, B.D. (2022), Simulation-based Inference of Reionization Parameters from 3D Tomographic 21 cm Light-cone Images, ApJ, 926, 151 (DOI: 10.3847/1538-4357/ac457d)

[88] Shimabukuro, H., Semelin, B. (2017), Analysing the 21 cm signal from the epoch of reionization with artificial neural networks, MNRAS 468, 3869 (DOI: 10.1093/mnras/stx734)

[89] Choudhury, M., Datta, A., Chakraborty, A. (2020), Extracting the 21 cm Global signal using artificial neural networks, MNRAS 491, 4031 (DOI: 10.1093/mnras/stz3107)

[90] Choudhury, M., Chatterjee, A., Datta, A., Choudhury, T. R. (2021), Using artificial neural networks to extract the 21-cm global signal from the EDGES data, MNRAS 502, 2815 (DOI: 10.1093/mnras/stab180)

[91] Choudhury, M., Datta, A., Majumdar, S. (2022), Extracting the 21-cm power spectrum and the reionization parameters from mock data sets using artificial neural networks, MNRAS 512, 5010 (DOI: 10.1093/mnras/stac736)

[92] Thompson, R.A., Moran, J.M., Swenson, G.W. (2011), Interferometry and Synthesis in Radio Astronomy, Astronomy and Astrophysics Library Series, Springer, Cham

[93] Taylor, G. B., Carilli, C. L., Perley, R. A. (1999), Synthesis Imaging in Radio Astronomy II. Synthesis Imaging in Radio Astronomy II Astronomical Society of the Pacific Conference Series, 180 (URL: https://ui.adsabs.harvard.edu/abs/1999ASPC..180.....T)

[94] Loeb, A.and Barkana, R. (2001), The Reionization of the Universe by the First Stars and Quasars, ARA&A 39(1), 19 (DOI 10.1146/annurev.astro.39.1.19)

[95] Morales, M. F., Wyithe, J. S.B. (2010), Reionization, Cosmology with 21-cm Fluctuations, ARA&A 48(1), 127 (DOI 10.1146/annurev-astro-081309-130936)

[96] Fan, X., Carilli, C.L., Keating, B. (2006), Observational Constraints on Cosmic Reionization, ARA&A 44(1), 415 (DOI: 10.1146/annurev.astro.44.051905.092514)

[97] Planck Collaboration (2019), Planck 2018 results. I. Overview and the cosmological legacy of Planck, A&A 641, A1 (DOI:10.1051/0004-6361/201833880)

[98] Field, George B. (1959a), The Spin Temperature of Intergalactic Neutral Hydrogen, ApJ 129, 536 (DOI: 10.1086/146653)

[99] Field, George B. (959b), The Time Relaxation of a Resonance-Line Profile, ApJ 129, 551, (DOI: 10.1086/146654)

[100] Parsons, A. R., Backer, D. C., Foster, G. S., Wright, M. C. H., Bradley, R. F., Gugliucci, N. E., Parashare, C. R., Benoit E. E., Aguirre, J. E., Jacobs, D. C., Carilli, C. L., Herne, D., Lynch, M. J., Manley, J. R., Werthimer, D. J. (2010), The Precision Array for Probing the Epoch of Re-ionization: Eight Station Results, AJ 139(4), 1468 (DOI: 10.1088/0004-6256/139/4/1468)

[101] DeBoer D.R., Parsons, A. R., Aguirre, J. E., Alexander, P., Ali, Z. S., Beardsley, A. P., Bernardi, G., Bowman, J. D., Bradley, R. F., Carilli, C. L., Cheng, C., de Lera Acedo, E., Dillon, J. S., Ewall-Wice, A., Fadana, G., Fagnoni, N., Fritz, R., Furlanetto, S. R., Glendenning, B., Greig, B., Grobbelaar, J., Hazelton, B., Hewitt, J. N., Hickish, J., Jacobs, D. C., Julius, A., Kariseb, M., Kohn, S. A., Lekalake, T., Liu, A., Loots, A., MacMahon, D., Malan, L., Malgas, C., Maree, M., Martinot, Z., Mathison, N., Matsetela, E., Mesinger, A., Morales, M. F., Neben, A. R., Patra, N., Pieterse, S., Pober, J. C., Razavi-Ghods, N., Ringuette, J., Robnett, J., Rosie, K., Sell, R., Smith, C., Syce, A., Tegmark, M., Thyagarajan, N., Williams, P. K.G., Zheng, H. (2017), Hydrogen Epoch of Reionization Array (HERA), PASP 129(974), 045001 (DOI: 10.1088/1538-3873/129/974/045001)

[102] Kolopanis, M., Jacobs, D. C., Cheng, C., Parsons, A. R., Kohn, S. A., Pober, J. A., Aguirre, J. E., Ali, Z. S., Bernardi, G., Bradley, R. F., Carilli, C. L., DeBoer D.R., Dexter, M. R., Dillon, J. S., Kerrigan, J., Klima, P., Liu, A., MacMahon, D. H. E., Moore, D. F., Thyagarajan, N., Nunhokee, C. D., Walburgh, W. P., Walker, A. (2019), A Simplified, Lossless Reanalysis of PAPER-64, ApJ 833(2), 133 (DOI: 10.3847/1538-4357/ab3e3a)

[103] Wayth, R. B., Tingay, S. J., Trott, C. M., Emrich, D., Johnston-Hollitt, M., McKinley, B., Gaensler, B. M., Beardsley, A. P., Booler, T., Crosse, B. (2018),

The Phase II Murchison Widefield Array: Design overview, PASA 35, e033 (DOI: 10.1017/pasa.2018.37)

[104] Li, W., Pober, J. C., Barry, N., Hazelton, B. J., Morales, M.F., Trott C. M., Lanman, A., M. Wilensky, M., Sullivan, I., Beardsley, A. P., Booler, T., Bowman, J. D., Byrne, R., Crosse, B., Emrich, D., Franzen, T. M. O., Hasegawa, K., Horsley, L., Johnston-Hollitt, M., Jacobs, D. C., Jordan, C. H., Joseph, R. C., Kaneuji, T., Kaplan, D. L., Kenney, D., Kubota, K., Line, J., Lynch C., McKinley, B., Mitchell, D. A., Murray, S., Pallot, D., Pindor, B., Rahimi, M., Riding, J., Sleap, G., Steele, K., Takahashi, K., Tingay, S. J., Walker, M., Wayth, R.B., Webster, R.L., Williams, A., Wu, C., Wyithe, J. S. B., Yoshiura, S., Zheng, Q. (2019) First Season MWA Phase II EoR Power Spectrum Results at Redshift 7, ApJ 887(2), 141 (DOI: 10.3847/1538-4357/ab55e4)

[105] Shaver, P. A., Windhorst, R. A., Madau, P., de Bruyn, A. G. (1999), Can the reionization epoch be detected as a global signature in the cosmic background?, A&A 345, 380 (URL: https://ui.adsabs.harvard.edu/abs/1999A&A...345..380S)

[106] Chakraborty, A., Datta, A., Mazumder, A. (2022), A Comparative Analysis to Deal with Missing Spectral Information Caused by RFI in Cosmological H I 21 cm Observations, ApJ 929, 104 (DOI: 10.3847/1538-4357/ac5cc5)

[107] Di Matteo, T., Perna, R., Abel, T., Rees, M. J. (2002), Radio Foregrounds for the 21 Centimeter Tomography of the Neutral Intergalactic Medium at High Redshifts, ApJ 564(2), 576 (DOI: 10.1086/324293)

[108] Di Matteo, T., Ciardi, B., Miniati, F. (2004), The 21-cm emission from the reionization epoch: extended and point source foregrounds, MNRAS 345(4), 1053, (DOI: 10.1111/j.1365-2966.2004.08443.x)

[109] Cooray, A., Furlanetto, S. R. (2004), Free-Free Emission at Low Radio Frequencies, ApJ 606(1), L5 (DOI: 10.1086/421241)

[110] Datta, A., Bhatnagar, S., Carilli, C.L. (2009), Detection of Signals from Cosmic Reionization, ApJ 703(2), 1851 (DOI: 10.1088/0004-637x/703/2/1851)

[111] Zaldarriaga, M., urlanetto, S. R., Hernquist, L. (2004), 21 Centimeter Fluctuations from Cosmic Gas at High Redshifts, ApJ 608(2), 622 (DOI: 10.1086/386327)

[112] Bharadwaj, S., Saiyad Ali, Sk. (2005), On using visibility correlations to probe the Hi distribution from the dark ages to the present epoch – I. Formalism and the expected signal, MNRAS 356(4), 1519 (DOI: 10.1111/j.1365-2966.2004.08604.x)

[113] 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), Foreground simulations for the LOFAR–epoch of reionization experiment, MNRAS 389(3), 1319 (DOI: 10.1111/j.1365-2966.2008.13634.x)

[114] Jelić V., Zaroubi S., Labropoulos P., Bernardi G., de Bruyn A. G., Koopmans L. V. E. (2010), Realistic simulations of the Galactic polarized foreground: consequences for 21-cm reionization detection experiments MNRAS, 409, 1647 (DOI: 10.1111/j.1365-2966.2010.17407.x)

[115] Chapman E., Bonaldi A., Harker G., Jelic V., Abdalla F. B., Bernardi G., Bobin J., Dulwich, F., Mort, B., Santos, M., Starck, J. L. (2015), Cosmic Dawn and Epoch of Reionization Foreground Removal with the SKA, Advancing Astrophysics with the Square Kilometre Array (AASKA14), April 2015 (DOI: 10.22323/1.215.0005)

[116] Chapman E., Abdalla F. B., Bobin J., Starck J.-L., Harker G., Jelić V., Labropoulos P., Zaroubi, S., Brentjens, M. A., de Bruyn, A. G., Koopmans, L. V. E. (2013), The scale of the problem: recovering images of reionization with Generalized Morphological Component Analysis MNRAS, 429, 165 (DOI: 10.1093/mnras/sts333)

[117] Datta A., Bowman J. D., Carilli C. L.(2010), Bright Source Subtraction Requirements for Redshifted 21 cm Measurements ApJ, 724, 526 (DOI: 10.1088/0004-637X/724/1/526)

[118] Ghosh A., Prasad J., Bharadwaj S., Ali S. S., Chengalur J. N.(2012), Characterizing foreground for redshifted 21 cm radiation: 150 MHz Giant Metrewave Radio Telescope observations MNRAS, 426, 3295 (DOI:10.1111/j.1365-2966.2012.21889.x)

[119] Koopmans L., Pritchard J., Mellema G., Aguirre J., Ahn K., Barkana R., van Bemmel I., Bernardi, G., Bonaldi, A., Briggs, F., de Bruyn, A. G., Chang, T. C., Chapman, E., Chen, X., Courty, B., Dayal, P., Ferrara, A., Fialkov, A., Fiore, F., Ichiki, K.., Illiev, I. T., Inoue, S., Jelic, V., Jones, M., Lazio, J., Maio, U., Majumdar, S., Mack, K. J., Mesinger, A., Morales, M. F., Parsons, A., Pen, U.L., Santos, M., Schneider, R., Semelin, B., de Souza, R. S., Subrahmanyan, R., Takeuchi, T., Vedantham, H., Wagg, J., Webster, R., Wyithe, S., Datta, K. K., Trott, C. (2015), The Cosmic Dawn and Epoch of Reionisation with SKA, Advancing Astrophysics with the Square Kilometre Array (AASKA14), April 2015 (DOI: 10.22323/1.215.0001)

[120] de Oliveira-Costa, A., Tegmark, M., Gaensler, B. M., Jonas, J., Landecker, T. L., Reich, P.(2008), A model of diffuse Galactic radio emission from 10 MHz to 100 GHz MNRAS 388, 247 (DOI: 10.1111/j.1365-2966.2008.13376.x)

[121] Wilman, R. J., Miller L., Jarvis M. J., Mauch T., Levrier F., Abdalla F. B., Rawlings S., Klöckner, H. -R.,Obreschkow, D., Olteanu,D., Young, S. (2008), A semi-empirical simulation of the extragalactic radio continuum sky for next generation radio telescopes MNRAS 388, 1335 (DOI: 10.1111/j.1365-2966.2008.13486.x)

[122] Bonaldi, A., Bonato, M., Galluzzi, V., Harrison, I., Massardi, M., Kay, S., De Zotti, G., Brown, M. L.(2019), The Tiered Radio Extragalactic Continuum Simulation (T-RECS), MNRAS 482, 2 (DOI: 10.1093/mnras/sty2603)

[123] Trott, C. M., Wayth, R. B., Tingay, S. J. (2012) The Impact of Point-source Subtraction Residuals on 21 cm Epoch of Reionization Estimation ApJ 757, 101 (DOI: 10.1088/0004-637X/757/1/101)

[124] Ghosh, A., Bharadwaj, S., Ali, S. S., Chengalur, J. N. (2011), GMRT observation towards detecting the post-reionization 21-cm signal, MNRAS 411, 2426 (DOI: 10.1111/j.1365-2966.2010.17853.x)

[125] Sims, P. H., Lentati, L., Alexander, P., Carilli, C. L.(2016), Contamination of the Epoch of Reionization power spectrum in the presence of foregrounds, MNRAS 462, 3069 (DOI: 10.1093/mnras/stw1768)

[126] Santos, M. G., Cooray, A., Knox, L. (2005), Multifrequency Analysis of 21 Centimeter Fluctuations from the Era of Reionization, ApJ 625, 575 (DOI: 10.1086/429857)

[127] Tasse, C., Hugo, B., Mirmont, M., Smirnov, O., Atemkeng, M., Bester, L., Hardcastle, M. J., Lakhoo, R., Perkins, S., Shimwell, T. (2018), Faceting for direction-dependent spectral deconvolution, A&A 611, 87 (DOI: 10.1051/0004-6361/201731474)

[128] Datta, K. K., Choudhury, T. R., Bharadwaj, S. (2007), The multifrequency angular power spectrum of the epoch of reionization 21-cm signal, MNRAS 378, 119 (DOI: 10.1111/j.1365-2966.2007.11747.x)

[129] Ali, S. S., Bharadwaj, S., Chengalur, J. N. (2008), Foregrounds for redshifted 21-cm studies of reionization: Giant Meter Wave Radio Telescope 153-MHz observations, MNRAS 385, 2166 (DOI: 10.1111/j.1365-2966.2008.12984.x)

