Instrumental Effects in 21 cm Cosmology: One-point Statistics and Power Spectrum with the HERA Interferometer

by

Honggeun Kim

B.S. Astronomy, Yonsei University, 2013
M.S. Astronomy, Yonsei University, 2016

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2024

Authored by: Honggeun Kim
Department of Physics
August 16, 2024

Certified by: Jacqueline N. Hewitt
Professor of Physics, Thesis Supervisor

Accepted by: Lindley Winslow
Professor of Physics
Associate Department Head of Physics

Instrumental Effects in 21 cm Cosmology: One-point Statistics and Power Spectrum with the HERA Interferometer

by

Honggeun Kim

Submitted to the Department of Physics
on August 16, 2024 in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

ABSTRACT

The epoch of reionization (EoR) signifies a critical phase in the universe's evolution, marking the shift from a predominantly neutral intergalactic medium to the ionized state observed today. A key aspect of studying the EoR involves observing the redshifted 21 cm line emission with radio telescopes. A significant challenge in this endeavor is isolating the faint 21 cm signals from bright foreground emissions and systematics. This collection of works focuses on understanding the impact of instrumental systematic effects on statistical measurements, such as the one-point statistics and power spectrum, using the Hydrogen Epoch of Reionization Array (HERA).

First, I investigate one-point statistics measured from image cubes based on HERA Phase I observations after foreground removal for the first time. I highlight the influence of systematics on these measurements, by measuring the second and third moments. These analyses show that, despite efforts to mitigate systematics, the residual systematics still cause deviations in the measurements from the expected values. In addition, I evaluate EoR models against observational data, suggesting the second moment measurements likely reject the cold reionization model characterized by inefficient X-ray heating. The third moment, which captures non-Gaussianity features of the signals, is significantly diminished by the instrument response and further reduced by the foreground removal process, making it challenging to probe non-Gaussianity. However, there remains the potential to detect some skewness at low redshifts.

One potential systematic for HERA involves calibration errors stemming from per-antenna perturbations due to feed misalignment. I have simulated these calibration errors by modeling realistic perturbed primary beams for HERA Phase II observations. The chromatic calibration errors are critical since they can cause foreground emission to contaminate the region of Fourier space expected to be dominated by cosmological signals. I then present the work focused on developing a method to mitigate the calibration errors and foreground leakage, thereby recovering the clean EoR window.

Thesis supervisor: Jacqueline N. Hewitt
Title: Professor of Physics

Acknowledgments

When I arrived at MIT, everything felt new and challenging—a different culture, a fresh environment, and new research projects. I vividly recall the effort it took to pass the written qualifying exams when I joined the physics department, followed by the oral qualifying exam. Each class was demanding, requiring the ability to quickly acquire knowledge and solve complex problems across various fields of physics. Though these experiences were tough, they were incredibly valuable, enriching, and enjoyable.

These experiences were made especially meaningful by the people around me—my supervisor, co-workers, and colleagues. First and foremost, I extend my deepest gratitude to my supervisor, Jackie. Your unwavering guidance, support, and mentorship have been pivotal in shaping my research journey. Your insightful feedback and constant encouragement pushed me to explore the boundaries of my work, for which I am truly thankful. I also want to express my appreciation to the committee members for their valuable inputs and support throughout this process, and for dedicating their time and effort to reviewing my thesis.

My research was largely supported by the HERA collaboration, and I am grateful for the collaborative and supportive research environment they provided. It has been a privilege to be a member of this group. I would like to give special thanks to several researchers: Bang Nhan, whose assistance with beam simulations was invaluable—without him, this project would not have been possible. Nick Kern, whose constructive feedback consistently enhanced the quality of my research. Zhilei Xu, a former postdoc at MKI, whose discussions and collaborations I thoroughly enjoyed. Josh Dillon, a leading scientist at HERA, who always offered insightful advice and guidance, especially in understanding the Phase I observational data. Aaron Ewall-Wice, another core researcher at HERA, who helped me get familiar with the HERA data when I first joined the group.

I also extend my gratitude to the faculty, staff, and my colleagues at MIT. I am proud to be part of the Kavli Institute at MIT (MKI). The faculty here are not only insightful and supportive but also enthusiastic in their research fields, setting an example of what it means to be an excellent researcher. Special thanks to my friends and colleagues at MKI and MIT—you made my PhD experience enjoyable and comfortable. I also want to thank my Korean friends, who provided a cultural connection that enriched my time here.

Lastly, I want to express my deepest appreciation to my family. My mother, who supported us after my father passed away when I was a child, has been my greatest source of strength. Your endless love, encouragement, and support have been the backbone of my

journey. I am also thankful to my sister and her family, who provided a comforting home in the U.S., helping me stay connected to my roots both physically and emotionally. Special thanks to my younger brother, whose enthusiasm and support have always been a great help to me.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The study of 21 cm cosmology is an emerging field that promises to revolutionize our understanding of the early universe. The 21 cm line, resulting from the hyperfine transition of neutral hydrogen atoms, offers a unique window into the Cosmic Dawn and the Epoch of Reionization (EoR). During these pivotal epochs, the first stars, galaxies, and black holes formed, ionizing the surrounding hydrogen and fundamentally altering the universe's structure. Observing the 21 cm signal allows us to trace these changes, providing insight into the processes that shaped the cosmos.

One of the most significant aspects of 21 cm cosmology is its ability to probe the intergalactic medium (IGM) across vast cosmic scales and times. Unlike other forms of radiation, the 21 cm signal can penetrate the clouds of gas and dust that often obscure other wavelengths, giving us a clearer picture of the early universe. By mapping the distribution and state of neutral hydrogen over time, we can reconstruct the history of cosmic reionization and understand the interplay between radiation and matter during the universe's formative years.

The observational efforts in 21 cm cosmology include several cutting-edge radio telescopes, such as the Hydrogen Epoch of Reionization Array (HERA; DeBoer et al., 2017), the Square Kilometre Array (SKA; Koopmans et al., 2015), and the Low-Frequency Array (LOFAR; van Haarlem et al., 2013). These instruments are designed to detect the faint 21 cm signals from high redshifts, pushing the boundaries of current technology. However, the challenges are significant, as the 21 cm signal is extremely weak and must be distinguished from foreground emission that is several orders of magnitude stronger. Advanced data processing techniques, such as foreground subtraction and statistical analyses, are essential to isolating the cosmological signal.

In parallel with observational efforts, theoretical modeling and simulations play a crucial

role in 21 cm cosmology. Tools like 21cmFAST (Mesinger et al., 2011) provide semi-numerical simulations of the 21 cm signal, incorporating astrophysical and cosmological parameters to predict the signal's characteristics under various scenarios. These simulations help interpret observational data, refine theoretical models, and guide future observations. By simulating the large-scale structures and the ionization history of the universe, we can generate three-dimensional maps of the 21 cm brightness temperature, offering valuable insights into the distribution of matter and the processes driving reionization.

By studying the 21 cm line, we can uncover the properties of the first stars and galaxies, the nature of dark matter, and the dynamics of cosmic evolution. This not only sheds light on the past but also informs our understanding of the present universe. The interplay between theory, simulation, and observation in 21 cm cosmology highlights the collaborative effort needed to address one of the most challenging and exciting frontiers in astrophysics.

1.1 Theoretical Overview of the 21 cm Cosmology

In the hydrogen atom, the hyperfine splitting occurs due to the interaction between the magnetic moment of the proton in the nucleus and the magnetic moment of the electron. The proton and the electron can have their spins aligned either parallel or antiparallel to each other. When the spins are parallel, the energy is slightly higher compared to when the spins are antiparallel. This small energy difference results in two distinct energy states.

The transition between these two hyperfine energy states in the hydrogen atom leads to the emission or absorption of a photon with a wavelength of approximately 21 cm, corresponding to a frequency of 1420.41 MHz. This is known as the 21 cm line. During the era of interest, the 21 cm line is redshifted according to $\nu = \nu_{21}/(1+z)$ where ν_{21} is the rest frequency of the 21 cm line and z is the redshift. This places the line within the low-frequency radio region of the electromagnetic spectrum.

This line is explored by using a spin temperature (T_S) that is determined by the ratio of populations between the two spin states,

$$\frac{n_1}{n_0} = 3 \exp\left(- \frac{h\nu_{21}}{k_B T_S} \right) \tag{1.1}$$

where the front factor of 3 indicates the degeneracy ratio of the states, n_1 is the number of atoms in the excited state, n_0 is the number of atoms in the ground hyperfine state, h is Plank's constant, and k_B is Boltzmann's constant.

The brightness temperature, which is our observable, is determined by the contrast be-

tween the spin temperature and the background radiation temperature. Specifically, the brightness temperature of 21 cm signals is given by Furlanetto et al. (2006),

$$\begin{aligned}\delta T_b(\nu) &= \frac{T_{\rm S}-T_\gamma}{1+z}\left(1-e^{-\tau_{21}}\right)\\ &\approx 27 x_{\rm H_I}(1+\delta_{\rm m})\left(\frac{H}{dv_\parallel/dr+H}\right)\left(1-\frac{T_\gamma}{T_{\rm S}}\right)\\ &\times\left(\frac{1+z}{10}\frac{0.15}{\Omega_{\rm M}h^2}\right)^{1/2}\left(\frac{\Omega_{\rm b}h^2}{0.023}\right)\,{\rm mK},\end{aligned} \tag{1.2}$$

where T_γ represents the background radio temperature (e.g., Cosmic Microwave Background (CMB) temperature), τ_{21} is the optical depth at the 21 cm frequency, $x_{\rm H_I}$ denotes the neutral fraction of the hydrogen gas. $\delta_{\rm m}=\rho/\bar{\rho}-1$ describes the matter density fluctuations, where ρ is the matter density and $\bar{\rho}$ is its mean. $dv_\parallel/dr$ is the gradient of the peculiar velocity along the line-of-sight. The detectability of the 21 cm signal primarily depends on the spin temperature. A signal is only observable if the spin temperature is different from the background temperature.

The spin temperature is determined by three primary processes:

1. Absorption/Emission of 21 cm Photons: This involves the interaction of 21 cm photons with the radio background, primarily the CMB. The exchange of photons between the hydrogen atoms and the background radiation can either raise or lower the spin temperature depending on the relative temperatures of the two.

2. Collisions: Collisions with other hydrogen atoms and with electrons can transfer energy and cause spin flips, influencing the spin temperature. The rate of these collisions and the density of hydrogen atoms and electrons in the medium play crucial roles in determining the overall spin temperature.

3. Resonant Scattering of Lyα Photons: The scattering of Lyα photons, which are emitted by early stars and galaxies, can excite hydrogen atoms to an intermediate state, leading to a spin flip. Because of large cross section of Lyα scattering, this process, known as the Wouthuysen-Field (WF) effect, can strongly couple the spin temperature to the kinetic temperature of the gas.

These processes collectively determine the spin temperature. In the limit of equilibrium, the spin temperature can be summarized as (e.g., Field, 1958),

$$T_{\rm S}^{-1}=\frac{T_\gamma^{-1}+x_\alpha T_\alpha^{-1}+x_c T_{\rm K}^{-1}}{1+x_\alpha+x_c}, \tag{1.3}$$

where T_α is the color temperature of the Lyα photon, $T_{\rm K}$ is the kinetic temperature of the gas, x_α and x_c are the coupling coefficients of the Lyα scattering and atomic collisions, respectively.

I want to describe a basic picture of the evolution of the signal briefly. After the recombination ($z \approx 1100$), the free electrons are coupled to the CMB photons through Compton scattering, or $T_{\rm K} = T_\gamma$. Since the spin temperature is determined by the collisional coupling, the spin temperature is the same as the photon temperature, resulting in $\delta T_b = 0$ and no detectable 21 cm signal. As the density decreases ($z \approx 200$), the gas is decoupled from the CMB photons, leading to adiabatic cooling of gas $\propto (1+z)^2$ compared to the photon cooling according to the universe expansion $\propto (1+z)$, which indicates the gas is cooling faster than the photon. Since the gas density is high enough, the spin temperature is coupled to the kinetic temperature, resulting in $T_{\rm S} < T_\gamma$. As the universe expands further, the gas density drops and the collisional coupling becomes inefficient, and the spin temperature is coupled to the CMB temperature, $T_{\rm S} = T_\gamma$. Once the first starts emerge, they emit Lyα photons, which couple the spin temperature to the kinetic temperature via the WF effect. They also emit X-rays that heat the surrounding material, turning into $T_{\rm S} > T_\gamma$ during the EoR. As ionized regions grow, the impact of the neutral gas fraction in Equation (1.2) becomes significant.

Current observational 21 cm cosmology is particularly interested in the Cosmic Dawn and the EoR. For example, the Experiment to Detect the Global Epoch of Reionization Signature (EDGES; Bowman et al., 2018) reported a detection of an absorption feature in global 21 cm signal centered around 78 MHz, which corresponds to a redshift of approximately 17. This signal suggests, if confirmed, that the hydrogen gas was much colder than expected or that there was an excess radio background during the Cosmic Dawn.

Different radio interferometers have invested substantial efforts in investigating the power spectrum of 21 cm signals during the EoR, leading to several upper limits. These studies include the Giant Metre Wave Radio Telescope (GMRT; Paciga et al., 2013), the Murchison Widefield Array (MWA; Barry et al., 2019, Beardsley et al., 2016, Dillon et al., 2014, Ewall-Wice et al., 2016a, Tingay et al., 2013, Trott et al., 2020), the Donald C. Backer Precision Array for Probing the Epoch of Reionization (PAPER; Cheng et al., 2018, Kolopanis et al., 2019, Parsons et al., 2010), LOFAR (Gehlot et al., 2019, Mertens et al., 2020, Patil et al., 2017, van Haarlem et al., 2013), and HERA (DeBoer et al., 2017, Dillon and Parsons, 2016). Recent studies employing HERA (The HERA Collaboration et al., 2022a,b, 2023) have reported the upper limit of the 21 cm power spectrum, which provide constraints on the X-ray heating of the IGM and indicate scenarios with inefficient X-ray heating are disfavored by a redshift of $z \sim 8$.

In this thesis, I utilize the HERA interferometer to investigate the impacts of instrumental effects on statistical measurements of 21 cm signals. The details on the HERA instrument are described in the following section.

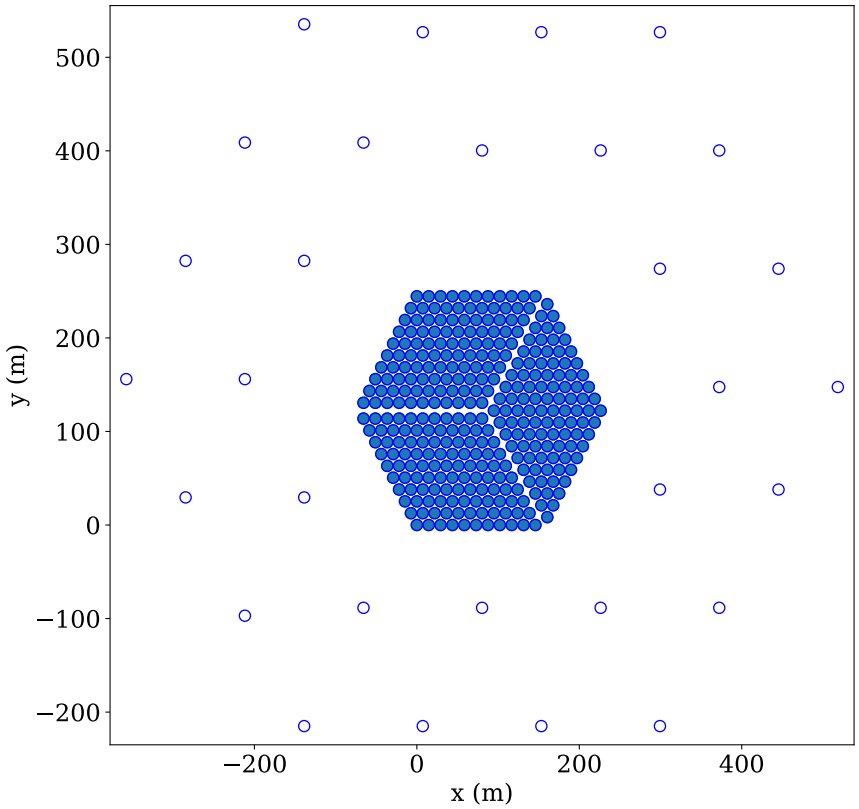

Figure 1.1: Full array configuration of HERA. The filled and open circles represent the core array (320 antennas) and outliers (30 antennas), respectively.

1.2 HERA Radio Interferometer

HERA, or the Hydrogen Epoch of Reionization Array (DeBoer et al., 2017), is a radio telescope interferometric array specifically designed to study the early universe by observing the 21 cm hydrogen line. HERA is located in the Karoo desert in South Africa, a site chosen for its radio-quiet environment, which minimizes interference from human-made radio signals. The primary aim of HERA is to measure the power spectrum of 21 cm fluctuations, which helps in understanding the timing, duration, and nature of reionization. This includes studying the first sources of light, such as first stars and galaxies, and their impact on the IGM.

HERA consists of a dense array of 14-meter parabolic dishes arranged in a hexagonal grid (Dillon and Parsons, 2016). The design emphasizes sensitivity to large-scale structures, crucial for detecting the faint 21 cm signal from the early universe. The compact and redundant array configuration is also optimized for precise calibration. Figure 1.1 depicts the full configuration of HERA array, composed of the core array (filled circles) and outriggers (open circles) with 350 antennas in total. The dense core array is crucial for measuring

the power spectrum, while the outriggers enhance imaging capabilities. HERA employs a redundant-baseline calibration strategy, which relies on the redundancy of measurements for a given antenna pair and does not require prior sky information. This approach should, in principle, yield precise calibration solutions as long as the measurements maintain the expected redundancy.

HERA utilizes two different systems; Phase I and Phase II. The Phase I system operated with repurposed PAPER's sleeved dipoles (Fagnoni et al., 2020), spanning a frequency range of 100–200MHz or a redshift range of 6–13. The HERA Collaboration et al. (2023) involves 94 night observations with about 40 operational antennas, releasing power spectrum results based on Phase I observations. In 2018, the instrument transitioned to the Phase II system, which features a new feed model, the Vivaldi feed (Fagnoni et al., 2021). This new feed covers a broader range of frequencies or redshifts, covering a redshift range of 5–30. This extension enables us to explore the Cosmic Dawn as well as EoR.

One important aspect of a radio inteferometer is the "foreground wedge" in Fourier space. This is the region where foreground emission, such as galactic synchrotron radiation and extragalactic point sources, dominates the observed signal. The spectrally smooth astrophysical foreground emission dominates at small $k_{\parallel}$ (line-of-sight wavenumbers). Meanwhile, the 21 cm signal, expected to be weak and have a different spectral structure than the foregrounds, extends to higher $k_{\parallel}$ values, forming the "EoR window". In this thesis, I focus on foreground systematics bleeding into the EoR window by investigating their effects on one-point statistics and the power spectrum.

1.3 Outline of Thesis

Throughout my thesis work, I have studied the impact of systematics on statistical measurements based on HERA observations. The Phase I observational data, with comprehensive calibration and system mitigation efforts, was thoroughly analyzed using the power spectrum. The power spectrum is optimal for studying the properties of underlying Gaussian random fields. However, highly non-Gaussian features are expected to arise from ionized bubbles during the EoR. Imaging-based analyses offer an alternative approach to capturing such non-Gaussianity. In Chapter 2, I investigate primary statistical measurements, such as variance and skewness, in image space and examined the effects of systematics on the measurements. I also forecast the detectability of non-Gaussianity using future HERA observations.

In Chapter 3, I focus on a key potential source of systematics: the per-antenna perturbation induced by feed misalignment. I extensively simulated radio primary beams by

perturbing the position of the feeds using an electromagnetic simulator and constructed mock Phase II HERA observations with the HERA core array (320 antennas). The simulated data was processed through the HERA calibration pipeline. I examined the effects of calibration errors caused by feed mispositioning and how these errors propagate into our power spectrum estimation, thereby contaminating the EoR window with foreground residuals. Chapter 4 addresses strategies for mitigating the calibration errors by employing three different methods: excluding antenna pairs with long separation during calibration, applying a temporal filter before the calibration and smoothing the gain solutions. I demonstrate combining these methods effectively suppresses the systematics, recovering the clean EoR window. These results set the requirement for feed positioning during HERA operations. Chapter 5 summarizes the thesis work and novel contributions.

Chapter 2

Exploring One-point Statistics in HERA Phase I Data: Effects of Foregrounds and Systematics on Measuring One-Point Statistics

This work is in preparation for publication. Author lists include Honggeun Kim, Jacqueline N. Hewitt, and Nicholas S. Kern.

Abstract

Measuring one-point statistics from redshifted 21 cm intensity maps offers a unique opportunity to explore non-Gaussian features of the early universe. In this study, we analyze images from the Hydrogen Epoch of Reionization Array (HERA) radio interferometer, focusing on one-point statistics after removing strong foreground emission and investigating the impacts of sysmtematics on the measurements. Using HERA Phase I observations over 94 nights, which are expected to be noise-dominated in the absence of foregrounds, we examine the the second (m_2) and third (m_3) moments, indicators of variance and skewness, respectively. Employing the `DAYENU`-filtering method, an advanced wedge-filtering technique, for foreground removal, we find that simulated foreground residuals are reduced to about 10% of the simulated 21 cm signal residuals, which is sufficiently low for probing the 21 cm intensity maps given the noiseless simulation. This result displays the `DAYENU`-filtering outperforms other foreground removal methods such as the standard wedge-masking method and the foreground subtraction technique based on principal component analysis. The amplitudes of one-point statistics measurements, such as m_2 and m_3, from the cosmological simulations are significantly reduced by the instrument response and further diminished by wedge-filtering, with a possible detection of skewness (S_3) remaining at low redshifts. Analyses with wedge-filtered observational data, along with expected noise simulations, show that

systematics can alter the probability distribution of the map pixels, affecting the m_2 and m_3 measurements. A likelihood analysis based on the observational data indicates that the m_2 measurements likely reject the cold reionization model characterized by inefficient X-ray heating. Small signals in m_3 due to the instrument response of the Phase I observation and wedge-filtering make it challenging to capture non-Gaussian nature of 21 cm models. Future forecasts with the full HERA core array indicate improved signal-to-noise ratios (S/N) for m_2 and m_3 with perfect foreground removal, though wedge-filtering reduces S/N significantly. The wedge-filtering has a large effect on reducing the non-Gaussian signals, and techniques need to be developed to make it possible to work within the wedge. Despite the low S/N after the wedge-filtering, one-point statistics from long observation periods can still yield tight constraints on model parameters using Fisher information matrix, contributing to our understanding of the distribution and evolution of neutral hydrogen.

2.1 Introduction

The epoch of reionization (EoR) stands as a pivotal phase in the evolution of the universe, marking the transition from a predominantly neutral intergalactic medium (IGM) to the ionized IGM that we see today. Studying the details of this cosmic transition provides crucial insights into the formation and evolution of the first luminous sources. One of the key observables during the EoR is the redshifted 21 cm line emission, emanating from the hyperfine splitting of ground-state neutral hydrogen. Radio interferometers, equipped to capture these faint signals, are expected to play an important role in unraveling the characteristics of the EoR.

In recent years, various radio interferometers have undertaken extensive efforts to probe the power spectrum of 21 cm fluctuations during the EoR, leading to several upper limits. These include the Giant Metre Wave Radio Telescope (GMRT; Paciga et al., 2013), the Murchison Widefield Array (MWA; Barry et al., 2019, Beardsley et al., 2016, Dillon et al., 2014, Ewall-Wice et al., 2016a, Tingay et al., 2013, Trott et al., 2020), the Donald C. Backer Precision Array for Probing the Epoch of Reionization (PAPER; Cheng et al., 2018, Kolopanis et al., 2019, Parsons et al., 2010), the Low Frequency Array (LOFAR; Gehlot et al., 2019, Mertens et al., 2020, Patil et al., 2017, van Haarlem et al., 2013), and the Hydrogen Epoch of Reionization Array (HERA; DeBoer et al., 2017, Dillon and Parsons, 2016).

More recently, The HERA Collaboration et al. (2022a, hereafter H22a) followed by The HERA Collaboration et al. (2023, hereafter H23) have published upper limits on the 21 cm power spectrum. The HERA Collaboration et al. (2022b, hereafter H22b) and H23 explored the astrophysical parameter space based on these upper limits, offering constraints on the X-ray heating of the IGM and disfavoring scenarios with inefficient X-ray heating by a redshift

of $z \sim 8$.

The power spectrum is a sufficient statistic for measuring Gaussian random fields, but the EoR 21 cm signal is expected to be highly non-Gaussian, particularly in the end stages of reionization where ionized bubbles dominate the observable signal (Kittiwisit et al., 2017, 2022, Watkinson and Pritchard, 2014, 2015). To explore non-Gaussian aspects, a simple way is to examine one-point statistics in the image domain, such as variance, skewness, and kurtosis, during the EoR, which offers a unique lens into the statistical properties of the observed 21 cm signals.

Theoretical investigations into one-point statistics using redshifted 21 cm signals aim to understand the relation between various astrophysical models and the evolution of variance and skewness measurements during reionization. For example, Wyithe and Morales (2007) examined the dependence of the variance and skewness on the mass of ionizing sources. Gluscevic and Barkana (2010) proposed that one-point statistics could potentially discern reionization mechanisms driven by large and small halos. Watkinson and Pritchard (2014) explored one-point statistics in the context of inside-out and outside-in models. There have been studies investigating the relationship between the X-ray heating and the variation in variance and skewness (Shimabukuro et al., 2015, Watkinson and Pritchard, 2015). Those theoretical efforts support the idea that one-point statistical quantities, serving as complementary statistics to power spectrum analysis, provide insights into the underlying astrophysical processes governing the reionization.

Wyithe and Morales (2007) highlighted that an MWA-500, composed of 500 dipoles, has the capability to detect skewness in 100-1000 hours and an MWA-5000, with approximately 10 times the collecting area of MWA-500, could map skewness in detail over a range of redshifts. Watkinson and Pritchard (2014) concluded that SKA observations with 1000 hr can distinguish their inside-out and outside-in models. Additionally, Kubota et al. (2016) investigated the detectability of variance and skewness based on MWA and LOFAR observations, revealing that they can achieve relatively large signal-to-noise (S/N) ratios within the redshift range between 7 and 9, assuming 1000 hr observations. These studies were conducted without accounting for the foreground effect.

Particularly, Kittiwisit et al. (2017) forecasted the measurability of one-point statistics at the EoR given the sensitivity of HERA instrument for 100-hr observations. They investigated the variance, skewness, and kurtosis considering various angular resolutions of the HERA instrument, incorporating a simple Gaussian kernel, and employing frequency binning. Their findings suggested that the full HERA array could potentially measure the one-point statistics measurements with high sensitivity assuming perfect foreground removal.

In practice, however, measuring one-point statistics is particularly challenging due to the presence of various systematic effects, including instrumental artifacts and foreground

contaminants. Especially, the foreground emissions, significantly brighter than the faint 21 cm signals by four to five orders of magnitude, need to be mitigated down to the 21 cm signal levels. Subtraction of foregrounds relying on prior knowledge is quite challenging due to the requirement of precise foreground modeling.

Due to the attributes of foregrounds having smooth frequency spectra, numerous techniques for foreground removal have been devised to isolate 21 cm signals from the foregrounds. For example, blind foreground subtraction methods based on Principal Component Analysis (PCA) have shown the effectiveness of the foreground subtraction in the image domain (e.g., Alonso et al., 2014, Cunnington et al., 2022, Spinelli et al., 2021). However, the blind methods essentially suffer from a signal loss as they remove the principal components shared by both foregrounds and 21 cm signals. An alternative method using Gaussian Progress Regression (GPR) is devised as a foreground subtraction technique in the visibility space (Mertens et al., 2018) and has been further developed with an effort to correct the signal loss (e.g., Kern and Liu, 2020, Mertens et al., 2020, Soares et al., 2021). Another commonly used method is a "foreground avoidance" method (e.g., Datta et al., 2010, Liu et al., 2014, Morales et al., 2012, Parsons et al., 2012, Pober et al., 2014, Thyagarajan et al., 2013, Trott et al., 2012, Vedantham et al., 2012). This method relies on the frequency Fourier transform of a smooth foreground spectrum, which is confined within low Fourier modes. By masking the Fourier modes occupied by foregrounds and extracting 21 cm signals, having fluctuating spectra, from unmasked Fourier modes, one may effectively filter out the foregrounds.

Under adequate foreground removal, there have been recent studies exploring detectability of non-Gaussian features in 21 cm signals. Greig et al. (2023) concluded that the SKA can detect wavelet scattering amplitudes in 100-degree field with 1000-hr integration with the foreground avoidance though it degrades the S/N of the detection. Raste et al. (2024) inferred that the SKA will be able to detect the bispectrum for $\sim$1000-hr observation at $z \sim 6$ for small k-modes, assuming optimistic foreground removal.

Kittiwisit et al. (2022) examined the impact of the foreground removal on the feasibility of measuring the one-point statistics based on HERA mock observations. They adopted a foreground avoidance approach to remove the foreground with various filtering widths. The results revealed a substantial reduction in the distinctive features of one-point statistics measurements by eliminating unique features related to overdense and ionized regions. Aggressive filtering results in more washed-out in intensity maps, accompanied by larger uncertainties. Nevertheless, they observed a degree of preservation in the increasing trends of skewness and kurtosis towards the end of reionization, especially when accounting for adequate integration time.

The previous papers that modeled detectability of non-Gaussian signals in the image domain have assumed that the maps are perfectly deconvolved and have a compact Gaussian restoring beam. In reality, the observed maps are corrupted by instrument response

or point-spread-function (PSF), forming "dirty maps" and correlated noise properties between map pixels. The classical image deconvolution algorithm CLEAN (Högbom, 1974) can be employed for the deconvolution but it only deconvolves the bright point sources. Multi-resolution CLEAN algorithms were considered to handle the deconvolution of diffuse emission and unresolved point sources (Cornwell, 2008, Wakker and Schwarz, 1988), but the deconvolution process can modify map noise properties and introduce artificial faint residual emission, making the statistics very difficult to interpret. Therefore, we choose instead to work directly with dirty maps, forward modeling the 21 cm signals and modeling the noise covariance matrix by propagating the measurement equation and the analysis procedure from true sky to measured image. This procedure fully takes into account the actual PSF of an interferometer.

In this study, we extend the work of Kittiwisit et al. (2022) by considering realistic primary beam models and instrument PSF, an advanced technique for foreground removal, and robust calculation of thermal noise through an optimal mapping technique. We utilize the foreground removal method developed by Ewall-Wice et al. (2020) based on Discrete Prolate Spheroidal Sequences (DPSS; Slepian, 1978) in the context of foreground avoidance, and show their effectiveness compared to the standard foreground avoidance approach and the blind PCA foreground subtraction method. We measure one-point statistics on the HERA Phase I observations after the foreground removal for the first time and examine the effect of instrument systematics residuals on the non-Gaussianity characteristics in the data. Furthermore, we evaluate 21 cm models based on one-point statistics measurements. Ultimately, we investigate the feasibility of measuring one-point statistics measurements for a fiducial 21 cm model, considering the observational sensitivity of future observations (HERA Phase II observations).

While higher-order statistics in the Fourier domain, such as the 21 cm bispectrum (e.g., Hutter et al., 2019, Kamran et al., 2021, Majumdar et al., 2018, 2020, Mondal et al., 2021, Raste et al., 2024, Shimabukuro et al., 2016, Watkinson et al., 2018, 2022, Yoshiura et al., 2015), offer a means to investigate non-Gaussianity, direct exploration of one-point statistics from tomographic images of the 21 cm signal has a power in its simplicity. We have chosen for this study to work in the image domain to explore instrumental and residual foreground effects, which are likely to be more straightforward to interpret in the image domain.

