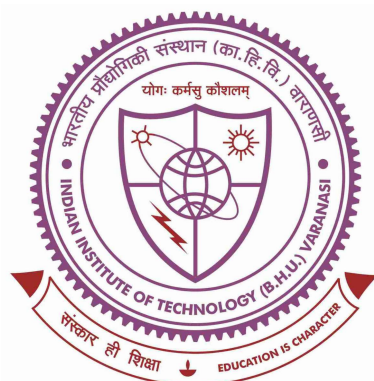


Calibration Requirements for EoR Observations - Effect of Time and Frequency Correlated Gains



**Thesis submitted in partial fulfillment
for the Award of Degree**

DOCTOR OF PHILOSOPHY

by

JAIS KUMAR

**DEPARTMENT OF PHYSICS
INDIAN INSTITUTE OF TECHNOLOGY
(BANARAS HINDU UNIVERSITY)
VARANASI - 221005**

ROLL NUMBER

16171010

YEAR OF SUBMISSION

2022

Certificate

It is certified that the work contained in the thesis titled “**Calibration Requirements for EoR Observations - Effect of Time and Frequency Correlated Gains**” by Jais Kumar has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

It is further certified that the student has fulfilled all the requirements of Comprehensive Examination, Candidacy and SOTA for the award of Ph.D. Degree.

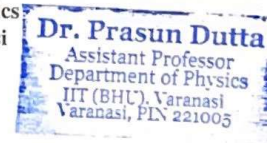
Date: *20/06/22*

Place: *Varanasi*

Supervisor

Prasun Dutta

Dr. Prasun Dutta
Assistant Professor
Department of Physics
IIT (BHU) Varanasi



Declaration

I, **Jais Kumar**, certify that the work embodied in this thesis is my own bonafide work and carried out by me under the supervision of **Dr. Prasun Dutta** from August 2016 to June 2022 at the **Department of Physics**, Indian Institute of Technology (BHU), Varanasi. The matter embodied in this thesis has not been submitted for the award of any other degree/diploma. I declare that I have faithfully acknowledged and given credits to the research workers whenever and wherever their works have been cited in my work in this thesis. I further declare that I have not wilfully copied any other's work, paragraphs, text, data, results, *etc.*, reported in journals, books, magazines, reports dissertations, theses, *etc.*, or available at websites and have not included them in this thesis and have not cited as my own work.

Date: 20/06/2022

Place: Varanasi



Signature of the Student

Jais Kumar

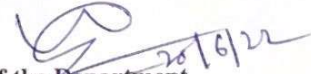
Certificate by the Supervisor

It is certified that the above statement made by the student is correct to the best of my knowledge.


Supervisor 20/06/22

Dr. Prasun Dutta

Dr. Prasun Dutta
Assistant Professor
Department of Physics
IIT (BHU), Varanasi
Varanasi, PIN 221005


Head of the Department
HEAD/विभागाध्यक्ष
भौतिकी विभाग/Deptt. of Physics
आ.प्रौ.सं./स.सं.वि.सं./IIT (BHU)
वाराणसी/Varanasi-221005

Copyright Transfer Certificate

Title of the Thesis: **Calibration Requirements for EoR Observations - Effect of Time and Frequency Correlated Gains**

Name of the Student : **Jais Kumar**

Copyright Transfer

The undersigned hereby assigns to the Indian Institute of Technology (Banaras Hindu University) Varanasi all rights under copyright that may exist in and for the above thesis submitted for the award of the *Doctor of Philosophy*.

Date: 20/06/2022

Place: Varanasi



Signature of the Student

Jais Kumar

Note: However, the author may reproduce or authorize others to reproduce material extracted verbatim from the thesis or derivative of the thesis for author's personal use provided that the source and the Institute's copyright notice are indicated.

To *The person who taught me
to draw the first symbols,
my grandfather, and my family.*

Life isn't about finding yourself. Life is about creating yourself.

- George Bernard Shaw

Acknowledgements

It has been an incredible journey at IIT (BHU) Varanasi towards working on this thesis. This Ph.D. has changed me so much in so many ways that it can hardly be described. There is a long list of people without whom the work in this thesis would not have been possible or at least much less enjoyable. I want to thank all the people who supervised, helped, inspired, and supported me during this period.