[130] Liu, A., Tegmark, M., Zaldarriaga, M. (2009), Will point sources spoil 21-cm tomography?, MNRAS 394, 1575 (DOI: 10.1111/j.1365-2966.2009.14426.x)

[131] Sims P. H., Pober J. C. (2019), Joint estimation of the Epoch of Reionization power spectrum and foregrounds, MNRAS 488, 2904 (DOI:10.1093/mnras/stz1888)

[132] Trott C. M., Pindor B., Procopio P., Wayth R. B., Mitchell D. A., McKinley B., Tingay S. J., Barry, N., Beardsley, A. P., Bernardi, G., Bowman, J. D., Briggs, F.,Cappallo, R. J.,Carroll, P., de Oliveira-Costa, A., Dillon, J. S., Ewall-Wice, A., Feng, L., Greenhill, L. J., Hazelton, B. J., Hewitt, J. N., Hurley-Walker, N., Johnston-Hollitt, M., Jacobs, Daniel C., Kaplan, D. L., Kim, H. S., Lenc, E., Line, J., Loeb, A., Lonsdale, C. J., Morales, M. F., Morgan, E., Neben, A. R., Thyagarajan, Nithyanandan, Oberoi, D., Offringa, A. R., Ord, S. M., Paul, S., Pober, J. C., Prabu, T., Riding, J., Udaya Shankar, N., Sethi, Shiv K., Srivani, K. S., Subrahmanyan, R., Sullivan, I. S., Tegmark, M., Webster, R. L., Williams, A., Williams, C. L., Wu, C., Wyithe, J. S. B (2016), CHIPS: The Cosmological H I Power Spectrum Estimator, ApJ 818, 139 (DOI: 10.3847/0004-637X/818/2/139)

[133] Intema, H. T., Jagannathan, P., Mooley, K. P., Frail, D. A. (2017), The GMRT 150 MHz all-sky radio survey. First alternative data release TGSS ADR1, A& A, 598 (DOI: 10.1051/0004-6361/201628536)

[134] Hurley-Walker, N., Callingham, J. R., Hancock, P. J., Franzen, T. M. O., Hindson, L., Kapińska, A. D., Morgan, J., Offringa, A. R., Wayth, R. B., Wu, C., Zheng, Q., Murphy, T., Bell, M. E., Dwarakanath, K. S., For, B., Gaensler, B. M., Johnston-Hollitt, M., Lenc, E., Procopio, P., Staveley-Smith, L., Ekers, R., Bowman, J. D., Briggs, F., Cappallo, R. J., Deshpande, A. A., Greenhill, L., Hazelton, B. J., Kaplan, D. L., Lonsdale, C. J., McWhirter, S. R., Mitchell, D. A, Morales, M. F., Morgan, E., Oberoi, D., Ord, S. M., Prabu, T., Shankar, N. Udaya, Srivani, K. S., Subrahmanyan, R., Tingay, S. J., Webster, R. L., Williams, A., Williams, C. L. (2017), GaLactic and Extragalactic All-sky Murchison Widefield Array (GLEAM) survey - I. A low-frequency extragalactic catalogue, MNRAS 464, 1146 (DOI: 10.1093/mnras/stw2337)

[135] Franzen, T. M. O., Vernstrom, T., Jackson, C. A., Hurley-Walker, N., Ekers, R. D., Heald, G., Seymour, N., White, S. V.(2019), Source counts and confusion at 72-231 MHz in the MWA GLEAM survey, PASA 36, e004 (DOI: 10.1017/pasa.2018.52)

[136] Williams, W. L., van Weeren, R. J., Röttgering, H. J. A., Best, P., Dijkema, T. J., de Gasperin, F., Hardcastle, M. J., Heald, G., Prandoni, I., Sabater, J., Shimwell, T. W., Tasse, C., van Bemmel, I. M., Brüggen, M., Brunetti, G., Conway, J. E., Enßlin, T., Engels, D., Falcke, H., Ferrari, C., Haverkorn, M., Jackson, N., Jarvis, M. J., Kapińska, A. D., Mahony, E. K., Miley, G. K., Morabito, L. K., Morganti, R., Orrú, E., Retana-Montenegro, E., Sridhar, S. S., Toribio, M. C., White, G. J., Wise, M. W., Zwart, J. T. L. (2016), LOFAR 150-MHz observations of the Boötes field: catalogue and source counts, MNRAS 460, 2385 (DOI: 10.1093/mnras/stw1056)

[137] Prandoni, I., Guglielmino, G., Morganti, R., Vaccari, M., Maini, A., Röttgering, H. J. A., Jarvis, M. J., Garrett, M. A., (2018), The Lockman

Hole Project: new constraints on the sub-mJy source counts from a wide-area 1.4 GHz mosaic, MNRAS 481, 4548 (DOI: 10.1093/mnras/sty2521)

[138] Hale, C. L., Williams, W., Jarvis, M. J., Hardcastle, M. J., Morabito, L. K., Shimwell, T. W., Tasse, C., Best, P. N., Harwood, J. J., Heywood, I., Prandoni, I., Röttgering, H. J. A., Sabater, J., Smith, D. J. B., van Weeren, R. J. (2019), LOFAR observations of the XMM-LSS field, A&A 622, A4 (DOI: 10.1051/0004-6361/201833906)

[139] Simpson, C., Martínez-Sansigre, A., Rawlings, S., Ivison, R., Akiyama, M., Sekiguchi, K., Takata, T., Ueda, Y., Watson, M. (2006), Radio imaging of the Subaru/XMM-Newton Deep Field - I. The 100-μJy catalogue, optical identifications, and the nature of the faint radio source population, MNRAS 372, 741 (DOI: 10.1111/j.1365-2966.2006.10907.x)

[140] Seymour, N., Dwelly, T., Moss, D., McHardy, I., Zoghbi, A., Rieke, G., Page, M., Hopkins, A., Loaring, N. (2008), The star formation history of the Universe as revealed by deep radio observations, MNRAS 386, 1695 (DOI:10.1111/j.1365-2966.2008.13166.x)

[141] Mignano, A., Prandoni, I., Gregorini, L., Parma, P., de Ruiter, H. R., Wieringa, M. H., Vettolani, G., Ekers, R. D. (2008), The ATESP 5 GHz radio survey. II. Physical properties of the faint radio population, A&A 477, 459 (DOI:10.1051/0004-6361:20078545)

[142] Smolčić, V., Schinnerer, E., Scodeggio, M., Franzetti, P., Aussel, H., Bondi, M., Brusa, M., Carilli, C. L., Capak, P., Charlot, S., Ciliegi, P., Ilbert, O., Ivezić, Ž, Jahnke, K., McCracken, H. J., Obrić, M., Salvato, M., Sanders, D. B., Scoville, N., Trump, J. R., Tremonti, C., Tasca, L., Walcher, C. J., Zamorani, G. (2008), A New Method to Separate Star-forming from AGN Galaxies at Inter-

mediate Redshift: The Submillijansky Radio Population in the VLA-COSMOS Survey, ApJS 177, 14 (DOI: 10.1086/588028)

[143] Padovani, P., Mainieri, V. , Tozzi, P., Kellermann, K. I., Fomalont, E. B., Miller, N., Rosati, P., Shaver, P. (2009), The Very Large Array Survey of the Chandra Deep Field South. IV. Source Population, ApJ 694, 235 (DOI: 10.1088/0004-637X/694/1/235)

[144] Padovani, P., Miller, N., Kellermann, K. I., Mainieri, V., Rosati, P., Tozzi, P. (2011), The VLA Survey of Chandra Deep Field South. V. Evolution and Luminosity Functions of Sub-millijansky Radio Sources and the Issue of Radio Emission in Radio-quiet Active Galactic Nuclei (ApJ) 740, 20 (DOI: 10.1088/0004-637X/740/1/20)

[145] Padovani, P., Bonzini, M., Kellermann, K. I., Miller, N., Mainieri, V., Tozzi, P. (2015), Radio-faint AGN: a tale of two populations, MNRAS 452, 1263 (DOI: 10.1093/mnras/stv1375)

[146] Lockman, F. J., Jahoda, K., McCammon, D. (1986), The Structure of Galactic H i in Directions of Low Total Column Density, ApJ 302, 432 (DOI: 10.1086/164002)

[147] Lonsdale, C. J., Smith, H. E., Rowan-Robinson, M., Surace, J., Shupe, D., Xu, C., Oliver, S., Padgett, D., Fang, F., Conrow, T., Franceschini, A., Gautier, N., Griffin, M., Hacking, P., Masci, F., Morrison, G., O'Linger, J., Owen, F., Pérez-Fournon, I., Pierre, M., Puetter, Ri., Stacey, G., Castro, S., Polletta, M., Farrah, D., Jarrett, T., Frayer, D., Siana, B., Babbedge, T., Dye, S., Fox, M., Gonzalez-Solares, E., Salaman, M., Berta, S., Condon, J. J., Dole, H., Serjeant, S. (2003), SWIRE: The SIRTF Wide-Area Infrared Extragalactic Survey, PASP 115, 897 (DOI: 10.1086/376850)

[148] de Ruiter, H. R., Zamorani, G., Parma, P., Hasinger, G., Hartner, G., Truemper, J., Burg, R., Giacconi, R., Schmidt, M. (1997), Deep radio observations of the "Lockman Hole", A&A 319, 7 (URL : https://ui.adsabs.harvard.edu/abs/1997A&A...319....7D)

[149] Ciliegi, P., Zamorani, G., Hasinger, G., Lehmann, I., Szokoly, G., Wilson, G. (2003), A deep VLA survey at 6 cm in the Lockman Hole A&A 398, 901 (DOI: 10.1051/0004-6361:20021721)

[150] Owen, F. N., Morrison, G. E. (2008) The Deep Swire Field. I. 20 cm Continuum Radio Observations: A Crowded Sky AJ 136, 1889 (DOI:10.1088/0004-6256/136/5/1889)

[151] Owen, F. N., Morrison, G. E., Klimek, M. D., Greisen, E. W. (2009) The Deep SWIRE Field. II. 90 cm Continuum Observations and 20 cm-90 cm Spectra ApJ 137, 4846 (DOI:10.1088/0004-6256/137/6/4846)

[152] Ibar, E., Ivison, R. J., Biggs, A. D., Lal, D. V., Best, P. N., Green, D. A. (2009) Deep multi-frequency radio imaging in the Lockman Hole using the GMRT and VLA - I. The nature of the sub-mJy radio population MNRAS 397, 281 (DOI: 10.1111/j.1365-2966.2009.14866.x)

[153] Condon, J. J., Cotton W. D., Fomalont E. B., Kellermann K. I., Miller N., Perley R. A., Scott D., Vernstrom, T., Wall, J.V. (2012), Resolving the Radio Source Background: Deeper Understanding through Confusion, ApJ 758, 23 DOI:10.1088/0004-637X/758/1/23

[154] Vernstrom, T., Scott, D., Wall, J. V., Condon, J. J., Cotton, W. D., Fomalont, E. B., Kellermann, K. I., Miller, N., Perley, R. A. (2014), Deep 3 GHz number counts from a P(D) fluctuation analysis MNRAS 440, 2791 (DOI: 10.1093/mnras/stu470)

[155] Vernstrom, T., Scott, D., Wall, J. V., Condon, J. J., Cotton, W. D., Perley, R. A. (2016), Deep 3-GHz observations of the Lockman Hole North with the Very Large Array - I. Source extraction and uncertainty analysis MNRAS 461, 2879 (DOI: 10.1093/mnras/stw1530)

[156] Vernstrom, T., Scott, Douglas, Wall, J. V., Condon, J. J., Cotton, W. D., Kellermann, K. I., Perley, R. A., (2016), Deep 3-GHz observations of the Lockman Hole North with the Very Large Array - II. Catalogue and μJy source properties, MNRAS 462, 2934 (DOI: 10.1093/mnras/stw1836)

[157] Garn, T., Green, D. A., Riley, J. M., Alexander, P. (2008), A 610-MHz survey of the Lockman Hole with the Giant Metrewave Radio Telescope - I. Observations, data reduction and source catalogue for the central 5 deg^2, MNRAS 387, 1037, (DOI: 10.1111/j.1365-2966.2008.13335.x)

[158] Garn, T. S., Green, D. A., Riley, J. M., Alexander, P. (2010), Further GMRT observations of the Lockman Hole at 610 MHz BASI 38, 103 (URL: https://ui.adsabs.harvard.edu/abs/2010BASI...38..103G)

[159] Mahony, E. K., Morganti, R., Prandoni, I., van Bemmel, I. M., Shimwell, T. W., Brienza, M., Best, P. N., Brüggen, M., Calistro Rivera, G., de Gasperin, F., Hardcastle, M. J., Harwood, J. J., Heald, G., Jarvis, M. J., Mandal, S., Miley, G. K., Retana-Montenegro, E., Röttgering, H. J. A., Sabater, J., Tasse, C., van Velzen, S., van Weeren, R. J., Williams, W. L., White, G. J. (2016), The Lockman Hole project: LOFAR observations and spectral index properties of low-frequency radio sources, MNRAS 463, 2997 (DOI: 10.1093/mnras/stw2225)

[160] Choudhuri, S., Bharadwaj, S., Ghosh, A., Ali, S. S. (2014) Visibility-based angular power spectrum estimation in low-frequency radio interferometric observations, MNRAS 445, 4351 (DOI: 10.1093/mnras/stu2027)

[161] Choudhuri, S., Bharadwaj, S., Chatterjee, S., Ali, S. S., Roy, N., Ghosh, A. (2016), The visibility-based tapered gridded estimator (TGE) for the redshifted 21-cm power spectrum, MNRAS 463, 4093 (DOI: 10.1093/mnras/stw2254)

[162] Swarup, G., Ananthakrishnan, S., Kapahi, V. K., Rao, A. P., Subrahmanya, C. R., Kulkarni, V. K. (1991), The Giant Metre-Wave Radio Telescope, Current Science 60, 95 (URL: https://ui.adsabs.harvard.edu/abs/1991CuSc...60...95S)