This paper is structured as follows: In Section 2.2, we provide an overview of the optimal mapping method utilized for mapmaking. Detailed information on observational and simulation data is presented in Section 2.3. Sections 2.4.1 outlines the advanced foreground removal technique and Section 2.4.2 discuss its efficacy in comparison to other methods. The forward modeling of noise data is discussed in Section 2.6. Section 2.7 provides details about the construction of data cubes used in our analysis. The discussion on the impact of systematics residuals in the observed map post-foreground removal is presented in Section 2.8. In

Section 2.9, we explore one-point statistics measurements for various 21 cm models in comparison with the HERA Phase I observational data. Additionally, Section 2.10.1 forecasts the feasibility of detecting one-point statistics and the uncertainty of model parameters in future observations. Finally, Section 2.11 summarizes the key findings of this study. Throughout the paper, we adhere to the cosmological parameters from Planck Collaboration et al. (2020): $\Omega_\Lambda = 0.6844$, $\Omega_\mathrm{m} = 0.3156$, $\Omega_\mathrm{b} = 0.04911$, and $H_0 = 67.27\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

2.2 Mapmaking with Direct Optimal Mapping

In this section, we outline the mathematical formalism employed in mapmaking with the Direct Optimal Mapping (DOM) technique (Dillon et al., 2014, 2015, Xu et al., 2022). DOM is considered optimal as it down-weights the impact of data with high noise variance. DOM offers two notable advantages: a robust calculation of statistical noise for each pixel and the ability to perform mapping without assuming a flat sky. The algorithm compensates for the interferometric "w-term" exactly, without the need to perform a w-projection (Cornwell et al., 2008), and it makes use of the actual (u, v) – spatial frequency – coordinates without the need to perform gridding and resampling. Nonzero w-terms can arise with wide-field mapping and with non-coplanar arrays. The HERA dishes deviate from coplanarity at the level of a few centimeters, and Xu et al. (2022) showed that neglecting such deviations result in map errors of about 5%.

In contrast to algorithms that work directly with visibility data for power spectrum estimation (Datta et al., 2010, Liu et al., 2014, Morales et al., 2012, Morales et al., 2018, Parsons et al., 2012, Pober et al., 2014, Thyagarajan et al., 2013, Trott et al., 2012, Vedantham et al., 2012) and with algorithms that compute a model of the sky in a Fourier basis ("m-modes"; Eastwood et al., 2018, Shaw et al., 2014), DOM requires a pixelized model of the sky which is the dominant source of infidelity in the images (assuming perfect calibration). For general applications of mapping, the direct inversion of DOM is expensive to compute. However, with HERA we can compute the maps more efficiently, because the instrumental design matrix (the $\mathbf{A}$ matrix defined below) does not change with time and can be pre-computed and stored, somewhat similar to the "Fast" implementation of Fast Holographic Deconvolution (Sullivan et al., 2012).

More specifically, a discretized version of the interferometric measurement (or visibility) equation is

$$V_{ij}(\nu, t) = \sum_{n=1}^{N_\mathrm{src}} B_{ij}(\hat{s}_n, \nu) I_\nu(\hat{s}_n, \nu) e^{-2\pi i \nu \mathbf{b}_{ij}(t)\cdot\hat{s}_n/c} \Delta\Omega, \tag{2.1}$$

where i and j denote an antenna pair, B_{ij} is a peak-normalized primary beam that is the

multiplication of far-field electric fields of two antennas, I_ν is the specific intensity of sky sources in Jansky per steradian (Jy/sr), $\hat{s}_n$ is the position vector of the celestial objects, and $\mathbf{b}(t)$ is the baseline (antenna pair) vector that changes with the Earth rotation. The exponential term is called the fringe term that represents the phase coming from cross-correlating signals measured by the antenna pair and associated with a geometric delay.

For a given frequency channel, Equation (2.1) can be equivalently put in a matrix form,

$$\boldsymbol{v} = \mathbf{A}\mathbf{x} + \mathbf{n}, \tag{2.2}$$

where $\mathbf{x}$ is a vector with $N_{\rm pix}$ elements of the sky flux density, and $\boldsymbol{v}$ and $\mathbf{n}$ are vectors containing $N_{\rm vis}$ elements of resultant visibility measurements and associated thermal noise, respectively. The $\mathbf{A}$ matrix includes the primary beam and the fringe term, described in Equation (2.1), with the shape of $N_{\rm vis} \times N_{\rm pix}$.

The mapmaking is a backward process of the interferometric measurement equation, and the optimal estimator has been studied extensively by Tegmark (1997) (see also Dillon et al. (2015) for review). The optimal estimator can be expressed as,

$$\hat{\mathbf{x}} = \mathbf{D}\mathbf{A}^\dagger\mathbf{N}^{-1}\boldsymbol{v}, \tag{2.3}$$

where $\hat{\mathbf{x}}$ is an optimal estimator, $\mathbf{D}$ is a normalization matrix with the shape of $N_{\rm pix} \times N_{\rm pix}$, and $\mathbf{N}^{-1}$ is the inverse of a noise matrix $\mathbf{N} = \langle \mathbf{n}\mathbf{n}^\dagger \rangle$ with the shape of $N_{\rm vis} \times N_{\rm vis}$. $\mathbf{A}^\dagger$ is a conjugate transpose of $\mathbf{A}$. An ensemble average of the estimator is $\langle \hat{\mathbf{x}} \rangle = \mathbf{D}\mathbf{A}^\dagger\mathbf{N}^{-1}\mathbf{A}\mathbf{x} = \mathbf{P}\mathbf{x}$ where $\mathbf{P} = \mathbf{D}\mathbf{A}^\dagger\mathbf{N}^{-1}\mathbf{A}$. The choice of $\mathbf{N}^{-1}$ is considered, as the weighting matrix makes the map "optimal" in the sense that it preserves information when estimating parameters from the images (Tegmark, 1997). Other criteria, such as a weighting that produces lower sidelobes in the synthesized beam, could be incorporated into the weighting matrix instead. Since we seek to measure faint cosmological signals, we choose the $\mathbf{N}^{-1}$ weighting.

To estimate the noise matrix $\mathbf{N}$ in Equation (2.3), we assume each visibility has uncorrelated thermal noise, which results in a diagonal matrix with the diagonal elements of σ_n^2,

$$\sigma_n^2 = \frac{V_{ii}V_{jj}}{\Delta\nu\Delta t}, \tag{2.4}$$

where V_{ii} indicates autocorrelation for the baseline i, and $\Delta\nu$ and Δt denote the width of frequency channel and integration time, respectively.

The choice of $\mathbf{D}$ matrix depends on what we pursue in mapmaking. For example, the least square estimator suggests $\mathbf{D} = [\mathbf{A}^\dagger\mathbf{N}^{-1}\mathbf{A}]^{-1}$ that deconvolves instrumental effects and recovers the true map. One problem of this approach is that $\mathbf{A}^\dagger\mathbf{N}^{-1}\mathbf{A}$ is usually ill-conditioned, which may cause a direct inverse that can magnify small perturbations in the visibility data

significantly. Pseudo-inverse can be an alternative but may require a proper regularization of noises, which is out of scope of this study.

We opt for the normalization scheme adopted by Xu et al. (2022). They defined a diagonal matrix $\mathbf{D}$ with $D_{ii} = \left(\sum_j |A^\dagger_{ij}| N^{-1}_{jj}\right)^{-1} = \left(\sum_j B_{ij} N^{-1}_{jj}\right)^{-1}$ where i and j indicate pixel and baseline-time indices, respectively. By assuming that the noise variance of visibility does not change significantly over time, which is approximately true for slowly varying autocorrelations, the matrix can be simplified to $D_{ii} \approx \bar{\sigma}^2_n / \left(\sum_t B_{it}\right)$ where $\bar{\sigma}^2_n = \left(\sum_b 1/\sigma^2_b\right)^{-1}$ with baseline b and B_{it} is the primary beam at pixel i and time t. This normalization is designed to preserve the flux density of a source at zenith. For example, the diagonal elements of the $\mathbf{P}$ matrix can be expressed as

$$P_{ii} = \sum_j D_{ii} (A^\dagger)_{ij} N^{-1}_{jj} A_{ji} \tag{2.5}$$

$$= D_{ii} \sum_j \frac{B^2_{ij}}{\sigma^2_j} \tag{2.6}$$

$$\approx \frac{\bar{\sigma}^2_n}{\sum_t B_{it}} \frac{\sum_t B^2_{it}}{\bar{\sigma}^2_n} = \frac{\sum_t B^2_{it}}{\sum_t B_{it}}. \tag{2.7}$$

This indicates a PSF-peak normalization at the center of primary beam for the peak-normalized primary beam. Note that signals in PSF-convolved maps will be attenuated by the primary beam as they are getting away from the zenith. We choose not to renormalize the map to remove the weighting by the primary beam (the "primary beam correction" of radio astronomy), because such a correction severely impacts the noise statistics in the map. For instance, introducing $D_{ii} = \left(\sum_j B^2_{ij} N^{-1}_{jj}\right)^{-1}$ can result in $P_{ii} = 1$ when N_{jj} does not change much over time, but this will diverge the noise variance at the pixel location far away from the beam center. See Xu et al. (2022) for more details.

An example of the PSF at zenith, which corresponds to a column of $\mathbf{P}$ matrix at the zenith location, is illustrated in Figure 2.1. The left panel is for the PSF of the HERA dipole interferometer and the right one is a Gaussian beam which is commonly assumed in other studies for simplicity. Due to the complexity, it is expected that convolving with the realistic PSF can result in correlation between pixels unlike the simple Gaussian beam. Hence, we adopt the forward modeling of simulation data, including 21 cm signals and noise simulations, to account for the effects of the correlations on one-point statistics and to ensure a fair comparison with the observational data.

The map obtained from Equation (2.3) is expressed in the unit of Jy/beam for the visibility in Jy. The map is then converted to mK units using the following equation,

$$\kappa(\nu) = \frac{c^2}{2 k_B \nu^2 \Omega_{\rm syn}} \times 10^{-23} \quad \left[\frac{\rm mK}{\rm Jy}\right], \tag{2.8}$$

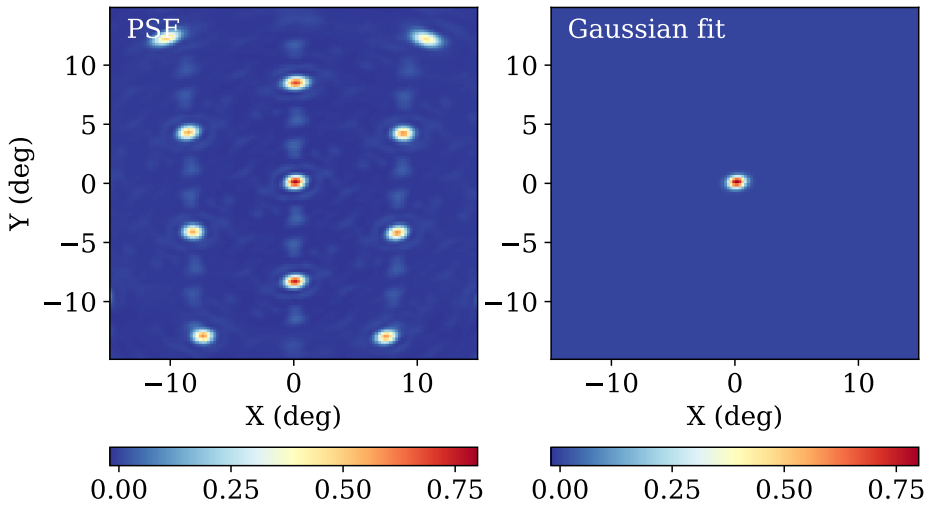

Figure 2.1: PSF of the HERA Phase I observation at 161.5 MHz (left) and its Gaussian fit (right). We forward model the simulated maps using the PSF to consider the correlated properties between pixels. The PSF in this example is integrated over 1 hour. The full width at half maximum of the Gaussian fit is 1.11° for the semi-major axis and 0.75° for the semi-minor axis.

where c is the speed of light, k_B is the Boltzmann constant and $\Omega_{\rm syn} = \Omega_{\rm ant}/{\rm N}_{\rm ant}$ is the solid angle of the synthesized beam which is equal to the solid angle of the primary beam divided by the number of antennas.

HERA observations provide east-east (EE) and north-north (NN) polarized visibilities. Maps are produced for each polarization and combined to form pseudo stokes-I maps, $\hat{\mathbf{x}}_I = (\hat{\mathbf{x}}_{EE} + \hat{\mathbf{x}}_{NN})/2$. This could potentially mitigate systematics observed at high delay modes, which may originate from Faraday-rotated pulsar emissions and uncorrelated systematics across polarizations (H22a). For simulated mock observations with no systematics, we simulate EE polarization only, assuming $\hat{\mathbf{x}}_I \approx \hat{\mathbf{x}}_{EE}$.

2.3 Data for Phase I Observation

In this section, we describe the observational data and simulations which are used for the one-point statistics study on the Phase I observations.

2.3.1 Observational Data

HERA is a draft-scanning interferometer located in South Africa. We utilize the calibrated data for Phase I HERA observations between October 2017 and March 2018 as in H23.

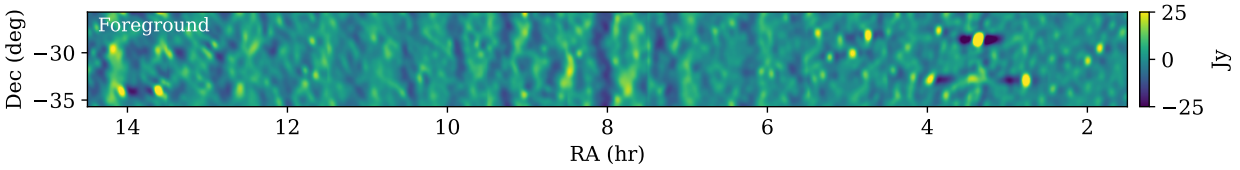

Figure 2.2: HERA stripe, displaying foreground emission, ranging from 1.5 to 14.5 hours centered at -30.7° in declination. The map is created by connecting sky patches which are made at every hour of LST with the size of 15×10 square degrees. Emissions at around 3 hr are grating lobes of Fornax A, which is one of the brightest foreground radio sources.

Based on nightly inspections of the data, 94 nights are chosen to be good quality data. Each night data is calibrated through the HERA calibration pipeline which uses redundant-baseline calibration (e.g., Dillon et al., 2020) followed by absolute calibration to solve for degenerate parameters remaining after the redundant calibration (e.g., Kern et al., 2020a, Li et al., 2018). Artificial Radio Frequency Interference (RFI) is then removed by excluding outliers in visibilities and calibration gain solutions. For LST grids with fixed bins of 21.4 s, all night observations are coherently averaged, which increases the signal-to-radios (S/N) significantly.

Two major systematics of the Phase I observation are a cable reflection and over-the-air crosstalk between antennas, which are mitigated through calibration using autocorrelations and fringe-rate filtering, respectively (Kern et al. 2019, 2020b, H22a, H23). The systematics correspond to high delay modes, which can appear in delay-filtered maps. Therefore, we investigate the effect of the systematics in our mapping before and after the mitigation in Section 2.8.

The final visibility product of the pipeline for power spectrum estimation is coherently time-averaged over 214 s, but we use data without this time averaging since our mapping scheme performs an appropriate coherent averaging over time. See H22a and H23 for more details about data reduction.

Out of 1024 frequency channels over 100–200 MHz, frequency ranges of 117.19–133.11 MHz and 152.25–167.97 MHz, corresponding to Band 1 and Band 2 of H23, are considered. These bands were chosen because they were relatively RFI-free. They target the redshift of ∼10.4 for the low band and ∼7.9 for the high band.

The observation was conducted with about 40 antennas, with the actual number varying at different LSTs based on the availability of antennas at each LST. Specifically, antennas that exhibit more than 50% good quality data over 94 nights at each LST (see "Unflagged antennas" in Figure 4 of H23) were included. We exclude baselines with projected east–west distances less than 14 m because the crosstalk mitigation is not applied to the baselines. Timestamps with less than 10 samples are removed due to potential large error contributions.

Instead of mapping the full sky at once, we construct a series of small patches of sky, each measuring 15×10 square degrees. These patches are centered at each hour of LST, ranging from 2 to 14 hr. Each patch consists of 66×44 pixels and is integrated over 1.6 hr.

Figure 2.2 shows a stitched map of the sky patches. We split the stitched map into four different regions for our one-point statistics analysis: Field I (1.5–2.5 hr), Field II (4.5–5.5 hr), Field III (6.5–9.5 hr), and Field IV (9.5–13.5 hr). Each field roughly corresponds to Field B, C, D, and E which are defined in H23. Field I and II are centered at 2 and 5 hr respectively, across 15 degrees, which are known to be foreground quiet radio sky. There are bright foregrounds in the 3-hr region which corresponds to grating lobes of Fornax A, which is a bright radio galaxy located at (RA, Dec) = (3.4^h, -37.2°). Thus, we do not include the region. Field III includes the anti-galactic center, centered at ∼8 hr, and its sidelobes. Field IV is the region between the anti-galactic center and galactic center.

2.3.2 Simulation Data

21 cm Signal Simulations

The brightness temperature of 21 cm signals is given by the contrast between the 21 cm spin temperature and the background radiation temperature (i.e., CMB temperature), which can be written as (e.g., Furlanetto et al., 2006),

$$\begin{aligned}\delta T_b(\nu) &= \frac{T_{\rm S}-T_\gamma}{1+z}\left(1-e^{-\tau_{21}}\right)\\ &\approx 27 x_{\rm H_I}(1+\delta_{\rm m})\left(\frac{H}{dv_\parallel/dr+H}\right)\left(1-\frac{T_\gamma}{T_{\rm S}}\right)\\ &\times\left(\frac{1+z}{10}\frac{0.15}{\Omega_{\rm M}h^2}\right)^{1/2}\left(\frac{\Omega_{\rm b}h^2}{0.023}\right)\,{\rm mK},\end{aligned} \tag{2.9}$$

where $T_{\rm S}$ is the gas spin temperature, T_γ is the CMB temperature, τ_{21} is the optical depth at the 21 cm frequency, $x_{\rm H_I}$ is the neutral fraction of the hydrogen gas, and $\delta_{\rm m}=\rho/\bar{\rho}-1$ is the matter density fluctuations where ρ and $\bar{\rho}$ are the matter density and its mean, respectively. $dv_\parallel/dr$ is the gradient of the peculiar velocity along the line-of-sight. This equation provides the baseline for computing the 21 cm brightness temperature.

To generate 21 cm brightness temperature cubes with associated astrophysical properties, we use a publicly available simulator, `21cmFAST` (Mesinger et al., 2011, Murray et al., 2020). We investigate the non-Gaussian characteristics of neutral hydrogen signals under different astrophysical conditions, by considering four different EoR scenarios. Those scenarios include the fiducial model studied by Park et al. (2019), and three derived models suggested by Greig

Table 2.1: Parameter values for different EoR models from Greig et al. (2022).

Model	$\log_{10}(f_{*,10})$	α_*	$\log_{10}(f_{\rm esc,10})$	$\alpha_{\rm esc}$	t_*	$\log_{10}(M_{\rm turn})$ $(M_\odot)$	$\log_{10}\frac{L_{X<2\,\rm keV}}{\rm SFR}$ $(\rm erg\,s^{-1}\,M_\odot^{-1}\,yr)$	E_0 (keV)	$\alpha_{\rm X}$
Fiducial	-1.30	0.50	-1.00	-0.50	0.50	8.7	40.50	0.50	1.00
Cold reionization	-1.30	0.50	-1.00	-0.50	0.50	8.7	38.00	0.50	1.00
Large halos	-0.70	0.50	-1.00	-0.50	0.50	9.9	40.50	0.50	1.00
Extended reionization	-1.65	0.50	-1.00	-0.50	0.50	8.0	40.50	0.50	1.00

et al. (2022) including cold reionization, large halos, and extended reionization models.

- Fiducial model: a baseline model which is constrained by the observed galaxy UV luminosity function at high redshift along with mock 21 cm observation (Park et al., 2019).
- Cold reionization: a model with restricted X-ray heating of IGM with enhanced 21 cm signal power compared to the fiducial model (Mesinger et al., 2014, Parsons et al., 2014). This model is disfavored by power spectrum studies (H22b, H23).
- Large halos: a model with a large characteristic (turnover) halo mass for quenching of star formation in halos. The reionization is mainly driven by bright star-forming galaxies.
- Extended reionization: a model with less efficient star formation rate which delays the reionization. The reionization is primarily driven by faint star-forming galaxies.

The key parameters that determine each model are described in Table 2.1 where $f_{*,10}$ is a normalization factor that determines the fraction of the gas in stars, $M_{\rm turn}$ is the turnover mass that represents a halo mass for quenching of star formation, and $\frac{L_{X<2\,\rm keV}}{\rm SFR}$ is the luminosity of soft X-ray per star formation rate, responsible for IGM heating. Other key parameters – including a power index associated with the gas fraction in stars (α_*), the characteristic star-forming timescale (t_*), parameters associated with the escape fraction of ionizing photons ($f_{\rm esc,10}, \alpha_{\rm esc}$), the minimum threshold energy of the X-ray spectrum (E_0), and the X-ray spectrum slope ($\alpha_{\rm X}$)– are fixed for the distinct models as detailed in Table 2.1. In this study, we do not account for the subgrid model of recombination proposed by Park et al. (2019) to enable faster simulations. Instead, we fix the mean free path parameter, denoted as `R_BUBBLE_MAX`, at 15 Mpc across all simulations.

The models are simulated over a redshift range of 6.5–13.0 with a cube size of 200^3 $\rm Mpc^3$ with 128 pixels per side. The coeval cubes are then interpolated to a whole sky map using Hierarchical Equal Area isoLatitude Pixelization (`HEALPix`; Gorski et al., 2005) at the redshift of interest, forming a lightcone data over redshifts. We interpolate the cubes into `nside` = 8192 to minimize interpolation artifacts and down-sample it to a computationally

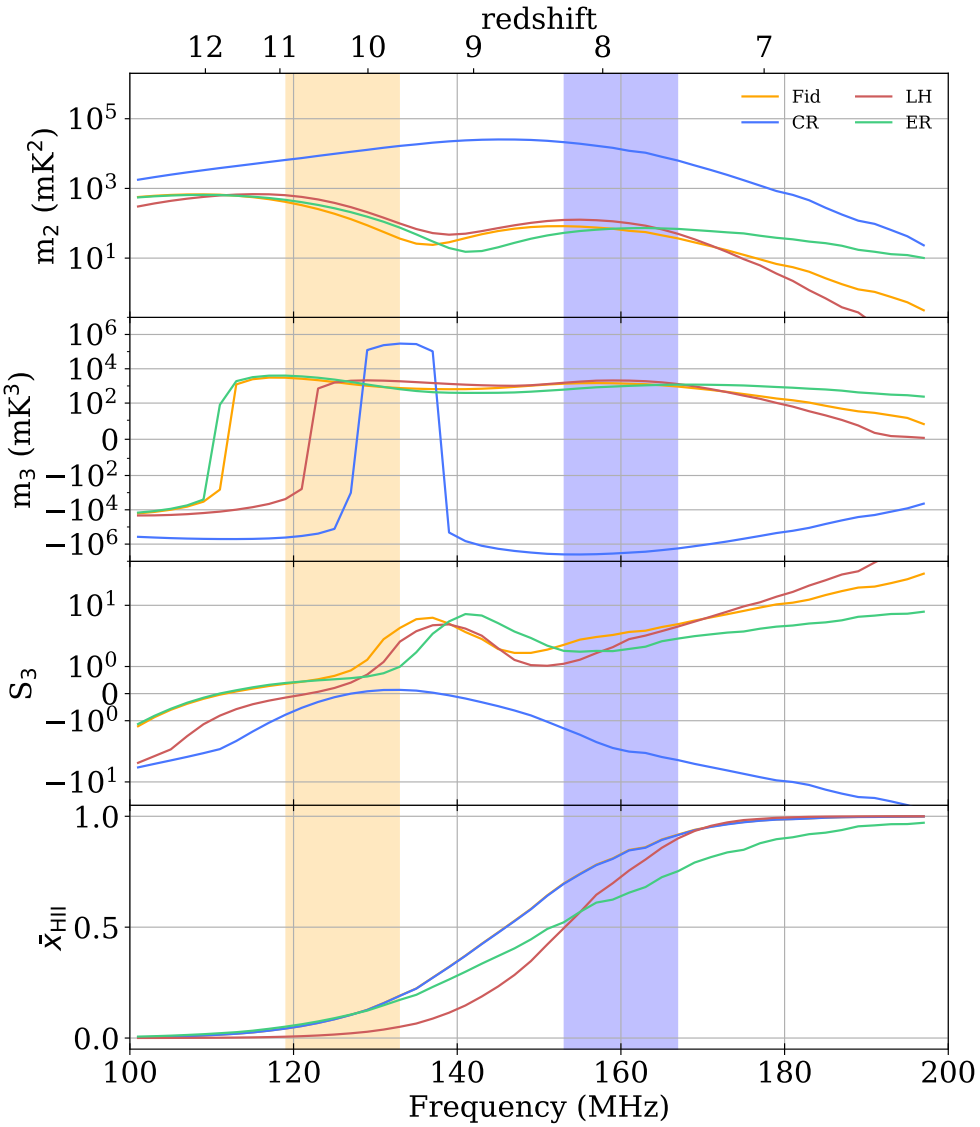

Figure 2.3: Evolutionary history of m_2 (equivalent to S_2), m_3, S_3, and mean ionization fraction measured from raw simulations for the four distinct 21 cm models. Each line indicates the fiducial (Fid), cold reionization (CR), large halos (LH), and extended reionization (ER) models. Band 1 and Band 2 are highlighted in orange and blue shaded regions.

feasible `nside` = 128. The details of this process are described in Appendix A of Kittiwisit et al. (2017).

A statistical moment is a quantitative measure that provides information about the shape and characteristics of a distribution of data. The p-th moment is defined as

$$m_p = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^p, \tag{2.10}$$

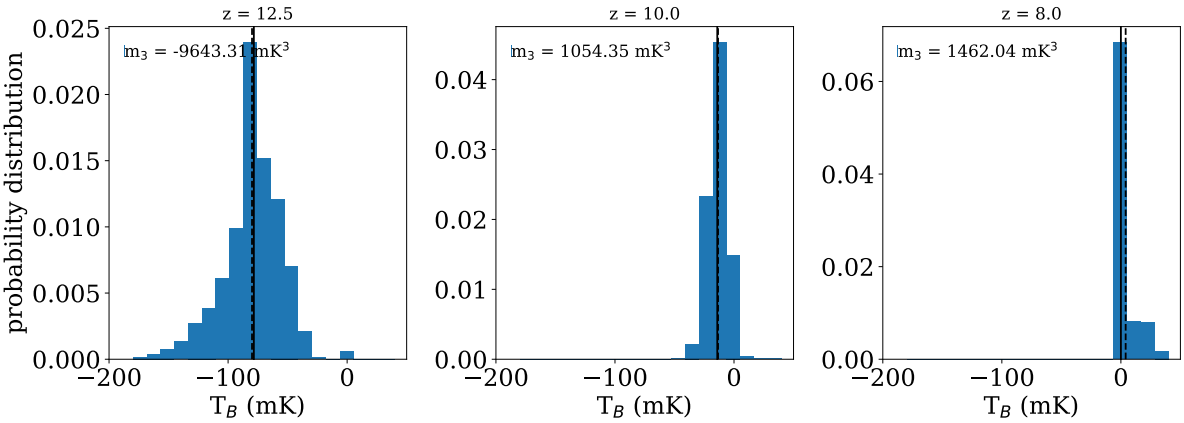

Figure 2.4: Probability distribution of the fiducial model at specific redshifts for raw simulations (i.e., no instrument effect). The dashed and solid vertical lines indicate the mean and median, respectively. At $z = 12.5$, the distribution exhibits an asymmetric shape due to the heating, leading to a negative m_3. As reionization advances, a prominent pileup at the zero bin becomes noticeable. Simultaneously, heating by star-forming galaxies raises the temperature of the IGM, resulting in a tail towards positive values and thus a positive m_3.

where N denotes the number of data, and μ represents the mean of the dataset.

Specific quantities we are interested in are variance and skewness. The variance measures the dispersion of the data which is closely related to a power spectrum. The skewness provides information about the asymmetry of the distribution that potentially captures non-Gaussianity. Mathematically, they are expressed as

$$S_2 = m_2, \tag{2.11}$$
$$S_3 = m_3/(m_2)^{3/2}, \tag{2.12}$$

where m_2 and m_3 are the second and third moments, and S_2 and S_3 represent variance and skewness, respectively.

Figure 2.3 shows the evolution of the second moment (first), third moment (second), skewness (third), and mean ionization fraction (last) measured from raw simulations (i.e., no instrumental effect) for each model. The large variance of cold reionization model compared to the others arises from the small spin temperature as inferred from Equation (2.9), which makes the model particularly interesting given our observation limit. The skewness, indicative of non-Gaussianity, diverges at the high frequency end as the ionization due to star-forming galaxies becomes significant. The cold reionization model exhibits a distinct skewness evolution toward negative values because the spin temperature of the model is relatively smaller than the background radiation temperature. The distinct reionization histories driven by different primary ionizing sources (large halos vs extended reionization) may be characterized by both variance and skewness measurements.

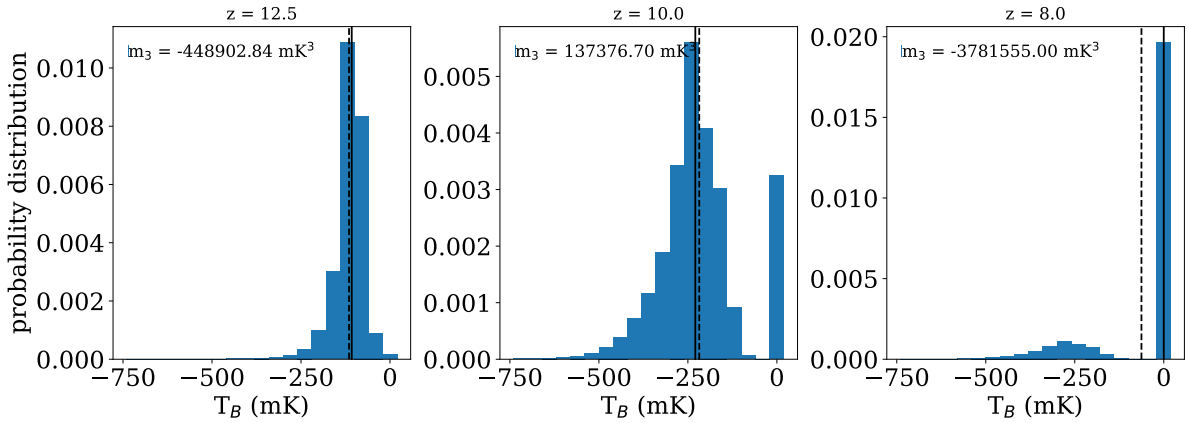

Figure 2.5: Probability distribution of the cold reionization model at specific redshifts for raw simulations (i.e., no instrument effect). The dashed and solid vertical lines indicate the mean and median, respectively. At $z = 12.5$, similar to the fiducial model, the cold reionization model shows a skewed distribution due to heating, resulting in a negative m_3. At $z = 10.0$, although the distribution centered around $T_B \sim -250$ mK exhibits negative tails, m_3 transitions to positive values due to the ionized IGM at the zero bin. As reionization progresses further, m_3 becomes negative again as a large fraction of the IGM is ionized, leaving behind some negative IGM regions.

More specifically, the fiducial model undergoes a sign flip in m_3 around $z \sim 12$. This behavior can be understood by examining the evolution of the brightness temperature probability distribution. Figure 2.4 shows the distribution for the fiducial model. At $z \sim 12.5$, the model exhibits an asymmetric distribution with a negative tail, resulting in a negative m_3. As the IGM heats up and becomes ionized, the distribution shifts in the opposite direction, developing a positive tail with the mean value centered around 0 mK, leading to a positive m_3.

An interesting feature is observed in the cold reionization model, shown in Figure 2.3, where m_3 undergoes two sign flips, unlike the other models. This can also be explained by the probability distribution at different redshifts. Figure 2.5 illustrates the brightness temperature distribution for the cold reionization model. Similar to the fiducial model, at $z \sim 12.5$, the model has a skewed distribution with a negative m_3. As the universe evolves, the IGM becomes reionized but remains cold due to inefficient heating. This reionization increases the height of the zero bin, resulting in a positive m_3. However, once a large fraction of the IGM is ionized, m_3 turns negative again due to some remaining negative bins.

The fiducial model is used to test foreground subtraction methods in Section 2.4 and all four models are employed to compare with observational data in Section 2.9.