The first one to thank would definitely be my supervisor, Dr. Prasun Dutta, who has always been there truly as a “Guide” in this journey. It is impossible to express my gratitude toward you in a small paragraph like this; still, I would like to express my heartfelt gratitude towards you for mentoring me on my academic path and helping me navigate this large academic landscape. I got to learn so many things from you, not only the things “To Do” but also the things “Not To Do” which I believe will always be helpful for me in academics as well as in everyday life. Thanks for mentoring me, tolerating me for so long, and forgiving me for the stupid mistakes I made during the research work. Although I have missed many deadlines, you were always patient and believed more than me that the thesis would be finished someday. This thesis would never have been possible without the unimaginable patience and constant encouragement from Prasun, whose role as my thesis guide kept propelling me forward toward reaching this end.

I want to express my sincere thanks to my research progress evaluation committee (RPEC) members, Dr. Somak Bhattacharyya, and Dr. Abhishek Kumar Srivastava for the invaluable inspiration and numerous insightful suggestions during the entire course of this research. I wish to express my deep regards to my collaborators Dr. Nirpuam Roy and Dr. Samir Choudhury, for their scientific discussion and invaluable suggestions during my research work.

I offer my special thanks to the Head of the Department of Physics, the DPGC chairman, and all other faculty members of the Department for their support, guidance, and motivation. It would be worth mentioning the name of Prof. Bhola Nath Dwivedi for his unique and philosophical approach towards the subject. I extend my thanks to all the research colleagues and staff of the Department of Physics, IIT (BHU), for their kind help whenever I required it. I want to thank UGC, the Government of India, for providing me the financial support through the fellowship scheme UGC-JRF. The support and the resources provided by the 'PARAM Shivay Facility' under the National Supercomputing Mission, Government of India, at the Indian Institute of Technology, Varanasi, are gratefully acknowledged. I would also like to acknowledge the computing facility provided by the Department of Physics and CIC IIT (BHU).

I want to thank all my labmates, Pavan, Meera, Urvashi, and Saumyadeep, for their help and great company in the lab. Heartly thanks to my beloved friends Kanchan, and Sudheer for their constant support and encouragement throughout the whole journey. The joyful life shared with Samir, Ajay, Jay Prakash, Manish, and other friends in the hostel is unforgettable, be the games of badminton, cricket or carrom, random talks or anything. The list of friends from campus also includes but is not limited to Pragati, Mahima, Vandna, Yamini, Bharat, and Anupam. Every moment spent with them in the campus has been so cheerful. Memories shared with you will always cherish in my nervous system.

This journey started not precisely from the date of registration at IIT (BHU) but well before that. There are a bunch of people who always supported and encouraged me to go after my dream. The impact of Ashok Pathak sir, Sugata Pratik Khastgir Sir, and Somnath Bharadwaj sir on my academic life can't be erased. The cooperation and moral support received from Nidhi, and the insightful discussions with Mohit Pal on any topic out of any world require special mention. I sincerely acknowledge the constant support and help received from Abinash Kumar Shaw during any technical or other related difficulties since

my MSc days. The list is neither complete nor ends here, along with these I also want to thank all other known or unknown persons who came into my life, played their part, and supported me in any way.

My entire family has been the biggest source of strength, whose love, and affection gave me inspiration at every step of my life. My gratitude to my grandparents Late Mr. Chhote Lal & Mrs. Ram Sakhi, parents Mr. Shiv Babu & Mrs. Amrawati Devi, uncle Late Mr. Mahesh Chandra & Mr. Shiv Murat, brothers Mr. Rajesh Kumar, Mr. Awadhesh Kumar, Mr. Dharmendra Kumar, and sister Kalpana, for the faith they have posed, the patience they have demonstrated, and the support they have been unquestioningly extending throughout my life. I dedicate this thesis to my family.