[163] Intema, H. T., van der Tol, S., Cotton, W. D., Cohen, A. S., van Bemmel, I. M., Röttgering, H. J. A. (2009), Ionospheric calibration of low frequency radio interferometric observations using the peeling scheme. I. Method description and first results A&A 501, 1185 (DOI: 10.1051/0004-6361/200811094)

[164] Intema, H. T. (2014), SPAM: Source Peeling and Atmospheric Modeling, Astrophysics Source Code Library, ascl:1408.006, (URL: https://ui.adsabs.harvard.edu/abs/2014ascl.soft08006I)

[165] Intema, H. T. (2014) SPAM: A data reduction recipe for high-resolution,low-frequency radio-interferometric observations, Astronomical Society of India Conference Series 13, 469 (URL: https://ui.adsabs.harvard.edu/abs/2014ASInC..13..469I)

[166] Greisen, E. W. (1998) Recent Developments in Experimental AIPS, Astronomical Data Analysis Software and Systems VII 145, 204 (URL: https://ui.adsabs.harvard.edu/abs/1998ASPC..145..204G)

[167] Greisen, E. W. (2003), AIPS, the VLA, and the VLBA, Information Handling in Astronomy - Historical Vistas 285, 109 (DOI: 10.1007/0-306-48080-8_7)

[168] Scaife, A. M. M. and Heald, G. H. (2012), A broad-band flux scale for low-frequency radio telescopes, MNRAS 423, L30 (DOI: 10.1111/j.1745-3933.2012.01251.x)

[169] Mohan, N. and Rafferty, D. (2015) PyBDSF: Python Blob Detection and Source Finder, Astrophysics Source Code Library, ascl:1502.007 (URL: https://ui.adsabs.harvard.edu/abs/2015ascl.soft02007M)

[170] Franzen, T. M. O., Banfield, J. K., Hales, C. A., Hopkins, A., Norris, R. P., Seymour, N., Chow, K. E., Herzog, A., Huynh, M. T., Lenc, E., Mao, M. Y., Middelberg, E. (2015), ATLAS - I. Third release of 1.4 GHz mosaics and component catalogues, MNRAS 453, 4020 (DOI: 10.1093/mnras/stv1866)

[171] Becker, R. H., White, R. L., Helfand, D. J. (1995), The FIRST Survey: Faint Images of the Radio Sky at Twenty Centimeters, ApJ 450, 559 (DOI: 10.1086/176166)

[172] Baars, J. W. M., Genzel, R., Pauliny-Toth, I. I. K., Witzel, A. (1977), The absolute spectrum of Cas A: an accurate flux density scale and a set of secondary calibrators, A&A 61, 99 (URL: https://ui.adsabs.harvard.edu/abs/1977A&A....61...99B)

[173] Smolčić, V., Novak, M., Bondi, M., Ciliegi, P., Mooley, K. P., Schinnerer, E., Zamorani, G., Navarrete, F., Bourke, S., Karim, A., Vardoulaki, E., Leslie, S., Delhaize, J., Carilli, C. L., Myers, S. T., Baran, N., Delvecchio, I., Miettinen, O., Banfield, J., Baloković, M., Bertoldi, F., Capak, P., Frail, D. A., Hallinan, G., Hao, H., Herrera Ruiz, N., Horesh, A., Ilbert, O., Intema, H., Jelić, V., Klöckner, H.-R., Krpan, J., Kulkarni, S. R., McCracken, H., Laigle, C., Middleberg, E., Murphy, E. J., Sargent, M., Scoville, N. Z., Sheth, K. (2017), The VLA-COSMOS 3 GHz Large Project: Continuum data and source catalog release A& 602, A1 (DOI: 10.1051/0004-6361/201628704)

[174] Chakraborty, A., Datta, A., Choudhuri, S., Roy, N., Intema, H., Choudhury, M., Datta, K. K, Pal, S., Bharadwaj, S., Dutta, P., Choudhury, T. R. (2019),

Detailed study of the ELAIS N1 field with the uGMRT - I. Characterizing the 325 MHz foreground for redshifted 21 cm observations, MNRAS 487, 4102, (DOI: 10.1093/mnras/stz1580)

[175] Chakraborty, A. Roy, N., Datta, A., Choudhuri, S., Datta, K. K., Dutta, P., Bharadwaj, S., Intema, H., Choudhury, M., Pal, S., Choudhury, T. R. (2019), Detailed study of ELAIS N1 field with the uGMRT - II. Source properties and spectral variation of foreground power spectrum from 300-500 MHz observations, MNRAS 490, 243 (DOI: 10.1093/mnras/stz2533)

[176] Ocran, E. F., Taylor, A. R., Vaccari, M., Ishwara-Chandra, C. H., Prandoni, I. (2020), Deep GMRT 610 MHz observations of the ELAIS N1 field: catalogue and source counts, MNRAS 491, 1127 (DOI: 10.1093/mnras/stz2954)

[177] Eddington, A. S. (1913), On a formula for correcting statistics for the effects of a known error of observation, MNRAS 73, 359 (DOI: 10.1093/mnras/73.5.359)

[178] Intema, H. T., van Weeren, R. J., Röttgering, H. J. A., Lal, D. V. (2011), Deep low-frequency radio observations of the NOAO Boötes field. I. Data reduction and catalog construction, A&A 535, A38 (DOI: 10.1051/0004-6361/201014253)

[179] Williams, W. L., Intema, H. T., Röttgering, H. J. A. (2013), T-RaMiSu: the Two-meter Radio Mini Survey. I. The Boötes Field, A&A 549, A55 (DOI: 10.1051/0004-6361/201220235)

[180] Hancock, P. J., Murphy, T., Gaensler, B. M., Hopkins, A., Curran, J. R. (2012), Compact continuum source finding for next generation radio surveys, MNRAS 422, 1812 (DOI: 10.1111/j.1365-2966.2012.20768.x)

[181] Hancock, P. J., Trott, C. M., Hurley-Walker, N. (2018), Source Finding in the Era of the SKA (Precursors): Aegean 2.0., PASA 35, e011 (DOI: 10.1017/pasa.2018.3)

[182] Windhorst, R. A., Miley, G. K., Owen, F. N., Kron, R. G., Koo, D. C. (1985), Sub-millijansky 1.4 GHz source counts and multicolor studies of weak rado galaxy populations, ApJ 289, 494 (DOI: 10.1086/162911)

[183] Choudhuri, S., Bharadwaj, S., Ali, S. S., Roy, N., Intema, H. T., Ghosh, A. (2017), The angular power spectrum measurement of the Galactic synchrotron emission in two fields of the TGSS survey, MNRAS 470, L11, (DOI: 10.1093/mnrasl/slx066)

[184] Bernardi, G., de Bruyn, A. G., Brentjens, M. A., Ciardi, B., Harker, G., Jelić, V., Koopmans, L. V. E., Labropoulos, P., Offringa, A., Pandey, V. N., Schaye, J., Thomas, R. M., Yatawatta, S., Zaroubi, S. (2009), Foregrounds for observations of the cosmological 21 cm line. I. First Westerbork measurements of Galactic emission at 150 MHz in a low latitude field, A&A 500, 965 (DOI: 10.1051/0004-6361/200911627)

[185] Iacobelli, M., Haverkorn, M., Orrú, E., Pizzo, R. F., Anderson, J., Beck, R., Bell, M. R., Bonafede, A., Chyzy, K., Dettmar, R.-J., Ensslin, T. A., Heald, G., Horellou, C., Horneffer, A., Jurusik, W., Junklewitz, H., Kuniyoshi, M., Mulcahy, D. D., Paladino, R., Reich, W., Scaife, A., Sobey, C., Sotomayor-Beltran, C., Alexov, A., Asgekar, A., Avruch, I. M., Bell, M. E., van Bemmel, I., Bentum, M. J., Bernardi, G., Best, P., Birzan, L., Breitling, F., Broderick, J., Brouw, W. N., Brüggen, M., Butcher, H. R., Ciardi, B., Conway, J. E., de Gasperin, F., de Geus, E., Duscha, S., Eislöffel, J., Engels, D., Falcke, H., Fallows, R. A., Ferrari, C., Frieswijk, W., Garrett, M. A., Griessmeier, J., Gunst, A. W., Hamaker, J. P., Hassall, T. E., Hessels, J. W. T., Hoeft, M., Hörandel, J., Jelic, V., Karastergiou, A., Kondratiev, V. I., Koopmans, L. V. E., Kramer, M., Kuper, G., van Leeuwen, J., Macario, G., Mann, G., McKean, J. P., Munk, H., Pandey-Pommier, M., Polatidis, A. G., Röttgering, H., Schwarz,

D., Sluman, J., Smirnov, O., Stappers, B. W., Steinmetz, M., Tagger, M., Tang, Y., Tasse, C., Toribio, C., Vermeulen, R., Vocks, C., Vogt, C., van Weeren, R. J., Wise, M. W., Wucknitz, O., Yatawatta, S., Zarka, P., Zensus, A. (2013), Studying Galactic interstellar turbulence through fluctuations in synchrotron emission. First LOFAR Galactic foreground detection, A&A 558, A72 (DOI: 10.1051/0004-6361/201322013)

[186] Choudhuri, S., Ghosh, A., Roy, N., Bharadwaj, S., Intema, H. T., Ali, S. S. (2020), All-sky angular power spectrum - I. Estimating brightness temperature fluctuations using the 150-MHz TGSS survey, MNRAS 494, 1936, (DOI: 10.1093/mnras/staa762)

[187] Haslam, C. G. T., Salter, C. J., Stoffel, H., Wilson, W. E. (1982), A 408-MHZ All-Sky Continuum Survey. II. The Atlas of Contour Maps, A&A Supplement Series 47, 1 (URL: https://ui.adsabs.harvard.edu/abs/1982A&AS...47....1H)

[188] La Porta, L., Burigana, C., Reich, W., Reich, P. (2008), The impact of Galactic synchrotron emission on CMB anisotropy measurements. I. Angular power spectrum analysis of total intensity all-sky surveys, A&A 479, 641 (DOI: 10.1051/0004-6361:20078435)

[189] Reich, P., Reich, W. (1988), A map of spectral indices of the galactic radio continuum emission between 408 MHz and 1420 MHz for the entire northern sky, A&A Supplement Series 74, 7 (URL: https://ui.adsabs.harvard.edu/abs/1988A&AS...74....7R)

[190] Liu, A., Shaw, J. R. (2020), Data Analysis for Precision 21 cm Cosmology PASP 132 (DOI:10.1088/1538-3873/ab5bfd)

[191] Chapman, E., Jelić, V.(2019), Chapter 6: Foregrounds and their mitigation, The Cosmic 21-cm Revolution, IOP Publishing (DOI: 10.1088/2514-3433/ab4a73ch6)

[192] Bernardi, G. (2019), Chapter 5: 21 cm Observations: Calibration, Strategies, Observables, The Cosmic 21-cm Revolution, IOP Publishing (DOI: 10.1088/2514-3433/ab4a73ch5)

[193] Haslam, C. G. T., Klein, U., Salter, C. J., Stoffel, H., Wilson, W. E., Cleary, M. N., Cooke, D. J., Thomasson, P. (1981), A 408 MHz all-sky continuum survey. I - Observations at southern declinations and for the North Polar region, A&A 100, 209 (URL: https://ui.adsabs.harvard.edu/abs/1981A&A...100..209H)

[194] Haslam, C. G. T., Salter, C. J., Stoffel, H., Wilson, W. E. (1982), A 408-MHZ All-Sky Continuum Survey. II. The Atlas of Contour Maps, A&AS 47, 1. (URL: https://ui.adsabs.harvard.edu/abs/1982A&AS...47....1H)

[195] Guzmán, A. E., May, J., Alvarez, H., Maeda, K. (2011), All-sky Galactic radiation at 45 MHz and spectral index between 45 and 408 MHz, A&A 525 (DOI:10.1051/0004-6361/200913628)

[196] Remazeilles, M., Dickinson, C., Banday, A. J., Bigot-Sazy, M.-A., Ghosh, T. (2015), An improved source-subtracted and destriped 408-MHz all-sky map, MNRAS 451, 4311 (DOI:10.1093/mnras/stv1274)

[197] Dowell, J., Taylor, G. B., Schinzel, F. K., Kassim, N. E., Stovall, K. (2017), The LWA1 Low Frequency Sky Survey, MNRAS 469, 4537 (DOI:10.1093/mnras/stx1136)

[198] Zheng, H., Tegmark, M., Dillon, J. S., Kim, D. A., Liu, A., Neben, A. R., Jonas, J. , Reich, P., Reich, W. (2017), An improved model of dif-

fuse galactic radio emission from 10 MHz to 5 THz, MNRAS 464, 3486 (DOI:10.1093/mnras/stw2525)

[199] Irfan, M. O., Bull, P., Santos, M.G., Wang, J., Grainge, K., Li, Y., Carucci, I.P., Spinelli, M., Cunnington, S. (2022), Measurements of the diffuse Galactic synchrotron spectral index and curvature from MeerKLASS pilot data, MNRAS 509, 4923 (DOI:10.1093/mnras/stab3346)

[200] Byrne, R. , Morales, M.F., Hazelton, B., Sullivan, I., Barry, N., Lynch, C., Line, J. L. B., Jacobs, D.C. (2022), A map of diffuse radio emission at 182 MHz to enhance epoch of reionization observations in the Southern hemisphere, MNRAS 510, 2011 (DOI:10.1093/mnras/stab3276)

[201] , Shimwell, T. W., Röttgering, H. J. A., Best, P. N., Williams, W. L., Dijkema, T. J., de Gasperin, F., Hardcastle, M. J., Heald, G. H., Hoang, D. N., Horneffer, A., Intema, H., Mahony, E. K., Mandal, S., Mechev, A. P., Morabito, L., Oonk, J. B. R., Rafferty, D., Retana-Montenegro, E., Sabater, J., Tasse, C., van Weeren, R. J., Brüggen, M., Brunetti, G., Chyży, K. T., Conway, J. E., Haverkorn, M., Jackson, N., Jarvis, M. J., McKean, J. P., Miley, G. K., Morganti, R., White, G. J., Wise, M. W., van Bemmel, I. M., Beck, R., Brienza, M., Bonafede, A., Calistro Rivera, G., Cassano, R., Clarke, A. O., Cseh, D., Deller, A., Drabent, A., van Driel, W., Engels, D., Falcke, H., Ferrari, C., Fröhlich, S., Garrett, M. A., Harwood, J. J., Heesen, V., Hoeft, M., Horellou, C., Israel, F. P., Kapińska, A. D., Kunert-Bajraszewska, M., McKay, D. J., Mohan, N. R., Orrú, E., Pizzo, R. F., Prandoni, I., Schwarz, D. J., Shulevski, A., Sipior, M., Smith, D. J. B., Sridhar, S. S., Steinmetz, M., Stroe, A., Varenius, E., van der Werf, P. P., Zensus, J. A., Zwart, J. T. L. (2017), The LOFAR Two-metre Sky Survey. I. Survey description and preliminary data release, A&A 598, A104 (DOI: 10.1051/0004-6361/201629313)