The simulated whole sky maps are then transformed into our mock radio observation using Equation (2.1)[1]. The mock observations are based on the antenna layout of the HERA

[1]We utilize a simulator provided by https://github.com/vispb/vispb.

Table 2.2: Summary of Observational and Simulation Data for mapmaking.

Data	Description	Field
Observation (Phase I)	With Systematics Mitigation	I, II, III, IV
	Without Systematics Mitigation	I, II, III, IV
Simulation (Phase I)		
Foreground + 21 cm (fiducial)	Foreground Subtraction Validation	II
21 cm (four models)	EoR Detection Test Model	I, II, III, IV
21 cm (cold reionization)	Cosmic Variance Estimation	I, II, III, IV
Thermal Noise	Noise limit & Uncertainty of Mock Observation	I, II, III, IV
Simulation (Phase II)		
21 cm (fiducial)	Forecasting Future Observation	3.5–9.5 hr
21 cm (perturbed)	Fisher Matrix Analysis	3.5–9.5 hr
Thermal Noise	Sensitivity on Prediction	3.5–9.5 hr

Phase I observation (H23) using a dipole feed beam for the primary beam (Fagnoni et al., 2020). The simulated EoR signals described above are employed for the intensity of sources. The visibility simulations cover a frequency range of Band 1 and Band 2, with a spacing of 97.7 kHz. The mock observations span a 0–16 hr LST window that covers the fields of interests as described in Section 2.3.1, at a time cadence of 21.4 seconds.

In order to account for the cosmic variance in one-point statistics measurements, we sample the sky at 30 different locations and construct mock visibilities for the cold reionization model.

Foreground Simulations

To assess the performance of foreground removal techniques in Section 2.4, we also create foreground visibilities by using Equation (2.1). We choose the LST of 5 hr field for the test of foreground removal. For the foreground sources, we include compact radio components and diffuse sky models. The former is drawn from the GaLactic and Extragalactic All-sky MWA (GLEAM; Hurley-Walker et al., 2017) survey, and the latter is from a Global Sky Model (GSM; Zheng et al., 2017).

For the point-source sky model, we incorporate both GLEAM I (Hurley-Walker et al., 2017) and GLEAM II (Hurley-Walker et al., 2019) catalogs, along with the peeled bright sources documented in Table 2 of Hurley-Walker et al. (2017) and Fornax A (e.g., Bernardi et al., 2013). The GLEAM I survey essentially spans the entire southern hemisphere and achieves 95% completeness at 160 mJy, excluding regions corresponding to the Large and Small Magellanic Clouds and the galactic plane stripe. The GLEAM II catalog aims to fill the gap within the diffuse galactic plane. While the second catalog is not entirely complete, observations at LST = 5 hr are minimally affected by this incompleteness, as the primary

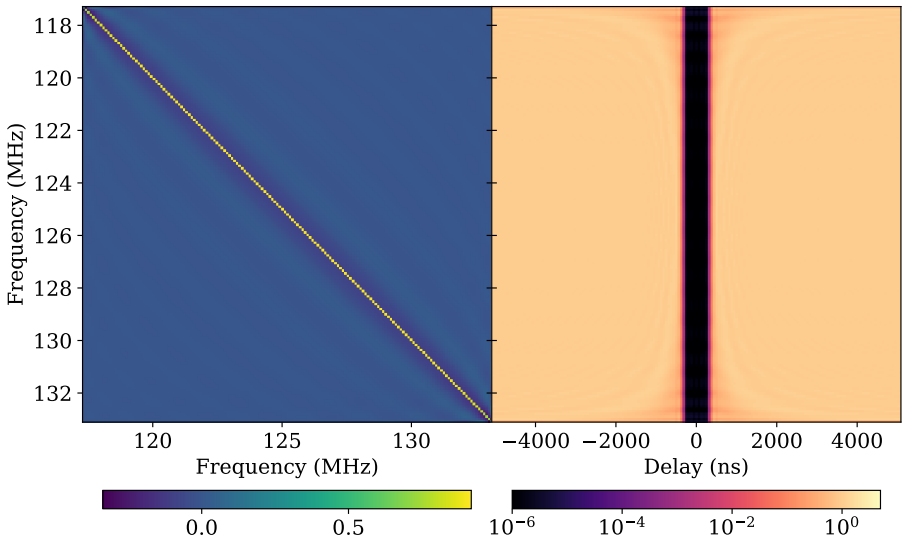

Figure 2.6: A filter matrix for a 25.3 m baseline in the frequency × frequency dimension (left) and in the frequency × delay dimension (right). The matrix on the left represents the filter to be applied to the visibility data. In the right panel, it is illustrated how the filter effectively suppresses input signals within the specified filter size at each frequency channel.

beam's main lobe is sufficiently distant from the galactic plane. In addition, we try to fill in the gap using a GSM described below.

Zheng et al. (2017) provides diffuse GSM maps with the bright compact sources removed from the sky model. This precaution helps minimize the risk of double-counting bright point sources when merging the point-source and diffuse sky models. The GSM offers high-resolution models in `HEALPix`, pixelized at `nside` = 1024 for each frequency. The maps are then downgraded to `nside` = 256 to ensure computational feasibility. The combination of the point source survey and GSM equips us adequately to investigate the impact of foregrounds on our foreground subtraction techniques, when the simulated foreground visibilities are added to the EoR visibility simulations.

2.4 Foreground Removal

2.4.1 Wedge-filtering using `DAYENU`

In this section, we discuss a foreground filtering method using DPSS within the context of wedge-filtering, followed by a comparative analysis of this method with other approaches, including a standard wedge-masking method based on Fourier basis and foreground subtraction using PCA in Section 2.4.2.

Wedge-filtering is a tool to avoid the foreground that occupies low Fourier modes and to extract cosmological signals located at high delay modes. Specifically, we employ the 2D power spectrum defined by the Fourier modes $k_\perp$ and $k_\parallel$, perpendicular and parallel to the line-of-sight, respectively. The foreground with smooth spectral structure tends to be located at low $k_\parallel$ in the 2D power spectrum. The size of the foreground region along the $k_\parallel$ axis in the power spectrum increases with $k_\perp$ due to the chromaticity of the interferometer, forming the foreground wedge. Because the 21 cm signals are expected to have complex spectral structure, they can exist outside the foreground wedge, called the EoR window (Datta et al., 2010, Liu et al., 2014, Morales et al., 2012, Morales et al., 2018, Parsons et al., 2012, Pober et al., 2014, Thyagarajan et al., 2013, Trott et al., 2012, Vedantham et al., 2012). The wedge-filtering is then selecting $k_\parallel$ modes in the EoR window, larger than a criterion. The criterion can be chosen based on the field-of-view (FoV) of an instrument that determines a maximum $k_\parallel$ or delay mode occupied by foregrounds,

$$\tau_{\rm max} = \frac{b \sin\theta}{c}, \tag{2.13}$$

where b is a separation between an antenna pair, θ is the angular extent of a dish, and c is the speed of light (Liu et al., 2014, Parsons et al., 2012). Note that $k_\parallel \propto \tau$.

For HERA 21 cm cosmology measurements, we must consider the FoV to be down to the horizon of the observer's frame, and thus the wedge-filtering size is subject to the horizon limit ($\tau_{\rm horizon} = b/c$). In addition, we may need to apply an extra buffer ($\tau_{\rm buffer}$) above the horizon limit due to the chromaticity of the primary beam, especially when there is strong horizon emission convolved with the primary beam (Thyagarajan et al., 2016) as well as systematics such as calibration errors (Kim et al., 2022, 2023, Orosz et al., 2019). For instance, $\tau_{\rm buffer} = 300$ ns is considered in H23 for their power spectrum estimation. In this study, we choose the same buffer size of 300 ns for the wedge-filtering.

More specifically, one way of wedge-filtering uses an inverse covariance matrix which is defined in the delay (Fourier) space by a top-hat filter as suggested by Ewall-Wice et al. (2020). They developed a linear foreground filtering method based on the DPSS basis which has two key characteristics: 1) it diagonalizes the covariance matrix in the frequency domain, which is in the form of sinc function (i.e., $\mathrm{sinc}\, x = \sin x/x$), and 2) it maximizes foreground powers within the top-hat filtering window in the Fourier space. These attributes contribute to the effectiveness and successful performance of a filtering process.

We utilize the DPSS Approximate lazY filtEriNg of foregroUnds (`DAYENU`; Ewall-Wice et al., 2020) method[2] to filter out foregrounds per baseline. The filtering process is governed by two essential parameters: the filtering size and a suppression factor for foregrounds within the filtering window. Figure 2.6 represents an example of the filter for a 25.3 m baseline. In

[2]The python library is publicly available in https://github.com/HERA-Team/uvtools.

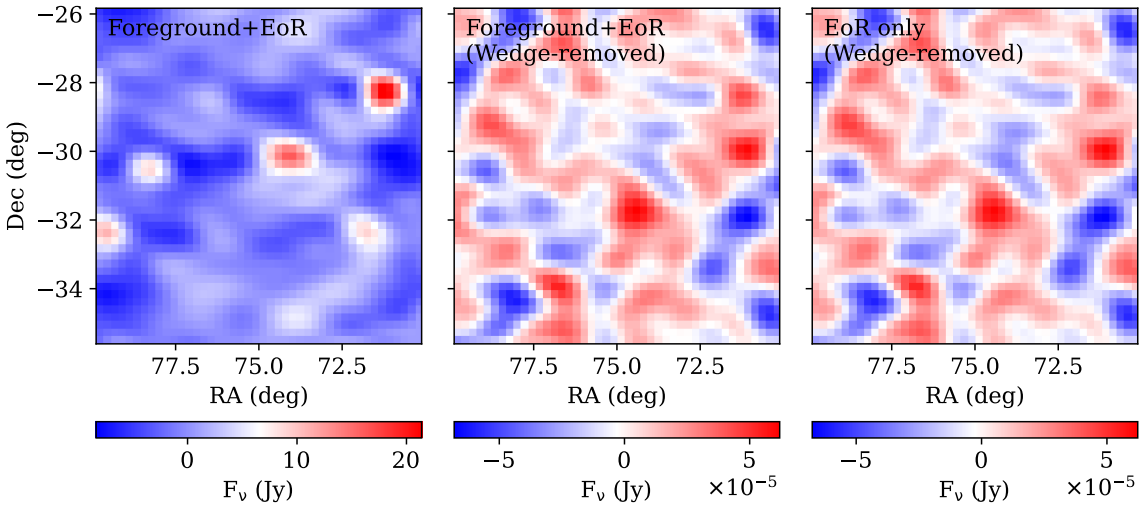

Figure 2.7: Unfiltered and filtered maps of noiseless simulations at 160.0 MHz using the `DAYENU` filter. The first panel presents a map of foreground and EoR models at LST of 5-hr field, The field is dominated by two point sources and it is hard to detect the EoR signal. In the second panel, wedge-filtered result of the first panel reveals that the amplitude of the map is reduced by a factor of 10^5. We found the filtered results for foreground + EoR (second panel) and EoR-only (third panel) simulations are consistent within $\sim$10% errors, which demonstrates foreground residuals are sufficiently suppressed.

the left panel, the filter to be applied to visibility data is depicted, and the right panel shows effective suppression of signals within the filter size to 10^{-6} in the Fourier domain. Due to the sub-optimal performance at the band edges, we exclude the first and last four channels in our analysis.

The wedge-filtered visibilities are turned into images using Equation (2.3). Figure 2.7 shows maps before and after wedge-filtering using `DAYENU` for the noiseless simulation data with foreground and fiducial EoR models at 160.0 MHz. We use the filtering size of $\tau_{\rm filter} = \tau_{\rm horizon} + 300$ ns and the foreground suppression factor $\epsilon = 10^{-9}$. In the first panel, we show the map before the foreground subtraction, displaying that the dominant signals come from the foregrounds and the 21 cm signals are barely visible. After wedge-filtering as shown in the second panel, the foreground sources shown in the first panel are filtered out and the amplitude of the map is decreased by 4 orders of magnitudes. The third panel presents a wedge-removed result for the EoR-only simulation, which aligns closely with the second panel. This indicates that the suppressed foreground residuals are small enough to detect EoR signals. Given this result, we apply the wedge-filtering to EoR-only simulations in Sections 2.9 and 2.10.1, assuming the foreground residuals in simulations are sufficiently suppressed.

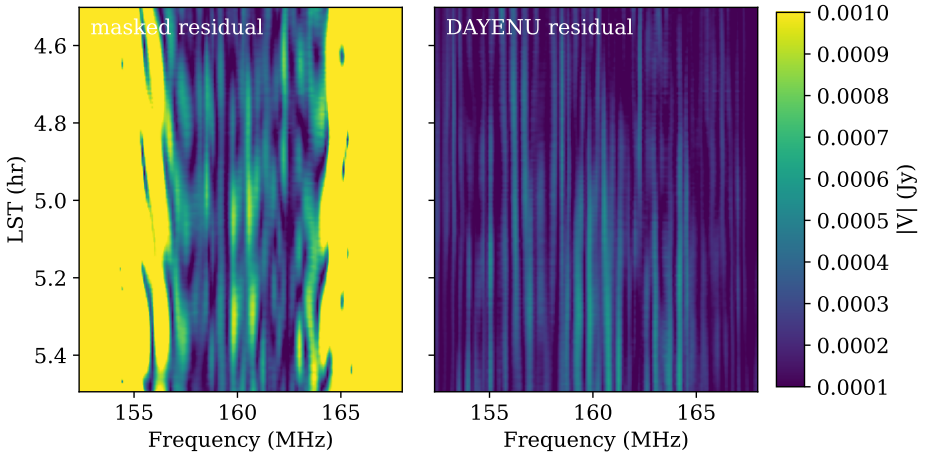

Figure 2.8: Residual waterfalls after the wedge-masking (left) and `DAYENU`-filtering (right) displayed as a function of frequency and LST. The original data includes both foregrounds and 21 cm signals and the residuals are expected to be dominated by filtered EoR signals. `DAYENU`-filtering results in a clean residual waterfall, reducing the foreground residuals below the 21 cm signals as seen in Figure 2.7. Meanwhile, the performance of the wedge-masking method is sub-optimal especially at the frequency edges due to the influence of the tapering function. Overall, the `DAYENU` method outperforms the wedge-masking approach.

2.4.2 Different Foreground Subtraction Approach

In this section, we explore different techniques for foreground subtraction and compare their performance to the `DAYENU`-filtering method based on simulation data.

A Standard Wedge-Masking Approach

Another technique for wedge-filtering involves masking low-delay modes containing foregrounds with the Fourier basis adopted by other studies (e.g., Kittiwisit et al., 2022, Prelogović et al., 2021). The procedure is as follows:

- Fourier transform visibilities tapered by a window function along the frequency direction. We use a Blackman-Harris window function. This window function helps reduce sidelobes of foreground emission in Fourier space that potentially contaminate the EoR window.
- Mask the data inside a top-hat filter with a filter size of $\tau_{\rm filter} = \tau_{\rm horizon} + 300$ ns.
- Inverse transform the masked data to the frequency domain and undo the tapering function.

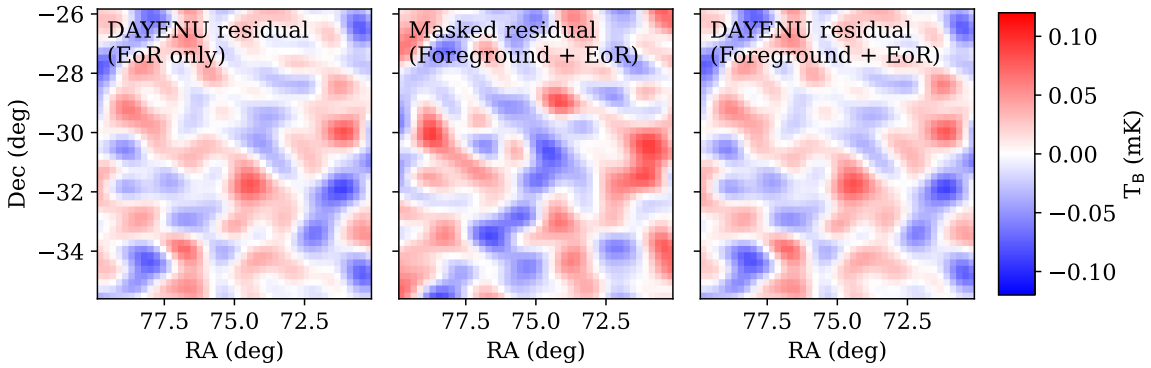

Figure 2.9: Comparison between DAYENU and wedge-masking methods. The first panel represents a filtered EoR model provided for a reference. The second and third panels show the filtered results for foreground + EoR models using wedge-masking and DAYENU-filtering methods. We pick a central frequency channel, 160.0 MHz. The wedge-masking result shows disparity compared to the reference, which indicates the masked result contains significant foreground residuals.

In Figure 2.8, the waterfall of simulated visibility as a function of frequency and LST is shown for the wedge-masking (left) and DAYENU-filtering (right) methods, respectively. We select the sky centered at LST = 5 hr and Band 2. Both results are filtered with the buffer size of 300 ns for the simulation data containing both foreground and 21 cm signals. The wedge-masked results exhibit a sub-optimal performance of filtering especially at the frequency band edges due to the tapering function, which is consistent with the result of Kittiwisit et al. (2022).

Figure 2.9 shows images for the filtered data at the central frequency channel (160.0 MHz) created through mapmaking process. The first panel, which is a DAYENU-filtered map for EoR-only signals, is given for reference. The second and third panels are the filtered results for the simulation data including both foregrounds and 21 cm signals using the wedge-masking and DAYENU-filtering methods, respectively. The DAYENU residual recovers the reference image as expected, while a noticeable disparity is observed in the case of the wedge-masking method compared to the first panel. This suggests that the DAYENU-filtering method outperforms the standard wedge-masking approach. In the remaining sections, the DAYENU-filtering method is chosen for our wedge-filtering technique.

PCA Subtraction Technique

In this section, we contrast the wedge-filtering technique with a foreground subtraction method based on PCA in the image domain, a widely employed approach in radio astronomy (e.g., Alonso et al., 2014, Cunnington et al., 2022, Spinelli et al., 2021). The analysis is performed with simulation data. This PCA method identifies spectrally smooth foreground components, which are subsequently subtracted from the original data to isolate subtle

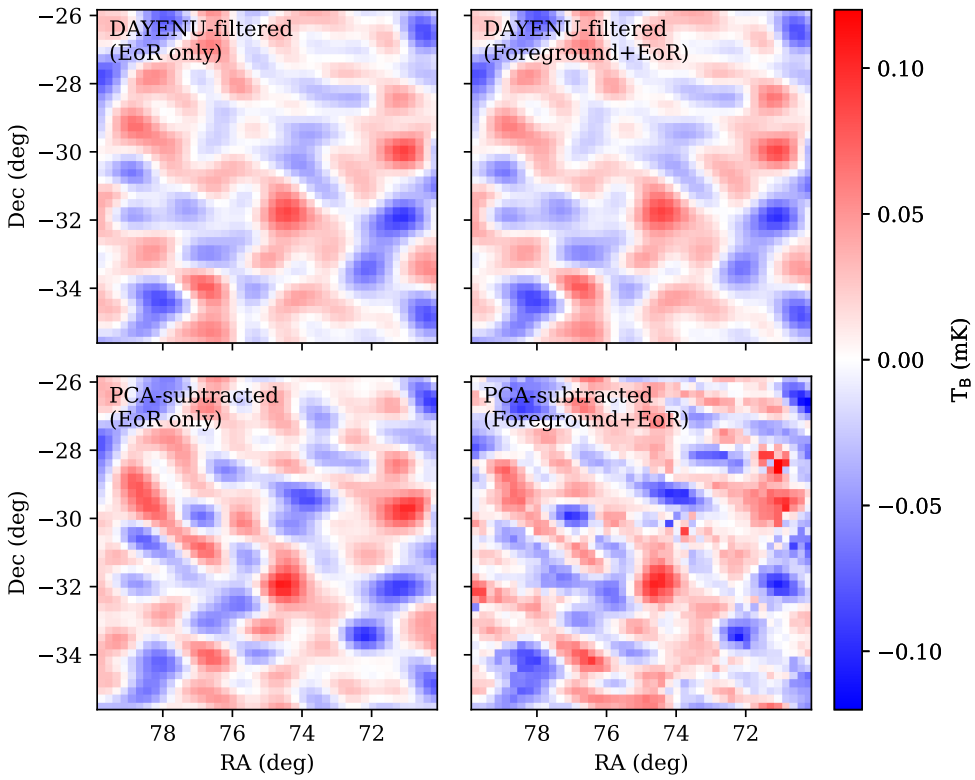

Figure 2.10: Comparison of wedge-filtering and PCA foreground subtraction at 160 MHz. Top and bottom rows represent `DAYENU`-filtered and PCA-subtracted results, respectively. The difference between the first and second panel presents the foreground residual after the foreground removal. Foreground residuals shown in the bottom right panel indicates the `DAYENU`-filtering outperforms the PCA-subraction.

cosmological signals.

The process of performing PCA to create an foreground estimate of $\hat{\mathbf{x}}^{\mathrm{FG}}$ can be succinctly summarized by the following steps:

- Centering the Data: subtract the mean of the map at each frequency channel, $\hat{\mathbf{x}}^{\mathrm{c}}(\nu) = \hat{\mathbf{x}}(\nu) - \langle\hat{\mathbf{x}}(\nu)\rangle_\theta$ where $\hat{\mathbf{x}} = \hat{\mathbf{x}}^{\mathrm{FG}} + \hat{\mathbf{x}}^{\mathrm{EoR}}$ and $\langle\cdots\rangle_\theta$ indicates averaging over pixels.
- Covariance Matrix: compute the covariance matrix of the centered data,

$$\mathbf{C} = \frac{\mathbf{X}^{\mathrm{T}}\mathbf{X}}{N_\theta - 1} \tag{2.14}$$

  where $\mathbf{X}$ is composed of the centered data with $N_\theta \times N_\nu$ and $N_\theta = N_{\mathrm{RA}} \times N_{\mathrm{Dec}}$.
- Eigenvalue and Eigenvector Calculation: calculate the eigenvalues (Λ) and corresponding eigenvectors ($\mathbf{U}$) of the covariance matrix, $\mathbf{C} = \mathbf{U}\Lambda\mathbf{U}^{\mathrm{T}}$. The eigenvectors represent

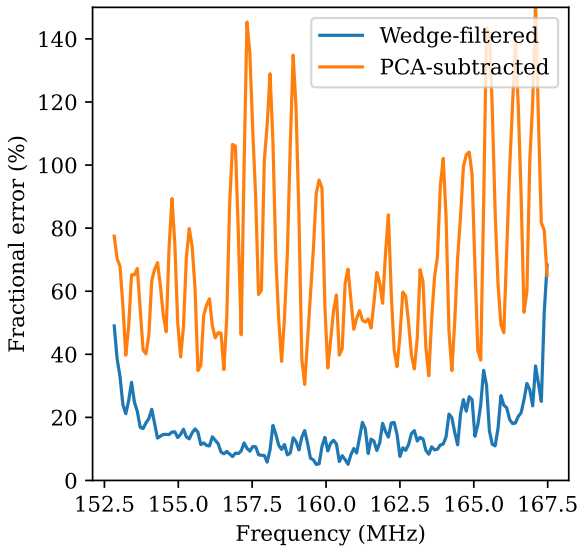

Figure 2.11: Comparison of performance of wedge-filtering (blue) and PCA foreground subtraction (orange) as a function of frequency. The fractional error associated with wedge-filtering is consistently lower than that of PCA subtraction across the frequency range. The mean error across frequencies suggests that wedge-filtering retains approximately 10% of foreground residuals.

the directions along which the data varies the most, while the eigenvalues represent the magnitude of the variance in these directions.

- Selecting Principal Components: rank the eigenvectors by their corresponding eigenvalues in decreasing order. These eigenvectors are often called "principal components". We analyzed the eigenvalue spectrum as a function of the component rank and identified a suitable PCA component rank, which serves as the transition point between the sharply decreasing and gradually changing segments in the spectrum. We found $N_{\rm PCA} = 20$ is enough to capture most features of foregrounds while minimizing the signal loss of the EoR, which gives $\mathbf{U}_{\rm FG}$ from the columns of $\mathbf{U}$ corresponding to $N_{\rm component}$.

- Projecting Data and Foreground Subtraction: The final step involves projecting the original data onto the selected principal components, $\hat{\mathbf{x}}^{\rm FG} = \mathbf{U}_{\rm FG}^{\rm T}\hat{\mathbf{x}}$. Then the residual is the cosmological signal that we want to extract, $\hat{\mathbf{x}}^{\rm EoR} = \hat{\mathbf{x}} - \hat{\mathbf{x}}^{\rm FG} = (\mathbf{I} - \mathbf{U}_{\rm FG}^{\rm T})\hat{\mathbf{x}}$.

The choice of $N_{\rm PCA} = 20$ is determined by identifying the component where the eigenvalue exhibits a sharp cut-off with the principal component and simultaneously ensuring that the fractional error, defined as the ratio between the PCA-subtracted foreground residuals and the PCA-subtracted EoR signals, is minimum.

Top and bottom panels of Figure 2.10 present the foreground subtraction results for `DAYENU`-filtering and PCA-subtraction, respectively. For the wedge-filtering, we consider

$\tau_{\text{buffer}} = 300$ ns. The first column is the outcomes for the EoR-only simulation, displaying both wedge-filtering and PCA-subtraction methods yield similar results.

However, the second column, which includes both foregrounds and EoR signals, exhibits noticeable differences in performance between the methods. The `DAYENU`-filtered map, even with the addition of foregrounds, remains consistent with the EoR-only simulation. In contrast, the PCA-subtracted map looks noisy, which means the map is significantly contaminated by foreground residuals as shown in the bottom right panel.

Fractional errors, defined by the median ratio of the magnitude of foreground residuals to that of the filtered EoR-only data over pixels, are presented in Figure 2.11 as a function of frequency. As observed in Figure 2.10, the `DAYENU`-filtering method, on average, yields fractional errors of about 10%, whereas the PCA-subtraction method results in the fractional errors larger than 50%, indicating filtered data contaminated by foreground residuals significantly across all frequencies. The sub-optimal performance of the PCA-subtraction method may be due to the presence of sidelobes in our dirty maps, which makes the spectral structure of the foregrounds less smooth in the image space.

While the contamination can be mitigated by assigning more than 20 PCA components to the foreground, it comes at the cost of a significant reduction in the EoR signal. Consequently, the PCA subtraction method may not be optimal for studying one-point statistics with our instrument. Based on the results for the `DAYENU` wedge-filtering method compared to the wedge-masking and PCA-subtraction methods, we adopt the `DAYENU` filter as our filtering method throughout the remaining sections.

2.5 Instrument Effects on One-point Statistics

In Section 2.4.1, I explored the effect of wedge-filtering on the resultant images. In this section, I examine how the PSF and wedge-filtering affect measuring one-point statistics of our theory models. Figure 2.12 shows m_2 (top), m_3 (middle), and S_3 (bottom) measurements for the (thermal) noiseless fiducial model. The dotted lines indicate raw EoR model which means it is free from the effect of instrument and foregrounds.

The blue lines represent cases where the instrument response or PSF is applied without accounting for foreground contamination. The solid lines and shaded regions denote the mean and sample variance measured across a 60×10 square degree field over 30 different sky realizations, respectively. Due to our map normalization scheme, which preserves surface brightness, the one-point statistics measurements are not conserved after PSF convolution, leading to a significant reduction in variance by a factor of 10^4. Although m_3 is affected by large sample variance, there remains a possibility of detecting non-zero m_3 signals at

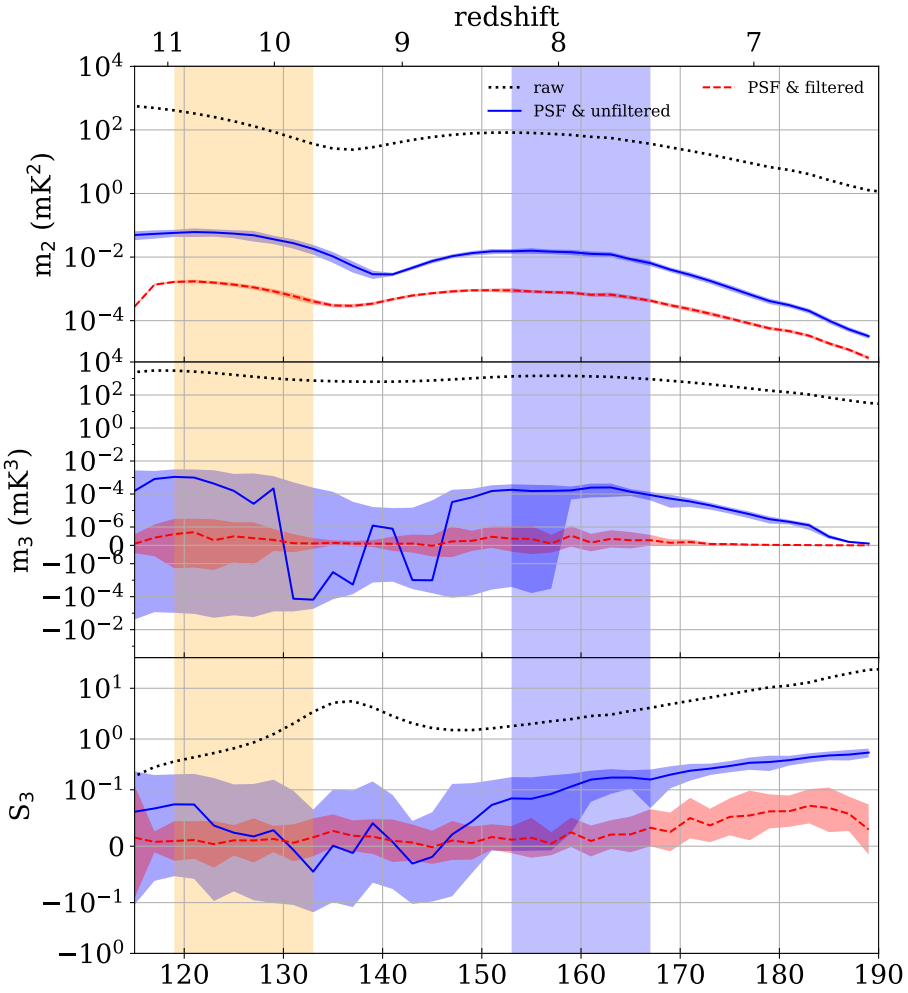

Figure 2.12: m_2 (top), m_3 (middle), and S_3 (bottom) for the fiducial model as a function of frequency given Phase I mock observations. The dotted line represents measurements from raw simulations without instrument effects. The blue solid line and shaded region indicate the mean and sample variance of each statistic after applying the PSF, assuming no foreground. The red dashed line and shaded region show the mean and sample variance after applying both the PSF and wedge-filtering. The PSF significantly reduces the magnitudes of the m_2 and m_3 measurements. For m_3 in particular, the wedge-filtering further diminishes the m_3 signals, making it challenging to detect non-Gaussian features. Before applying wedge-filtering, the median sample variance (shaded region) relative to the mean amplitude (solid line) is 19%, 132%, and 110% for m_2, m_3, and S_3, respectively. After wedge-filtering, these median values turn into 13%, 192%, and 181% for m_2, m_3, and S_3, respectively. For m_3, the scale is linear between -10^{-6} and 10^{-6}, while outside this range, it transitions to a logarithmic scale. For S_3, the scale is linear between -10^{-1} and 10^{-1}, and otherwise it is a logarithmic scale.

higher frequency bands. The reduced sample variance beyond 160 MHz may result from the increased reionization of neutral hydrogen, causing different sky fields to appear more similar. Additionally, S_3 (bottom panel) suggests that meaningful detection at $z > 8$ is challenging, but there is an increasing trend in S_3 toward lower redshifts, despite its magnitude being

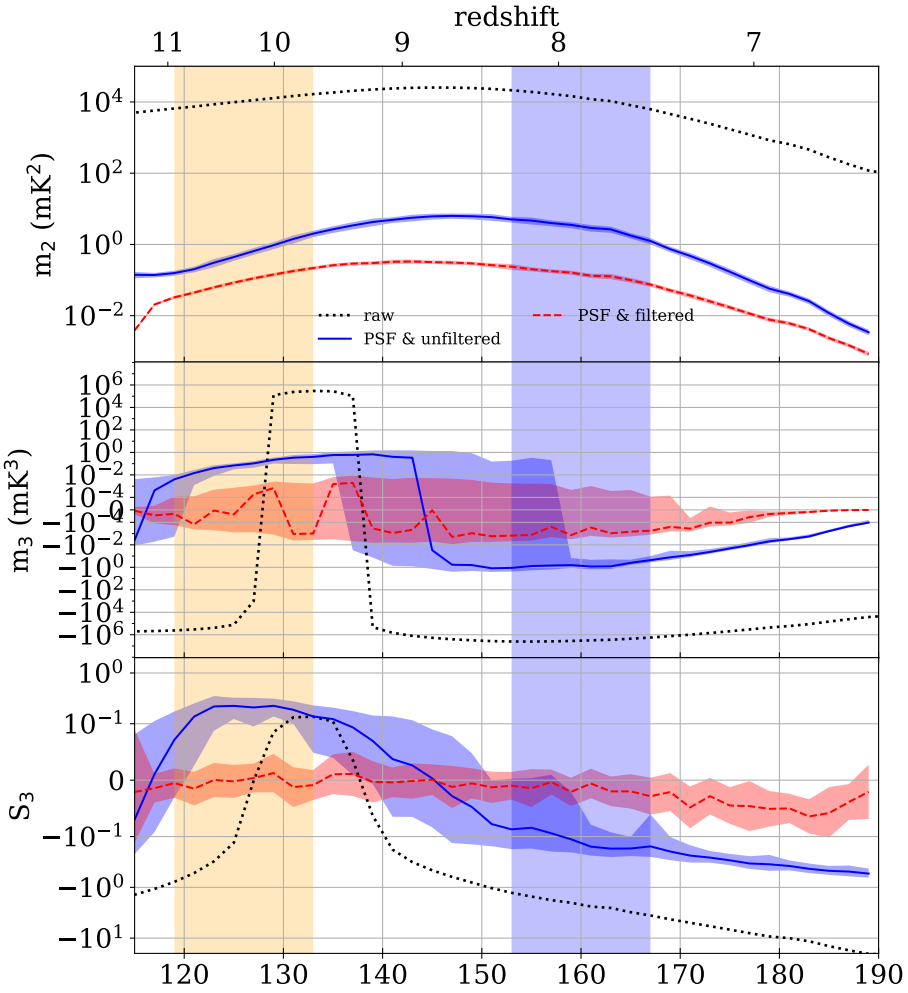

Figure 2.13: m_2 (top), m_3 (middle), and S_3 (bottom) for the cold reionization model as a function of frequency given Phase I mock observations. The format is the same as that of Figure 2.12 but for the cold reionization model. Similar to the fiducial model, the variance is reduced by a factor of 10^4 when the PSF is applied, and it is further reduced by two orders of magnitude after wedge-filtering. As with the fiducial model, detecting non-Gaussianity, whether through m_3 or S_3, becomes challenging due to the combined effects of the PSF and wedge-filtering. Before applying wedge-filtering, the median sample variance (shaded region) relative to the mean amplitude (solid line) is 20%, 76%, and 60% for m_2, m_3, and S_3, respectively. After applying wedge-filtering, these median values change to 13%, 238%, and 244% for m_2, m_3, and S_3, respectively. For m_3, the scale is linear between -10^{-4} and 10^{-4}, transitioning to a logarithmic scale outside this range. For S_3, the scale is linear between -10^{-1} and 10^{-1}, with a logarithmic scale applied beyond these limits.

reduced by roughly an order of magnitude compared to S_3 derived from the raw simulation.