Finally, I bow with reverence and gratitude to thank the Almighty, who has enriched me with such an excellent opportunity and infused the power in my mind to fulfill the work assigned to me. Without the blessings of Lord Shiva, it would not have been possible to travel so far on this journey. Har Har Mahadev



Jais Kumar

Department of Physics

IIT (BHU) Varanasi

Table of contents

Acknowledgements	xi
List of figures	xix
List of tables	xxvii
List of symbols	xxviii
Abstract	xxxii
1 Introduction	1
1.1 Structure Formation and Evolution of the Universe	2
1.2 H I 21- cm signal as a Probe of EoR	3
1.3 Intensity mapping and 21- cm Power Spectrum	5
1.4 Foregrounds	7
1.5 Current status of CD/EoR Experiments	8
1.5.1 Effect of Gain Errors	9
1.6 Aim of this thesis	12
1.7 Technical details	15
2 H I Power Spectrum and Foregrounds	17
2.1 The H I Power Spectrum and Dark Matter Power Spectrum	17

2.2	Interferometers and visibility measurements	21
2.3	Power Spectrum Estimator	24
2.4	Foregrounds	30
2.4.1	Extragalactic point sources	32
2.4.2	Galactic Diffuse Synchrotron Emission	33
2.5	Foregrounds Mitigation	39
2.5.1	Foreground Avoidance	39
2.5.2	Foreground Subtraction	40
2.5.3	Foreground supression	40
2.6	Discussion	42
3	Radio Interferometric Observations: Gain Errors and Calibration	45
3.1	Interferometric Observation and Gain	45
3.1.1	Time and frequency dependence of the Gain	47
3.1.2	Decomposing the complex gain into antenna-based gains	49
3.2	Calibration	50
3.2.1	Errors in the Calibration	53
3.3	Residual Gain Errors and Its Effects	56
3.3.1	Analytical Results	56
3.3.2	Simulation Results	58
3.3.3	Discussion and Conclusion	62
4	Effect of Time Correlated Gains	65
4.1	Gain error model	65
4.2	Power spectrum estimation and bias	68
4.3	A toy model of baseline pair fractions	72
4.4	Residual Gain Errors in presence of extragalactic point sources	75

4.4.1	Point source sky model	76
4.4.2	Simulation method	77
4.4.3	Results	80
4.5	Discussion and Conclusion	84
5	Analytical Estimates of Bias and Variance of Power Spectrum	91
5.1	Gain Error Model	92
5.2	Analytical Estimates of Bias and Variance of the Power Spectrum	95
5.2.1	Effect of gain errors in visibility correlation	95
5.2.2	Bias and Variance of the power spectrum	96
5.3	Comparing analytical expression of bias and variance with simulation	100
5.4	Results: Different effects of gain errors	107
5.5	Discussion and Conclusion	116
6	Effect of Frequency Correlated Bandpass Errors	119
6.1	Time and Frequency-Dependent Gains	119
6.1.1	Bandpass Errors	120
6.1.2	Modelling the Frequency Correlated Gain Errors	121
6.2	Analytical Estimates of Bias and Variance	123
6.2.1	Type of Baseline Pair Fractions	124
6.2.2	Bias and Variance	127
6.3	Analysing the Results: Effects of Time and Frequency-Dependent Gain Errors	130
6.3.1	Baseline pair fractions for SKA-1 Low	130
6.3.2	Foreground Model and Other Observational Details	131
6.4	Discussion and Conclusion	138
7	Summary and Future Scope	141

References	145
Appendix A Empirical Fit of the Maximum Baseline Contours	171
Appendix B Calculation of the Bias and Variance of the Power Spectrum	173
List of Publications	181

List of figures

- 2.1 This shows the linear power spectrum of dark matter density fluctuations at the present epoch. A baseline u at a frequency ν will probe the power spectrum at all Fourier modes $k \geq k_{min} = (2\pi/r_\nu)U$. The values of k_{min} are shown at different frequencies for $U = 100$. This figure can be used to determine the length scale probed by any baselines U . The figure is adapted from Bharadwaj and Ali (2005). 19
- 2.2 The sample uv-coverage of the uGMRT at 130 MHz, for 8 hours of observation. This shows the points in uv-plane where visibilities are sampled. Here, (u, v) are the antenna separations in wavelength units at the observing frequency of 150 MHz. 23
- 2.3 This cartoon schematically shows the gridding in uv-space. The black dots in the grid under consideration represents the measured visibilities which will be correlated to estimate the power spectrum at a baseline U 27