[202] Lynch, C. R. Galvin, T. J., Line, J. L. B., Jordan, C. H., Trott, C. M., Chege, J. K., McKinley, B., Johnston-Hollitt, M., Tingay, S. J. (2021), The MWA long baseline Epoch of reionisation survey—I. Improved source catalogue for the EoR 0 field, PASA 38 (DOI:10.1017/pasa.2021.50)

[203] Bernardi, G., de Bruyn, A. G., Harker, G., Brentjens, M. A., Ciardi, B., Jelić, V., Koopmans, L. V. E., Labropoulos, P., Offringa, A., Pandey, V. N., Schaye, J., Thomas, R. M., Yatawatta, S., Zaroubi, S. (2010), Foregrounds for observations of the cosmological 21 cm line. II. Westerbork observations of the fields around 3C 196 and the North Celestial Pole, A&A 522 (DOI:10.1051/0004-6361/200913420)

[204] Lenc, E., Gaensler, B. M., Sun, X. H.,Sadler, E. M., Willis, A. G., Barry, N., Beardsley, A. P., Bell, M. E., Bernardi, G., Bowman, J. D., Briggs, F., Callingham, J. R., Cappallo, R. J., Carroll, P., Corey, B. E., de Oliveira-Costa, A., Deshpande, A. A., Dillon, J. S., Dwarkanath, K. S., Emrich, D., Ewall-Wice, A., Feng, L., For, B. -Q., Goeke, R., Greenhill, L. J., Hancock, P., Hazelton, B. J., Hewitt, J. N., Hindson, L., Hurley-Walker, N., Johnston-Hollitt, M., Jacobs, D. C., Kapińska, A. D., Kaplan, D. L., Kasper, J. C., Kim, H. -S., Kratzenberg, E., Line, J., Loeb, A., Lonsdale, C. J., Lynch, M. J., McKinley, B., McWhirter, S. R., Mitchell, D. A., Morales, M. F., Morgan, E., Morgan, J., Murphy, T., Neben, A. R., Oberoi, D., Offringa, A. R., Ord, S. M., Paul, S., Pindor, B., Pober, J. C.,Prabu, T., Procopio, P., Riding, J., Rogers, A. E. E., Roshi, A., Udaya Shankar, N., Sethi, S. K., Srivani, K. S., Staveley-Smith, L., Subrahmanyan, R., Sullivan, I. S., Tegmark, M., Thyagarajan, N., Tingay, S. J., Trott, C., Waterson, M., Wayth, R. B., Webster, R. L., Whitney, A. R., Williams, A., Williams, C. L., Wu, C., Wyithe, J. S. B., Zheng, Q. (2016,), Low-frequency Observations of Linearly Polarized Structures in the Interstellar Medium near the South Galactic Pole, ApJ 830, 38 (DOI: 10.3847/0004-637X/830/1/38)

[205] Hardcastle, M. J., Croston, J. H. (2020), Radio galaxies and feedback from AGN jets, New AR 88, (DOI: 10.1016/j.newar.2020.101539)

[206] Mandal, S., Prandoni, I., Hardcastle, M. J., Shimwell, T. W., Intema, H. T., Tasse, C., van Weeren, R. J., Algera, H., Emig, K. L., Röttgering, H. J. A., Schwarz, D. J., Siewert, T. M., Best, P. N., Bonato, M., Bondi, M., Jarvis, M. J., Kondapally, R., Leslie, S. K., Mahatma, V. H., Sabater, J., Retana-Montenegro, E., Williams, W. L. (2021), Extremely deep 150 MHz source counts from the LoTSS Deep Fields, A&A 648, A5 (DOI: 10.1051/0004-6361/202039998)

[207] Tegmark, M., Eisenstein, D. J., Hu, W., de Oliveira-Costa, A. (2000), Foregrounds and Forecasts for the Cosmic Microwave Background, ApJ 530, 133 (DOI: 10.1086/308348)

[208] Planck Collaboration (2016), Planck 2015 results. X. Diffuse component separation: Foreground maps, A&A 594 (DOI:10.1051/0004-6361/201525967)

[209] Chapman, E., Abdalla, F.B., Bobin, J., Starck, J.-L., Harker, G., Jelić, V., Labropoulos, P., Zaroubi, S., Brentjens, M.A., de Bruyn, A. G., Koopmans, L. V. E. (2012), The scale of the problem: recovering images of reionization with Generalized Morphological Component Analysis, MNRAS 429, 165 (DOI: 10.1093/mnras/sts333)

[210] Chapman, Emma, Abdalla, Filipe B., Harker, Geraint, Jelić, Vibor, Labropoulos, Panagiotis, Zaroubi, Saleem, Brentjens, Michiel A., de Bruyn, A. G., Koopmans, L. V. E. (2012), Foreground removal using fastica : a showcase of LOFAR-EoR, MNRAS 423, 2518 (DOI: 10.1111/j.1365-2966.2012.21065.x)

[211] Mertens, F.G., Ghosh, A., Koopmans, L.V.E (2018), Statistical 21-cm signal separation via Gaussian Process Regression analysis, MNRAS 478, 3640 (DOI: 10.1093/mnras/sty1207)

[212] Seldner, M., Peebles, P. J. E. (1981), Clustering of 4C radio sources, MNRAS 194, 251 (DOI: 10.1093/mnras/194.2.251)

[213] Shaver, P. A., Pierre, M. (1989), Large-scale anisotropy in the sky distribution of extragalactic radiosources, A&A 220, 35 (URL: https://ui.adsabs.harvard.edu/abs/1989A&A...220...35S)

[214] Condon, J. J., Cotton, W. D., Greisen, E. W., Yin, Q. F., Perley, R. A., Taylor, G. B., Broderick, J. J. (1998), The NRAO VLA Sky Survey, AJ 115, 1693 (DOI: 10.1086/300337)

[215] Condon, J. J. (1989), The 1.4 GHz Luminosity Function and Its Evolution, ApJ 338, 13 (DOI: 10.1086/167176)

[216] Afonso, J., Georgakakis, A., Almeida, C., Hopkins, A. M., Cram, L. E., Mobasher, B., Sullivan, M. (2005), The Phoenix Deep Survey: Spectroscopic Catalog, ApJ 624, 135–154) (DOI:10.1086/428923)

[217] Bonzini, M. Padovani, P., Mainieri, V., Kellermann, K. I., Miller, N., Rosati, P., Tozzi, P., Vattakunnel, S. (2013), The sub-mJy radio sky in the Extended Chandra Deep Field-South: source population, MNRAS 436, 3759, (DOI:10.1093/mnras/stt1879)

[218] Fanaroff, B. L., Riley, J. M. (1974), The morphology of extragalactic radio sources of high and low luminosity, MNRAS 167, 31 (DOI:10.1093/mnras/167.1.31P)

[219] Padovani, P. (2016), The faint radio sky: radio astronomy becomes mainstream A&AR 29, 1 (DOI: 10.1007/s00159-016-0098-6)

[220] Blake, C., Wall, J. (2002), Quantifying angular clustering in wide-area radio surveys, MNRAS 337, 993 (DOI: 10.1046/j.1365-8711.2002.05979.x)

[221] Lindsay, S. N. , Jarvis, M. J., Santos, M. G., Brown, M. J. I., Croom, S. M., Driver, S. P., Hopkins, A. M., Liske, J., Loveday, J., Norberg, P., Robotham, A. S. G. (2014), Galaxy and Mass Assembly: the evolution of bias in the radio source population to z∼1.5, MNRAS 440, 1527 (DOI: 10.1093/mnras/stu354)

[222] Hale, C. L., Jarvis, M J, Delvecchio, I, Hatfield, P W, Novak, M, Smolčić, V, Zamorani, G. (2018) The clustering and bias of radio-selected AGN and star-forming galaxies in the COSMOS field, MNRAS 474, 4133 (DOI: 10.1093/mnras/stx2954)

[223] Best, P. N., Heckman, T. M. (2012), On the fundamental dichotomy in the local radio-AGN population: accretion, evolution and host galaxy properties, MNRAS 421, 1569 (DOI: 10.1111/j.1365-2966.2012.20414.x)

[224] Ineson, J., Croston, J. H., Hardcastle, M. J., Kraft, R. P., Evans, D. A., Jarvis, M. (2015), The link between accretion mode and environment in radio-loud active galaxies, MNRAS 453, 2682, (DOI: 10.1093/mnras/stv1807)

[225] Hardcastle, M. J., Gürkan, G., van Weeren, R. J., Williams, W. L., Best, P. N., de Gasperin, F., Rafferty, D. A., Read, S. C., Sabater, J., Shimwell, T. W., Smith, D. J. B., Tasse, C., Bourne, N., Brienza, M., Brüggen, M., Brunetti, G., Chyży, K. T., Conway, J., Dunne, L., Eales, S. A., Maddox, S. J., Jarvis, M. J., Mahony, E. K., Morganti, R., Prandoni, I., Röttgering, H. J. A., Valiante, E., White, G. J. (2016), LOFAR/H-ATLAS: a deep low-frequency survey of the Herschel-ATLAS North Galactic Pole field, MNRAS 462, 1910 (DOI: 10.1093/mnras/stw1763)

[226] Calistro Rivera, G., Williams, W. L., Hardcastle, M. J., Duncan, K., Röttgering, H. J. A., Best, P. N., Brüggen, M., Chyży, K. T., Conselice, C. J., de Gasperin, F., Engels, D., Gürkan, G., Intema, H. T., Jarvis, M. J., Mahony, E. K., Miley, G. K., Morabito, L. K., Prandoni, I., Sabater, J., Smith,

D. J. B., Tasse, C., van der Werf, P. P., White, G. J. (2017), The LOFAR window on star-forming galaxies and AGNs - curved radio SEDs and IR-radio correlation at $0<z<2.5$, MNRAS 469, 3468, (DOI: 10.1093/mnras/stx1040)

[227] Williams, W. L., Calistro Rivera, G., Best, P.N., Hardcastle, M.J., Röttgering, H.J.A., Duncan, K.J., de Gasperin, F., Jarvis, M.J., Miley, G.K., Mahony, E.K., Morabito, L.K., Nisbet, D.M., Prandoni, I., Smith, D.J.B., Tasse, C., White, G.J. (2018), LOFAR-Boötes: properties of high- and low-excitation radio galaxies at $0.5 < z < 2.0$, MNRAS 475, 3429 (DOI: 10.1093/mnras/sty026)

[228] Press, W. H., Schechter, P. (1974), Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation, ApJ 187, 425 (DOI: 10.1086/152650)

[229] Lacey, C., Cole, S. (1993), Merger rates in hierarchical models of galaxy formation, MNRAS 262, 627, (DOI: 10.1093/mnras/262.3.627)

[230] Lacey, C., Cole, S. (1994), Merger Rates in Hierarchical Models of Galaxy Formation - Part Two - Comparison with N-Body Simulations, MNRAS 271, 676 (DOI: 10.1093/mnras/271.3.676)

[231] Sheth, R. K., Tormen, G. (1999), Large-scale bias and the peak background split, MNRAS 308, 119 (DOI: 10.1046/j.1365-8711.1999.02692.x)

[232] Peebles, P. J. E. (1980), The large-scale structure of the universe Princeton University Press

[233] Peacock, J. A., Cole, S., Norberg, P., Baugh, C.M., Bland-Hawthorn, J., Bridges, T., Cannon, R.D., Colless, M., Collins, C., Couch, W., Dalton, G., Deeley, K., De Propris, R., Driver, S.P., Efstathiou, G., Ellis, R.S., Frenk, C.S., Glazebrook, K., Jackson, C., Lahav, O., Lewis, I., Lumsden, S., Maddox, S., Percival, W.J., Peterson, B.A., Price, I., Sutherland,

W., Taylor, K. (2001), A measurement of the cosmological mass density from clustering in the 2dF Galaxy Redshift Survey, Nature 410, 169 (URL: https://ui.adsabs.harvard.edu/abs/2001Natur.410..169P)

[234] Percival, W. J., Baugh, C.M., Bland-Hawthorn, J., Bridges, T., Cannon, R., Cole, S., Colless, M., Collins, C., Couch, W., Dalton, G., De Propris, R., Driver, S.P., Efstathiou, G., Ellis, R.S., Frenk, C. S., Glazebrook, K., Jackson, C., Lahav, O., Lewis, I., Lumsden, S., Maddox, S., Moody, S., Norberg, P., Peacock, J.A., Peterson, B.A., Sutherland, W., Taylor, K. (2001), The 2dF Galaxy Redshift Survey: the power spectrum and the matter content of the Universe, MNRAS 327, 1297, (DOI: 10.1046/j.1365-8711.2001.04827.x)

[235] Norberg, P., Baugh, C.M., Hawkins, E., Maddox, S., Peacock, J.A., Cole, S., Frenk, C.S., Bland-Hawthorn, J., Bridges, T., Cannon, R., Colless, M., Collins, C., Couch, W., Dalton, G., De Propris, R., Driver, S.P., Efstathiou, G., Ellis, R.S., Glazebrook, K., Jackson, C., Lahav, O., Lewis, I., Lumsden, S., Madgwick, D., Peterson, B.A., Sutherland, W., Taylor, K. (2001), The 2dF Galaxy Redshift Survey: luminosity dependence of galaxy clustering, MNRAS 328, 64 (DOI: 10.1046/j.1365-8711.2001.04839.x)