When foregrounds are accounted for and wedge-filtering is applied to mitigate them, the amplitude of the one-point statistics measurements is further reduced. The dashed red line in Figure 2.12 shows that m_2 decreases by an order of magnitude, and the signal in m_3 approaches zero, indicating that most of the non-Gaussianity information is washed out due

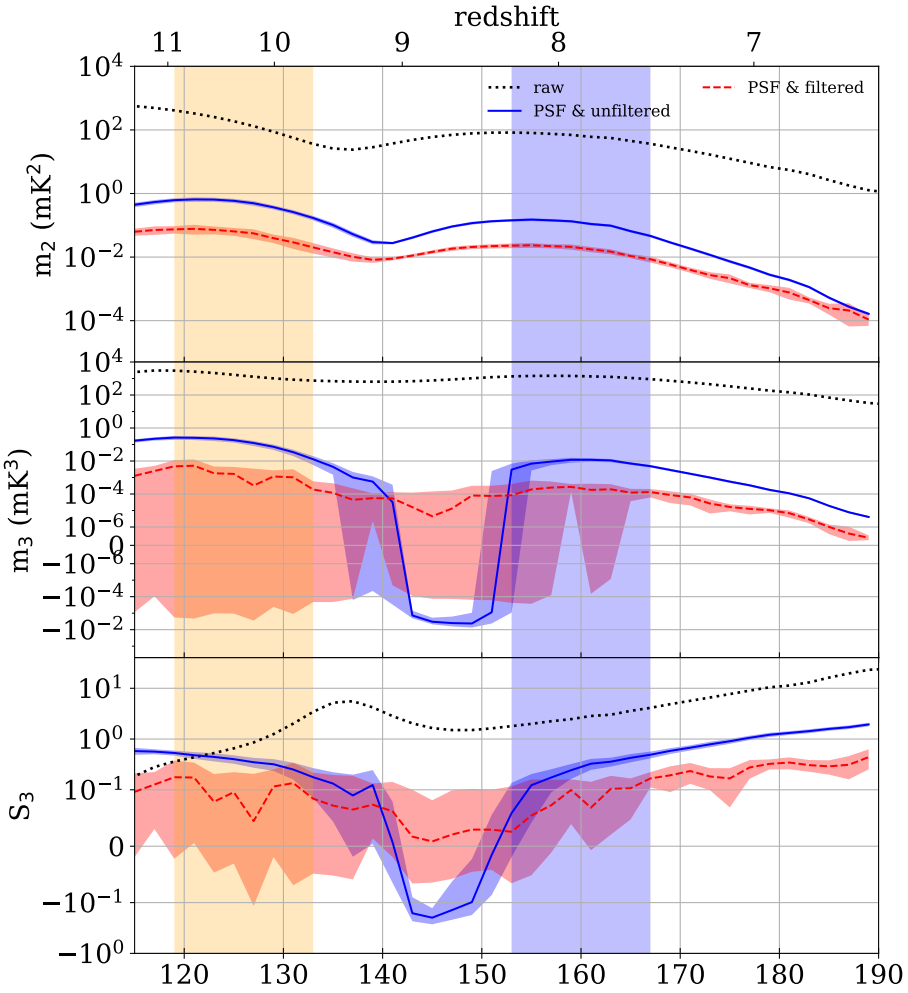

Figure 2.14: m_2 (top), m_3 (middle), and S_3 (bottom) for the fiducial model as a function of frequency with HERA-320. The format is the same as that of Figure 2.12 but for the HERA-320 configuration. The PSF effect reduces m_2 by a factor of 10^3 assuming no foreground, with further reduction through wedge-filtering by an order of magnitude. In the blue curve, m_3 exhibits a distinct trend compared to Figure 2.12, with relatively smaller sample variance. The red line shows that wedge-filtering decreases m_3 signals and increases the sample variance, making it challenging to detect significant m_3 signals at $z > 7.5$. However, there may still be potential to detect deviations in m_3 and S_3 measurements from the zero baseline at low redshifts. Before applying wedge-filtering, the median sample variance (shaded region) relative to the mean amplitude (solid line) is 10%, 29%, and 15% for m_2, m_3, and S_3, respectively. After applying wedge-filtering, these median values change to 23%, 37%, and 112% for m_2, m_3, and S_3, respectively

to the wedge-filtering. The S_3 plot suggests that detecting non-zero skewness in Band 1 and Band 2 is challenging, although there is a slight increasing trend toward lower redshifts.

Figure 2.13 presents the results for the cold reionization model. Similar to the fiducial model, m_2 decreases by a factor of 10^4 after the instrument response is applied and is further reduced by an additional order of magnitude after the wedge-filtering is applied.

In Section 2.3.2, we discussed the sign flips in m_3 observed in the raw simulations of the cold reionization model, which are revisited by the black dotted line in Figure 2.13. The sign flips were previously interpreted based on the idea that the skewed shape of the brightness temperature distribution, combined with the pileup in the zero bin due to reionization, determines the sign of m_3, as illustrated in Figure 2.5. When considering the PSF effect with no foreground (blue solid line), sign flips still occur but at different locations compared to the black dotted line, possibly due to additional effects like grating and/or side lobes. The blue shaded region indicates that the uncertainty in the location of these sign flips spans 140—160 MHz. However, once wedge-filtering is applied, the m_3 and S_3 signals are mostly removed.

In addition to the Phase I observation, we explore the predicted one-point statistics for the HERA-320 array configuration. The PSF of HERA-320 is narrower compared to that of the Phase I observation, allowing us to probe smaller scales. Figure 2.14 shows the results for the fiducial model with 320 antennas. The mean and sample variance, represented by lines and shaded regions respectively, are estimated across a 60×10 square degree field from 30 random independent sky locations. Due to the ability to capture additional signals from smaller scales, the variance after applying the PSF (blue line) is larger than in Figure 2.12, though still reduced by a factor of 10^3 compared to the raw simulations. The wedge-filtering continues to reduce the variance by an order of magnitude, as indicated by the red curve.

Unlike the Phase I observation, which utilized around 40 antennas, HERA-320 predicts that non-zero m_3 signals (assuming no foregrounds, as shown by the blue line) are detectable across redshifts, with sign flips being relatively well-identified due to the smaller sample variance compared to Figure 2.12. However, when wedge-filtering is applied, detecting non-zero m_3 at $z > 7.5$ becomes challenging because of the large sample variance. Nonetheless, there remains some potential for detecting non-zero m_3 and S_3 at lower redshifts in the red curves. To fully assess detectability for future observations, it is necessary to account for thermal noise in addition to sample variance. In Section 2.10.1, we explore the detectability with HERA-320 at low redshifts, considering different observation time.

2.6 Noise Simulations

If the observational data is free from systematics and dominated by thermal noise after foreground removal, noise simulations provide theoretical expectation of measured one-point statistics (Sections 2.8 and 2.9). Additionally, noise simulations are used to indicate the sensitivity of future observations, allowing us to forecast one-point statistics for the full core array of HERA (Section 2.10.1).

Specifically, thermal noise in the visibility data per baseline follows

$$\mathbf{n} \sim \frac{1}{\sqrt{2}}\mathcal{N}(0, \sigma_n^2) + \frac{i}{\sqrt{2}}\mathcal{N}(0, \sigma_n^2), \tag{2.15}$$

where $\mathcal{N}(0, \sigma_n^2)$ is a normal distribution with a zero mean and variance specified by Equation (3.6). We draw samples of the noise visibility based on Equation (2.15) using a Monte-Carlo simulation and perform forward modeling to form noise images as follows:

$$\hat{\mathbf{x}}_{\mathrm{N}} = \mathbf{D}\mathbf{A}^{\dagger}\mathbf{N}^{-1}\mathbf{n}. \tag{2.16}$$

The noise properties in the image domain can be understood through the map noise covariance matrix, representing the correlated noise between pixels, which can be compute as follows,

$$\mathbf{C}_{\mathrm{N}} = \mathbf{D}\mathbf{A}^{\dagger}\mathbf{N}^{-1}\mathbf{A}\mathbf{D}^{\mathrm{T}} = \mathbf{P}\mathbf{D}^{\mathrm{T}}. \tag{2.17}$$

The noise variance, diagonal elements of $\mathbf{C}_{\mathrm{N}}$, can be explicitly written as,

$$C_{ii} = P_{ii}D_{ii} \tag{2.18}$$

$$\approx \frac{\sum_t B_{it}^2}{\sum_t B_{it}} \frac{\bar{\sigma}_n^2}{\sum_t B_{it}} \tag{2.19}$$

$$= \bar{\sigma}_n^2 \frac{\sum_t B_{it}^2}{\left(\sum_t B_{it}\right)^2}, \tag{2.20}$$

For a single time integration, the variance of a pixel is equal to σ_n^2, which indicates the noise variance is unaffected by the beam and independent of the pixel's location. For longer integration, the noise is attenuated by the primary beam at pixels far away from the zenith. However, because the signals also experience attenuation at these off-zenith pixel locations due to the primary beam, S/N ratio may not see improvement at such pixel locations. For example, a naive estimation of S/N ratio using Equations (2.7) and (2.20) suggest S/N $\propto \sqrt{\sum_t B_{it}^2}$, indicating the decrease of S/N at pixel locations distant from the beam center.

Mapmaking processes for noise data of the Phase I observation and future observation are more detailed in Sections 2.7 and 2.10.1, respectively.

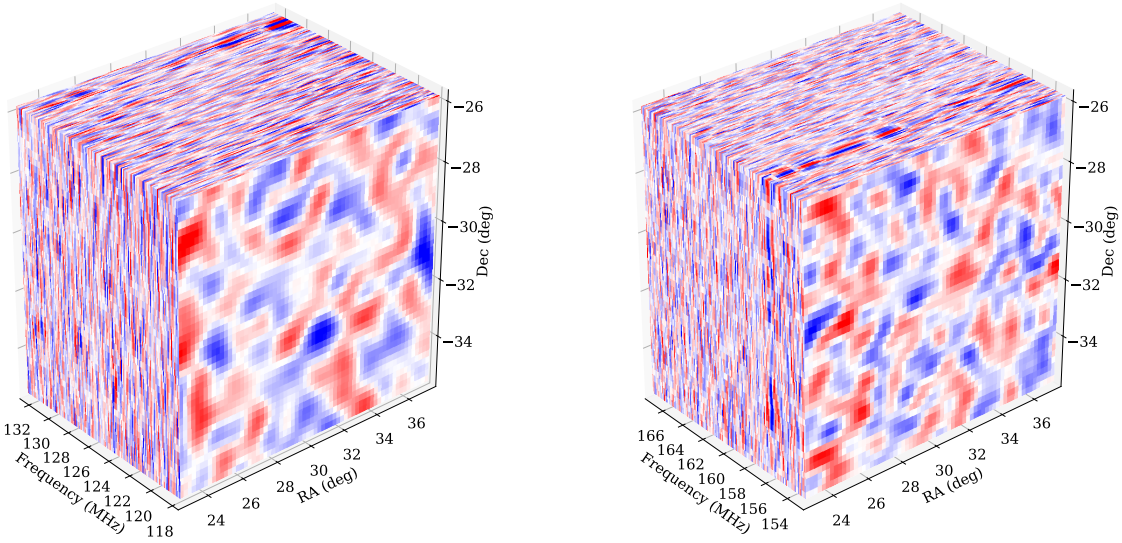

Figure 2.15: Image cubes for wedge-filtered observational data with comprehensive systematics removal. The foreground subtraction is performed using the DAYENU-filtering method. The wedge-filtered data is expected to be noise-limited. Left and right panels represent Band 1 and Band 2 results, respectively. The maps are centered at 2 hr (Field I). Those cubes are employed to measure statistical quantities at each frequency or redshift. Bright slices, appearing at 131 and 157 MHz, correspond to systematics residuals discussed in Section 2.8.2.

2.7 Wedge-filtered Data Cube Construction for Phase I Observation

In this section, we provide a description of the data products used for measuring one-point statistics for the Phase I observation. As we are interested in measuring one-point statistics at each redshift, data cubes in the dimension of RA $\times$ Dec $\times$ frequency are constructed for observational and simulation data.

As described in Section 2.3.1, we use calibrated observational visibility data for Band 1 and Band 2 for each EE- and NN-polarizations. In order to remove noise biases for variance measurements (note that skewness is an unbiased measurement), we divide the visibility data into even and odd time groups by selecting every other timestamp. If thermal noise in the maps remains uncorrelated across different time groups, it is expected that when cross-multiplying maps from these time groups, the resulting variance will be free from biases introduced by noise. We utilize observational data with and without systematics mitigation to investigate the effect of systematics in measuring one-point statistics in the image domain.

The visibility datasets undergo the DAYENU filtering with the buffer size of 300 ns and the foreground suppression factor of 10^{-9}. The filtered data are converted into images through the mapmaking process (Section 2.2) for each frequency channel, polarization, and time

group (even and odd). Pseudo stokes-I maps are made by averaging two polarization maps. Image cubes for Band 1 and Band 2 are constructed for each hour of LST, spanning 15×10 square degrees in RA × Dec, which are then stitched together to cover each field as defined in Section 2.3.1.

An example for the wedge-filtered data cubes of observational data is presented in Figure 2.15. Left and right panels correspond to the image cube of Band 1 and Band 2, respectively, centered at the 2 hr LST field (Field I) for a 15×10 square degree field. The cubes are generated from observational data, with comprehensive systematics mitigation, implying that they should exhibit random Gaussian fields from thermal noise. However, even after the systematics were subtracted, some systematics residuals may still be present, as indicated by bright slices appearing at around 131 and 157 MHz. We will discuss statistical measurements on the image cubes in Section 2.8.2.

HERA observations can achieve higher resolution along the line-of-sight (or frequency) direction, approximately 2 cMpc, compared to the spatial direction, which is around 60 cMpc for the entire HERA core array. The finer resolution along the frequency axis than the spatial direction is illustrated in Figure 2.15.

For the 21 cm simulation data described in Section 2.3.2, we convert the simulated visibility data into data cubes for each band after the `DAYENU` filtering with the same buffer size and the suppression factor as the observational data. Though the 21 cm simulation is noiseless, we split the visibility data into different time groups to be consistent with the process of the observational data. All four distinct EoR models undergo the same process. Mock visibilities to estimate the cosmic variance of the cold reionization model are also converted to image cubes in the same way, which are used in Section 2.9.

As described in Section 2.6, thermal noise is realized in visibility space using a Monte-Carlo simulation according to the specification of the Phase I observation. We simulate 1000 sets of random noise visibilities, spanning a LST range of 0–15 hr for EE- and NN-polarizations. We divide each visibility data into even and odd timestamps. The visibility data of each time group is then `DAYENU`-filtered and transformed to pseudo stokes-I maps as the observational data. The noise image cubes are used for statistical inference of observational data compared to expected noise levels in Sections 2.8 and 2.9.

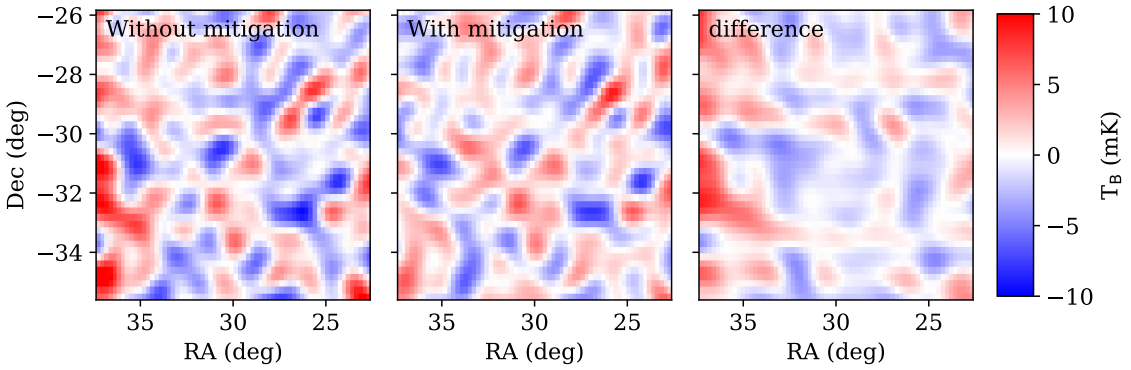

Figure 2.16: Comparison of maps before (first panel) and after systematics mitigation (second panel) at 126.0 MHz in Field I. The last panel shows the difference between them, which is of a similar order of magnitude as the mitigated result, indicating the data before mitigation is systematics-limited.

2.8 Effects of Systematics on Measuring One-Point Statistics

2.8.1 Systematics Residual in Wedge-Filtered Data

Major components of systematics that were mitigated in the HERA Phase I observation are the cable reflection and crosstalk (H22a, H23). The reflection in the coaxial cable may occur due to an impedance mismatch at the termination of the cable. This reflection results in duplicates of signals appearing at a certain delay, equivalent to twice the light travel time along the cable. This systematic effect is addressed through reflection calibration, a process that involves fitting reflection coefficients using the autocorrelation visibility within a delay range of 150–1500 ns (Kern et al., 2019, 2020b). This calibration method effectively removes the reflection terms through a direction-independent calibration process.

Crosstalk represents another significant systematic effect in Phase I observations, potentially manifesting across a wide range of delays. The hypothesis suggests that feed-to-feed reflections may cause autocorrelation visibilities to be copied into cross-correlations at higher delays. The autocorrelations typically exhibit a relatively gradual evolution over a primary beam crossing timescale. With this attribute, in order to mitigate the corruption, a modeling technique using Singular Value Decomposition (SVD) is employed for each baseline to identify time and delay modes that are influenced by the crosstalk. This technique, coupled with low-pass filtering along the time axis through GPR, aids in suppressing the systematics significantly (Kern et al., 2019, 2020b). The full systematic mitigation on the observational data we use was implemented by H23.

Figure 2.16 exhibits the wedge-filtered maps without and with the systematics mitigation

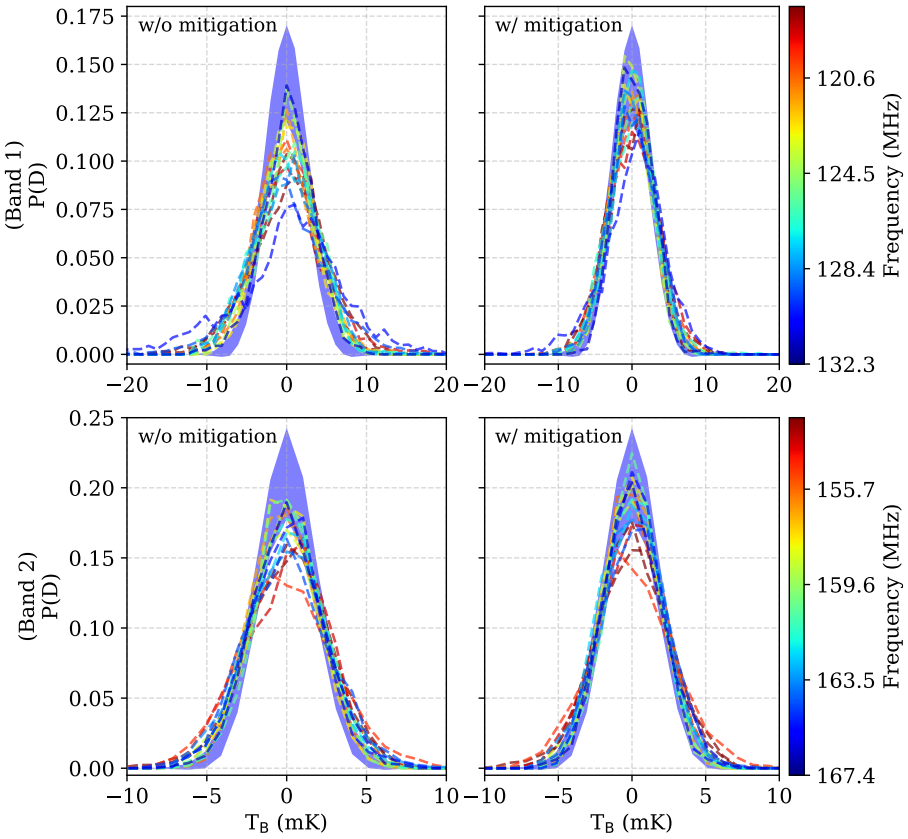

Figure 2.17: P(D) of wedge-filtered maps for data before the systematics mitigation (left column) and after the mitigation (right column) for Band 1 (top) and Band 2 (bottom). The 95% confidence interval of noise simulation distributions is presented as a blue shaded region for reference. The deviation in P(D) from the reference distribution for the data without systematics mitigation is reduced after applying the systematics mitigation, although some outliers still remain.

described above. Notably, certain prominent diffuse structures evident in the first panel are no longer present after the mitigation process (second panel). In the absence of appropriate mitigation, wedge-filtered maps might retain foregrounds introduced by systematic effects. The last column, illustrating the difference between the two maps, potentially indicates the presence of foreground residuals leaking into high delay modes in the absence of systematics mitigation.

To investigate the effects of systematics on map properties, we employ a P(D) analysis (Condon, 1974, Scheuer, 1957). P(D) represents a density profile of pixel values in brightness temperature, where random thermal noise is expected to follow a Gaussian distribution. The results for Field I are shown in Figure 2.17. The top and bottom panels correspond to Band 1 and Band 2, respectively. Every 10th channel from each band is selected for illustration purposes. Maps from even and odd time groups are averaged. To examine how systematics affect the distribution shape, P(D) is measured from maps of the wedge-removed data before

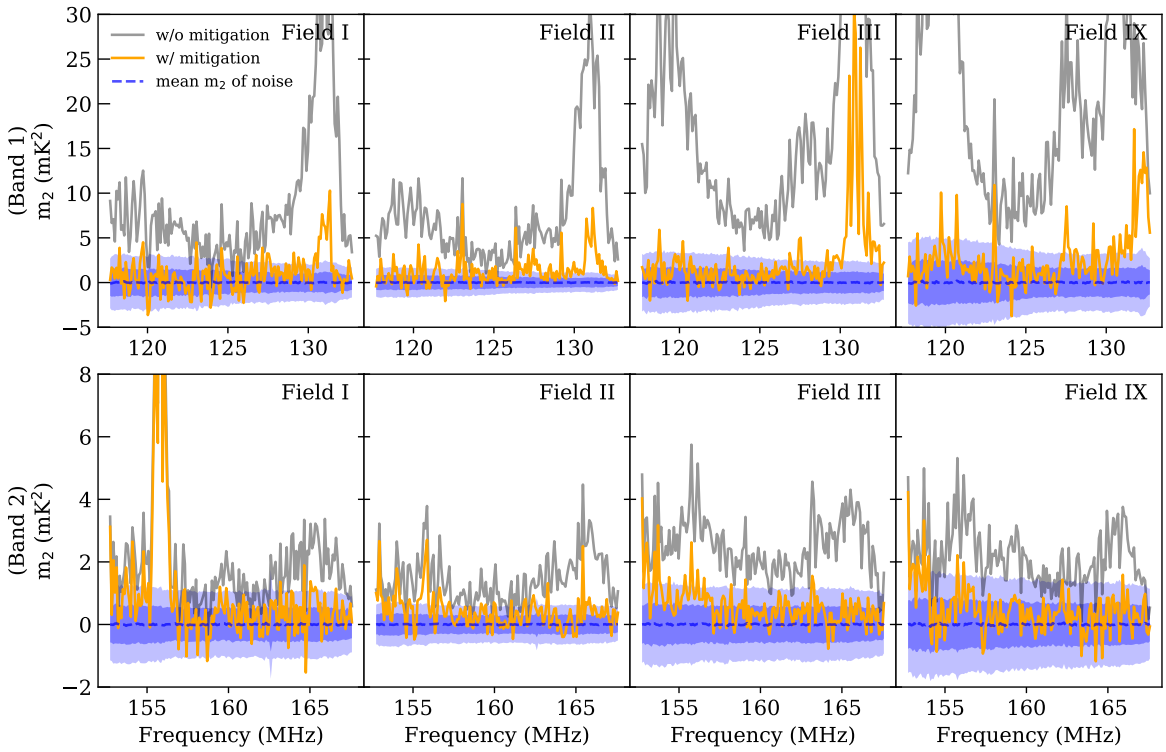

Figure 2.18: m_2, variance measurements, of wedge-filtered observation and noise maps as a function of frequency at different fields. Gray and orange curves represent the data before and after systematics mitigation, respectively. The dark and light blue shaded regions, corresponding to 1σ and 2σ confidence, respectively, are defined by the 1000 random noise realizations. Both Band 1 (top row) and Band 2 (bottom row) results show the variance is significantly reduced with the systematic mitigation though there are still systematics observed beyond the theoretical noise expectation even after the systematics mitigation for all LSTs.

(left) and after (right) systematic mitigation. The 95% confidence intervals of 1000 noise simulations are shown as blue shaded regions for reference.

In the absence of systematics removal, for most frequency channels, noticeable deviations from a Gaussian profile, such as broadening, asymmetry, and tailedness, are evident in the P(D), indicative of non-Gaussianity caused by systematics. After systematics subtraction in the HERA Phase I observation, where the data is expected to be noise-limited, the P(D) for most channels are aligned with the expected noise profile, but there are still some outliers. We further investigate the effects of systematics residuals on one-point statistics measurements quantitatively over a range of frequency in the following section.

2.8.2 Statistical Measurements for Wedge-removed Observational Data

As observed in Section 2.8.1, some residual systematics remain in the wedge-filtered map even after the mitigation efforts. In this section, we examine the systematics effects on one-point statistics of observational data in comparison to expected theoretical noise across a range of frequencies.

Variance is a primary quantity in one-point statistics which is related to the power spectrum,

$$\sigma^2 = \int \frac{d^3k}{(2\pi)^3} P(\mathbf{k}), \tag{2.21}$$

where $P(\mathbf{k})$ is a power spectrum of signals. If there are systematic residuals in the data, particularly foreground residuals leaking into the EoR window in the power spectrum, the variance tends to increase and becomes sensitive to the presence of these systematic residuals.

To eliminate noise bias in variance measurements, we cross-multiply the even and odd wedge-removed maps and average the product across pixels at each frequency channel for both observational data and noise simulations for each field.

The variance measurements, or m_2, are presented in Figure 2.18 for Band 1 (top) and Band 2 (bottom) at each field. The y-axis represents the variance after removing the noise bias. The dashed line represents the mean, while the dark and light blue shaded regions indicate the 1σ and 2σ confidence intervals, respectively, based on 1000 noise simulations. The noise simulation fluctuates around a zero mean, as expected. A smaller root-mean-square (RMS) of the m_2 measurements is achieved in Field II, as it is a relatively quiet radio sky in terms of foreground noise.

In both Band 1 and Band 2, there are substantial excess of variance in the presence of systematics (gray curve) as expected. The oscillating patterns have a period of approximately 1 MHz, corresponding to 1000 ns, which is a characteristic delay of crosstalk. After implementing systematics mitigation, the excess of variance is reduced, with a more pronounced effect observed in Band 1. However, the orange curves still show variance excess, indicating remaining systematics, especially more prominent in Field II with deeper observations. This excess of variance could arise from systematics residuals of cable reflections, crosstalk, faint RFIs, and/or chromatic gain errors introducing foreground leakage into high delay modes (H22a, H23).

The removal of strong and broad RFIs around 137 MHz (associated with ORBCOMM satellites) and 150 MHz during calibration results in significant discontinuities in the spec-

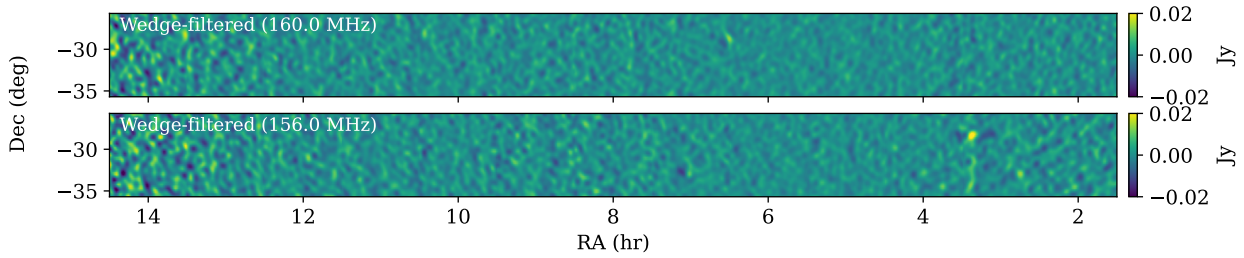

Figure 2.19: HERA stripe for wedge-removed observational data. The top and bottom panels display maps at 160.0 MHz and 156.0 MHz, respectively. The top panel shows a map that is relatively free of systematics, whereas the bottom panel exhibits a distinct foreground residual feature at 3.3 hr, which are residuals of the grating lobes of Fornax A.

trum (H22a H23). These gaps may interact with the data reduction process, including systematics mitigation, inpainting of removed data points, and smoothing of calibration gains. This interaction might explain unsuccessful removal of systematic residuals including the observed peaks around 131 MHz and 156 MHz. Although these features are attenuated in power spectrum estimation due to a tapering function, they can be manifest in our imaging analysis.