- 2.4 the angular power spectrum plot for various components of foregrounds at $z=9.2$ ($\nu = 140$ MHz), along with the 21- cm power spectrum (thick blue line). The red thin solid line is for the point sources and the red thin dotted line represents the Galactic Synchrotron emission. The extragalactic free-free and Galactic free-free components are shown with a red thin dashed line and a red thin long-dashed line, respectively. The green dot-dashed line represents the CMB power spectrum. The figure is reproduced from Santos et al. (2005). 37
- 2.5 A schematic picture of an expected 2D power spectrum showing the region of Fourier space that is accessible to the radio interferometric observations to measure the redshifted 21- cm signal. The high modes along the line of sight are limited by the spectral resolution and the low modes, for all the modes perpendicular to the line of sight, are essentially dominated by the spectrally smooth foregrounds. The chromatic response of the interferometers results in the leakage of these foregrounds to higher modes along the line of sight, leading to the foreground wedge. The primary field of view line and the horizon line are contamination limits dependent on how far off-axis sources are in the sky. The remaining region in Fourier space, which is free from the foreground contamination is termed the EoR window and is a promising region in which to pursue the first detection of the 21- cm power spectrum. The most sensitive, foreground-free measurement modes are expected to be in the lower, left-hand corner of the EoR window. The figure is reproduced from Barry et al. (2016). 41
- 3.1 Variation of the real part of the visibility with residual phase gain error for different amounts of residual amplitude gain errors σ_δ 59

3.2	Variation of the imaginary part of the visibility with phase gain error for different amounts of residual amplitude gain errors σ_δ	60
3.3	Effect of amplitude gain error on the noise in the real part of visibility for various residual phase gain errors σ_ϕ	61
3.4	Effect of phase gain error on the noise in the imaginary part of visibility for various residual amplitude gain errors σ_δ	62
4.1	Models for different baseline pair distributions are shown as a function of baseline for an array with a maximum baseline of $12\text{ k}\lambda$. The quantities n_i correspond to a fraction of baseline pairs of a particular kind as discussed in section 4.2. The horizontal grey line corresponds to the case when all types of baseline pairs have the same proportions.	73
4.2	Power spectral gain is plotted as a function of baselines (solid lines) for two types of baseline pair fractions discussed in the text. The values in the y axis are normalized by the standard deviation of the gain errors. The horizontal dashed line corresponds to the case when there is no time correlation in the residual gain errors. The dot-dashed line is parallel to the auto-correlation spectra of the gain errors (see text for details). A significant change in the power spectra is expected for such time-correlated gain errors.	74
4.3	The point source foreground. x and y are the x positions and y positions of the sources from the pointing center in arcmin.	77
4.4	Variation of baseline distribution in uv-space with the different directions in the sky. Baseline density, the number of baselines per unit area at baseline U doesn't change much for different declinations.	79

-
- 4.5 Residual visibility correlation plotted with baseline U . σ_δ is in percentage and σ_ϕ is in degree. U_{max} is the point where bias in the power spectrum exceeds the EoR power spectrum, shown by the vertical dotted line. . . . 81
- 4.6 This contour plot is showing the variation of U_{max} as a function of residual amplitude and phase gain errors. The maximum allowed baseline length U_{max} in kilo wavelengths is written over contour plots. Here amplitude gain errors σ_δ and σ_ϕ (radian) are given in percent and degree respectively. The values of $\alpha_\delta = \alpha_\phi = 0.97$ 83
- 4.7 Contour plot showing the variation of U_{max} in kilo wavelengths with the α parameters. The grey solid line corresponds to results from the simulation and the dark dashed lines are an empirical fit to the simulation. Values of the σ parameters fixed at 0.02. 84
- 4.8 The ratio of the bias to the power spectrum $\mathcal{G}(U)$ for point source sky for different residual gain errors. 85
- 4.9 Power spectrum for different components of the foreground. The dot-dashed line is the Poisson component of point sources used in this paper, the dashed line corresponds to the galactic diffuse synchrotron emission (GDSE) (Jelić et al., 2008). The solid curve integrates the effect of both foreground components. 87

4.10 Bias in the power spectrum with a compact and diffuse foreground model in the presence of different residual gain error models and observation time. The α parameters are kept at a fiducial value of 0.97 for all gain error models. The black dashed line corresponds to σ parameters of unity and an observation time of 1000 hours. The black dot-dashed line corresponds to the same gain error model but with 10000 hours of observation time. The grey-dashed line corresponds to σ parameter set to 0.1 and 128 hours of observation only. The grey solid line is the redshifted 21-cm signal (Bharadwaj and Ali, 2005) expected at 130 MHz. 88