[236] Eisenstein, D. J., Zehavi, I., Hogg, D.W., Scoccimarro, R., Blanton, M.R., Nichol, RC., Scranton, R, Seo, H., Tegmark, M., Zheng, Z., Anderson, S.F., Annis, J., Bahcall, N., Brinkmann, J., Burles, S., Castander, F.J., Connolly, A., Csabai, I., Doi, M., Fukugita, M., Frieman, J.A., Glazebrook, K., Gunn, J.E., Hendry, J.S., Hennessy, G., Ivezić, Z., Kent, S., Knapp, G.R., Lin, H., Loh, Y., Lupton, R.H., Margon, B., McKay, T.A., Meiksin, A., Munn, J.A., Pope, A., Richmond, M.W., Schlegel, D., Schneider, D.P., Shimasaku, K., Stoughton, C., Strauss, M.A., SubbaRao, M., Szalay, A.S., Szapudi, I., Tucker, D.L., Yanny, B., York, D.G. (2005) Detection of the Baryon Acoustic Peak in the Large-

Scale Correlation Function of SDSS Luminous Red Galaxies, ApJ 633, 560, (DOI: 10.1086/466512)

[237] Wang, Y., Brunner, R. J., Dolence, J. C. (2013), The SDSS galaxy angular two-point correlation function, MNRAS 432, 1961 (DOI: 10.1093/mnras/stt450)

[238] de Simoni, F., Sobreira, F., Carnero, A., Ross, A.J., Camacho, H.O., Rosenfeld, R., Lima, M., da Costa, L.A.N., Maia, M.A.G. (2013), Large-scale analysis of the SDSS-III DR8 photometric luminous galaxies angular correlation function, MNRAS 435, 3017, (DOI: 10.1093/mnras/stt1496)

[239] Shi, F., Yang, X., Wang, H., Zhang, Y., Mo, H. J., van den Bosch, F.C., Li, S., Liu, C., Lu, Yi, Tweed, D., Yang, L. (2016), Mapping the Real-space Distributions of Galaxies in SDSS DR7. I. Two-point Correlation Functions, ApJ 833, 241 (DOI: 10.3847/1538-4357/833/2/241)

[240] Carvalho, G. C., Bernui, A., Benetti, M., Carvalho, J. C., Alcaniz, J. S. (2016), Baryon acoustic oscillations from the SDSS DR10 galaxies angular correlation function, Ph.Rev.D 93 (DOI: 10.1103/PhysRevD.93.023530)

[241] Camacho, H. and DES Collaboration (2019), Dark Energy Survey Year 1 results: measurement of the galaxy angular power spectrum, MNRAS 487, 3870 (DOI: 10.1093/mnras/stz1514)

[242] Ross, N. P.,daÂngela, J., Shanks, T., Wake, D.A., Cannon, R.D., Edge, A. C., Nichol, R. C.,Outram, P. J., Colless, M., Couch, W.J., Croom, S.M., De Propris, R., Drinkwater, M.J., Eisenstein, D.J., Loveday, J., Pimbblet, K.A., Roseboom, I. G., Schneider, D.P., Sharp, R.G., Weilbacher, P. M. (2007), The 2dF-SDSS LRG and QSO Survey: the LRG 2-point correlation function and redshift-space distortions, MNRAS 381, 573 (DOI: 10.1111/j.1365-2966.2007.12289.x)

[243] Heinis, S., Budavári, T., Szalay, A.S., Arnouts, S., Aragón-Calvo, M.A., Wyder, T.K., Barlow, T.A., Foster, K., Peter, F.G., Martin, D. C., Morrissey, P., Neff, S. G., Schiminovich, D., Seibert, M., Bianchi, L., Donas, J., Heckman, T.M., Lee, Y., Madore, B.F., Milliard, B., Rich, R.M, Yi, S.K. (2009), Spatial Clustering from GALEX-SDSS Samples: Star Formation History and Large-Scale Clustering, ApJ 698, 1838 (DOI: 10.1088/0004-637X/698/2/1838)

[244] Salazar-Albornoz, S., Sánchez, A.G., Grieb, J. N., Crocce, M., Scoccimarro, R., Alam, S., Beutler, F., Brownstein, J.R., Chuang, C., Kitaura, F., Olmstead, M.D., Percival, W.J., Prada, F., Rodríguez-Torres, S., Samushia, L., Tinker, J., Thomas, D., Tojeiro, R., Wang, Y., Zhao, G. (2017), The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: angular clustering tomography and its cosmological implications, MNRAS 468, 2938, (DOI: 10.1093/mnras/stx633)

[245] Hao, C.-N., Kennicutt, R. C., Johnson, B. D., Calzetti, D., Dale, D. A., Moustakas, J.(2011), Dust-corrected Star Formation Rates of Galaxies. II. Combinations of Ultraviolet and Infrared Tracers, ApJ 741, 124 (DOI: 10.1088/0004-637X/741/2/124)

[246] Cucciati, O. Tresse, L., Ilbert, O., Le Fèvre, O., Garilli, B., Le Brun, V., Cassata, P., Franzetti, P., Maccagni, D., Scodeggio, M., Zucca, E., Zamorani, G., Bardelli, S., Bolzonella, M., Bielby, R. M., McCracken, H. J., Zanichelli, A., Vergani, D. (2012), The star formation rate density and dust attenuation evolution over 12 Gyr with the VVDS surveys, A&A 539, A31 (DOI: 10.1051/0004-6361/201118010)

[247] Singh, V., Beelen, A., Wadadekar, Y., Sirothia, S., Ishwara-Chandra, C. H., Basu, A., Omont, A., McAlpine, K., Ivison, R. J., Oliver, S., Farrah, D., Lacy, M. (2014), Multiwavelength characterization of faint ultra steep spectrum ra-

dio sources: A search for high-redshift radio galaxies, A&A 569, A52 (DOI: 10.1051/0004-6361/201423644)

[248] Jarvis, M., Taylor, R., Agudo, I., Allison, J. R., Deane, R. P., Frank, B., Gupta, N., Heywood, I., Maddox, N., McAlpine, K., Santos, M., Scaife, A. M. M., Vaccari, M., Zwart, J. T. L., Adams, E., Bacon, D. J., Baker, A. J., Bassett, B. A., Best, P. N., Beswick, R., Blyth, S., Brown, M. L., Bruggen, M., Cluver, M., Colafrancesco, S., Cotter, G., Cress, C., Davé, R., Ferrari, C., Hardcastle, M. J., Hale, C. L., Harrison, I., Hatfield, P. W., Klockner, H. R., Kolwa, S., Malefahlo, E., Marubini, T., Mauch, T., Moodley, K., Morganti, R., Norris, R. P., Peters, J. A., Prandoni, I., Prescott, M., Oliver, S., Oozeer, N., Rottgering, H. J. A., Seymour, N., Simpson, C., Smirnov, O., Smith, D. J. B. (2016), The MeerKAT International GHz Tiered Extragalactic Exploration (MIGHTEE) Survey, MeerKAT Science: On the Pathway to the SKA. (URL: https://ui.adsabs.harvard.edu/abs/2016mks..confE...6J)

[249] Saxena, A., Röttgering, H. J. A., Rigby, E. E. (2017), Modelling the luminosities and sizes of radio sources: radio luminosity function at z = 6, MNRAS 469, 4083 (DOI: 10.1093/mnras/stx1150)

[250] DeBoer, D. R., Gough, R. G., Bunton, J. D., Cornwell, T. J., Beresford, R. J., Johnston, S., Feain, I. J., Schinckel, A. E., Jackson, C. A., Kesteven, M. J., Chippendale, A., Hampson, G. A., O'Sullivan, J. D., Hay, S. G., Jacka, C. E., Sweetnam, T. W., Storey, M. C., Ball, L., Boyle, B. J (2009), Australian SKA Pathfinder: A High-Dynamic Range Wide-Field of View Survey Telescope, IEEE Proceedings 97, 1507 (DOI: 10.1109/JPROC.2009.2016516)

[251] van Haarlem, M. P. et al. (2013), LOFAR: The LOw-Frequency ARray, A&A 556, A2 (DOI: 10.1051/0004-6361/201220873)

[252] Bell, E. F. (2003), Estimating Star Formation Rates from Infrared and Radio Luminosities: The Origin of the Radio-Infrared Correlation, ApJ 586, 794 (DOI: 10.1086/367829)

[253] Jarvis, M. J., Smith, D. J. B., Bonfield, D. G., Hardcastle, M. J., Falder, J. T., Stevens, J. A., Ivison, R. J., Auld, R., Baes, M., Baldry, I. K., Bamford, S. P., Bourne, N., Buttiglione, S., Cava, A., Cooray, A., Dariush, A., de Zotti, G., Dunlop, J. S., Dunne, L., Dye, S., Eales, S., Fritz, J., Hill, D. T., Hopwood, R., Hughes, D. H., Ibar, E., Jones, D. H., Kelvin, L., Lawrence, A., Leeuw, L., Loveday, J., Maddox, S. J., Michałowski, M. J., Negrello, M., Norberg, P., Pohlen, M., Prescott, M., Rigby, E. E., Robotham, A., Rodighiero, G., Scott, D., Sharp, R., Temi, P., Thompson, M. A., van der Werf, P., van Kampen, E., Vlahakis, C., White, G. (2010), Herschel-ATLAS: the far-infrared-radio correlation at $z < 0.5$, MNRAS 409, 92 (DOI: 10.1111/j.1365-2966.2010.17772.x)

[254] Davies, L. J. M., Huynh, M. T., Hopkins, A. M., Seymour, N., Driver, S. P., Robotham, A. G. R., Baldry, I. K., Bland-Hawthorn, J., Bourne, N., Bremer, M. N., Brown, M. J. I., Brough, S., Cluver, M., Grootes, M. W., Jarvis, M., Loveday, J., Moffet, A., Owers, M., Phillipps, S., Sadler, E., Wang, L., Wilkins, S., Wright, A. (2017) Galaxy And Mass Assembly: the 1.4 GHz SFR indicator, SFR-M_* relation and predictions for ASKAP-GAMA, MNRAS 466, 2312 (DOI: 10.1093/mnras/stw3080)

[255] Delhaize, J., Smolci´c, V., Delvecchio, I., Novak, M., Sargent, M., Baran, N., Magnelli, B., Zamorani, G., Schinnerer, E., Murphy, E. J., Aravena, M., Berta, S., Bondi, M., Capak, P., Carilli, C., Ciliegi, P., Civano, F., Ilbert, O., Karim, A., Laigle, C., Le F`evre, O., Marchesi, S., McCracken, H. J., Salvato, M., Seymour, N., Tasca, L. (2017) The VLA-COSMOS 3 GHz Large Project: The

infrared-radio correlation of star-forming galaxies and AGN to z $\lesssim 6$, A&A 602, A4 (DOI: 10.1051/0004-6361/201629430)

[256] Gürkan, G., Hardcastle, M.J., Smith, D.J.B., Best, P.N., Bourne, N., Calistro-Rivera, G., Heald, G., Jarvis, M.J., Prandoni, I., Röttgering, H.J.A., Sabater, J., Shimwell, T., Tasse, C., Williams, W.L. (2018) LOFAR/H-ATLAS: the low-frequency radio luminosity-star formation rate relation, MNRAS 475, 3010 (DOI: 10.1093/mnras/sty016)

[257] Magliocchetti, M., Popesso, P., Brusa, M., Salvato, M., Laigle, C., McCracken, H. J., Ilbert, O. (2017) The clustering properties of radio-selected AGN and star-forming galaxies up to redshifts z ~ 3, MNRAS 464, 3271 (DOI: 10.1093/mnras/stw2541)

[258] Chakraborty, A., Dutta, P., Datta, A., Roy, N. (2020), The study of the angular and spatial distribution of radio-selected AGNs and star-forming galaxies in the ELAIS N1 field, MNRAS 494, 3392 (DOI: 10.1093/mnras/staa945)

[259] Camera, S., Santos, M.G., Bacon, D.J., Jarvis, M.J., McAlpine, K., Norris, R.P., Raccanelli, A., Röttgering, Huub (2012), Impact of redshift information on cosmological applications with next-generation radio surveys, MNRAS 427, 2079 (DOI: 10.1111/j.1365-2966.2012.22073.x)

[260] Raccanelli, A. Zhao, G., Bacon, D.J., Jarvis, M.J., Percival, W.J., Norris, R.P., Röttgering, H., Abdalla, F.B., Cress, C.M., Kubwimana, J., Lindsay, S., Nichol, R.C., Santos, M.G., Schwarz, D.J. (2012), Cosmological measurements with forthcoming radio continuum surveys, MNRAS 424, 801 (DOI: 10.1111/j.1365-2966.2012.20634.x)

[261] Planck Collaboration (2014), Planck 2013 results. XVII. Gravitational lensing by large-scale structure, A&A 571, A17 (DOI: 10.1051/0004-6361/201321543)

[262] Allison, R. and 27 colleagues 2015. The Atacama Cosmology Telescope: measuring radio galaxy bias through cross-correlation with lensing, MNRAS 451, 849 (DOI: 10.1093/mnras/stv991)

[263] Donoso, E., Yan, L., Stern, D., Assef, R. J. (2014), The Angular Clustering of WISE-selected Active Galactic Nuclei: Different Halos for Obscured and Unobscured Active Galactic Nuclei, ApJ 789, 44 (DOI: 10.1088/0004-637X/789/1/44)

[264] Magnelli, B., Ivison, R. J., Lutz, D., Valtchanov, I., Farrah, D., Berta, S., Bertoldi, F., Bock, J., Cooray, A., Ibar, E., Karim, A., Le Floc´h, E., Nordon, R., Oliver, S. J., Page, M., Popesso, P., Pozzi, F., Rigopoulou, D., Riguccini, L., Rodighiero, G., Rosario, D., Roseboom, I., Wang, L., Wuyts, S. (2015), The far-infrared/radio correlation and radio spectral index of galaxies in the SFR-M_* plane up to z~2 A&A 573, A45 (DOI: 10.1051/0004-6361/201424937)