Figure 2.19 illustrates an example of systematics residuals, or foreground residuals, in image space. Wedge-removed maps at different frequency channel are presented. Each panel has the same dimension of Figure 2.2. At 160.0 MHz (top panel), the foreground-removed map displays a relatively clean map with minimal foreground contamination. In contrast, at 156.0 MHz (bottom panel), there are noticeable features of foreground residuals leaking into the EoR window at 3.3 hr, corresponding to the residuals of grating lobes from Fornax A. We find the presence of foreground residuals across a range of frequencies, which may explain the variance excess shown in Figure 2.18. The increasing amplitude of noise at around 14 hr in both maps is due to the sky approaching the bright galactic center.

To separate the behaviors of m_2 and m_3, instead of the standardized third moment S_3, we use the m_3 as a statistical indicator of non-Gaussianity. As evident in Figure 2.17, some frequency channels affected by systematics exhibit pronounced non-Gaussian features, even with the mitigation efforts. In Figure 2.20, we depict the m_3 as a function of frequency for Band 1 and Band 2 in each field. Because m_3 is an unbiased measurement, the maps of even and odd time groups are averaged before we measure the m_3. The blue dashed lines represent the mean, and the dark and light shaded regions denote the 1σ and 2σ confidence intervals of m_3 respectively, as calculated from 1000 noise simulations. As expected the thermal noise drawn from a random Gaussian distribution has a zero mean skewness.

The results for the data after the mitigation are shown in orange lines. The m_3 measurements oscillate around the zero mean, which is largely consistent with the noise simulations. However, there are peaks surpassing the 2σ boundary across all fields, particularly noticeable

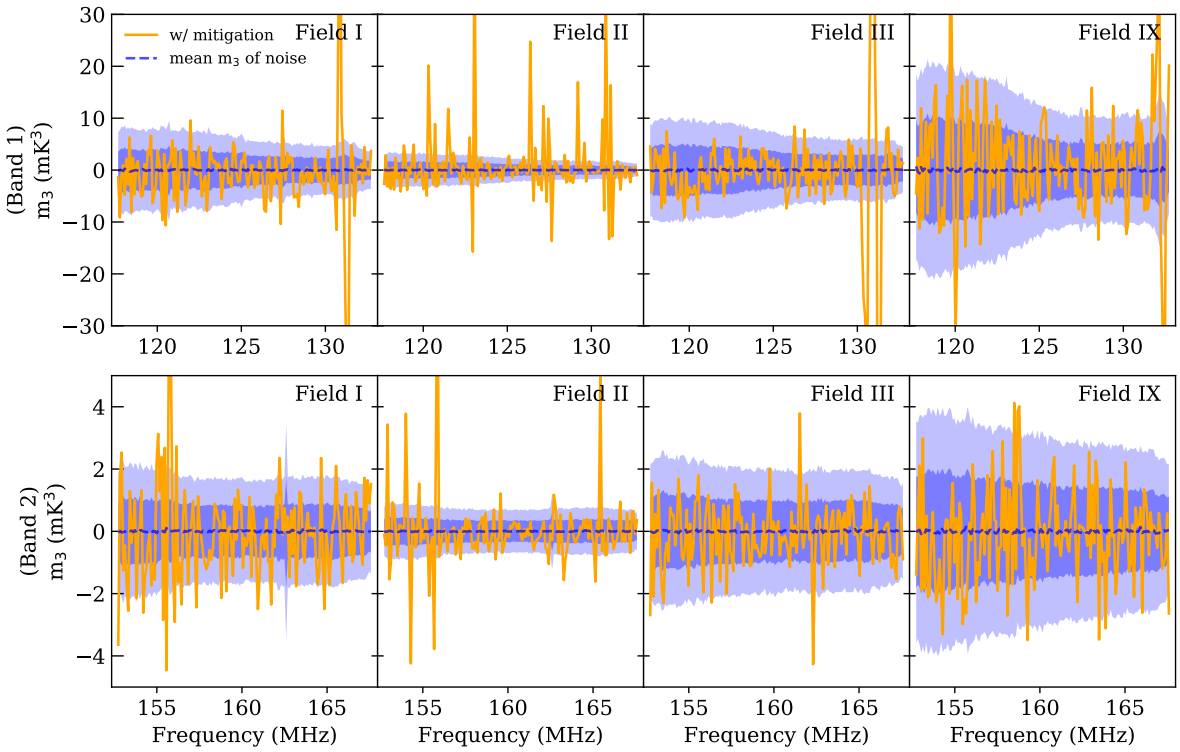

Figure 2.20: m_3 measured from the wedge-filtered observation, along with the noise simulations denoted as the mean, 1σ, and 2σ confidence intervals, at different fields. Orange curves represent the data with systematics mitigation. Though there are some outliers observed beyond the 2σ boundary, m_3 measurements for large fraction of frequency channels are confined within the 2σ range except for Field II.

in Field II as we see in Figure 2.18. Unlike the variance, systematics effects in m_3 can be either positive or negative values. The results from m_2 and m_3 underscore the significance of mitigating systematics to a deeper level for exploring one-point statistics of 21 cm signals, especially in future observations with HERA.

2.9 Comparison with 21 cm Simulations

As detailed in Section 2.3.2, we simulate four distinct EoR models to investigate one-point statistics measurements to assess their detectability. In this section, we perform a likelihood analysis for m_2 measurements and explore the detectability of m_3 measurements based on Phase I observations compared to the 21 cm model.

Previous studies adopted analytic calculations to estimate the errors associated with the statistical quantities (e.g., Kittiwisit et al., 2017, Watkinson and Pritchard, 2014). In this study, rather than relying on an analytic form, we employ 1000 noise Monte Carlo simulations

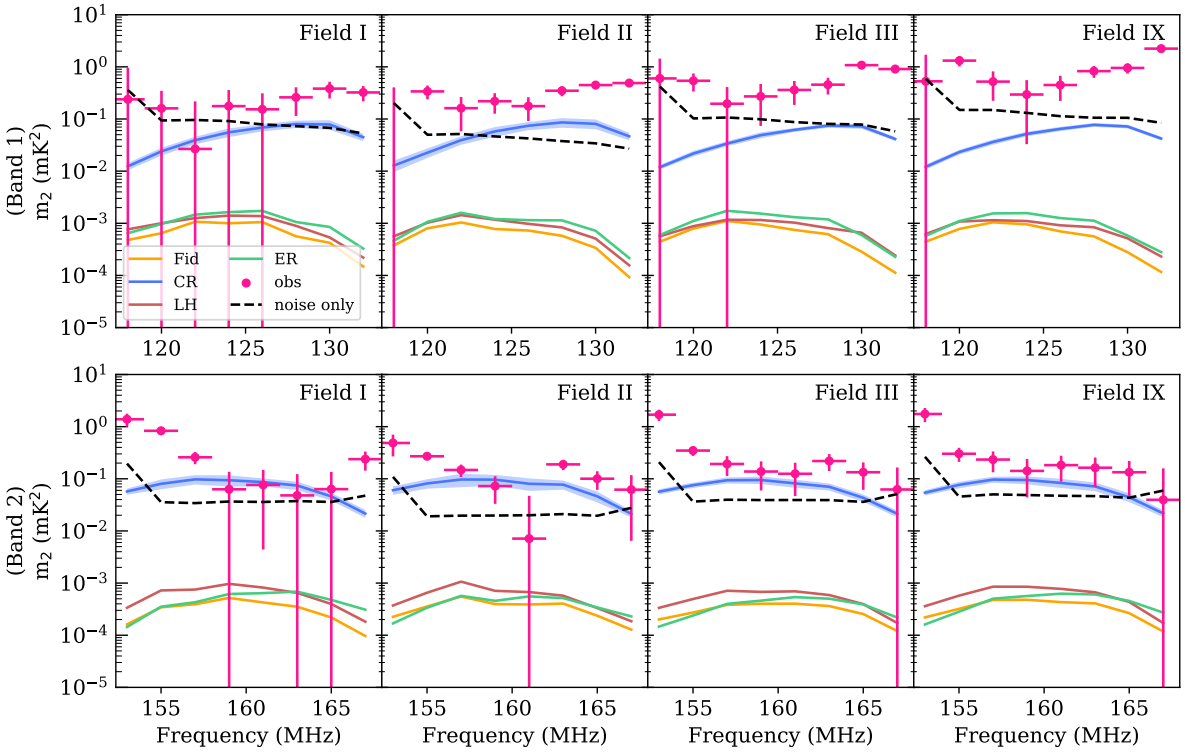

Figure 2.21: m_2 measurements based on wedge-filtered data in comparison to simulations for each field and each band. Frequency averaging over 0.5 MHz bin is performed to improve the sensitivity. Variance is measured from every 2 MHz sliced volume. The 1σ errors computed from noise simulations are indicated by the black dashed lines, providing the theoretical expectation. The m_2 values of the observational data are represented by pink symbols with 2σ error bars, derived from the same noise simulations, along the y-axis. The x-axis error bar represents the bin width. The lowest upper limit, considering the 2σ error bar, is 0.047 mK^2 at 161 MHz in Field II. Additionally, four 21 cm models are compared to the observational data points. These simulations do not include thermal noise. Blue shaded regions correspond to the cosmic variance of the cold reionization model derived from 20 samples.

with a brute-force approach. This method allows us to estimate the uncertainties of the mean of statistical quantities and directly calculate the error bars on the observational data. It accounts for correlations between pixels induced by instrumental effects and correlations between frequency channels introduced by wedge-filtering.

2.9.1 Statistical Tests on m_2

While variance does not measure non-Gaussianity, it is useful for evaluating the likelihood of an EoR model based on the fluctuations of 21 cm signal of each model relative to observational data, analogous to the power spectrum analysis. Figure 2.21 presents the m_2 measurements for the four 21 cm models in comparison to the observational data points

in each field. We use the dataset with systematics mitigated. To increase sensitivity in our observations, we coherently average the intensity maps over every 5 channels, roughly corresponding to a 0.5 MHz bin size, which improves sensitivity in m_2 by about 5 times. We then bin the image cube into sub-band data volumes using a 2 MHz interval between 117–133 MHz for Band 1 and between 152–168 MHz for Band 2. The variance is measured in each sub-band volume, further enhancing sensitivity in m_2 by about 2 times. As in the previous section, we cross-multiply image volumes from even and odd time groups to remove noise bias. The same process is also applied to the maps of wedge-filtered noise and 21 cm simulations.

We use the noise simulations to estimate the uncertainty of the mean of m_2, or RMS of the noise simulations, represented by black dashed lines in Figure 2.21, which provides the theoretical expectation when the foreground-removed observational data is noise-dominated. This corresponds to dark shaded region in Figure 2.18 but with deeper noise level as we average intensity maps and measure the variance in a volume as described above. Additionally, the noise simulations are used to estimate the observational error bars for each sub-band, without accounting for the effects of systematics on these error bars.

We show the m_2 measurements for Band 1 (top row) and Band 2 (bottom row) in Figure 2.21, where the observational data is denoted as pink circles with 2σ error bars from the noise simulations. The data points with m_2 smaller than zero are set to zero and the upper limit is set by the 2σ error bars from the zero. The data points with positive values have a upper limit of $m_2 + 2\sigma$. In addition, we illustrate the m_2 measurements for 21 cm simulation data including fiducial, cold reionization, large halos, and extended reionization models. Cosmic variance is illustrated as a shaded region for the cold reionization model, accounting for roughly 15% of the amplitude of m_2.

It is apparent that some frequency bins are heavily contaminated by systematics residuals, leading to a large variance beyond the 2σ error bar compared to the expectation. The variance excess is more evident across frequency bins than we see in Figure 2.18 because the thermal noise is integrated down through averaging the images along the frequency while the systematics may not be integrated down. Morales et al. (2018) studied the effects of calibration errors on an imaging power spectrum in comparison with a delay-spectrum power spectrum. They found that the imaging power spectrum is more sensitive to frequency-independent calibration errors than the delay-spectrum power spectrum because calibration errors from different baselines introduce spectral structure in the image space. The integral of the imaging power spectrum corresponds to the variance we measure, which may explain why our variance measurements are sensitive to systematics.

If the data is free from systematics and aligned with the thermal noise expectation, Band 2 may be employed to reject the cold reionization, especially in Field II as the data points are lower than the model. However, in the presence of the systematics residuals, it is

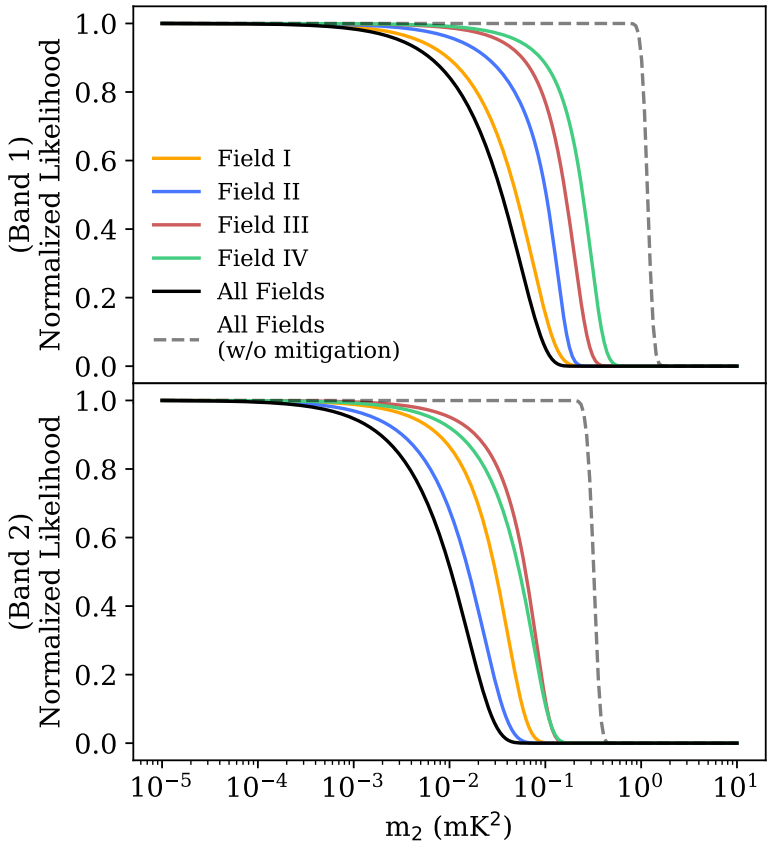

Figure 2.22: Normalized likelihood for m_2, marginalized over systematics defined in Equation (2.23). We present the normalized likelihood for each field for Band 1 (top) and Band 2 (bottom), using all observational data points shown in Figure 2.21. The black solid line is the likelihood considering all data points across all fields. To see the impact of systematics mitigation, we provide the likelihood without the mitigation, based on all fields (gray dashed line). We assume the variance model of $m_2 \propto \nu^0$. The cold reionization model has m_2 about 0.07 mK^2 on average in Band 2, which implies the model is likely to be rejected by the likelihood analysis.

not trivial to make such conclusion. For Band 2, there is one frequency bin in Field II where the m_2 value of the observational data, when considering the 2σ error bar, is lower than that of the cold reionization model. The upper limits could help constrain certain model parameters.

To understand the impact of the variance measurements on evaluating 21 cm models, we employ a likelihood analysis, similar to the approach used in H22b and H23. The likelihood is a probability of the observed data given different values of the model parameters. Assuming random Gaussian thermal noise, the likelihood is written as,

$$\mathcal{L}(\mathbf{d}|\theta, \mathcal{M}, \mathbf{u}) \propto \exp\left(-\frac{1}{2}\mathbf{r}(\theta, \mathbf{u})^{\mathrm{T}}\Sigma^{-1}\mathbf{r}(\theta, \mathbf{u})\right), \tag{2.22}$$

where $\mathbf{r}(\theta, \mathbf{u}) = \mathbf{d} - \mathbf{u} - \mathbf{W}\mathbf{m}(\theta)$, $\mathbf{d}$ is the m_2 measurement of observational data, θ are parameters of a model $\mathcal{M}$, and $\mathbf{u}$ is the systematics. $\mathbf{W}$ is a window function that maps a theoretical model to observable. Because we forward model our measurements, our 21 cm measurements account for the effect of the window function. Σ represents a covariance matrix of the observational data and models.

Assuming the systematics are positive, which is supported by Figure 2.18, a marginalized likelihood over systematics can be obtained by integrating Equation (2.22) with the constraint $\mathbf{u} \geq 0$,

$$\mathcal{L}(\mathbf{d}|\theta, \mathcal{M}) \propto \prod_{i}^{N_\mathrm{d}} \frac{1}{2}\left(1 + \mathrm{erf}\left[\frac{t_i}{\sqrt{2}\sigma_i}\right]\right), \tag{2.23}$$

where erf is the error function, t_i is ith element of $\mathbf{t} = \mathbf{d} - \mathbf{W}\mathbf{m}$, and N_d is the number of frequency bin. σ_i is the square root of the quadrature sum of the variance derived from noise simulations and the cosmic variance corresponding to 15% of the model variance. In this derivation, we assume each bin is uncorrelated. The details on the derivation are described in H22b.

In Figure 2.22, we displays the normalized likelihood, computed for all data points of each field given in Figure 2.21. The "All Fields" result (black solid line) is based on all data points from all four fields. Top and bottom panels represent Band 1 and Band 2 results, respectively. We assume the model variance of $m_2 \propto \nu^0$ for both bands. The shape of the normalized likelihood serves as an upper limit to constrain feasible model variance. We also present the normalized likelihood before systematics mitigation, indicating a significant improvement after the removal of systematics.

The likelihood for "All Fields" in Band 2 is the most effective in constraining models. The cold reionization model has about 0.07 mK^2 in m_2 averaged over the frequency band of Band 2. The likelihood at $m_2 \sim 0.07$ mK^2 is $< 10^{-5}$, suggesting that the cold reionization model is likely to be rejected, consistent with findings from power spectrum analyses in H22b and H23. To investigate other models, we need an improvement in variance by at least two orders of magnitude.

2.9.2 Comparison with 21 cm Model based on m_3

Investigating EoR models through skewness measurements, considering the sensitivity of the observational data, instrumental effects, and signal loss due to foreground removal, is quite challenging. In this section, we show the m_3 measurements of the observational data in comparison with the cold reionization model, which exhibits a larger variance than other 21 cm models.

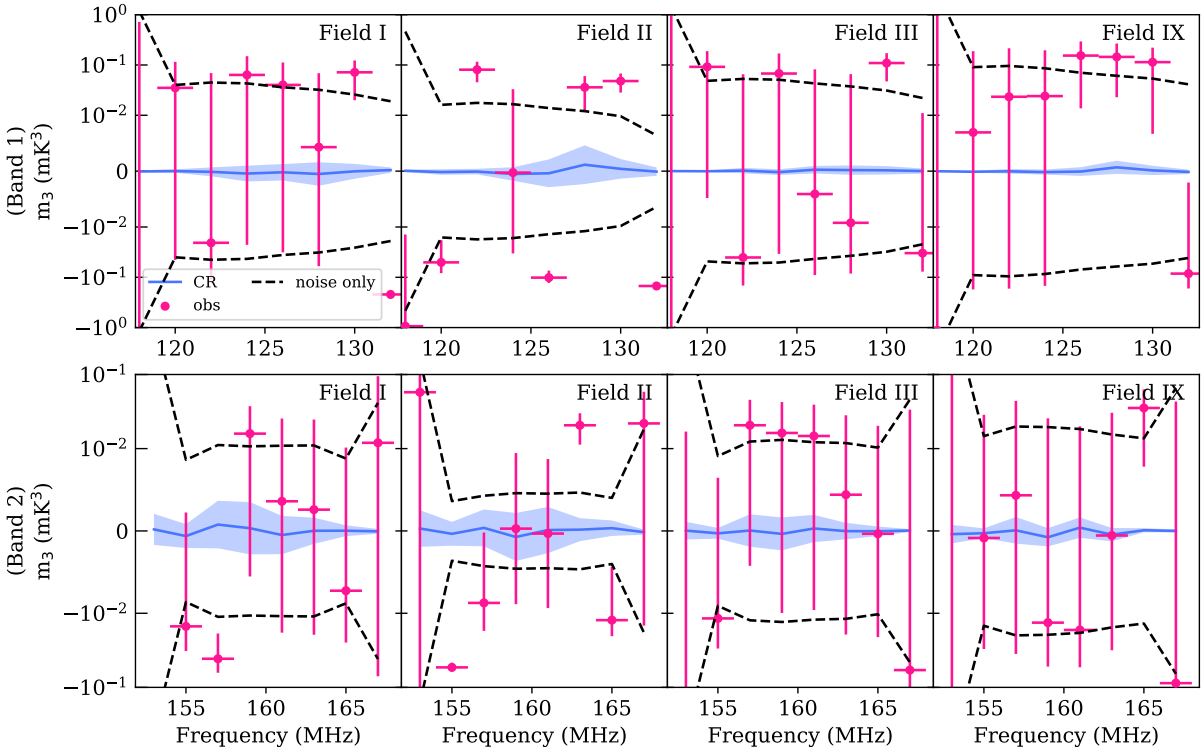

Figure 2.23: m_3 measurements for wedge-removed observational data (pink circles) in comparison to simulations for Band 1 (top) and Band 2 (bottom) across all fields. The 1σ errors in m_3 derived from noise simulations are represented by black dashed lines. The observational data is along with 2σ error bars, which are derived from the same noise simulations. Though there are some outliers especially in Field II, the observational data is largely consistent with the noise simulations. Additionally, the noiseless cold reionization model, characterized by larger variance than other 21 cm models, is depicted with blue lines including the cosmic variance shown in shaded regions. Due to the instrument resolution and wedge-filtering, the model effectively shows zero m_3. A close-up view of this model is presented in Figure 2.24.

Similar to Section 2.8.2, we make image cubes by averaging maps from even and odd timestamps. We average the images for 5 channels and bin the frequency-averaged data into 2 MHz sliced volumes to measure the m_3 as we do for the m_2 measurements.

We show m_3 measurements of wedge-filtered observational data, along with the noise simulations and the noiseless cold reionization model in Figure 2.23. The black dashed lines are the 1σ error derived from the noise simulations. Unlike the variance, skewness measurements of 21 cm model can be either positive or negative, and thus we present m_3 values of noise simulations in both axes.

The observational data, represented by pink symbols, is closely aligned with the black dashed lines within the 2σ error bars, which are also derived from the noise simulations. There are still some frequency bins that are heavily affected by systematics, particularly

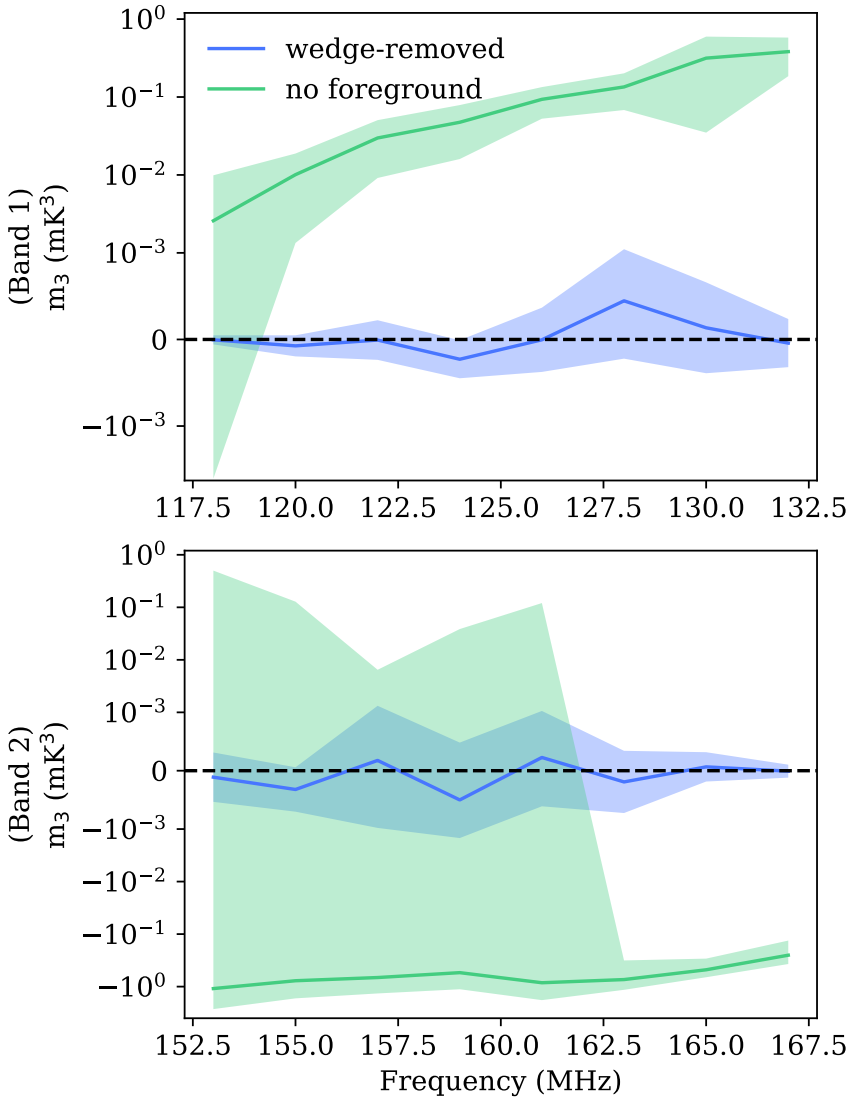

Figure 2.24: m_3 measurements for the noiseless cold reionization model for each band. The blue solid lines and shaded regions are the mean m_3 and 1σ cosmic variance of wedge-filtered data over 20 samples described in Section 2.3, respectively. The m_3 measurements assuming no foreground are also depicted for comparison.

in Field II of Band 1, but the overall features of m_3 are less affected by the systematics compared to the m_2 results shown in Figure 2.21. This could be due to the fact that the systematics at each frequency channel can be either positive or negative in m_3, and as a result, the effects of these systematics may be smeared out when measuring m_3 in 2 MHz volumes. But as we go deeper in our observation, the systematics will be crucial in evaluating 21 cm models using m_3 and the systematics should be mitigated in a deeper level.

The m_3 of the cold reionization model after the wedge-filtering becomes close to 0, which is shown in blue lines in Figure 2.23. Given the amplitude of the signals and the sensitivity

of the observation, drawing a meaningful conclusion from this result is challenging. Furthermore, because the systematics can be either positive or negative, the likelihood defined in Equation (2.23) cannot be used and it is hard to use the likelihood analysis in constraining the model parameters without any constraints on the systematics.

The reduction of m_3 values in 21 cm models is partially because the convolution of the instrument response (PSF) tends to wash out small-scale features responsible for non-Gaussianity (Wyithe and Morales, 2007). An additional decrease arises from the wedge-filtering, expected for the removal of large-scale frequency structures corresponding to small delay modes (Harker et al., 2009, Kittiwisit et al., 2022), as discussed in Section 2.5. Figure 2.24 shows a close-up version of the cold reionization shown in Figure 2.23. We use all fields in measuring the m_3 assuming there is no thermal noise. We present two cases for the m_3 measurements: the wedge-filtered data (blue) and no foreground data (green). Shaded regions represent the cosmic variance measured from 20 samples. The no foreground case, or perfect foreground-removal scenario, assumes that there is no signal loss of 21 cm signals when removing the foreground emission. In the top panel, the green line deviates from the zero baseline (black dashed line), which means we may be able to detect some non-Gaussianity if the foreground is removed without a signal loss of 21 cm signals and there is no thermal noise. Meanwhile, the amplitudes of the m_3 measurements are significantly reduced after applying the wedge-filtering method, making it difficult to detect m_3 even for the noiseless simulation. The significant cosmic variance shown in the bottom panel (green shaded region) might be due to m_3 being influenced by the presence of large-scale ionized bubbles, which emerge during the late stage of reionization. In the next section, we examine the prediction of m_2 and m_3 measurements assuming full HERA core array.

2.10 Forecasting Future Observations for One-point Statistics

As we discuss in Section 2.9.2, the relatively poor image resolution of Phase I observations tends to remove the non-Gaussianity information arising from small-scale structure. We extend the analysis to forecast m_2 and m_3 measurements considering a better resolution of the instrument for future observations using the full HERA core array. The full HERA core array (HERA-320; Dillon and Parsons, 2016) consists of 320 antennas with approximately 0.4° resolution. We simulate mock observations using the array configuration, with the fiducial 21 cm model described in Section 2.3.2. A Vivaldi feed beam (Fagnoni et al., 2021) is used for the predicted observation.

We consider 7-hr observation per night, spanning 3–10 hr in LST at a cadence of 30 seconds. The mock observation is simulated across 157.88–169.97 MHz with a frequency channel

width of 122.1 kHz, the same width of HERA Phase II observation. The frequency band covers a redshift range of 7.4–8.0, corresponding to the period when large-scale ionized bubbles begin to emerge. The mock visibilities are wedge-filtered and transformed into images using the same method applied to HERA Phase I observations. For the wedge-filtering, we consider an optimistic scenario where the foreground is well confined in the foreground wedge with buffer = 0 ns. In addition, we include the perfect foreground removal case for comparison. We repeat the processes 10 times by choosing different sky positions to estimate the cosmic variance in m_2 and m_3 measurements.

For the noise estimation, we utilize the system temperature, $T_{\rm sys} = T_{\rm rec} + T_{\rm sky} = 80 + 180\big(\nu/180\,{\rm MHz}\big)^{-2.6}\,{\rm K}$. The sky temperature ($T_{\rm sky}$) is driven from a relatively quiet zone on the sky (Haslam et al., 1982) and the receiver temperature ($T_{\rm rec}$) is adopted from the mean receiver temperature studied by Razavi-Ghods et al. (2017). The system temperature is used to construct the noise variance in the visibility space in Jy as follows:

$$\sigma_n = \frac{2k_B\nu^2\Omega_p}{c^2}\frac{T_{\rm sys}}{\sqrt{\Delta\nu\Delta t}}, \tag{2.24}$$

where Ω_p is the solid angle of the primary beam. The noise visibilities are created by using Equation (2.15). Similar to the approach used for Phase I observations, we simulate 30 random noise realizations assuming 300 and 1000 night observations and transform them into images after applying `DAYENU`-filtering with the same filtering configuration of the 21 cm simulation data as described above. To remove the noise bias in m_2, we split the visibility data into even and odd time groups and create the maps separately, as we do for Phase I observations.

The intensity maps for the 21 cm fiducial model and noise simulations are generated over a 90×10 square degree area by stitching together maps centered at each hour from 4 to 9 hrs, each map being 15×10 square degrees in size. Since we simulate EE-polarization only, we convert the noise maps to pseudo stokes-I maps by multiplying $1/\sqrt{2}$, assuming the effect of each polarization on noise map is approximately symmetric and each polarization map is uncorrelated. This enhances the sensitivity of statistical measurements. The 21 cm maps remain the same, assuming $m_{EE} \approx m_{NN} \approx m_I$.

Because our wedge-filtering and mapmaking processes are linear, we simply combine the 21 cm maps and noise maps to create noisy 21 cm maps. The m_2 and m_3 are measured from the noisy image cubes that are binned into 2 MHz sliced volume after being averaged over every 5 frequency channels.