5.1 Variation of χ with baseline U for different correlation times T_{corr} and uv-grid size ΔU for an integration time of 16 seconds. 101

5.2 Variation of baseline pair fractions with baseline U for GMRT, 8 hours of observation. Integration time is 16 sec and uv-grid size is $0.004 k\lambda$. The fraction of baseline pairs of type II is zero for the above-used parameters for GMRT. 102

5.3 Comparison of contribution from different terms in the expression for the variance of the power spectrum σ_p^2 . The result shown here is for 8 hours of the uGMRT observation with an integration time of 16 sec and bandwidth of 256 kHz. We have chosen 5% gain error in both real and imaginary parts of the visibilities and $T_{corr} = 16$ seconds. 104

- 5.4 We compare the analytical estimates of \mathcal{B}_P (thick black) and σ_P (thick grey) with their estimate from simulation (thin black and thin grey respectively for \mathcal{B}_P and σ_P). On the left side, we show the comparison plots for \mathcal{B}_P , and the comparison plots for σ_P are shown on the right side. The plots are shown for eight hours of observing time and a bandwidth of 64 kHz. In the top panel we keep $\sigma_R = 1\%$ and $T_{corr} = 16$ sec and show \mathcal{B}_P in (a) and σ_P in (b), for different values of σ_I . In the bottom panel we fix $\sigma_R = 1\%$ and $\sigma_I = 5\%$ and show the bias \mathcal{B}_P in (c) and σ_P in (d), for $T_{corr} = 8, 16, 64$ seconds. 105
- 5.5 Variation of \mathcal{B}_P (thin black) and σ_P (thin grey) with baseline is shown for 100 and 400 days of observing time. Here, we consider a bandwidth of 16 MHz, and a correlation time of $T_{corr} = 16$ seconds. Both the quantities σ_R and σ_I are set to $\sigma_R = \sigma_I = \sigma_g$. The thick black line corresponds to the expected EoR power spectrum (Bharadwaj and Ali, 2005). 107
- 5.6 Variation of \mathcal{B}_P (thin black) and σ_P (thin grey) with $\sigma_R = \sigma_I = \sigma_g$ for correlation time $T_{corr} = 16, 64$ seconds. Here, we consider a bandwidth of 16 MHz, and an observation time of 400 days. The plots are made at a baseline of 300λ . The thick black line corresponds to the expected EoR power spectrum (Bharadwaj and Ali, 2005) 110
- 5.7 Variation of \mathcal{B}_P (thin black) and σ_P (thin grey) against different correlation time with $\sigma_R = \sigma_I = \sigma_g = 0.1, 1.0\%$. Here, we consider a bandwidth of 16 MHz and an observation time of 400 days. The plots are made at a baseline of 300λ . The thick black line corresponds to the expected EoR power spectrum (Bharadwaj and Ali, 2005). 112

- 5.8 Variation of \mathcal{B}_P (thin black) and σ_P (thin grey) against different integration time $\Delta\tau$ with $\sigma_g = 0.1, 1.0\%$. Here, we consider a bandwidth of 16 MHz, and an observation time of 400 days. The plots are made at a baseline of 300λ . The correlation time for the gain errors are kept the same as $\Delta\tau$. The thick black line corresponds to the expected EoR power spectrum (Bharadwaj and Ali, 2005). The thick grey lines show the values of Risk \mathcal{R}_P as defined in the eqn 5.11 113
- 5.9 Variation of \mathcal{R}_P against different integration time $\Delta\tau$ for the best possible case with the uGMRT at 150 MHz. Here, we consider a bandwidth of 16 MHz, and observation times of 100, 200, and 400 days. The plots are made at a baseline of 300λ . The thick black line corresponds to the expected EoR power spectrum (Bharadwaj and Ali, 2005). 115
- 6.1 Variation of baseline pair fractions with baseline U for SKA-1 Low, 8 hours of observation. Integration time is 4 sec and uv-grid size is $0.004k\lambda$. The fraction of baseline pairs is dominated by type 4 for the above-used parameters. 131
- 6.2 Variation of \mathcal{B}_{C_l} (black) and σ_{C_l} (gray) with frequency separation $\Delta\nu$ for 0.1 and 1.0 % bandpass error (σ_b) and 0.5 and 1.0 MHz of correlation length in frequency (ν_c). Integration time and correlation time are 4 sec, and uv-grid size is $0.004k\lambda$. The gain error value σ_g is kept constant at 0.1%. The bandwidth of observation is 8 MHz. The total observation time is 8 hours (1 day of observation). 133
- 6.3 Variation of \mathcal{B}_{C_l} (black) and σ_{C_l} (gray) with bandpass error σ_b , for 0.5 and 1.0 MHz of correlation length in frequency (ν_c) and frequency separation $\Delta\nu = 1.0$ and 4.0 MHz. The total observation time is 1 day. 135