[265] Delvecchio, I. Daddi, E., Sargent, M. T., Jarvis, M. J., Elbaz, D., Jin, S., Liu, D., Whittam, I. H., Algera, H., Carraro, R., D´Eugenio, C., Delhaize, J., Kalita, B. S., Leslie, S., Molnár, D. Cs., Novak, M., Prandoni, I., Smolci´c, V., Ao, Y., Aravena, M., Bournaud, F., Collier, J. D., Randriamampandry, S. M., Randriamanakoto, Z., Rodighiero, G., Schober, J., White, S. V., Zamorani, G. (2021), The infrared-radio correlation of star-forming galaxies is strongly $M_\star$-dependent but nearly redshift-invariant since $z \sim 4$, A&A 647, A123 (DOI: 10.1051/0004-6361/202039647)

[266] Bonato, M. and 16 colleagues 2021. The LOFAR Two-metre Sky Survey Deep Fields: A new analysis of low-frequency radio luminosity as a star-formation tracer in the Lockman Hole region, A&A 656, A48 (DOI: 10.1051/0004-6361/202141286)

[267] Kaiser, N. (1984), On the spatial correlations of Abell clusters, ApJ 284, L9 (DOI: 10.1086/184341)

[268] Bardeen, J. M., Bond, J. R., Kaiser, N., Szalay, A. S. (1986), The Statistics of Peaks of Gaussian Random Fields, ApJ 304, 15 (DOI: 10.1086/164143)

[269] Desjacques, V., Jeong, D., Schmidt, F.(2018), Large-scale galaxy bias, Phy.Reps. 733, 1 (DOI: 10.1016/j.physrep.2017.12.002)

[270] Blake, C. A., Abdalla, F. B., Bridle, S. L., Rawlings, S. (2004), Cosmology with the SKA, NewAR 48, 1063 (DOI: 10.1016/j.newar.2004.09.045)

[271] Carilli, C. L., Rawlings, S. (2004), Motivation, key science projects, standards and assumptions, NewAR 48, 979 (DOI: 10.1016/j.newar.2004.09.001)

[272] Seljak, U. (2009), Extracting Primordial Non-Gaussianity without Cosmic Variance, PRL 102, 21302 (DOI: 10.1103/PhysRevLett.102.021302)

[273] Raccanelli, A., Doré, O., Bacon, D.J., Maartens, R., Santos, M.G., Camera, S., Davis, T.M., Drinkwater, M.J., Jarvis, M., Norris, R., Parkinson, D. (2015), Probing primordial non-Gaussianity via iSW measurements with SKA continuum surveys, JCAP 2015, 42 (DOI: 10.1088/1475-7516/2015/01/042)

[274] Abdalla, F. B., Bull, P., Camera, S., Benoit-Lévy, A., Joachimi, B., Kirk, D., Kloeckner, H. R., Maartens, R., Raccanelli, A., Santos, M. G., Zhao, G. B (2015), Cosmology from HI galaxy surveys with the SKA, Advancing Astrophysics with the Square Kilometre Array (AASKA14)

[275] Cress, C. M., Helfand, D. J., Becker, R. H., Gregg, M. D., White, R. L. (1996), The Angular Two-Point Correlation Function for the FIRST Radio Survey, ApJ 473, 7 (DOI: 10.1086/178122)

[276] Overzier, R. A., Röttgering, H. J. A., Rengelink, R. B., Wilman, R. J. (2003), The spatial clustering of radio sources in NVSS and FIRST: implications for galaxy clustering evolution, A&A 405, 53, (DOI: 10.1051/0004-6361:20030527)

[277] Rana, S., Bagla, J. S. (2019), Angular clustering of point sources at 150 MHz in the TGSS survey, MNRAS 485, 5891, (DOI: 10.1093/mnras/stz831)

[278] Siewert, T. M., Hale, C., Bhardwaj, N., Biermann, M., Bacon, D. J., Jarvis, M., Röttgering, H. J .A., Schwarz, D. J., Shimwell, T., Best, P. N., Duncan, K. J., Hardcastle, M. J., Sabater, J., Tasse, C., White, G. J., Williams, W. L. (2020), One- and two-point source statistics from the LOFAR Two-metre Sky Survey first data release, A&A 643. (DOI: 10.1051/0004-6361/201936592)

[279] Bonato, M., Prandoni, I., De Zotti, G., Brienza, M., Morganti, R., Vaccari, M. (2021), New constraints on the 1.4 GHz source number counts and luminosity functions in the Lockman Hole field, MNRAS 500, 22 (DOI: 10.1093/mnras/staa3218)

[280] Tiwari, P., Ghosh, S., Jain, P. (2019), The Galaxy Power Spectrum from TGSS ADR1 and the Effect of Flux Calibration Systematics, ApJ 887, 175 (DOI: 10.3847/1538-4357/ab54c8)

[281] Norris, R. P. et al. (2013), Radio Continuum Surveys with Square Kilometre Array Pathfinders, PASA 30, e020 (DOI: 10.1017/pas.2012.020)

[282] Dunlop, J. S., Peacock, J. A. (1990), The redshift cut-off in the luminosity function of radio galaxies and quasars, MNRAS 247, 19 (URL: https://ui.adsabs.harvard.edu/abs/1990MNRAS.247...19D)

[283] Shimwell, T. W. et al. (2019), The LOFAR Two-metre Sky Survey. II. First data release, A&A 622, A1 (DOI: 10.1051/0004-6361/201833559)

[284] Tasse, C., Shimwell, T., Hardcastle, M. J., O´Sullivan, S. P., van Weeren, R., Best, P. N., Bester, L., Hugo, B., Smirnov, O., Sabater, J., Calistro-Rivera, G., de Gasperin, F., Morabito, L. K., Röttgering, H., Williams, W. L., Bonato, M., Bondi, M., Botteon, A., Brüggen, M., Brunetti, G., Chyzy, K. T., Garrett, M. A., Gürkan, G., Jarvis, M. J., Kondapally, R., Mandal, S., Prandoni, I., Repetti, A., Retana-Montenegro, E., Schwarz, D. J., Shulevski, A., Wiaux, Y. (2021), The LOFAR Two-meter Sky Survey: Deep Fields Data Release 1. I. Direction-dependent calibration and imaging, A&A 648, A1 (DOI: 10.1051/0004-6361/202038804)

[285] Shirley, R., Roehlly, Y., Hurley, P.D., Buat, V., Campos Varillas, M.C., Duivenvoorden, S., Duncan, K.J., Efstathiou, A., Farrah, D., González Solares, E., Malek, K., Marchetti, L., McCheyne, I., Papadopoulos, A., Pons, E., Scipioni, R., Vaccari, M., Oliver, S. (2019), HELP: a catalogue of 170 million objects, selected at 0.36-4.5 μm, from 1270 deg^2 of prime extragalactic fields, MNRAS 490, 634, (DOI: 10.1093/mnras/stz2509)

[286] York, D. G. et al. (2000), The Sloan Digital Sky Survey: Technical Summary, AJ 120, 1579 (DOI: 10.1086/301513)

[287] Eisenstein, D. J. et al. (2011), SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way, and Extra-Solar Planetary Systems, AJ 142, 72 (DOI: 10.1088/0004-6256/142/3/72)

[288] Blanton, M. R. et al. (2017), Sloan Digital Sky Survey IV: Mapping the Milky Way, Nearby Galaxies, and the Distant Universe, AJ 154, 28 (DOI: 10.3847/1538-3881/aa7567)

[289] Alam, S. et al. (2021), Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Cosmological implications from two decades of spectro-

scopic surveys at the Apache Point Observatory, Phy.Rev.D 103,083533 (DOI: 10.1103/PhysRevD.103.083533)

[290] Duncan, K. J., Brown, M. J. I., Williams, W.L., Best, P.N., Buat, V., Burgarella, D., Jarvis, M.J., Małek, K.a, Oliver, S. J., Röttgering, H.J. A., Smith, D.J. B. (2018), Photometric redshifts for the next generation of deep radio continuum surveys - I. Template fitting, MNRAS 473, 2655 (DOI: 10.1093/mnras/stx2536)

[291] Duncan, K. J., Jarvis, M. J., Brown, M. J. I., Röttgering, H. J. A. (2018), Photometric redshifts for the next generation of deep radio continuum surveys - II. Gaussian processes and hybrid estimates, MNRAS 477, 5177 (DOI: 10.1093/mnras/sty940)

[292] Rowan-Robinson, M., Babbedge, T., Oliver, S., Trichas, M., Berta, S., Lonsdale, C., Smith, G., Shupe, D., Surace, J., Arnouts, S., Ilbert, O., Le Févre, O., Afonso-Luis, A., Perez-Fournon, I., Hatziminaoglou, E., Polletta, M., Farrah, D., Vaccari, M. (2008), Photometric redshifts in the SWIRE Survey, MNRAS 386, 697, (DOI: 10.1111/j.1365-2966.2008.13109.x)

[293] Rowan-Robinson, M., Gonzalez-Solares, E., Vaccari, M., Marchetti, L. (2013), Revised SWIRE photometric redshifts, MNRAS 428, 1958 (DOI: 10.1093/mnras/sts163)

[294] Lewis, J. R., Bunclark, P. S., Irwin, M. J., McMahon, R. G., Walton, N. A. (2000), The Wide Field Survey on the Isaac Newton Telescope, Astronomical Data Analysis Software and Systems IX 216, 415 (URL: https://ui.adsabs.harvard.edu/abs/2000ASPC..216..415L)

[295] Hildebrandt, H., Choi, A., Heymans, C., Blake, C., Erben, T., Miller, L., Nakajima, R., van Waerbeke, L., Viola, M., Buddendiek, A., Harnois-Déraps,

J., Hojjati, A., Joachimi, B., Joudaki, S., Kitching, T. D., Wolf, C., Gwyn, S., Johnson, N., Kuijken, K., Sheikhbahaee, Z., Tudorica, A., Yee, H. K. C. (2016), RCSLenS: The Red Cluster Sequence Lensing Survey, MNRAS 463, 635 (DOI: 10.1093/mnras/stw2013)

[296] Chambers, K.C. et al. (2019), The Pan-STARRS1 Surveys, arXiv eprint : 1612.05560 (URL: https://arxiv.org/abs/1612.05560)

[297] Muzzin, A., Wilson, G., SpARCS Collaboration (2007), The Spitzer Adaptation of the Red-Sequence Cluster Survey, American Astronomical Society Meeting Abstracts 211, 78.05, (URL: https://ui.adsabs.harvard.edu/abs/2007AAS...211.7805M)

[298] Lawrence, A., Warren, S. J., Almaini, O., Edge, A. C., Hambly, N. C., Jameson, R. F., Lucas, P., Casali, M., Adamson, A., Dye, S., Emerson, J. P., Foucaud, S., Hewett, P., Hirst, P., Hodgkin, S. T., Irwin, M. J., Lodieu, N., McMahon, R. G., Simpson, C., Smail, I., Mortlock, D., Folger, M. (2007), The UKIRT Infrared Deep Sky Survey (UKIDSS), MNRAS 379, 1599 (DOI: 10.1111/j.1365-2966.2007.12040.x)

[299] Mauduit, J.-C. et al. (2012), The Spitzer Extragalactic Representative Volume Survey (SERVS): Survey Definition and Goals, PASP 124, 714 (DOI: 10.1086/666945)

[300] Dye, S., Lawrence, A., Read, M. A., Fan, X., Kerr, T., Varricatt, W., Furnell, K. E., Edge, A. C., Irwin, M., Hambly, N., Lucas, P., Almaini, O., Chambers, K., Green, R., Hewett, P., Liu, M. C., McGreer, I., Best, W., Zhang, Z., Sutorius, E., Froebrich, D., Magnier, E., Hasinger, G., Lederer, S. M., Bold, M., Tedds, J. A. (2017), The UKIRT Hemisphere Survey: definition and J-band data release, MNRAS 473, 5113 (DOI: 10.1093/mnras/stx2622)

[301] Magliocchetti, M., Maddox, S. J., Lahav, O., Wall, J. V. (1998), Variance and skewness in the FIRST survey, MNRAS 300, 257 (DOI: 10.1046/j.1365-8711.1998.01904.x)

[302] Williams, W. L. et al. (2019), The LOFAR Two-metre Sky Survey. III. First data release: Optical/infrared identifications and value-added catalogue, A&A 622, A2 (DOI: 10.1051/0004-6361/201833564)

[303] Galvin, T. J., Huynh, M.T., Norris, R.P., Wang, X.R., Hopkins, E., Polsterer, K., Ralph, N.O., O'Brien, A.N., Heald, G.H. (2020), Cataloguing the radio-sky with unsupervised machine learning: a new approach for the SKA era, MNRAS 497, 2730 (DOI: 10.1093/mnras/staa1890)

[304] Blake, C., Wall, J. (2002), Measurement of the angular correlation function of radio galaxies from the NRAO VLA Sky Survey, MNRAS 329, L37 (DOI: 10.1046/j.1365-8711.2002.05163.x)

[305] Oort, M. J. A.(1987), Radio galaxies at very low flux levels, Ph.D. Thesis (URL: https://ui.adsabs.harvard.edu/abs/1987PhDT........40O)

[306] Huynh, M. T., Jackson, C. A., Norris, R. P., Prandoni, I. (2005), Radio Observations of the Hubble Deep Field-South Region. II. The 1.4 GHz Catalog and Source Counts, AJ 130, 1373 (DOI: 10.1086/432873)

[307] Magliocchetti, M., Lutz, D., Rosario, D., Berta, S., Le Floc'h, E., Magnelli, B., Pozzi, F., Riguccini, L., Santini, P. (2014), The PEP survey: infrared properties of radio-selected AGN, MNRAS 442, 682 (DOI: 10.1093/mnras/stu863)

[308] Magliocchetti, M., Lutz, D., Santini, P., Salvato, M., Popesso, P., Berta, S., Pozzi, F. (2016), The PEP survey: evidence for intense star-forming activity in the majority of radio-selected AGN at z $\gtrsim$ 1, MNRAS 456, 431 (DOI: 10.1093/mnras/stv2645)

[309] McAlpine, K., Jarvis, M. J., Bonfield, D. G. (2013), Evolution of faint radio sources in the VIDEO-XMM3 field, MNRAS 436, 1084, (DOI: 10.1093/mnras/stt1638)

[310] Smolčić, V., Novak, M., Delvecchio, I., Ceraj, L., Bondi, M., Delhaize, J., Marchesi, S., Murphy, E., Schinnerer, E., Vardoulaki, E., Zamorani, G. (2017), The VLA-COSMOS 3 GHz Large Project: Cosmic evolution of radio AGN and implications for radio-mode feedback since z $\sim$5, A&A 602, A6 (DOI: 10.1051/0004-6361/201730685)

[311] Ocran, E. F., Taylor, A.R., Vaccari, M., Ishwara-Chandra, C.H., Prandoni, I., Prescott, M., Mancuso, C. (2021), The evolution of the low-frequency radio AGN population to z$\cong$ 1.5 in the ELAIS N1 field, MNRAS 500, 4685 (DOI: 10.1093/mnras/staa3538)

[312] Prescott, M., Mauch, T., Jarvis, M. J., McAlpine, K., Smith, D. J. B., Fine, S., Johnston, R., Hardcastle, M. J., Baldry, I. K., Brough, S., Brown, M. J. I., Bremer, M. N., Driver, S. P., Hopkins, A. M., Kelvin, L. S., Loveday, J., Norberg, P., Obreschkow, D., Sadler, E. M. (2016) Galaxy And Mass Assembly (GAMA): the 325 MHz radio luminosity function of AGN and star-forming galaxies, MNRAS 457, 730 (DOI: 10.1093/mnras/stv3020)

[313] Ocran, E. F., Taylor, A.R., Vaccari, M., Ishwara-Chandra, C.H., Prandoni, I., Prescott, M., Mancuso, C. (2020), Cosmic evolution of star-forming galaxies to z $\cong$1.8 in the faint low-frequency radio source population, MNRAS 491, 5911 (DOI: 10.1093/mnras/stz3401)

[314] Szokoly, G. P., Bergeron, J., Hasinger, G., Lehmann, I., Kewley, L., Mainieri, V., Nonino, M., Rosati, P., Giacconi, R., Gilli, R., Gilmozzi, R., Norman, C., Romaniello, M., Schreier, E., Tozzi, P., Wang, J. X., Zheng, W., Zirm, A.