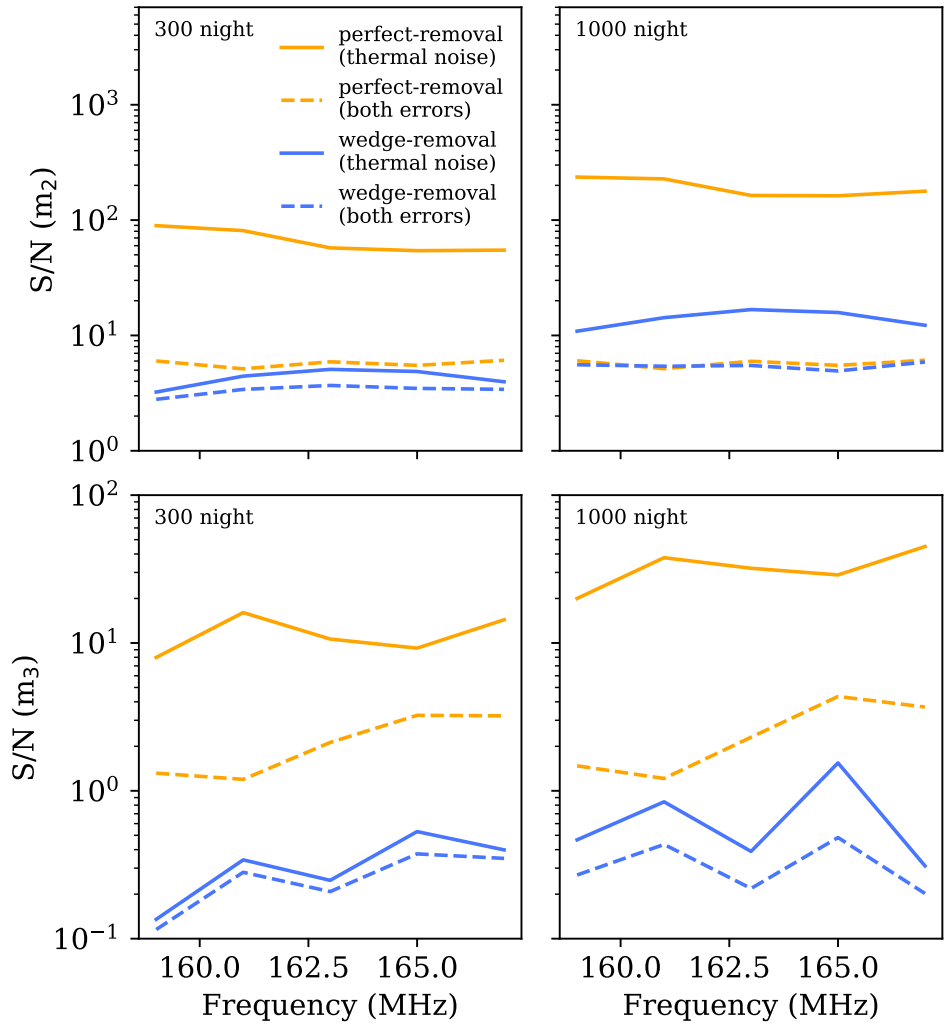

Figure 2.25: Predictions of S/N for m_2 (top) and m_3 (bottom) measurements assuming 300 (left) and 1000 (right) night observations, as forecasted with HERA (HERA-320) using the fiducial EoR model. We present the S/N considering different foreground removal methods including the perfect foreground removal and wedge removal, represented by orange and blue lines, respectively. In addition, we illustrate the case when considering the thermal noise only (solid lines) and the thermal noise and cosmic variance added in quadrature (dashed lines).

2.10.1 Detectability of One-point Statistics with HERA-320

We investigate the detectability of m_2 and m_3 for observation periods of 300 and 1000 nights. For each observation run, we estimate the thermal noise and cosmic variance in m_2 and m_3 measurements using the noisy 21 cm data cubes. The magnitudes of mean values of m_2 and m_3 across different samples and noise realizations are used as our signal values in computing S/N. Additionally, we estimate the S/N using two different foreground removal strategies: perfect foreground removal and wedge removal.

Figure 2.25 illustrates the S/N of m_2 (top panels) and m_3 (bottom panels) for different observation time. The perfect foreground removal and wedge removal methods are presented by orange and blue lines, respectively. In computing the noise, we account for two noise sources: thermal noise and cosmic variance. The solid lines represent the scenario considering only thermal noise from the noise simulations. The dashed lines show the results when both thermal noise and cosmic variance are combined in quadrature.

If the thermal noise is the only noise source, the S/N of m_2 measurements reaches around 100 for the 300-night observation, assuming the perfect foreground removal. When the wedge-filtering is applied, the S/N drops to around 5 for the 300-night observation and to 10 for the 1000-night observation. However, when the cosmic variance is considered in addition to the thermal noise, the S/N is limited by the cosmic variance, reducing to about 6 even for the 1000-night observation, regardless of the foreground removal methods. Note that this S/N can be improved if we take into account a larger sky coverage for mapmaking, which increases the number of independent samples and reduces the cosmic variance.

The m_3 measurements have a lower S/N compared to the m_2. With the optimistic case when considering the perfect foreground removal with the thermal noise only, the S/N can be around 10 for the 300-night observation. This gets worse when the cosmic variance is included, decreasing the S/N to about 2. It gets even worse once we apply the wedge-removal, and the S/N becomes less than one even for the 1000-night observation. This result is cosmic variance limited. Though the S/N is small, using all information from all frequency bins of m_2 and m_3 may provide some insight in constraining model parameters. In the following section, we examine the uncertainty of model parameters reconstructed from m_2 and m_3 measurements based on the uncertainty of observational data using Fisher information matrices.

2.10.2 Model Confidence Interval with Fisher Information Matrix

The Fisher matrix is a powerful tool in statistical estimation theory to assess the precision with which model parameters can be estimated from data (Fisher, 1935). The Fisher matrix is widely used in fields such as cosmology to forecast the performance of an instrument based on the power spectrum (Ewall-Wice et al., 2016b, Liu and Parsons, 2016, Pober et al., 2014). It is a square matrix that encapsulates information about how the likelihood function of a set of parameters behaves around its maximum likelihood (ML) estimate. Assuming the errors in data are Gaussian, the ln-likelihood is expressed as,

$$\ln \mathcal{L}(\mathbf{x}, \boldsymbol{\theta}) = -\sum_{\beta} \frac{1}{2\sigma_\beta^2}(x_\beta - m_\beta(\boldsymbol{\theta}))^2, \tag{2.25}$$

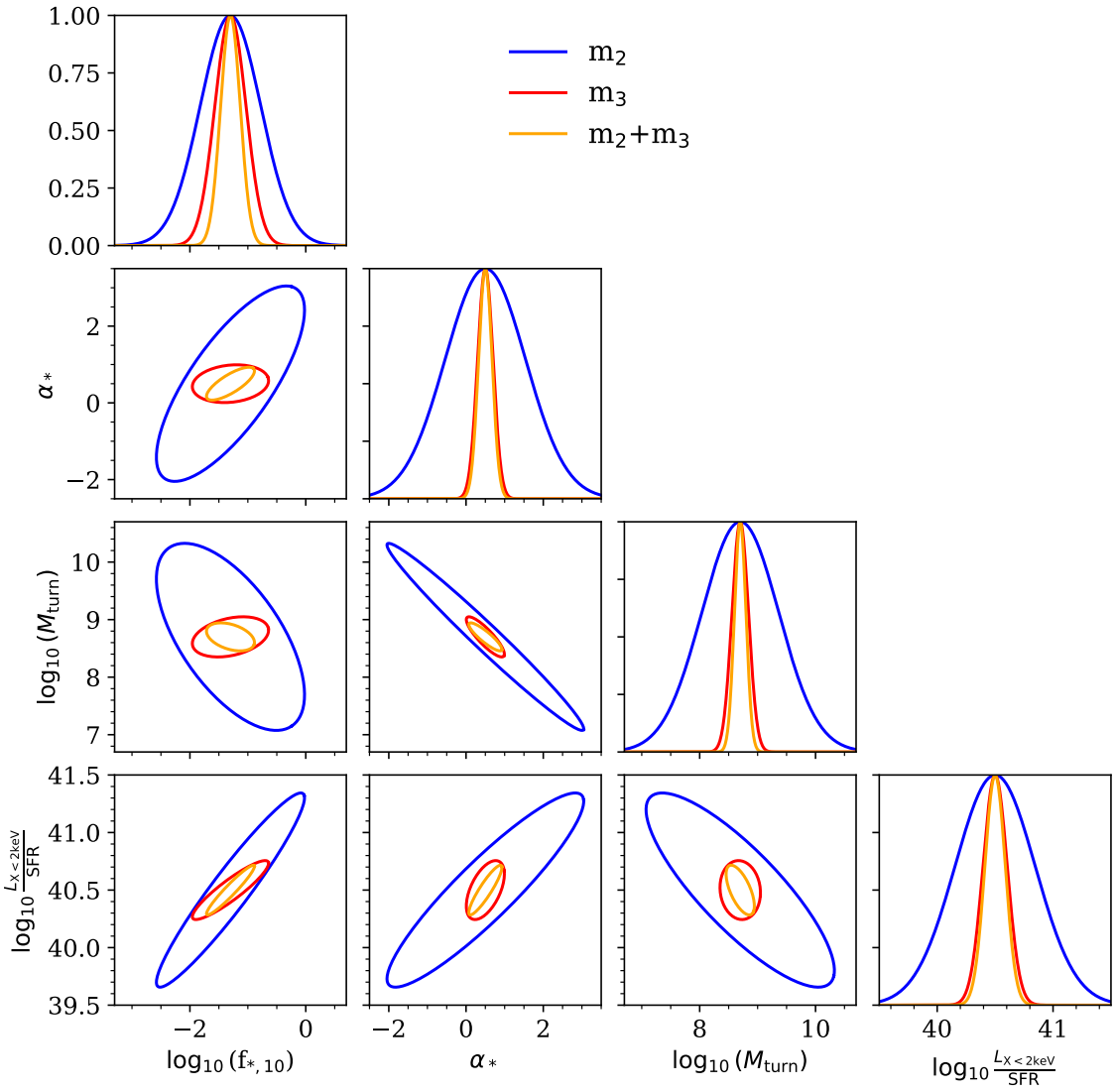

Figure 2.26: The 95% confidence interval derived from the m_2 and m_3 measurements, assuming the perfect foreground removal over the 300-night mock observation with the fiducial EoR model.

where β indexes indicate the frequency bins, x_β is the one-point statistics measurement of observed data, σ_β^2 is the quadrature sum of the thermal noise and cosmic variance we explore in the previous section, and m_β is the one-point statistics measurement varying with model parameters $\boldsymbol{\theta}$.

The elements of the Fisher matrix are defined as the expected value of the second derivative (the Hessian) of the ln-likelihood function with respect to the parameters. For the simple

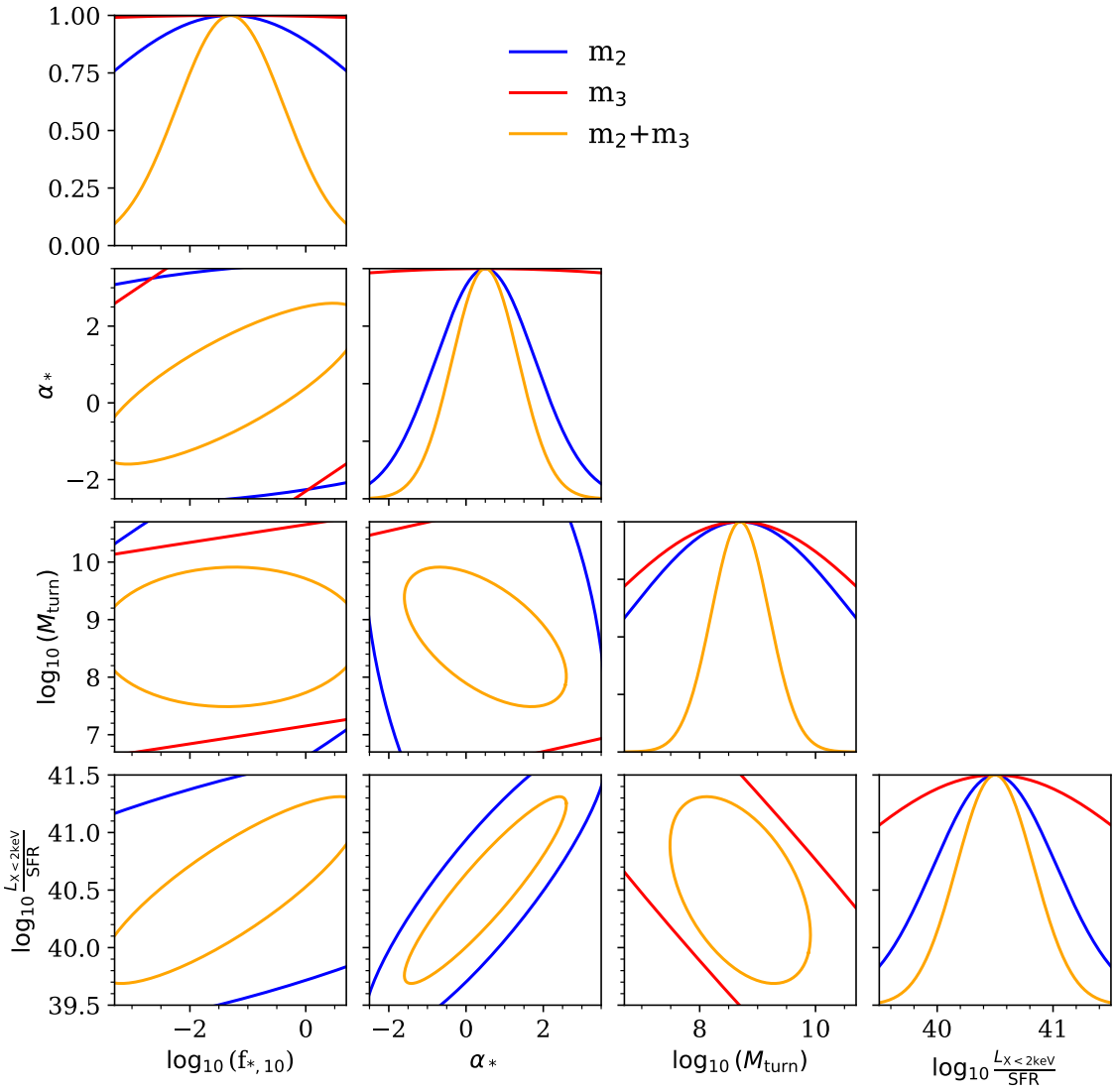

Figure 2.27: Similar to Figure 2.26 but with the wedge-filtered data. Due to the low S/N of m_2 and m_3, the parameters are loosely constrained.

Gaussian likelihood in Equation (2.25), the Fisher matrix is given by,

$$F_{ij} = -\left\langle \frac{\partial^2 \ln \mathcal{L}}{\partial\theta_i \partial\theta_j} \right\rangle \tag{2.26}$$

$$= \sum_\beta \frac{1}{\sigma_\beta^2} \frac{\partial m_\beta}{\partial\theta_i} \frac{\partial m_\beta}{\partial\theta_j}. \tag{2.27}$$

If the likelihood function is Gaussian, the covariance matrix of the parameters is the inverse of the Fisher matrix,

$$C_{ij} = (\mathbf{F}^{-1})_{ij}, \tag{2.28}$$

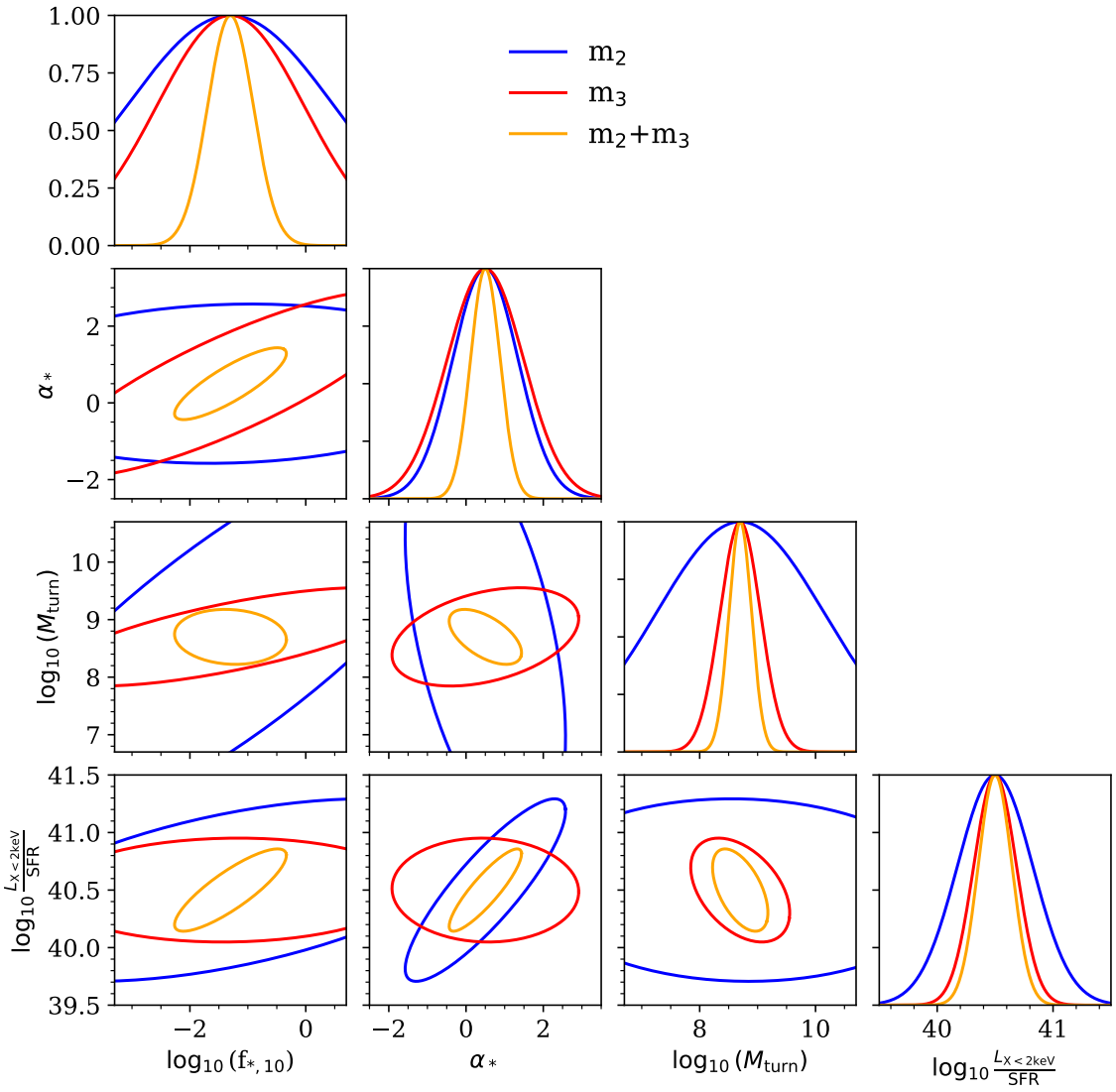

Figure 2.28: Similar to Figure 2.27 but for the 1000-night mock observation with the fiducial EoR model. With the improved S/N compared to the 300-night observation, we have a better precision in parameter estimation, suggesting that we are able to discern model parameters of the four models that are presented in Section 2.9.

which provides confidence intervals of the parameters. To compute the first derivative in Equation (2.27), we run 21cmFAST by slightly perturbing the model parameters from the fiducial model, forward model image cubes from visibility simulations for 10 samples at different sky locations, calculate the mean of m_2 and m_3 over the samples, and use the mean values to compute the first derivative. We focus on $f_{*,10}$, σ_*, $M_{\rm turn}$, and $L_{\rm X<2keV}/{\rm SFR}$ which are key parameters that determine the four models discuss in Section 2.9. For the rest of the parameters, we assume their values are fixed as indicated in Table 2.1.

Assuming the fiducial EoR model is at the ML point, we show the results of covariance

Table 2.3: The 1σ confidence intervals derived from the Fisher matrix with different types of observations and measurements. The small difference between 300 and 1000 night observations for the perfect foreground removal is because the S/N of m_2 and m_3 measurements is determined by the cosmic variance rather than the thermal noise. With the sufficient observation time (1000 nights), $m_2 + m_3$ measured from the wedge-filtered data is able to constrain the model parameters enough to distinguish four models given in Table 2.1.

	$\log_{10}(f_{*,10})$	α_*	$\log_{10}(M_{\rm turn})$ $(M_\odot)$	$\log_{10}\frac{L_{X<2\,\rm keV}}{\rm SFR}$ $(\rm erg\,s^{-1}\,M_\odot^{-1}\,yr)$
Fiducial model	-1.30	0.50	8.70	40.50
Perfect removal (300 nights)				
m_2	±0.52	±1.04	±0.66	±0.34
m_3	±0.27	±0.20	±0.14	±0.10
$m_2 + m_3$	±0.17	±0.18	±0.10	±0.09
Perfect removal (1000 nights)				
m_2	±0.52	±1.04	±0.66	±0.34
m_3	±0.27	±0.19	±0.14	±0.10
$m_2 + m_3$	±0.16	±0.16	±0.09	±0.08
Wedge removal (300 nights)				
m_2	±2.70	±1.28	±1.92	±0.53
m_3	±15.20	±15.54	±2.47	±1.43
$m_2 + m_3$	±0.92	±0.86	±0.50	±0.33
Wedge removal (1000 nights)				
m_2	±1.78	±0.85	±1.44	±0.32
m_3	±1.27	±0.99	±0.35	±0.18
$m_2 + m_3$	±0.39	±0.38	±0.20	±0.15

matrix for the perfect foreground removal with the 300-night observation in Figure 2.26. The contours indicate the 95% percentile confidence intervals. The confidence interval constructed from m_2 reveals degeneracy between model parameters. The m_3 provides tighter constraints in different directions from the results of m_2, which helps break the degeneracy. The best precision is achieved when considering both m_2 and m_3 together. The 1σ errors of the 1D profiles in the diagonal panels are indicated in Table 2.3

In practice, it is difficult to make a perfect foreground removal without a signal loss of 21 signals. Figure 2.27 displays the results with the 300 night observation when the wedge-filtering is applied. As we see in Figure 2.25, the S/N of m_2 and m_3 is significantly reduced with the wedge removal, which make it challenging to constrain the model parameters using the Fisher matrix. Nevertheless, using both m_2 and m_3 provides a constraint for $M_{\rm turn}$ tight enough to discern the fiducial model from large halos and extended reionization models.

Lastly, we examine the wedge-filtered case with a long observation run (Figure 2.28). For the 1000-night observation, there is an improvement in the S/N of m_2 and m_3, although not substantial. However, the precision of parameter estimation is significantly enhanced

compared to the 300-night observation. Because m_2 and m_3 provide orthogonal information in certain parameter spaces, combining $m_2 + m_3$ results in tight confidence intervals, offering high precision capable of distinguishing between models described in Table 2.1. This indicates that skewness, along with variance, may serve as a promising measurement for estimating model parameters with high precision.

2.11 Summary

In this study, we examined the properties of images constructed by the HERA radio interferometer, after removing bright foreground emissions. Using HERA Phase I observations, previously employed for power spectrum analyses (H23), we measured one-point statistics. We forward-modeled the observational data, 21 cm simulations, and noise simulations to account for pixel correlation effects caused by the instrument response and frequency correlations arising from the foreground removal technique.

The observational data includes 94 nights of observations for Band 1 and Band 2. We analyzed the data before and after systematics mitigation to study the impact of systematics on one-point statistics such as the second (m_2) and third (m_3) moments, corresponding to variance and skewness, respectively. The radio interferometric images were constructed using the direct optimal mapping method, which optimally down-weights visibility data with large noise variance. We mapped four different fields of the sky.

We adopted a foreground avoidance method to remove foregrounds. To evaluate the efficacy of the foreground avoidance strategy using the DPSS basis, we compared the performance of three different foreground removal methods: `DAYENU`-filtering, the standard wedge-masking method, and the PCA foreground subtraction method. Based on simulation results, we found that `DAYENU`-filtering outperforms the other two methods, resulting in foreground residuals at about 10% of the 21 cm signal residuals. Consequently, we chose `DAYENU`-filtering for our foreground removal method.

In Section 2.5, we discussed the impact of instrument effects on the one-point statistics of cosmological models. The instrument response or PSF from the Phase I observation filters out small-scale signals, significantly reducing the amplitude of m_2 and m_3 measurements. Further reduction occurs due to wedge-filtering, which removes large-scale fluctuations, leaving near-zero m_3 and S_3 signals, except at low redshifts where weak non-zero signals might appear. We also examined the predicted one-point statistics for HERA-320, which utilizes 320 antennas. The results indicate that, by capturing additional small-scale information, HERA-320 offers a better chance to detect non-Gaussian features in m_3 and S_3 across redshifts, with relatively small sample variance. However, wedge-filtering continues to cause substantial signal loss, rendering the measurements indistinguishable from the zero baseline

due to the sample variance. Despite this, S_3 shows a distinct rising trend toward lower redshifts.

By comparing the observational data with and without systematics mitigation, we observed that systematics can introduce non-Gaussianity into the pixel distributions of the wedge-filtered maps (Section 2.8). The data with systematics mitigation efforts is expected to be noise-limited if there are no systematics; however, they still contain some systematics residuals that lead to an excess of variance (m_2) compared to the expected noise level. In Figure 2.19, evident features of foreground residuals are observed, which may arise from calibration errors or other remaining systematics effects coupled with foreground emissions. The m_3 measurements from the observational data also show outliers beyond the theoretical expectation drawn from the noise simulations.

In Section 2.9, we present the one-point statistics of observational data compared with 21 cm and noise simulations. The lowest m_2 for the noise data is achieved in Field II, showing significant discrepancies between the observational data and theoretical expectation for most of the frequency bins due to the systematics contamination. A likelihood analysis using m_2 from the observational data was conducted to evaluate 21 cm models. We found that the m_2 measurements from all fields in Band 2 likely reject the cold reionization model, characterized by larger variance than other 21 cm models, which is consistent with the power spectrum results (H22b, H23). Due to the large size of the synthesized beam and the signal loss caused by wedge-filtering, detecting non-zero m_3 measurements of the 21 cm model is not feasible.

Finally, we forecast future observations with the full HERA core array. The improved array configuration provides better image resolution, reducing the wash-out of non-Gaussianity from small-scale features. We investigated the S/N of m_2 and m_3 after wedge-filtering compared to perfect foreground removal without 21 cm signal loss. Assuming 300-night observations, the S/N with the perfect foreground removal scenario achieves 100 for m_2 and 10 for m_3 if the mock observation only contains thermal noise. When the cosmic variance is accounted for, the S/N reduces to 6 for m_2 and 2 for m_3. This indicates the S/N is restricted by the cosmic variance than the thermal noise. Wedge-filtering significantly reduces S/N, resulting in S/N < 1 for m_3 measurements even for long observation periods (1000 nights). If we combine all frequency bins, however, Fisher matrix analysis suggests that, with wedge-filtering applied over 1000-night observations, m_3 along with m_2 measurements yield sufficiently tight reconstructed confidence levels of model parameters, allowing for exploration of the parameters described in Table 2.1. This may indicate that higher-order statistics, which represent non-Gaussianity, can be a crucial tool for studying model parameters associated with ionized bubbles.

Since wedge-filtering makes it challenging to detect non-Gaussian features with a high S/N, minimizing its effect is crucial. Achieving a smaller foreground wedge size may be

possible with an image-based power spectrum rather than the delay-based power spectrum (Morales et al., 2018). There are also efforts to reduce the wedge size by optimizing the array configuration (Murray and Trott, 2018). A smaller wedge size may improve the detection of non-Gaussianity in 21 cm signals. Additionally, efforts are underway to minimize 21 cm signal loss by isolating the signal from the foreground using GPR (Kern and Liu, 2020, Mertens et al., 2020, Soares et al., 2021). Recently, attempts have been made to recover 21 cm signals within the wedge after foreground removal using machine learning techniques (e.g., Gagnon-Hartman et al., 2021). With these advanced techniques, future observations for HERA may lead to more optimistic results.

Chapter 3

The Impact of Beam Variations on Power Spectrum Estimation for 21 cm Cosmology I: Simulations of Foreground Contamination for HERA

This work is published in The Astrophysical Journal. Author lists include Honggeun Kim, Bang D. Nhan, Jacqueline N. Hewitt, Nicholas S. Kern, Joshua S. Dillon, Eloy de Lera Acedo, Scott B. C. Dynes, Nivedita Mahesh, Nicolas Fagnoni, and David R. DeBoer.

Abstract

Detecting cosmological signals from the Epoch of Reionization (EoR) requires high-precision calibration to isolate the cosmological signals from foreground emission. In radio interferometery, perturbed primary beams of antenna elements can disrupt the precise calibration, which results in contaminating the foreground-free region, or the EoR window, in the cylindrically averaged power spectrum. For Hydrogen Epoch of Reionization Array (HERA), we simulate and characterize the perturbed primary beams induced by feed motions such as axial, lateral, and tilting motions, above the 14-meter dish. To understand the effect of the perturbed beams, visibility measurements are modeled with two different foreground components, point sources and diffuse sources, and we find different feed motions present a different reaction to each type of sky source. HERA's redundant-baseline calibration in the presence of non-redundant antenna beams due to feed motions introduces chromatic errors in gain solutions, which produces foreground power leakage into the EoR window. The observed leakage from vertical feed motions comes predominately from point sources around zenith. Furthermore, the observed leakage from horizontal and tilting feed motion comes predominately from the diffuse components near the horizon. Mitigation of chromatic gain errors will be necessary for robust detection of the EoR signals with minimal foreground

bias, and this will be discussed in the subsequent paper.

3.1 Introduction

The Epoch of Reionization (EoR), when luminous galaxies formed and interacted with the surrounding intergalactic medium, is of particular interest for better understanding of the universe's history. One promising approach for probing this period is to measure 21-cm emission originating from the hyperfine transition of neutral hydrogen at high redshift. This observation is feasible thanks to abundant neutral hydrogen in the early universe and the transparency of the forbidden transition line (see Furlanetto et al., 2006, Liu and Shaw, 2020, McQuinn, 2016, for reviews).

Radio interferometric experiments designed to detect such cosmological signals include the Giant Metre Wave Radio Telescope (GMRT; Paciga et al., 2013), the Murchison Widefield Array (MWA; Barry et al., 2019, Beardsley et al., 2016, Dillon et al., 2014, Ewall-Wice et al., 2016a, Tingay et al., 2013, Trott et al., 2020), the Donald C. Backer Precision Array for Probing the Epoch of Reionization (PAPER; Cheng et al., 2018, Kolopanis et al., 2019, Parsons et al., 2010), the Low Frequency Array (LOFAR; Gehlot et al., 2019, Mertens et al., 2020, Patil et al., 2017, van Haarlem et al., 2013), and the Hydrogen Epoch of Reionization Array (HERA; DeBoer et al., 2017, Dillon and Parsons, 2016, The HERA Collaboration et al., 2022a). They have placed upper limits on the power spectrum, but none has yet made a robust detection of the EoR signal.

In redshifted 21-cm observations for the EoR, the removal of the foreground, which is 4 to 5 orders of magnitude brighter, is crucial. One strategy for separating the cosmological signal from the foreground is to take advantage of the 2D power spectrum defined by the baseline length and time delay of an interferometer. A radio interferometer measures the sky signal by cross-correlating the voltage outputs received by a pair of antennas. Spectral smoothness of foregrounds confines a frequency Fourier transform of the measurements to low time delays or line-of-sight cosmological modes $k_\parallel$. Chromaticity of the interferometer can introduce spectral structure into the foreground emission, and cause foregrounds to leak into high $k_\parallel$ (Datta et al., 2010, Liu et al., 2014, Morales et al., 2012, Morales et al., 2018, Parsons et al., 2012, Pober et al., 2014, Thyagarajan et al., 2013, Trott et al., 2012, Vedantham et al., 2012). The maximal $k_\parallel$ is set by the delay of a source at the horizon which increases with increasing baseline length. This forms a "foreground wedge" in the 2D power spectrum and leaves the foreground confined in the wedge. In contrast, fluctuation in brightness of cosmological signals in the line-of-sight direction or frequency makes the measurement spread over a wide range of cosmological modes in the Fourier space, allowing a foreground-free "EoR window" outside the wedge.