-
- 6.4 Bias and Variance for SKA-1 Low ($\mathcal{B}_{SKA}, \sigma_{SKA}$) and uGMRT ($\mathcal{B}_{GM}, \sigma_{GM}$) with $\Delta\nu$ for $\sigma_g = \sigma_b = 0.1$, and $\nu_c = 0.5$ MHz. Here we compare the performance of SKA-1 Low and uGMRT for their optimum correlation and integration time. 136
- 6.5 Variation of \mathcal{B}_{C_l} (black) and σ_{C_l} (gray) for SKA-1 Low with frequency separation $\Delta\nu$ for $\sigma_b = 0.1$ and 1.0 % and $\nu_c = 0.5$ and 1.0 MHz, same as figure 1 but after excluding the effect of gain and bandpass errors coming through the baseline pairs of type 1 and 3. 137
- 6.6 This figure shows the bias in the cylindrical averaged power spectrum estimate for $\sigma_b = 1\%$ and the frequency correlation width $\nu_c = 0.5$ MHz in $(k_{\perp}, k_{\parallel})$ plane. The gain error value σ_g is kept at 0.1%. The bias is normalized to the highest value. The blue line shows the wedge boundary. 138

List of tables

2.1	The values of the fiducial parameters used in equations 6.14 and 2.13, at $l = 1000$ and 130MHz. This table is reproduced from Santos et al. (2005).	38
2.2	Pros and Cons of the various foreground removal strategies used in experiments aiming to detect the 21- cm signal from EoR.	43
A.1	Values of the fitting parameters defined in equation 23. Values are given in the second row and the corresponding errors in the parameter are in the third row.	171

List of Symbols

Symbols	Definition
\tilde{g}_A	Complex gain for antenna A
δ_A	Residual amplitude gain of antenna A
ϕ_A	Residual phase gain of antenna A
δ_{AR}	Real part of residual gain of antenna A
δ_{AI}	Imaginary part pf residual gain of antenna A
σ_δ^2	Variance of residual amplitude gain
σ_ϕ^2	Variance of residual phase gain
σ_R^2	Variance of real part of residual gain
σ_I^2	Variance imaginary part pf residual gain
$\Delta\tau$	Integration time
T_D	Time taken by an antenna pair to cross a grid
$\tilde{b}_A(\nu)$	Bandpass response of antenna A
b_{AR}	Real part of bandpass response of antenna A
b_{AI}	Imaginary part pf bandpass response of antenna A
σ_b^2	Variance of bandpass response
$\xi(\tau)$	Normalized two-point correlation function of time-correlated gain errors
$\xi(\Delta\nu)$	Normalized two-point correlation function of frequency correlated bandpass errors
T_{Corr}	Correlation time (sec)
χ	Average effect of time-correlation over a grid
$\Delta\nu$	Frequency separation between two channels
$\Delta\nu_c$	Frequency channel width (MHz)
ν_c	Frequency correlation length

Symbols	Definition
n_i	Baseline pair fractions of Type i (used in Chapter 5)
n'_i	Baseline pair fractions of Type i (used in Chapter 6)
σ_g	Residual gain error for the case $\sigma_R = \sigma_I$
N	Total number of antennae of an array
N_d	Number of days of observation
N_b	Number of baselines in a grid
N_g	Number of baseline pairs in a grid
N_B	Number of baseline pairs in an annulus at baseline U
N_G	Number of grids in an annulus at baseline U
$\tilde{N}_i(\mathbf{U}_i)$	System noise in the i^{th} visibility measurement
σ_N	Noise rms in real or imaginary part of the visibility
P_{HI}	HI 21-cm power spectrum (Jy^2)
P_F	Foreground power spectrum (Jy^2)
\mathcal{B}_P	Bias of the H I 21- cm power spectrum
σ_P^2	Variance of the H I 21- cm power spectrum
$C_l(\Delta\nu)$	Multifrequency angular power spectrum of foregrounds at frequency separation $\Delta\nu$ (mK^2)
$I_l(\nu_1, \nu_2)$	Foreground frequency decorrelation function
α	Mean spectral index
ζ	Foreground frequency decorrelation factor
β	power law index for foreground power spectrum
\mathcal{B}_{Cl}	Bias of the H I 21- cm angular power spectrum estimate
σ_{Cl}^2	Variance of the H I 21- cm power spectrum
ν_e	Rest frame frequency of H I 21- cm signal
r_ν	Comoving distance corresponding to ν
r'_ν	Derivative of r_ν