(2004), The Chandra Deep Field-South: Optical Spectroscopy. I., ApJS 155, 271 (DOI: 10.1086/424707)

[315] Donley, J. L. Koekemoer, M. Brusa, P. Capak, C. N. Cardamone, F. Civano, O. Ilbert, C. D. Impey, J. S. Kartaltepe, T. Miyaji, M. Salvato, D. B. Sanders, J. R. Trump, G. Zamorani (2012), Identifying Luminous Active Galactic Nuclei in Deep Surveys: Revised IRAC Selection Criteria, ApJ 748, 142 (DOI: 10.1088/0004-637X/748/2/142)

[316] Magliocchetti, M. Maddox, S.J., Jackson, C.A., Bland-Hawthorn, J., Bridges, T., Cannon, R., Cole, S., Colless, M., Collins, C., Couch, W., Dalton, G., De Propris, R., Driver, S.P., Efstathiou, G., Ellis, R.S., Frenk, C.S., Glazebrook, K., Lahav, O., Lewis, I., Lumsden, S., Peacock, J.A., Peterson, B.A., Sutherland, W., Taylor, K. (2002), The 2dF Galaxy Redshift Survey: the population of nearby radio galaxies at the 1-mJy level, MNRAS 333, 100 (DOI: 10.1046/j.1365-8711.2002.05386.x)

[317] Kerscher, M., Szapudi, I., Szalay, A. S (2000), A Comparison of Estimators for the Two-Point Correlation Function, ApJ 535, L13 (DOI: 10.1086/312702)

[318] Landy, S. D., Szalay, A. S. (1993), Bias and Variance of Angular Correlation Functions, ApJ 412, 64 (DOI: 10.1086/172900)

[319] Hamilton, A. J. S (1993), Toward Better Ways to Measure the Galaxy Correlation Function, ApJ 417, 19 (DOI: 10.1086/173288)

[320] Ling, E. N., Frenk, C. S., Barrow, J. D. (1986), Uncertainties in the cluster-cluster correlation function, MNRAS 223, 21 (DOI: 10.1093/mnras/223.1.21P)

[321] Jarvis, M., Bernstein, G., Jain, B. (2004), The skewness of the aperture mass statistic, MNRAS 352, 338 (DOI: 10.1111/j.1365-2966.2004.07926.x)

[322] Limber, D. N. (1953), The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field, ApJ 117, 134 (DOI: 10.1086/145672)

[323] Peacock, J. A., Smith, R. E (2000), Halo occupation numbers and galaxy bias, MNRAS 318, 1144 (DOI: 10.1046/j.1365-8711.2000.03779.x)

[324] Eisenstein, D. J., Hu, W., Tegmark, M. (1999), Cosmic Complementarity: Joint Parameter Estimation from Cosmic Microwave Background Experiments and Redshift Surveys, ApJ 518, 2 (DOI: 10.1086/307261)

[325] Magliocchetti, M., Maddox, S.J., Hawkins, E., Peacock, J.A., Bland-Hawthorn, J., Bridges, T., Cannon, R., Cole, S., Colless, M., Collins, C., Couch, W., Dalton, G., Propris, R. de, Driver, S.P., Efstathiou, G., Ellis, R.S., S. Frenk, C., Glazebrook, K., Jackson, C.A., Jones, B., Lahav, O., Lewis, I., Lumsden, S., Norberg, P., Peterson, B.A., Sutherland, W., Taylor, K. (2004), The 2dF galaxy redshift survey: clustering properties of radio galaxies, MNRAS 350, 1485 (DOI: 10.1111/j.1365-2966.2004.07751.x)

[326] Gilli, R., Daddi, E., Zamorani, G., Tozzi, P., Borgani, S., Bergeron, J., Giacconi, R., Hasinger, G., Mainieri, V., Norman, C., Rosati, P., Szokoly, G., Zheng, W. (2005), The spatial clustering of X-ray selected AGN and galaxies in the Chandra Deep Field South and North, A&A 430, 811 (DOI: 10.1051/0004-6361:20041375)

[327] Gilli, R. et al. (2009), The spatial clustering of X-ray selected AGN in the XMM-COSMOS field, A&A 494, 33 (DOI: 10.1051/0004-6361:200810821)

[328] Gilli, R., Daddi, E., Chary, R., Dickinson, M., Elbaz, D., Giavalisco, M., Kitzbichler, M., Stern, D., Vanzella, E. (2007), The spatial clustering of mid-IR selected star forming galaxies at z ~1 in the GOODS fields, A&A 475, 83 (DOI: 10.1051/0004-6361:20077506)

[329] Dolley, T., Brown M.J.I., Weiner B.J., Brodwin M., Kochanek C. S., Pimbblet K.A., Palamara D.P., Jannuzi B.T., Dey A., Atlee D.W. , Beare R. (2014), The Clustering and Halo Masses of Star-forming Galaxies at z < z 1, ApJ 797, 125 (DOI: 10.1088/0004-637X/797/2/125)

[330] Starikova, S., Berta S., Franceschini A., Marchetti L., Rodighiero G., Vaccari M., Vikhlinin A. (2012), Clustering of Star-forming Galaxies Detected in Mid-infrared with the Spitzer Wide-area Survey, ApJ 751, 126 (DOI: 10.1088/0004-637X/751/2/126)

[331] Magliocchetti, M., Popesso, P., Rosario, D., Lutz, D., Aussel, H., Berta, S., Altieri, B., Andreani, P., Cepa, J., Castañeda, H., Cimatti, A., Elbaz, D., Genzel, R., Grazian, A., Gruppioni, C., Ilbert, O., Le Floc'h, E., Magnelli, B., Maiolino, R., Nordon, R., Poglitsch, A., Pozzi, F., Riguccini, L., Rodighiero, G., Sanchez-Portal, M., Santini, P., Förster Schreiber, N. M., Sturm, E., Tacconi, L., Valtchanov, I. (2013), The Herschel-PEP survey: evidence for downsizing in the hosts of dusty star-forming systems, MNRAS 433, 127 (DOI: 10.1093/mnras/stt708)

[332] Jarvis, M. J., Rawlings, S., Eales, S., Blundell, K.M., Bunker, A.J., Croft, S., McLure, R.J., Willott, C.J. (2001), A sample of 6C radio sources designed to find objects at redshift z>4 - III. Imaging and the radio galaxy K-z relation, MNRAS 326, 1585 (DOI: 10.1111/j.1365-2966.2001.04730.x)

[333] Mendez, A. J. and Coil, A.L. and Aird, J. and Skibba, R.A. and Diamond-Stanic, A.M., and Moustakas, J. and Blanton, M.R. and Cool, R.J. and Eisenstein D.J., and Wong K.C., and Zhu G. (2016), PRIMUS + DEEP2: Clustering of X-Ray, Radio, and IR-AGNs at z~0.7, ApJ 821, 55 (DOI: 10.3847/0004-637X/821/1/55)

[334] Sabater, J., Best, P. N., Tasse, C., Hardcastle, M. J., Shimwell, T. W., Nisbet, D., Jelic, V., Callingham, J. R., Röttgering, H. J. A., Bonato, M., Bondi, M., Ciardi, B., Cochrane, R. K., Jarvis, M. J., Kondapally, R., Koopmans, L. V. E., O´Sullivan, S. P., Prandoni, I., Schwarz, D. J., Smith, D. J. B., Wang, L., Williams, W. L., Zaroubi, S. (2021), The LOFAR Two-meter Sky Survey: Deep Fields Data Release 1. II. The ELAIS-N1 LOFAR deep field, A&A 648, A2 (DOI: 10.1051/0004-6361/202038828)

[335] Norris, R. P. et al. (2011), EMU: Evolutionary Map of the Universe, PASA 28, 215 (DOI: 10.1071/AS11021)

[336] Norris, R. P. et al. (2021), The Evolutionary Map of the Universe pilot survey, PASA 38, e046 (DOI: 10.1017/pasa.2021.42)

[337] Dayal, P., Ferrara, A. (2018), Early galaxy formation and its large-scale effects, Physics Reports 780, 1 (DOI: 10.1016/j.physrep.2018.10.002)

[338] Bowman, J. D., Rogers, A. E. E., Monsalve, R. A., Mozdzen, T. J., Mahesh, N. (2018.), An absorption profile centred at 78 megahertz in the sky-averaged spectrum, Nature 555, 67 (DOI: 10.1038/nature25792)

[339] Sokolowski, M., Tremblay, S.E., Wayth, R.B., Tingay, S.J., Clarke, N., Roberts, P., Waterson, M., Ekers, R.D., Hall, P., Lewis, M., Mossammaparast, M., Padhi, S., Schlagenhaufer, F., Sutinjo, A., Tickner, J. (2015), BIGHORNS - Broadband Instrument for Global HydrOgen ReioNisation Signal, PASA 32, e004 (DOI: 10.1017/pasa.2015.3)

[340] Voytek, T. C., Natarajan, A., Jáuregui García, J. M., Peterson, J. B., López-Cruz, O. (2014), Probing the Dark Ages at z ~20: The SCI-HI 21 cm All-sky Spectrum Experiment, ApJ 782, L9 (DOI: 10.1088/2041-8205/782/1/L9)

[341] Singh, S., Subrahmanyan, R. (2019), The Redshifted 21 cm Signal in the EDGES Low-band Spectrum, ApJ 880, 26 (DOI: 10.3847/1538-4357/ab2879)

[342] Price, D. C., Greenhill, L. J., Fialkov, A., Bernardi, G., Garsden, H., Barsdell, B. R., Kocz, J., Anderson, M. M., Bourke, S. A., Craig, J., Dexter, M. R., Dowell, J., Eastwood, M. W., Eftekhari, T., Ellingson, S. W., Hallinan, G., Hartman, J. M., Kimberk, R., Lazio, T. Joseph W., Leiker, S., MacMahon, D., Monroe, R., Schinzel, F., Taylor, G. B., Tong, E., Werthimer, D., Woody, D. P. (2018), Design and characterization of the Large-aperture Experiment to Detect the Dark Age (LEDA) radiometer systems, Monthly Notices of the Royal Astronomical Society 478, 4193 (DOI: 10.1093/mnras/sty1244)

[343] Zahn, O., Mesinger, A., McQuinn, M., Trac, H., Cen, R., Hernquist, L. E. (2011), Comparison of reionization models: radiative transfer simulations and approximate, seminumeric models, MNRAS 414, 727 (DOI: 10.1111/j.1365-2966.2011.18439.x)

[344] de Lera Acedo, E., Trott, C.M., Wayth, R.B., Fagnoni, N., Bernardi, G., Wakley, B., Koopmans, L.V.E., Faulkner, A.J., bij de Vaate, J.G. (2017), Spectral performance of SKA Log-periodic Antennas I: mitigating spectral artefacts in SKA1-LOW 21 cm cosmology experiments, MNRAS 469, 2662 (DOI: 10.1093/mnras/stx904)

[345] Trott, C. M., de Lera Acedo, E., Wayth, R.B., Fagnoni, N., Sutinjo, A.T., Wakley, B., Punzalan, C.I.B. (2017), Spectral performance of Square Kilometre Array Antennas - II. Calibration performance, MNRAS 470, 455, (DOI: 10.1093/mnras/stx1224)

[346] Joseph, R. C., Trott, C. M., Wayth, R. B. (2018), The Bias and Uncertainty of Redundant and Sky-based Calibration Under Realistic Sky and Telescope Conditions, ApJ 156, 285 (DOI: 10.3847/1538-3881/aaec0b)

[347] Li, W. et al. (2018), Comparing Redundant and Sky-model-based Interferometric Calibration: A First Look with Phase II of the MWA, ApJ 863, 170 (DOI: 10.3847/1538-4357/aad3c3)

[348] Offringa, A. R. et al. (2015), The Low-Frequency Environment of the Murchison Widefield Array: Radio-Frequency Interference Analysis and Mitigation, PASA 32, e008 (DOI: 10.1017/pasa.2015.7)

[349] Barry, N., Hazelton, B., Sullivan, I., Morales, M. F., Pober, J. C. (2016), Calibration requirements for detecting the 21 cm epoch of reionization power spectrum and implications for the SKA, MNRAS 461, 3135 (DOI: 10.1093/mnras/stw1380)

[350] Trott, C. M., Wayth, R. B. (2016), Spectral Calibration Requirements of Radio Interferometers for Epoch of Reionisation Science with the SKA, PASA 33, e019 (DOI: 10.1017/pasa.2016.18)