For successful foreground removal, high-precision calibration is essential to prevent foreground contamination from leaking into the EoR window due to chromatic errors from inaccurate calibration. There are calibration techniques adopted by past and current interferometric experiments, including sky-based calibration (Pearson and Readhead, 1984, Rau et al., 2009) and redundant-baseline calibration (Dillon et al., 2020, Liu et al., 2010, Wieringa, 1992). The former relies on precise prior information of the sky, and an inaccurate or incomplete sky model can cause artificial frequency structure in calibration solutions (Barry et al., 2016, Ewall-Wice et al., 2017, Gehlot et al., 2021, Mouri Sardarabadi and Koopmans, 2018). The latter is accompanied by a key assumption of redundancy in measurements, and nonredundancy resulting from non-uniform primary beams can be a source of the chromatic gain error (Byrne et al., 2019, Choudhuri et al., 2021, Orosz et al., 2019). In addition, redundant calibration cannot solve for all degrees of freedom (dofs) of the direction-independent gain term, and must be followed by an absolute calibration step, which is also subject to the biases described above for sky calibration (Byrne et al., 2019, Kern et al., 2020a). In this study, we focus on the impact of feed motion on redundant-baseline calibration as well as its absolute calibration step as designed for the HERA instrument.

Orosz et al. (2019) studied the impact of non-uniform primary beams on redundant-baseline calibration and resulting power spectrum estimate. They simulated interferometric measurements with the HERA array configuration and an analytic Airy beam by perturbing the beam pointing angle and width. They found perturbed primary beams can introduce spectral structure into calibration solutions which are then responsible for the foreground in the wedge leaking into the EoR window. The actual HERA antennas have more complex structure in their beams and the result based on the Airy beam is hard to map to the real system of HERA. In addition, the simulation was performed with about 100 bright point sources and the effect of a diffuse sky model on the power spectrum analysis was not addressed. These motivate us to take one step further, by using a physical antenna model and a more representative sky model. Shaw et al. (2014) also studied the impact of per-antenna primary beam deviations on foreground removal for low-frequency 21-cm intensity mapping, finding that per-antenna beam deviations need to be constrained below 10% for robust signal detection.

In this study, we employ a realistic primary beam obtained from a physical HERA antenna model using the Computer Simulation Technology (CST) electromagnetic simulation software. In the model, we perturb the positions of the antenna feed relative to the dish and compute the corresponding primary beam patterns. We focus on the HERA Phase II system with a new broadband (50–250 MHz) Vivaldi feed that is designed to cover the epoch of Cosmic Dawn at lower frequencies as well as the EoR (Fagnoni et al., 2021). We choose a mid frequency band, 160–180 MHz, corresponding to $z \sim 7.4$ that is a midpoint of the late EoR model (Douspis, M. et al., 2015, Greig and Mesinger, 2016, Millea and Bouchet, 2018, Mitra et al., 2015, Qin et al., 2020) consistent with the constraints on the electron scattering

optical depth of the Cosmic Microwave Background observation (Planck Collaboration et al., 2020). For a representative foreground model, a diffuse sky model as well as point sources are taken into account. We simulate visibilities with the perturbed primary beams and the sky model, and pass them through HERA calibration pipelines. The effects of nonredundancy in the visibilities on the calibration gains and power spectrum estimation are explored. Further investigations into mitigation of the chromatic gain errors and foreground leakage will be discussed in the subsequent paper.

This paper is organized as follows. In Section 3.2, we detail the configuration of primary beam simulations with feed motions and present characteristics of the simulated beams. In Section 4.2, the visibility measurements are simulated with the perturbed beams, and nonredundancy of the visibilities are investigated as a function of baseline lengths, feed motions, and the type of sky model. In Section 3.4, we show results from redundant-baseline calibration as well as absolute calibration, representing chromatic gain errors. In Section 4.5, the effects of the feed motions on power spectrum estimation are discussed. Throughout this study, we adopt the cosmological parameters from Planck Collaboration et al. (2016): $\Omega_\Lambda = 0.6844$, $\Omega_\mathrm{m} = 0.3156$, $\Omega_\mathrm{b} = 0.04911$, and $H_0 = 67.27\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

3.2 Antenna Feed Perturbation Simulations

In this section we summarize the CST simulation descriptions and results of the antenna far-field electric field pattern in the presence of positional offsets between the Vivaldi feed and the 14-meter parabolic dish. The physical 3D model was derived from that of Fagnoni et al. (2021). The model consists of all the major components making up a single HERA antenna, including metal, PVC, and fiberglass components, coaxial cables, an enclosure for the front-end module (FEM) with its CAT7 connector, a cylindrical ferrite radio frequency (RF) choke for each coaxial cable, metal mounting hardware, suspension cables, a concrete hub, and a soil slab[1] beneath the dish.

3.2.1 CST Simulation Configuration

Due to the wind load and relaxation in the feed suspension system, the feed alignment can move away from the desired feed position at the focal point of the dish. Field measurement of the feed alignment using laser plum bobs on a subset of HERA elements across a time span of 8 weeks in 2020 determined that the standard deviation of the feed offset could be

[1]The material is assumed to be sandy dry soil with relative permitivity of $\epsilon_r = 2.55$ and loss tangent of $\tan\delta = 0.016$.

as large as several centimeters (Rath et al. HERA Memo #95[2]). Separate observations on tilts of feeds measured by accelerometers installed in the FEM indicate a root mean square (rms) of the tip-tilt of a few degrees.

Due to computational limitations, a simplified 3D Vivaldi feed model was derived from the detailed model developed by Fagnoni et al. (2021). We removed the smaller mechanical features, such as the screw mounting holes, coaxial cables, suspension cables, and the FEM. For a subset of the feed motions, we found that the overall far-field beam characteristics for the simplified model and the detailed model were consistent, so the simplification does not have a large impact on our study. The primary beam of each antenna element is computed in an isolated simulation domain bounding box, and we ignore mutual interactions between antennas, such as the cross-coupling effect in the HERA system. However, we expect that cross-coupling might degrade calibration solutions (see, for example, Josaitis et al., 2022, Kern et al., 2020b); calibration procedures that include a model of cross-coupling are under development. We defer a quantitative study of the interaction of beam imperfections and cross-coupling to future work.

The antenna dish model consists of 24 faceted parabolic panels, as in the deployed instrument. However, instead of using actual galvanized steel wire mesh panels, the dish model assumes solid aluminum panels. This simplification circumvents the challenge in properly simulating the wire mesh and significantly reduce the number of simulation meshcells. Fabrication imperfections such as surface smoothness, potential gaps between dish panels, or structural sagging due to gravity are not considered in this study. Due to the randomness in nature, those defects may lower the antenna gain due to signal loss but they are expected to have much less impact on the overall main beam's characteristics than the feed position perturbations.

The Vivaldi feed position was perturbed in the lateral xy-plane and along the boresight direction in the z-axis, separately, in the CST model. We set the range of feed offsets along the z-axis to be ± 7 cm, with increment of 0.5 cm between -2 and 2 cm and of 1 cm otherwise, relative to the fiducial feed height position of $(x, y, z) = (0.0, 0.0, 5.0)$ m where the z-position is the distance between the reference point on the feed (see the origin in Figure 2 of Fagnoni et al., 2021) and the vertex of the dish. The offsets in the xy-plane spanning ± 6 cm are sampled in a regular grid with 0.5 cm interval between -2 and 2 cm and 1 cm interval otherwise. We include an additional 9 points at the corner and the midpoint of each side of a 7-cm long square to deal with potential outliers.

The feed can also tilt relative to the dish as the suspending cables relax over time. We simulated tilts at the fiducial position only, without any additional transverse offset. The feed is tilted in the range of $\pm 6°$, with increment of $0.5°$ between $\pm 1°$ and of $1°$ otherwise,

[2] http://reionization.org/science/memos/

where 0° tilt is equivalent to the feed pointing straight down at the dish's vertex. The azimuthal motion is equally spaced at 10° for each tilt at the fiducial position.

In total, we have three different classes of the feed motions for the CST beam simulations: one set of CST simulations produced 19 perturbed beam patterns for the vertical displacement, another set produced 209 perturbed beams for all the horizontal offsets in the xy plane, and the other set produced 295 perturbed beam patterns for all the tilts at the fiducial position.

Far-field electric fields are simulated by exciting the East-West port over 160–180 MHz, with a frequency channel width of 0.125 MHz[3]. This frequency channel width is close to that of the HERA Phase II observation, and gives the large Nyquist limit needed for quantifying the amount of contamination in the EoR window at high $k_{\parallel}$.

The full-wave far-field simulation results were computed using CST's transient time solver based on Finite-Difference Time-Domain (FDTD) solver method, with open boundary conditions[4]. Benchmark tests suggest that CST's time solver with GPU acceleration[5] is more efficient for the required large number of beam simulations. A subset of the far-field results was compared with those of CST's Finite Element Method frequency solver, which is not compatible with GPU acceleration. The general beam characteristics due to feed motions are found to be consistent between the two solvers.

3.2.2 Characteristics of Primary Beams along the Feed Motion

Feeds which move away from their fiducial position produce perturbed far-field electric fields which are different from those at the fiducial position. This perturbed primary beam pattern can corrupt interferometric measurements and calibration pipelines. In this section, we present characteristics of the primary beams as a function of feed positions and frequency. We examine the shift and full width at half maximum (FWHM) of the main lobe of the power beam with feed motion along the three Cartesian axes or with tilt. In addition, we explore the spatial integral of the perturbed beam with respect to the unperturbed one to capture overall features of the perturbed beams with feed motion and frequency.

[3]The far-field results were exported with high precision to prevent round-off errors using CST's far-field source (FFS) export instead of the normal far-field export macro function. This is essential to achieve a large dynamic range in power spectrum estimation.

[4]The perfectly matched layer (PML) boundary conditions were applied at the the boundary interfaces of the simulation bounding box to emulate an infinite simulation domain for the electromagnetic wave to propagate outward with negligible reflection at the interfaces.

[5]Combining with local meshing scheme for small antenna components, the simplified antenna model with $\sim 7.5 \times 10^7$ volumetric meshcells took an average of 3–5 hours to simulate and export each beam set on a computer server equipped with two Tesla V100-PCIE-32GB GPUs.

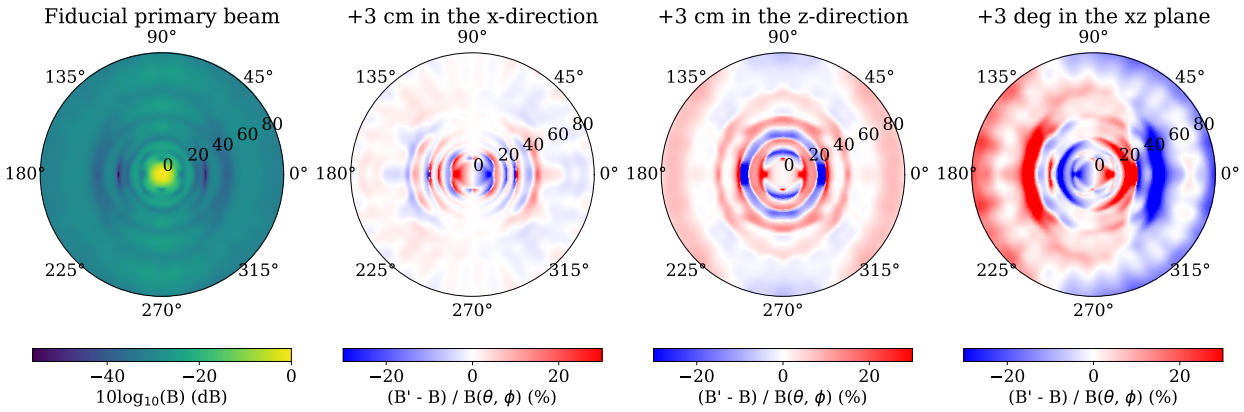

Figure 3.1: The fiducial beam pattern (far left) and the difference of the primary beam models (B') relative to the fiducial model (B) for feed displacement by +3 cm in the x-direction (center left), by +3 cm in the z-direction (center right), and tilt by +3° in the xz-plane (far right) at 165 MHz. The difference is normalized by the amplitude of the fiducial model. The different feed motion leads to the different beam pattern.

The far-field electric fields simulated by CST of a linear feed polarization p are defined in a basis of unit vectors $\hat{\theta}$ and $\hat{\phi}$ in spherical coordinate,

$$\mathbf{E}^p(\theta, \phi, \nu) = E^p_\theta(\theta, \phi, \nu)\hat{\theta} + E^p_\phi(\theta, \phi, \nu)\hat{\phi}, \tag{3.1}$$

where θ and ϕ are the zenith angle and the azimuthal angle, respectively. Equation 3.1 can also be thought of as a row in the antenna Jones matrix (e.g., Equation (5) of Kohn et al., 2019). The power beams are then calculated by multiplying each component of far-field electric fields assuming unpolarized sky emissions (Kohn et al., 2019),

$$B^{pp}_{ij}(\theta, \phi, \nu) = E^p_{i,\theta} {E^p_{j,\theta}}^* + E^p_{i,\phi} {E^p_{j,\phi}}^*, \tag{3.2}$$

where pp is a feed polarization pair, and i and j indicate a pair of antennas. Throughout this paper, we consider a single feed polarization ("east-oriented") in our CST and visibility simulations.

Figure 3.1 displays the fiducial power beam (far left) and the difference between the perturbed feed and the fiducial one at 165 MHz. The feed moving horizontally (center left) or tilting (far right) yields a shift of the beam pattern with that from the tilting motion moving opposite to that from the x-direction motion. The feed moving vertically changes the beam width, producing relatively isotropic differences (center right).

To quantify the characteristics of the main lobe, we first locate the maximum of the power beam and fit a one-dimension Gaussian profile to the first null in the xz plane (or in east-west direction, EW) or in the yz plane (or in north-south direction, NS) at boresight of the beam to measure the peak location and the beam width of the main lobe.

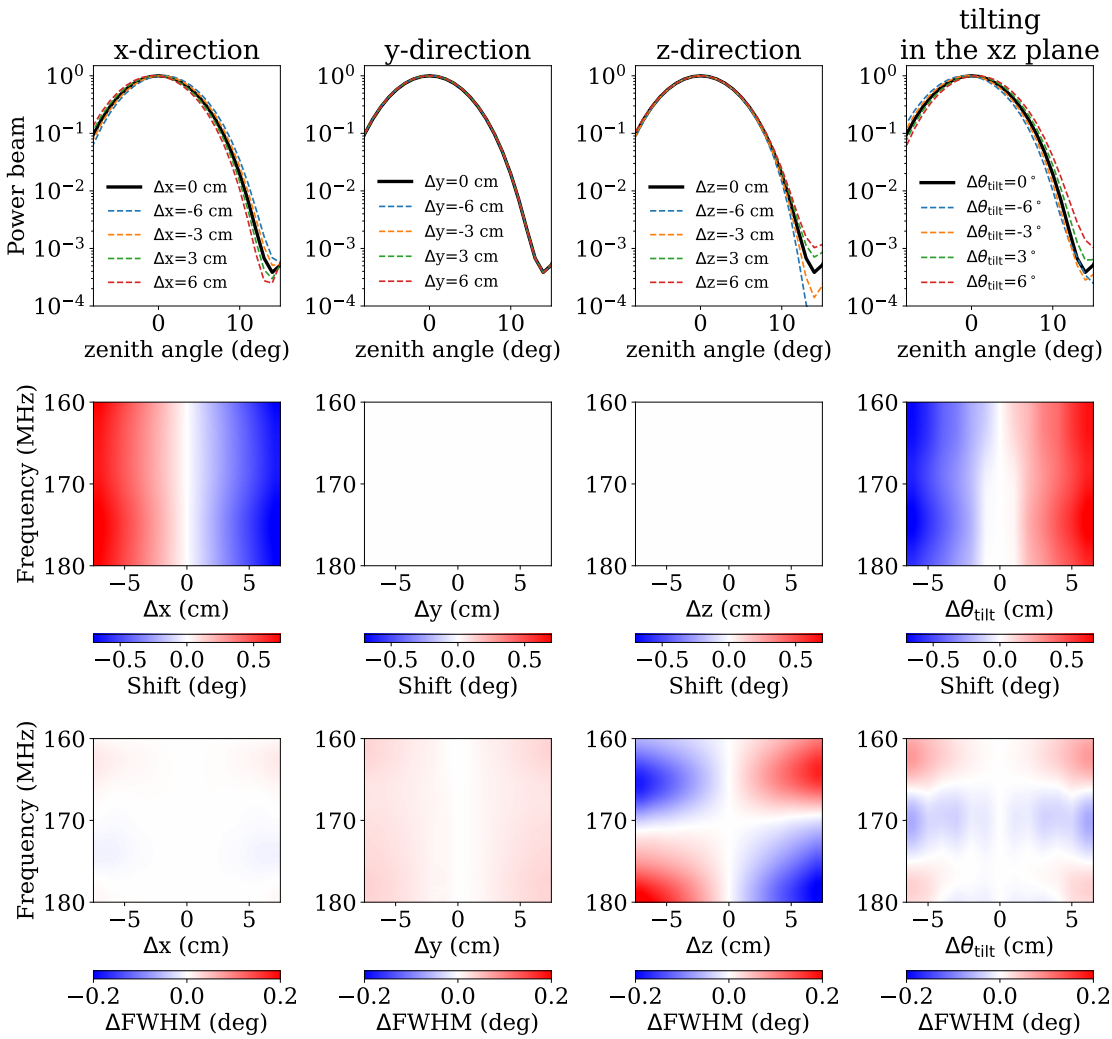

Figure 3.2: Behaviors of the main lobe of the power beam in the xz plane or in the EW direction with different types of feed motions: x-direction (far left), y-direction (center left), z-direction (center right), and tilting feed motions in the xz plane (far right). The top row shows line profiles of power beams that are normalized to the peaks at 165 MHz. Based on line profiles, we measured the shift of maximum of the main lobe (middle row) and the FWHM of the main lobe (bottom row) as a function of feed motion and frequency. When the feed moves in x-direction or tilts in the xz plane, the main lobe shifts in the xz plane as expected. The beam width changes relatively more when the feed moves in the z-direction.

Figure 3.2 shows the behaviors of the main lobe of the primary beam with the feed motion. In the top row, we show line-cut profiles of main lobes in the xz plane for the x-direction (first column), y-direction (second column), z-direction (third column), and tilting (last column) feed motion. If we look at the first panel, the line profile of the power beam shifts in the xz plane at different feed positions moving along the x-axis. The measured shift of the peak of the main lobe in the xz plane is presented in the middle row as a function of feed offset and frequency. In the first panel of the middle row, there are three interesting things to note.

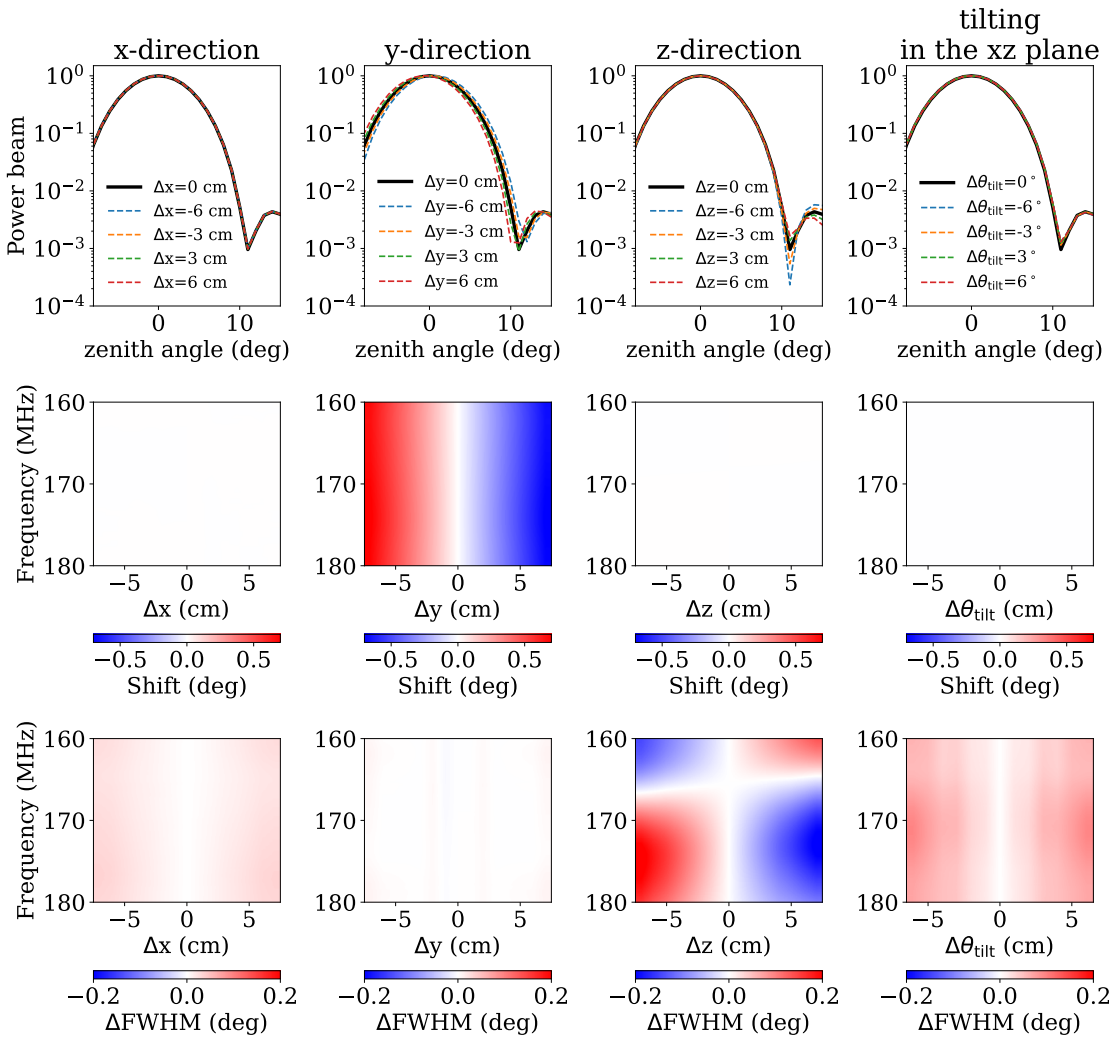

Figure 3.3: The same format as that of Figure 3.2 but for the behaviors of the primary beam in the yz plane or in the NS direction. The shift of the peak location in the main lobe is observed for the y-direction feed motion, while the change to the beam width mainly arises from the vertical motion or the tilting motion.

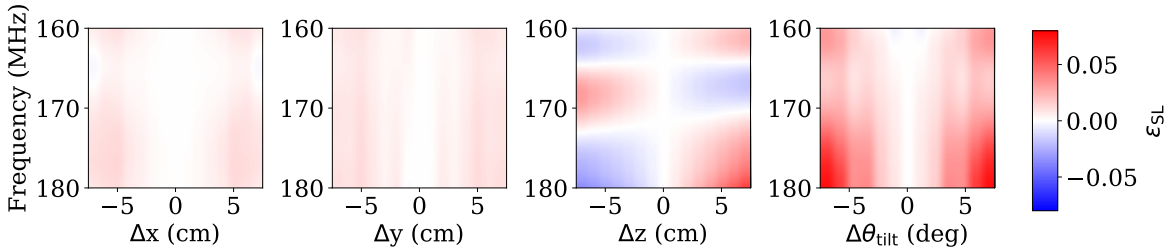

Figure 3.4: Integrated primary beams outside the main lobes of the perturbed beam with respect to the unperturbed beam. For the translational motions shown in the first three panels, the feed displacement is labeled from -7 to 7 cm. For the tilting motion, we consider the motion in the xz plane with tilt measured from zenith spanning from -6° to 6°.

First, the direction of the beam shift is opposite to that of the feed displacement. Second, the shifts of the peaks are frequency independent. Finally, there is a nearly linear relationship between the feed offsets and the beam shifts. These characteristics can be understood from geometrical optics, $\theta \approx -\Delta x/h$ [rad] $= -0.11\,(\Delta x/\mathrm{cm})$ [deg], where Δx is the offset of an object from the fiducial point, and h is the normal distance between the object and a lens. In the second equality, we consider our feed model where Δx is the offset of the feed in the x-direction and $h = 500$ cm, resulting in the relation consistent with the trend we observe. When the feed moves along the x-axis, the FWHM of the main lobe is weakly dependent on the feed motion (far left panel of the bottom row of Figure 3.2) and this may be because of a defocusing effect when the phase center deviates from the focal point with the feed motion. If the feed moves in the y-direction (center left panels of Figure 3.2), the profile remains nearly the same in the xz plane as expected.

The change in the beam shape with the feed offset in the z-axis is different from that with the lateral feed motion. Unlike the lateral motion, the peak of the main lobe stays unchanged with the feed motion (the third column of Figure 3.2), because change in the beam shape is symmetric and thus the beam is still pointing along the z-axis.. Rather, the axial motion leads to the change of the beam width as shown in the center right of the bottom row. This may be associated with the relation between the phase center and the focal point. That is, when the phase center, moving with the feed height, gets farther away from the focal point, the beam goes through the defocusing effect and its width grows broader (Baars, 2007). The phase center is frequency-dependent, so this effect is different for different frequencies.

When the feed tilts, the beam is also expected to tilt in the same direction. In the last column of Figure 3.2, the primary beam shifts in the EW direction when the feed tilts in the xz plane. Because the main lobe tilts with the feed motion, the beam does not just shift but its width also changes as a function of tilting angle.

In Figure 3.3, we show the same feed motions but the line-cut profiles in the yz plane. Feed offsets in the y-direction cause the shifts of the main lobe in the yz plane, while the feed motions along the x-direction or the tilting motions in the xz plane result in consistent line profiles regardless of the size of the feed displacement. The peak location of the main lobe still stays unchanged with the vertical feed motion while there is a clear trend of the change to the beam width as a function of the feed motion and frequency. The FWHM of the main lobe becomes slightly broader in the NS direction with the tilting motion.

In general, the amplitudes of the side lobes are a few order of magnitudes smaller than that of the main lobe. However, integrated properties of the side lobes can be crucial when a sky model contains diffuse bright sources in the side lobes. To measure the variation of

side lobes due to feed motion compared to the fiducial model, we define

$$\varepsilon_{\rm SL} = \frac{\iint B_{\rm pert}(\theta,\phi)\, d\Omega}{\iint B_{\rm unpert}(\theta,\phi)\, d\Omega} - 1. \tag{3.3}$$

Here $B_{\rm unpert}(\theta,\phi)$ and $B_{\rm pert}(\theta,\phi)$ are unperturbed and perturbed power beam, respectively. The integral is over $60^\circ \leq \theta < 90^\circ$ and $0^\circ \leq \phi < 360^\circ$. Anticipating our visibility simulations, the integral range is chosen so that it includes the region where the galactic plane lies at a local sidereal time (LST) of interest, 2.25 hours. The result is shown in Figure 3.4. The largest change in $\varepsilon_{\rm SL}$ is observed for the tilting motion and there is little change for the lateral feed motions.

3.3 Visibility Simulations with Perturbed Primary Beams

3.3.1 Foreground Visibility Simulations

A visibility for two antennas is the interferometric response with the amplitude proportional to the beam power pattern along with the flux density of the sky and the phase depending on the frequency and the geometric delay. This can be described in a discretized form as

$$V_{ij}(\nu) = \sum_{n=1}^{N_{\rm src}} B_{ij}(\hat{s}_n,\nu) S(\hat{s}_n,\nu) \exp\left(-\frac{2\pi i \nu}{c}\mathbf{b}_{ij}\cdot\hat{s}_n\right), \tag{3.4}$$

where B_{ij} is the primary beam in Equation (3.2), S is the flux density of a source, $\mathbf{b}_{ij}$ is a baseline vector of two antennas, and $\hat{s}_n$ is a pointing vector in the direction of a source.

In this study, we simulate visiblities using 320 antennas which make up the core of the HERA array (HERA-320; Dillon and Parsons, 2016). The configuration of the array is shown in Figure 3.5. There are 51,360 baselines in total and 1,502 unique baselines. This highly redundant-baseline configuration enables us to achieve high precision calibration with HERA if all antennas have uniform antenna responses. However, the distinct antenna response among the array elements due to feed displacement invalidates the redundancy assumption thus compromising the calibration. We account for this by assigning a different primary beam to each antenna in Equation (3.4) and seeing the effect of non-uniform antenna responses on the calibration.

In practice, the feed position of each antenna can be perturbed in any direction with random distances with respect to the fiducial position. As mentioned in Section 3.2.1, the sizes of feed offsets are expected to be a few centimeters for translation motions and a few degrees for tilts. For simplicity, we consider random Gaussian feed motions for 320 antennas

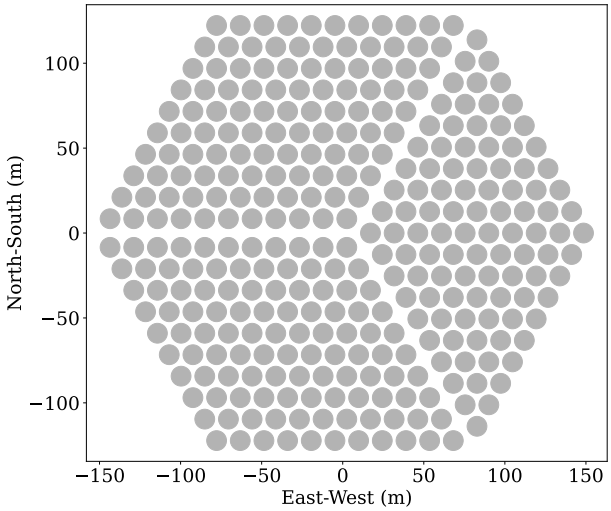

Figure 3.5: Configuration of HERA-320. This highly redundant array configuration is designed to achieve high-precision calibration and deep sensitivity to the 21-cm power spectrum.

with zero mean and a standard deviation (σ_{feed}) of 1, 2, 3, and 4 cm for lateral and axial feed motions, respectively, and of 1, 2°, and 3° for tilting motions.

Since the primary beams are simulated at grid points of feed motions described in Section 3.2.1, the primary beam patterns used in Equation (3.4) need to be interpolated at the feed positions drawn from the Gaussian distribution. Interpolating the primary beams along the feed motion direction can introduce spectral artifacts. We found, however, such an error is a much smaller effect than the beam variations due to the feed motion. Appendix 3.7.1 describes the interpolation and presents the interpolation error relative to the beam error driven by the feed motion.

In addition to the interpolation along the feed motion direction, the primary beam needs to be evaluated at the frequency of interest and angular positions of sources, requiring additional interpolation. Interpolating the primary beam along frequency at a finer resolution than the beam resolution may introduce unwanted spectral structure which depends on the interpolation methods (Lanman et al., 2020). To avoid this effect, the visibility needs to be simulated with a frequency channel width larger than or equal to the frequency resolution of the primary beam simulation. Large frequency channel widths, however, will shrink the size of the observable EoR window which is proportional to $1/(2\Delta\nu)$ where $\Delta\nu$ is the channel width. Based on the result of Orosz et al. (2019), in order to quantify the foreground contamination in the EoR window, the frequency resolution of our primary beam, 0.125 MHz, is fine enough to cover a large range of $k_{\parallel}$ and thus used for the simulation instead of interpolation along frequency.

The CST simulated beams are resolved at 1° in the spherical coordinate along θ and ϕ. Based on the sampled grid points, we interpolated far-field electric fields defined in

Equation (3.1) at angular positions of sky sources using bivariate cubic spline interpolation, and formed the power beam by following Equation (3.2). It is unclear how the instrument responds to emission near the horizon. Bassett et al. (2021) have investigated the effect of the horizon shape on extracting 21cm signals from observation, and found the signal extraction can be significantly improved when they include accurate information of the topological horizon such as nearby vegetation, buildings, and mountains. Modeling an accurate horizon can be crucial, but it is beyond the scope of this study and future work can explore more details. For simplicity, we applied a sharp cut-off to the observable sky and considered only sources above the horizon. For more accurate cubic spline interpolation, we allowed a 5-degree buffer below the horizon for the beam interpolation.

In constructing sky source models, we considered two types of foreground sources: galactic and extragalactic point sources, and diffuse synchrotron emission from our Galaxy. For the point source model, we adopted the GaLactic and Extragalactic All-sky MWA (GLEAM) survey which consists of GLEAM I (Hurley-Walker et al., 2017) and GLEAM II (Hurley-Walker et al., 2019). The former covers about 25,000 square degrees of sky outside the galactic plane between -90° and 30° in declination and the latter covers compact sources in the galactic plane. The main lobe of our primary beam at 160–180 MHz has the FWHM of about 10° centered at the declination of -30.7°, which means the field of view of the HERA main beam falls into the coverage of the GLEAM I survey and the sky model is good enough to provide a representative set of point sources for our analysis. Point sources with spectral indices derived from single power-law fits provided by the GLEAM catalogs are used for the simulations, giving about 260,000 point sources. We also restored peeled bright sources noted in Table 2 of Hurley-Walker et al. (2017) and Fornax A which has two bright radio lobes (Bernardi et al., 2013).