Symbols	Definition
H_0	Hubble parameter at present(Km/s/Mpc)
$H(z)$	Hubble parameter at redshift z (Km/s/Mpc)
h	$H_0/100$ km/s/Mpc
h_p	Planck's constant
Ω_b	Baryon density parameter
Ω_m	Matter density parameter
T_s	Spin temperature of hydrogen gas
τ_e	CMB electron scattering optical depth
T_γ	CMB temperature (K)
$\delta T_b(\mathbf{n}, z)$	Excess 21- cm brightness temperature along a direction \hat{n}
x_{HI}	Neutral hydrogen fraction
$\bar{x}_{HI}(z)$	Spatially averaged HI fraction
ρ_{HI}	Neutral Hydrogen Density
$\Delta\rho_{HI}$	Fluctuation in the neutral hydrogen density
$\bar{\rho}_H$	Mean Hydrogen Density
$\eta_{HI}(\mathbf{n}, z)$	21-cm radiation efficiency in redshift space
$\eta_{HI}(\mathbf{k}, z)$	Fourier transform of $\eta_{HI}(\mathbf{n}, z)$ (Mpc^{-3})
δ	Dark matter density contrast
$P(z, k)$	Dark matter power spectrum (Mpc^{-3})
$\delta_{21}(\mathbf{x})$	Fractional Perturbation to the H I brightness temperature
$\delta_{21}(\mathbf{k})$	Fourier Transform of $\delta_{21}(\mathbf{x})$
$\delta_D(x)$	Dirac delta function
T_γ	CMB temperature (K)

Abstract

In the cosmic history of the matter, at the end of the Epoch of recombination, tiny fluctuations in the matter densities were present. With the growth of these tiny fluctuations due to gravitational instabilities, the first luminous objects in the universe emerged. The era during which these first objects formed is termed as the Cosmic Dawn. The radiation from these luminous objects started ionizing the surrounding neutral hydrogen in the universe. As time progressed, the bubbles of ionized hydrogen increasingly overlapped, resulting in an entirely ionized universe. This process of phase transition of the universe from a neutral state to an almost completely ionized state is known as reionization. The Epoch of Re-ionization is one of the most important but less understood eras in the history of the universe's evolution. As neutral hydrogen was the dominant component of the universe during Cosmic Dawn, EoR, and post-EoR era, the radiation emerging from the forbidden spin-flip transition of neutral hydrogen provides excellent probe of these important stages of the universe. This radiation resulting from the line transition of the neutral hydrogen peaked at a frequency of 1420 MHz or equivalently at the 21-cm in wavelength units, hence getting its famous name "the 21-cm signal". Observing this 21-cm signal at redshifted frequencies can solve many mysteries, including our understanding of structure formation, the evolution, and the radiation mechanism of these first luminous objects. The direct imaging of the redshifted 21-cm signal is quite difficult due to the presence of very strong radiation coming from the foregrounds in similar observational frequency ranges, collectively known as foregrounds. Apart from the presence of orders of magnitude larger foregrounds in the observed frequency range, the instrumental effects of the interferometers, combined with the ionospheric effects, present a considerable challenge in the extraction of 21-cm signals from strong foregrounds. The combination of these instrumental and ionospheric effects, which is termed as "gain", is complex in nature and depends both on time and frequency.