[351] Ewall-Wice, A., Dillon, J. S., Liu, A., Hewitt, J. (2017), The impact of modelling errors on interferometer calibration for 21 cm power spectra, MNRAS 470, 1849 (DOI: 10.1093/mnras/stx1221)

[352] Dillon, J. S. and Kohn, S.A. and Parsons, A.R. and Aguirre, J.E. and Ali, Z.S. and Bernardi, G. and Kern, N.S. and Li, W. and Liu, A. and Nunhokee, C.D. and Pober, J.C. (2018), Polarized redundant-baseline calibration for 21 cm cosmology without adding spectral structure, MNRAS 477, 5670 (DOI: 10.1093/mnras/sty1060)

[353] Kern, N. S., Parsons, A. R., Dillon, J. S., Lanman, A. E., Fagnoni, N., de Lera Acedo, E. (2019), Mitigating Internal Instrument Coupling for 21 cm Cosmology. I. Temporal and Spectral Modeling in Simulations, ApJ 884, (DOI: 10.3847/1538-4357/ab3e73)

[354] Kumar, J., Dutta, P., Roy, N. (2020), Calibration requirements for epoch of reionization 21-cm signal observations - I. Effect of time-correlated gains, MNRAS 495, 3683 (DOI: 10.1093/mnras/staa1371)

[355] Chapman, E., Zaroubi, S., Abdalla, F.B., Dulwich, F., Jelić, V.,, Mort, B. (2016), The effect of foreground mitigation strategy on EoR window recovery, MNRAS 458, 2928 (DOI: 10.1093/mnras/stw161)

[356] Vedantham, H., Udaya Shankar, N., Subrahmanyan, R. (2012), Imaging the Epoch of Reionization: Limitations from Foreground Confusion and Imaging Algorithms, ApJ 745, 176 (DOI: 10.1088/0004-637X/745/2/176)

[357] Thyagarajan, N. et al. (2015), Foregrounds in Wide-field Redshifted 21 cm Power Spectra, ApJ 804, 14 (DOI: 10.1088/0004-637X/804/1/14)

[358] Thyagarajan, N. et al. (2015), Confirmation of Wide-field Signatures in Redshifted 21 cm Power Spectra, ApJ 807, L28 (DOI: 10.1088/2041-8205/807/2/L28)

[359] Hothi, I., Chapman, E., Pritchard, J.R., Mertens, F.G., Koopmans, L.V.E, Ciardi, B., Gehlot, B.K., Ghara, R., Ghosh, A., Giri, S.K., Iliev, I.T., Jelić, V., Zaroubi, S. (2021), Comparing foreground removal techniques for recovery of the LOFAR-EoR 21 cm power spectrum, MNRAS 500, 2264, (DOI: 10.1093/mnras/staa3446)

[360] Trott, C. M., Jordan, C. H., Murray, S. G., Pindor, B., Mitchell, D. A., Wayth, R. B., Line, J., McKinley, B., Beardsley, A., Bowman, J., Briggs, F., Hazelton, B. J., Hewitt, J., Jacobs, D., Morales, M. F., Pober, J. C., Sethi, S., Shankar, U., Subrahmanyan, R., Tegmark, M., Tingay, S. J., Webster, R. L., Wyithe, J. S. B. (2018), Assessment of Ionospheric Activity Tolerances for Epoch of

Reionization Science with the Murchison Widefield Array, ApJ 867, 15 (DOI: 10.3847/1538-4357/aae314)

[361] Jordan, C. H., Murray, S., Trott, C. M., Wayth, R. B., Mitchell, D. A., Rahimi, M., Pindor, B., Procopio, P., Morgan, J. (2017), Characterization of the ionosphere above the Murchison Radio Observatory using the Murchison Widefield Array, MNRAS 471, 3974 (DOI: 10.1093/mnras/stx1797)

[362] Offringa, A. R., Mertens, F., van der Tol, S., Veenboer, B., Gehlot, B. K., Koopmans, L. V. E., Mevius, M. (2019), Precision requirements for interferometric gridding in the analysis of a 21 cm power spectrum, A&A 631, A12 (DOI: 10.1051/0004-6361/201935722)

[363] Barry, N., Beardsley, A. P., Byrne, R., Hazelton, B., Morales, M. F., Pober, J. C., Sullivan, I. (2019), The FHD/ϵppsilon Epoch of Reionisation power spectrum pipeline, PASA 36, e026 (DOI: 10.1017/pasa.2019.21)

[364] Parsons, A., Pober, J., McQuinn, M., Jacobs, D., Aguirre, J. (2012), A Sensitivity and Array-configuration Study for Measuring the Power Spectrum of 21 cm Emission from Reionization, ApJ 753, 81 (DOI: 10.1088/0004-637X/753/1/81)

[365] Liu, A., Parsons, A. R., Trott, C. M. (2014), Epoch of reionization window. I. Mathematical formalism, P.R.-D 90, 023018 (DOI: 10.1103/PhysRevD.90.023018)

[366] Morales, M. F., Hewitt, J. (2004), Toward Epoch of Reionization Measurements with Wide-Field Radio Observations, ApJ 615, 7 (DOI: 10.1086/424437)

[367] Morales, M. F., Bowman, J. D., Hewitt, J. N. (2006), Improving Foreground Subtraction in Statistical Observations of 21 cm Emission from the Epoch of Reionization, ApJ 648, 767 (DOI: 10.1086/506135)

[368] Morales, M. F., Bowman, J. D., Hewitt, J. N. (2006), Improving Foreground Subtraction in Statistical Observations of 21 cm Emission from the Epoch of Reionization, ApJ 648, 767 (DOI: 10.1086/506135)

[369] Liu, A., Tegmark, M. (2011), A method for 21 cm power spectrum estimation in the presence of foregrounds, PR-D 83, 103006 (DOI: 10.1103/PhysRevD.83.103006)

[370] Morales, M. F., Hazelton, B., Sullivan, I., Beardsley, A. (2012), Four Fundamental Foreground Power Spectrum Shapes for 21 cm Cosmology Observations, ApJ 752, 137 (DOI: 10.1088/0004-637X/752/2/137)

[371] Pober, J. C., Parsons, A.R., Aguirre, J.E., Ali, Z., Bradley, R.F., Carilli, C.L., DeBoer, D., Dexter, M., Gugliucci, N.E., Jacobs, D.C., Klima, P.J., MacMahon, D., Manley, J., Moore, D.F., Stefan, I.I., Walbrugh, W.P. (2013), Opening the 21 cm Epoch of Reionization Window: Measurements of Foreground Isolation with PAPER, ApJ 768, L36 (DOI: 10.1088/2041-8205/768/2/L36)

[372] Dillon, J. S., Liu, A., Tegmark, M. (2013), A fast method for power spectrum and foreground analysis for 21 cm cosmology, PR-D 87, 043005 (DOI: 10.1103/PhysRevD.87.043005)

[373] Byrne, R. and Morales, M.F. and Hazelton, B. and Li, W. and Barry, N. and Beardsley, A.P. and Joseph, R. and Pober, J. and Sullivan, I. and Trott, C. (2019), Fundamental Limitations on the Calibration of Redundant 21 cm Cosmology Instruments and Implications for HERA and the SKA, ApJ 875, 70 (DOI: 10.3847/1538-4357/ab107d)

[374] Byrne, R., Morales, M. F., Hazelton, B. J., Wilensky, M. (2021), A unified calibration framework for 21 cm cosmology, MNRAS 503, 2457 (DOI: 10.1093/mnras/stab647)

[375] Dulwich, F., Mort, B. J., Salvini, S., Zarb Adami, K., Jones, M. E. (2009), OSKAR: Simulating Digital Beamforming for the SKA Aperture Array., Wide Field Astronomy & Technology for the Square Kilometre Array (URL: https://ui.adsabs.harvard.edu/abs/2009wska.confE..31D)

[376] Jaeger, S. (2008), The Common Astronomy Software Application (CASA), Astronomical Data Analysis Software and Systems XVII 394, 623 (URL: https://ui.adsabs.harvard.edu/abs/2008ASPC..394..623J)

[377] SKAO (2021), SKA Phase 1 Construction Proposal (URL: https://www.skatelescope.org/wp-content/uploads/2021/02/22380_Construction-Proposal_DIGITAL_v3.pdf)

[378] Dewdney, P. and Braun, R. (2016), SKA1-LOW CONFIGURATION COORDINATES – COMPLETE SET (URL: https://astronomers.skatelescope.org/wp-content/uploads/2016/09/SKA-TEL-SKO-0000422_02_SKA1_LowConfigurationCoordinates-1.pdf)

[379] Wayth, R. B., Tingay, Steven J., Trott, Cathryn M., Emrich, David, Johnston-Hollitt, Melanie, McKinley, Ben, Gaensler, B. M., Beardsley, A. P., Booler, T., Crosse, B., Franzen, T. M. O., Horsley, L., Kaplan, D. L., Kenney, D., Morales, M. F., Pallot, D., Sleap, G., Steele, K., Walker, M., Williams, A., Wu, C., Cairns, I.H., Filipovic, M. D., Johnston, S., Murphy, T., Quinn, P., Staveley-Smith, L., Webster, R., Wyithe, J. S. B. (2018), The Phase II Murchison Widefield Array: Design overview, PASA 35, e033 (DOI: 10.1017/pasa.2018.37)

[380] Mondal, R., Bharadwaj, S., Majumdar, S. (2017), Statistics of the epoch of reionization (EoR) 21-cm signal - II. The evolution of the power-spectrum error-covariance, MNRAS 464, 2992 (DOI: 10.1093/mnras/stw2599)

[381] Majumdar, S., Mellema, G., Datta, K.K., Jensen, H., Choudhury, T.R, Bharadwaj, S., Friedrich, M.M. (2014), On the use of seminumerical simulations in predicting the 21-cm signal from the epoch of reionization, MNRAS 443, 2843 (DOI: 10.1093/mnras/stu1342)

[382] Choudhury, T. R., Haehnelt, M. G., Regan, J. (2009), Inside-out or outside-in: the topology of reionization in the photon-starved regime suggested by Lyα forest data, MNRAS 394, 960 (DOI: 10.1111/j.1365-2966.2008.14383.x)

[383] Mesinger, A., Furlanetto, S., Cen, R. (2011), 21CMFAST: a fast, seminumerical simulation of the high-redshift 21-cm signal, MNRAS 411, 955 (DOI: 10.1111/j.1365-2966.2010.17731.x)

[384] Murray, S., Greig, B., Mesinger, A., Muñoz, J., Qin, Y., Park, J., Watkinson, C. (2020), 21cmFAST v3: A Python-integrated C code for generating 3D realizations of the cosmic 21cm signal, JOSS 5, 2582 (DOI: 10.21105/joss.02582)

[385] Majumdar, S., Jensen, H., Mellema, G., Chapman, E., Abdalla, F.B., Lee, K., Iliev, I.T., Dixon, K.L., Datta, K.K., Ciardi, B., Fernandez, E.R., Jelić, V., Koopmans, L.V.E., Zaroubi, S. (2016), Effects of the sources of reionization on 21-cm redshift-space distortions, MNRAS 456, 2080 (DOI: 10.1093/mnras/stv2812)

[386] Liu, A., Tegmark, M., Morrison, S., Lutomirski, A., Zaldarriaga, M. (2010), Precision calibration of radio interferometers using redundant baselines, MNRAS 408, 1029 (DOI: 10.1111/j.1365-2966.2010.17174.x)

[387] Braun, R. (2017), ANTICIPATED SKA1 SCIENCE PERFORMANCE, (URL: https://astronomers.skatelescope.org/wp-content/uploads/2017/10/SKA-TEL-SKO-0000818-01_SKA1_Science_Perform.pdf)

[388] Högbom, J. A. (1974), Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines, A&ASup.Ser. 15, 417 (URL: https://ui.adsabs.harvard.edu/abs/1974A&AS...15..417H)

[389] Hogg, D. W. (1999), Distance measures in cosmology (URL: https://ui.adsabs.harvard.edu/abs/1999astro.ph..5116H)

[390] Pober, J. C., Parsons, A.R., DeBoer, D.R., McDonald, P., McQuinn, M., Aguirre, J.E., Ali, Z., Bradley, R.F., Chang, T., Morales, M.F. (2013), The Baryon Acoustic Oscillation Broadband and Broad-beam Array: Design Overview and Sensitivity Forecasts, AJ 145, 65 (DOI: 10.1088/0004-6256/145/3/65)

[391] Hurley-Walker, N. et al. (2019), GaLactic and Extragalactic All-sky Murchison Widefield Array (GLEAM) survey II: Galactic plane 345345° ¡ l ¡ 67°, 180° ¡ l ¡ 240° PASA 26, e047 (DOI: 10.1017/pasa.2019.37)

[392] wijnholds Lonsdale, C. J. (2005), Configuration Considerations for Low Frequency Arrays, From Clark Lake to the Long Wavelength Array: Bill Erickson's Radio Science 345, 399 (URL: https://ui.adsabs.harvard.edu/abs/2005ASPC..345..399L)

[393] Wijnholds, S., van der Tol, S., Nijboer, R., van der Veen, A.-J. (2010), Calibration challenges for future radio telescopes, IEEE Sig.Proc. 27, 30 (DOI: 10.1109/MSP.2009.934853)

[394] Trott, C. M. (2014), Comparison of Observing Modes for Statistical Estimation of the 21 cm Signal from the Epoch of Reionisation, PASA 31, e026 (DOI: 10.1017/pasa.2014.23)

[395] Patwa, A. K., Sethi, S., Dwarakanath, K. S. (2021), Extracting the 21 cm EoR signal using MWA drift scan data, MNRAS 504, 2062 (DOI: 10.1093/mnras/stab989)

[396] Kern, N. S. et al. (2020), Absolute Calibration Strategies for the Hydrogen Epoch of Reionization Array and Their Impact on the 21 cm Power Spectrum, ApJ 890, 122 (DOI: 10.3847/1538-4357/ab67bc)

[397] Dillon, J. S. et al. (2020), Redundant-baseline calibration of the hydrogen epoch of reionization array, MNRAS 499, 5840 (DOI: 10.1093/mnras/staa3001)