The gains are estimated during the data preparation using various calibration procedures. Even after investing a good amount of effort, there are some residual errors in the gain. Although these residual gain errors are small in magnitude, these gain errors become important in high dynamic range observations with the presence of strong foregrounds. The systematic effects of these time and frequency-correlated residual gain errors originating from the measurement process introduce a bias and enhance the variance of the power spectrum measurements. Our work studies the effect of time and frequency-correlated residual gain errors in estimating the power spectrum of the sky brightness distribution in high dynamic range observations, such as EoR observations.

In the first project included in this thesis, we study the effect of residual gain errors in primary calibration. Modeling the residual gain errors as time-independent Gaussian random variables with zero mean, we analytically calculated the amplitude and variance of the visibility. Calculations showed that the visibility amplitude decreases with the phase gain error. Also, the variance of the visibility increases with both amplitude and phase gain error. For a single point source at the center of the field of view, we study the statistics of simulated visibilities and find that the simulated visibilities show similar behavior as predicted by analytical results.

In the next project, we discuss a methodology to estimate the bias in the power spectrum estimates of the redshifted 21-cm signal from neutral hydrogen in the presence of strong foregrounds. Considering time correlation in residual gain, we model the two-point correlation function of the residual gain errors using a power law and study its effect on the power spectrum. Through simulated observations using GMRT baseline configuration, we show that the presence of time-correlated residual gain errors introduces a bias in the power spectrum estimator that uses the correlation of visibilities in nearby baselines. We consider the contributions from various types of baseline pairs involved in the visibility measurement to be used in visibility correlation. We find that for the visibility correlation-

based power spectrum estimator, the bias in the power spectrum arises mainly from those types of baseline pairs that have at least one antenna in common. The bias also depends on the foreground model and the correlation properties of the gain errors. We define the ratio of the bias to the foreground power spectrum as the “Power spectral gain”. This quantity was later used to calculate the bias for the foreground model, including the bright point sources and the galactic diffuse synchrotron emission. Our results show that more than 10,000 hours of observation will be required using GMRT baseline configuration for the calibration accuracy of 1% for the full foreground. In this work, we do not consider the frequency dependence of the gain. Also, to check the effect of gain errors explicitly, we do not consider the thermal noise in our visibility measurements.

In the next project, we included this measurement noise in the estimation of bias and variance in the power spectrum with residual gain errors. We propagate errors in the presence of thermal noise and present a method to produce analytic estimates of the bias and variance in the power spectrum. We model the residual gain errors as time-correlated Gaussian random variables with zero mean. To incorporate the time correlation of the gain errors, we model the two-point correlation function of the gain errors using a Gaussian function with correlation time given as the half-width of the Gaussian. We study the effect of time-correlated residual gain errors in presence of strong foregrounds and thermal noise analytically and provide the mathematical expressions for bias and variance in the power spectrum measurements. We used simulated observations using GMRT baseline configuration in the presence of point source foregrounds. We observe that the simulation results are in quite good agreement with our analytical estimates. Later we use this analytical framework to understand various effects of the correlated gain errors. As the standard deviation in the residual gain errors increases, the bias in the estimation supersedes the variance. It is observed that an optimal choice of the time over which the gain solutions are estimated minimizes the combination of bias and variance, collectively known as the

risk. We also find that the interferometers with higher baseline densities are preferred instruments for these studies.

As the complex antenna gains are both time and frequency-dependent, we incorporate the frequency dependence of the gains in our analytical framework as a part of our final project. Similar to the time correlation of the gain errors, we model the frequency correlation of the gains also using a Gaussian function. Our analytical estimates show that the bias is more at low-frequency separations and increases with an increase in the correlation frequency. On the other hand, the variance is dominated by the contribution of thermal noise, and it shoots up at higher frequency separations. This is because, at higher frequency separations, the number of independent estimates of the visibility correlation will be less. We compare the performance of GMRT and SKA - 1 Low baseline configurations and find that SKA -1 Low baseline configuration will require ~ 50 times less observation time than the GMRT. This is primarily because of the baseline configuration and the larger collecting area of the SKA- 1 Low.

Our work sheds light on the importance of assessing the effect of time and frequency-correlated gain errors in high dynamic range observations, such as EoR experiments. Although the mathematical estimates of the bias and variance depend upon the array configuration, the methodology we present in this work is rather general and can be used to estimate the bias and variance of the power spectrum estimates for any telescope, given the antenna gain properties are known. We believe this would add and aid the quest of detecting the redshifted 21- cm signal from the Epoch of reionization.