For the diffuse sky model, the Global Sky Model (GSM) provided by Zheng et al. (2017) is used with a python package `PyGSM`. They removed 1% of the highest peak of residual after fitting maps through their iterative Principal Component Analysis (PCA) algorithm, which helps to remove some bright point sources from the diffuse map. In this study the brightness temperature of the diffuse sky map was generated in `HEALPix`, Hierarchical Equal Area isoLatitude Pixelization (Gorski et al., 2005), with an `Nside` of 256 yielding 786,432 pixels, then converted to flux density, assuming that the brightness temperature is constant in each pixel area. We then treated each pixel area as a point source and fed the sky model to our visibility calculations. For the frequencies we simulate and the size of the HERA array, we found that an `Nside` 256 `HEALPix` map resulted in enough resolutions for the proper power spectrum estimation, which is consistent with the results of Lanman et al. (2020).

We chose an LST of 2.25 hr for our single LST simulation, a relatively foreground-free zone but also good for point source calibration for HERA (Kern et al., 2020a). In addition, by choosing the LST, the gap of the GLEAM catalog present at RA$\sim$8 hr has a small

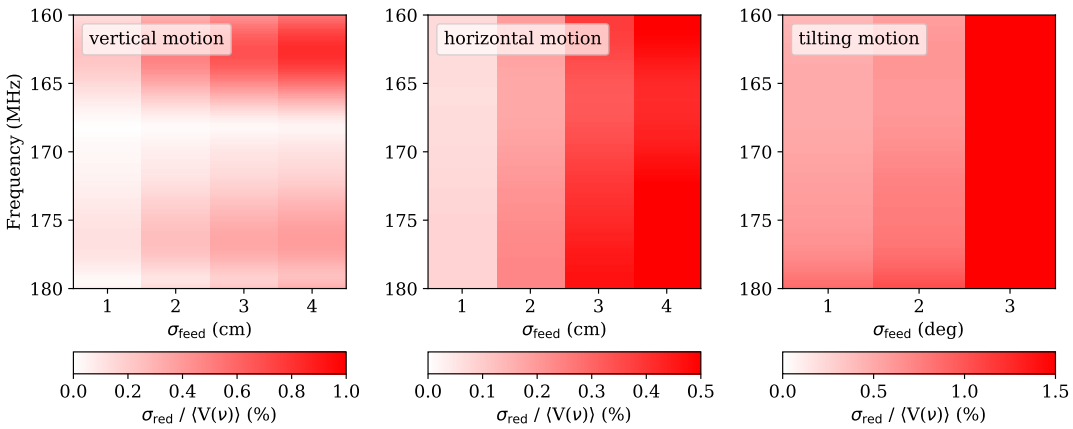

Figure 3.6: Relative σ_{red} of auto-correlations for vertical (left), horizontal (middle), and tilting feed motions (right) with feed offset and frequency.

effect on our visibility calculation as it is outside the main lobe of the primary beam. A radio interferometer has a different response to the point sources and the diffuse source. To better understand the role of each sky model on the visibility, we consider the two sky models separately if necessary. Otherwise, we account for those sky models together as the combined sky model.

3.3.2 Nonredundancy in Raw Visibility Measurements

When the feed positions are perturbed, redundant baselines are expected to produce different visibilities because of their different primary beam responses. This effect can be different depending on the configuration of the baseline (e.g., length and direction) and the type of the sky model (e.g., point sources or diffuse sources). One metric to evaluate the effect of the nonredundancy in the primary beams on visibilities is to measure a standard deviation of visibility measurements with the same baseline separation,

$$\sigma_{\mathrm{red},\alpha} = \sqrt{\frac{\sum_{\{i,j\}_\alpha} |V_{ij}(\nu) - \langle V_{ij}(\nu)\rangle_\alpha|^2}{N_{\mathrm{rbl}}}}, \tag{3.5}$$

where α indicates an index of a redundant-baseline group, N_{rbl} is the number of baselines in the redundant-baseline group and $\langle V_{\mathrm{ij}}(\nu)\rangle_\alpha$ is the mean visibility for the redundant baselines. Here the visibility is raw data for which antenna gains and calibration are not applied. We employed $\sigma_{\mathrm{red},\alpha}$ for auto- and cross-correlations to examine the nonredundancy in the raw visibility.

In general, the foreground spectrum is smooth, which means characteristic features in σ_{red}

with frequency are mainly due to the features imprinted in the primary beams. According to Figure 3.2 and 3.3, while at some frequencies the beam width gets either narrower or wider as the feed moves upward, at around 170 MHz the beam width is nearly constant with the feed motion. Similar patterns are observed in $\sigma_{\rm red}$ of the visibility measurements with the axial feed motion as shown in the first panel of Figure 3.6. The auto-correlation $\sigma_{\rm red}$ normalized by the averaged visibility amplitude over all antennas shows there is a null at around 170 MHz and the nonredundancy gets larger as it approaches the band edge. This value increases with the feed offset but the effect is small.

As we see in Figure 3.2 and 3.3, when the feed is displaced from the fiducial position in the xy plane, the pointing angle of the main beam also shifts which is proportional to the feed displacement but frequency-independent. The middle panel of Figure 3.6 shows $\sigma_{\rm red}$ which is largely consistent with the result of the beam characteristics of lateral feed motions, showing larger variations in the auto-correlations with larger perturbation in the feed motion but nearly constant with frequency.

Similarly, as discussed in Section 3.2.2, pointing errors of beams due to the tilting motion is roughly linear to the tip-tilt of the feed. In addition, unlike the lateral feed motion, tilting produces non-negligible changes in the beam width. Combining the two effects results in a slightly different trend (right panel) compared to the middle panel.

The standard deviation computed from cross-correlations is more interesting and important in the sense that 1) responses of short and long baselines to point sources and the diffuse source are different and 2) cross-correlations are involved with redundant-baseline calibration and power spectrum estimation.

We derived $\sigma_{{\rm red},\alpha}$ for each unique baseline using Equation (3.5) which is then averaged over the frequency band. Figure 3.7 visualizes $\sigma_{\rm red}$ of cross-correlation in the baseline space with combined (left), GLEAM (middle), and GSM (right) sky models. From top to bottom, vertical, horizontal, and tilting motions are given. We chose $\sigma_{\rm feed} = 3$ cm for axial and lateral motions and $\sigma_{\rm feed} = 3^\circ$ for the tilting motion. Though the visibility simulation is noiseless, it is useful to compare $\sigma_{\rm red}$ with the expected thermal noise defined as,

$$\sigma_{ij} = \sqrt{\frac{V_{ii}V_{jj}}{\Delta\nu\Delta t}}, \tag{3.6}$$

where V_{ii} is the visibility autocorrelation of antenna i, $\Delta\nu = 0.125$ MHz is the frequency channel width, and Δt $= 10$ s is the integration time. The variations of auto-correlations due to feed motion are of order 1% or less (Figure 3.6), which means the thermal noise levels derived from Equation (3.6) are similar for all feed motions. For simplicity, we took a single value from the fiducial model that is $\sigma_{\rm th} = \langle\sigma_{ij}\rangle = 2.5$ Jy averaged over the frequency band. $\sigma_{\rm red}$ for the GLEAM and the GSM is also divided by the same value for consistency.

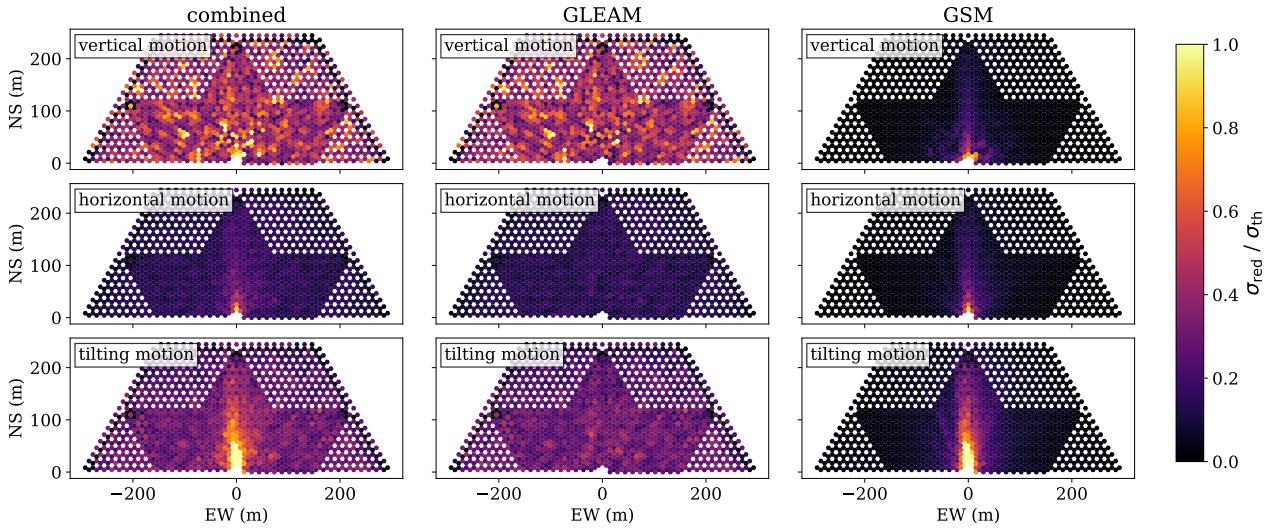

Figure 3.7: Standard deviation of visibilities within a nominally redundant group of each baseline vector for the vertical (top row), horizontal (middle row), and tilting feed motion (bottom row). The first column is the result with the combined sky model which is separated into the GLEAM (second column) and GSM (third column) sky model. The color scale indicates σ_{red} normalized by the thermal noise of the visibility (σ_{th}). The same value for σ_{th} which is 2.5 Jy from the fiducial model averaged over the frequency band is used for all panels. We consider visibilities with $\sigma_{\mathrm{feed}} = 3$ cm for the feed translation and $\sigma_{\mathrm{feed}} = 3°$ for the tilting. We see diffuse emission introduces strong nonredundancy for particular baseline types, while point sources generate nonredundancy across all baselines.

The ratio of the nonredundancy in visibilities to the thermal noise is a key quantity to determine whether the redundant-baseline calibration works properly when there are beam perturbations. In most panels, the color scale less than 1 indicates the thermal noise is larger than the nonredundancy error and thus the effect of the nonredundancy on redundant-baseline calibration seems not significant. However, antenna gains which are solved for by the calibration come from at least two or more visibility measurements, making the noise levels in antenna gains smaller compared to those in visibilities and the nonredundancy effect on the calibrated data visible. Indeed, we show reduced χ^2 from the redundant-baseline calibration including random thermal noises varies with the feed motion in Section 4.3.

In the first column of Figure 3.7, the nonredundancy defined in the baseline space is different for different feed motion. While the vertical feed motion shows the evidence of nonredundancy across all baselines, the feature is concentrated at short EW baselines for the tilting motion. The horizontal motion presents similar patterns to those from the tilting motion but the overall amplitude is smaller. These different trends of σ_{red} can be thought as the results of different responses of deformed primary beams due to the feed motions to GLEAM and GSM sky models.

For the GLEAM-only sky model (middle column), σ_{red} is uniform across the baselines

regardless of the feed motion. This is because a point source has a constant visibility amplitude for all baselines. One notable thing is $\sigma_{\rm red}$ is largest for the vertical motion and smallest for the horizontal motion, which is attribute to the different characteristics of the main lobe of the primary beam for different feed motions as discussed in Section 3.2.2. For example, when the feed moves along the vertical axis, the main beam widens/shrinks and collects more/less fluxes from the sky, which results in variation of the overall amplitude of visibilities and thus large $\sigma_{\rm red}$. In comparison, when the feed moves in the horizontal plane, the pointing angle of the main lobe shifts but with the nearly constant beam width. This shift of the beam makes the instrument observe a different part of sky but the total amount of received flux stays the same if the sky changes slowly in flux, which keeps $\sigma_{\rm red}$ small. The change of the main lobe width for the tilting motion is moderate and thus $\sigma_{\rm red}$ is between those of vertical and horizontal motions.

The diffuse model, however, behaves in a different way. Because the response of the interferometer to an extended source is strongest at the shortest baseline and declines with baseline length, $\sigma_{\rm red}$ is expected to be largest at short baselines as shown in the last column of Figure 3.7 for all three types of the feed motions. At LST = 2.25 hours when the galactic plane is located at about 80° away from zenith, the overwhelming power from the galactic plane with large solid angles around the horizon makes the emission from diffuse sources near the horizon even stronger than that along zenith despite considerable attenuation of the primary beam outside the main lobe (Thyagarajan et al., 2015a,b). This means, unlike the GLEAM case when $\sigma_{\rm red}$ is largely governed by the properties of main lobes, the features of side lobes play a role to characterize nonredundancy in visibilities arising from the diffuse sources. In this context, large $\sigma_{\rm red}$ in the tilting motion shown in the bottom right panel can be understood based on the large variation of $\epsilon_{\rm SL}$ defined in Equation (3.3) (Figure 3.4).

Another interesting feature of the GSM-only model is the pattern in $\sigma_{\rm red}$ elongated along the NS direction which may originate from directional response of the interferometer. The geometric delay, $\tau_{\rm g}$, between two antennas forming a baseline is constant along the direction normal to the baseline vector on the sky. Therefore, the fringe defined by the geometric delay, $f \sim \exp\left(-2\pi i\nu\tau_{\rm g}\right)$, has a constant value along the path orthogonal to the baseline vector. On the sky we are looking at, the galactic plane is stretched in the EW direction, which means the constant fringe in the EW direction for the NS baseline helps add up the flux from the galactic plane constructively and enhance the amplitudes of visibility measurements, showing the elongated pattern along the NS direction in $\sigma_{\rm red}$. Because the galactic plane moves around the sky as the Earth rotates, this effect is LST-dependent. For example, we found that at LST = 1 hour when the galactic center is located at the South-West corner of the sky aligned in the North-West direction, the pattern in $\sigma_{\rm red}$ is featured in the North-East direction as expected.

If we focus on the combined model again, now we can see the uniform trend of $\sigma_{\rm red}$ in

the vertical feed motion is due to the strong response of the motion to the point source sky while large $\sigma_{\rm red}$ characterized at short EW baselines in the tilting motion is due to the strong response of the motion to the diffuse sky. Overall, the nonredundancy for the vertical and tilting motions is larger than that for the horizontal motion. Orosz et al. (2019) showed that intrinsic chromaticity on long baselines is in charge of introducing chromatic gain errors when nonredundancy is present in redundant-baseline calibration. HERA with the compact array layout has more short baselines than long baselines, which means the large nonredundancy featured at short baselines resulting from the GSM component may have considerable impact on the calibration as well.

3.4 Redundant-baseline Calibration

In this section, we describe the antenna gains that are applied to the raw visibility measurements and need to be calibrated out in real observations. We also describe the results of the calibration in the presence of nonredundancy in the raw visibility.

3.4.1 Model of Input True Antenna Gains

The incident electric field on the feed and the output voltage are related by the system voltage response, $\mathbf{H}$, defined as

$$V_{\rm out}(\nu,\theta,\phi) = \mathbf{H}(\nu,\theta,\phi)\cdot\mathbf{E}_{\rm in}(\nu,\theta,\phi). \tag{3.7}$$

The system, consisting of the antenna, FEM, coaxial cables, and post-amplifier module, has $\mathbf{H}$ expressed as

$$\mathbf{H}(\nu,\theta,\phi) = \frac{Z_{\rm ant}(\nu)+Z_0}{\nu Z_0}\mathbf{E}_{\rm pat}(\nu,\theta,\phi)S_{21}(\nu), \tag{3.8}$$

where $Z_{\rm ant}$ is the complex impedance of the antenna, $Z_0 = 100\,\Omega$ is a termination impedance which is appropriate for a differential signal, $\mathbf{E}_{\rm pat}$ is the electric-field pattern, and S_{21} is a scattering parameter (Fagnoni et al., 2020). The S_{21} scattering parameter, describing the power transferred from port 1 to port 2 in the receiver, comes from lab measurements, and $Z_{\rm ant}$ and $\mathbf{E}_{\rm pat}$ are obtained from CST simulations. This $\mathbf{H}$ represents the antenna gain relating sky signals to instrument measurements, via the antenna and RF chain. Because we take into account a direction-independent gain in the following analysis, we pick the system voltage response at boresight where the beam response is strongest, $\mathbf{H}(\nu,0,0)$. The antenna gain is normalized to the peak for the full band, and the result is presented in Figure 3.8, with the highlighted frequency band of interest. Detailed derivation of the antenna gain is

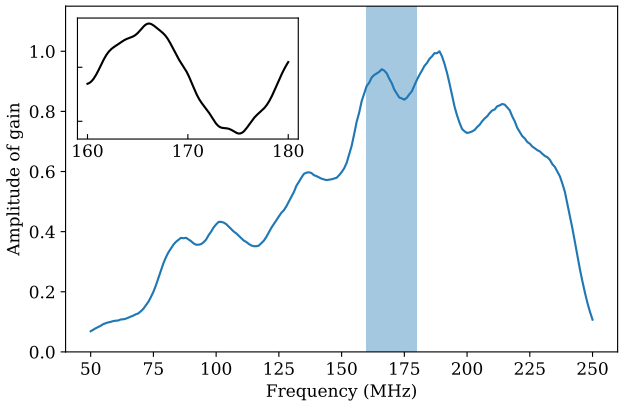

Figure 3.8: The amplitude of the gain derived from Equation (3.8). The response is normalized to the peak of the full band and has an arbitrary linear unit. The shaded region indicates the frequency band of 160–180 MHz which we use for visibility simulations. The inset panel shows the spectral features of the smoothed gain in the band.

addressed in Fagnoni et al. (2020) and Hewitt et al. 2021 (HERA memo #105[2]).

The S_{21} measurements were sampled at 1 MHz cadence and thus must be interpolated to the finer resolution of 0.125 MHz. To prevent unwanted fine-scale spectral artifacts from the interpolation, the gain amplitudes at 160–180 MHz were smoothed and interpolated by a Gaussian process regression (GPR) model with a fixed kernel size of 1 MHz after subtracting a 5th-order polynomial fit. The GPR may smooth out potential frequency structure smaller than 1 MHz but is found to be effective in suppressing interpolation artifacts. More details about the smoothing method are described in Lanman et al. (2020). The result of the smoothed gain is shown in the inset panel of Figure 3.8. Though the $Z_{\rm ant}(\nu)$, $\mathbf{E}_{\rm pat}$, and S_{21} vary with the feed displacement, we ignore this effect and model the same gain from the unperturbed case, regardless of the feed motion for simplicity.

To account for the realistic effect of the amplitude attenuator and the cable delay, we randomized amplitudes and phases of antenna gains over 320 antennas. We assume the attenuation η_i for each antenna follows a random Gaussian distribution with zero mean and a standard deviation of 0.2 and the cable delay τ_i also follows a random Gaussian distribution with zero mean and a standard deviation of 20 ns corresponding to 6-m long cable. The smoothed gain was then multiplied by $\exp(\eta_i + 2\pi i \tau_i \nu)$ and the final "true" gains were applied to the raw visibility to derive the measured visibility.

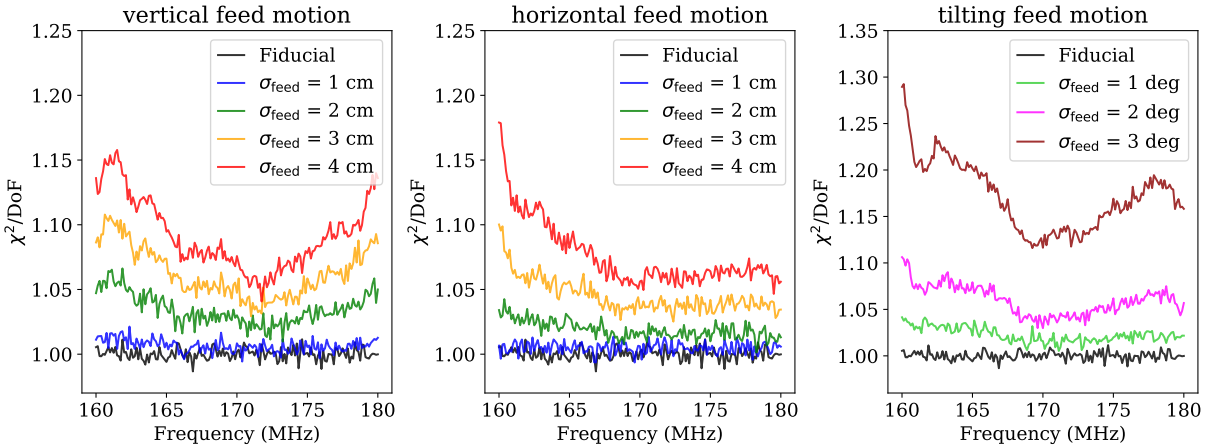

Figure 3.9: χ^2/dof of the redundant-baseline calibration for simulated visibilities with perturbed primary beams. From left to right, vertical, horizontal, and tilting feed motions are presented. The fiducial model yields χ^2/dof close to 1 as expected, which indicates a perfect calibration. When the feed is displaced from the fiducial position, χ^2/dof also deviates from 1. This effect is largest for the tilting motion, which is consistent with $\sigma_{\rm red}$ of raw visibilities shown in Figure 3.7. The noisy features are due to the thermal noise that is included only for the χ^2 statistics.

3.4.2 Effects of nonredundancy on Redundant-baseline Calibration

The general relation between the measured visibility $V_{ij}^{\rm obs}$ and the true visibility $V_{ij}^{\rm true}$ is assumed to be associated with antenna gains and thermal noise,

$$V_{ij}^{\rm obs} = g_i g_j^* V_{ij}^{\rm true} + n_{ij}, \tag{3.9}$$

where g_i and g_j are complex per-antenna gains and n_{ij} is the Gaussian thermal noise. The antenna gain represents direction-independent effects such as amplifiers and the delay offset due to the light-travel time delay along the path. The antenna-to-antenna variation of primary beams, which is our interest, is one common source of the direction-dependent effect (Smirnov, 2011). Equation (3.9), however, does not account for the direction-dependent correction, which may leave chromatic errors in the antenna gains. To minimize the potential spectral structure from the direction-independent calibration, some techniques such as smoothing the antenna gains can be employed. More details about mitigation with the techniques will be addressed in Kim et al. (in prep).

With all identical primary beam models, antenna pairs with the same separation are supposed to measure the same true visibility. The redundant-baseline calibration uses the redundancy of the visibility in the redundant baseline group rather than prior information of the sky to obtain the true visibility. More specifically, the redundant-baseline calibration

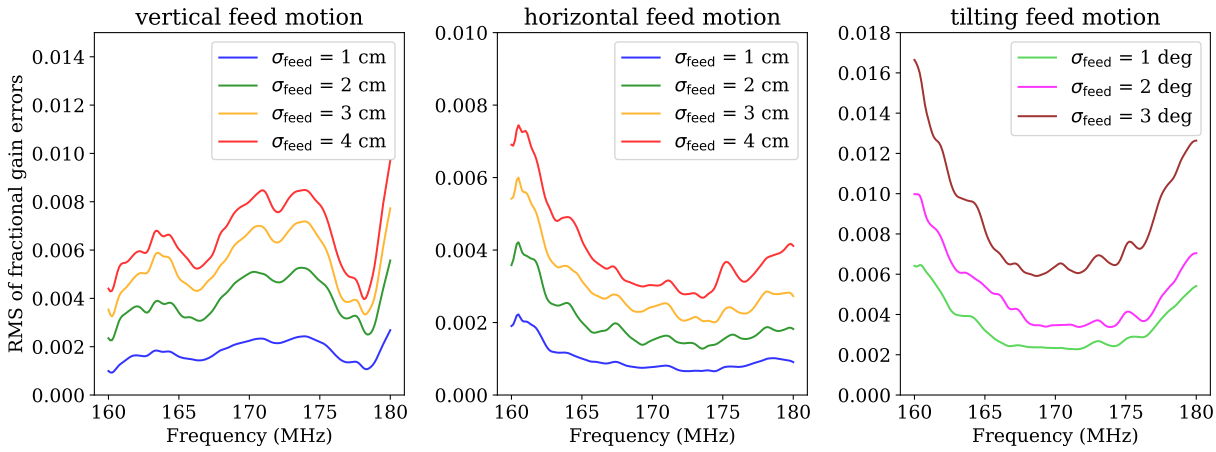

Figure 3.10: The rms of the fractional gain errors for the vertical (left), horizontal (middle), and tilting (right) feed motions. As expected, larger feed motion perturbations result in larger rms of the fractional gain errors. Although these gain errors are largely kept to about 1%, when convolved with bright foregrounds this could easily swamp the intrinsic 21-cm signal, if the gain errors have sufficient structure.

solves for the unique visibility along with the antenna gains as free parameters,

$$V_{ij}^{\rm obs} = g_i g_j^* V_{i-j}^{\rm sol}, \tag{3.10}$$

where $V_{i-j}^{\rm sol}$ is the unique visibility solution of the redundant baseline group with the same baseline separation. One approach to find the antenna gain and the visibility solution is to minimize χ^2,

$$\chi^2 = \sum_{i<j} \frac{|V_{ij}^{\rm obs} - g_i g_j^* V_{i-j}^{\rm sol}|^2}{\sigma_{ij}^2}, \tag{3.11}$$

where σ_{ij}^2 is the variance of n_{ij}. This is performed per polarization, per frequency, and per time. More comprehensive discussion about how to minimize Equation (4.2) is presented in Dillon et al. (2020). The calibration was performed with the publicly available software library, `hera_cal`[6].

In an ideal case when redundant baselines share the same true visibility, χ^2 is expected to be equal to the degree of freedom (dof). The dof for the redundant-baseline calibration is dof $= N_{\rm bl} - N_{\rm ubl} - N_{\rm ant} + 2$ where $N_{\rm bl}$ is the total number of baselines, $N_{\rm ubl}$ is the number of unique baselines, and $N_{\rm ant}$ is the number of antennas (Dillon et al., 2020). Figure 3.9 shows overall χ^2/dof as a function of frequency for vertical (left), horizontal (middle) and tilting (right) feed motions[7]. The measured visibility is simulated with the combined sky

[6] https://github.com/HERA-Team/hera_cal

[7] Even though our simulations are noise-free throughout the rest of our analysis, for appropriate χ^2 calculation, visibility measurements to derive Figure 3.9 include thermal noise generated with autocorrelations

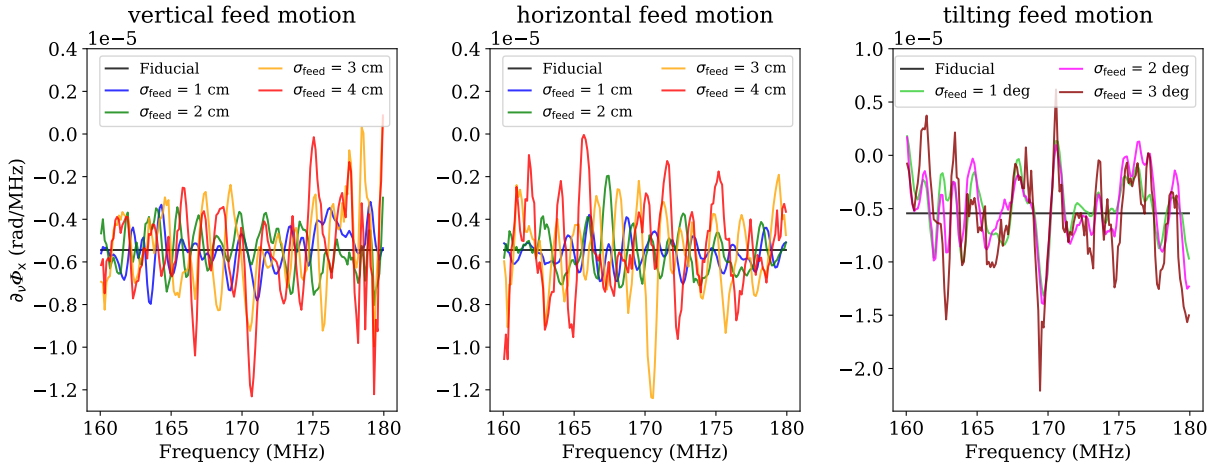

Figure 3.11: Slope of the x-direction phase gradient that is solved for by absolute calibration for vertical (left), horizontal (middle), and tilting (right) feed motions. $\partial_\nu \Phi_{\rm x}$ represents the derivative of the linear phase gradient in the x-direction with respect to frequency. We see the slope is nearly constant for the fiducial model, while there is high-frequency structure for the perturbed model which introduces additional chromatic gain errors.

model. As expected, the fiducial model yields χ^2/dof close to 1, which indicates a perfect calibration. When the feeds move away from the fiducial position and angle, χ^2/dof deviates from 1 and the deviation gets larger with the degree of perturbation. For a given $\sigma_{\rm feed}$, vertical feed motions result in a similar level of χ^2/dof to horizontal feed motions, but the former has a slightly larger mean value. Our choice of tilts, however, presents rather larger χ^2/dof compared to the translational motions. This is largely consistent with the results of $\sigma_{\rm red}$ for raw visibilities, showing the larger nonredundancy error arising from the tilting motion compared to that from the horizontal motion (Figure 3.7).

Another metric to quantify the effectiveness of the redundant-baseline calibration against perturbed beams with feed motions is the rms fractional gain errors. The fractional gain error of an antenna is defined as

$$f_{g,i} = \frac{|g_{{\rm pert},i}| - |g_{{\rm unpert},i}|}{|g_{{\rm unpert},i}|}. \tag{3.12}$$

Here $|g_{{\rm unpert},i}|$ and $|g_{{\rm pert},i}|$ indicate the gain amplitudes of the unperturbed and perturbed beams for antenna i, respectively. The rms of fractional gain errors is then calculated over all 320 antennas per frequency. In Figure 3.10, the vertical (left) and horizontal (middle) feed motions show $\lesssim$ 1% rms errors, while the tilting motion (right) displays larger errors, about 1–2 % rms errors, which is consistent with the results of χ^2/dof. In other words, larger chromatic gain errors indeed correlate to larger feed motion perturbations.

using Equation (3.6). We consider $\Delta\nu$ = 125 kHz and Δt = 10 seconds.

Though the highly redundant array configuration of HERA yields enough number of measurements to solve for the unknown parameters of redundant-baseline calibration, there are degeneracies between the antenna gains and the visibility solution that keep $g_i g_j^* V_{i-j}^{\rm sol}$ unchanged. The four degenerate parameters are the overall amplitude and phase, along with the EW and NS tip-tilt (Byrne et al., 2019, Dillon et al., 2018, Kern et al., 2020a, Li et al., 2018, Liu et al., 2010, Zheng et al., 2014). The last two are due to directional phase factors in x- and y-direction in the antenna gain which can be cancelled by rephasing the unique visibility solution (e.g., $g_i \rightarrow g_i e^{i\Phi_x x_i}$ and $V_{i-j}^{\rm sol} \rightarrow V_{i-j}^{\rm sol} e^{-i\Phi_x(x_i - x_j)}$, where Φ_x is a linear phase gradient in the x-direction). The overall phase is the degeneracy between antenna gains ($g_i \rightarrow e^{i\psi} g_i$, $g_j^* \rightarrow e^{-i\psi} g_j^*$) and is set arbitrarily. The three degeneracies other than the overall phase are solved for by an additional process referred as absolute calibration using a sky model.

We implemented absolute calibration with the model visibility simulated using the unperturbed primary beam model. As described in Section 3.2.2, the pointing angle of the primary beam can shift due to feed motion, resulting in shifted field of view of the observed sky relative to the desired fiducial pointing. Sky-based calibration with an inaccurate sky model is known to introduce frequency-dependent calibration errors (Barry et al., 2016, Byrne et al., 2019, Ewall-Wice et al., 2017, Gehlot et al., 2021). In other words, the difference in the sky observed by unperturbed and perturbed primary beams is an additional potential source of chromatic gain errors (Barry and Chokshi, 2022, Orosz et al., 2019). Figure 3.11 is an example showing the linear phase gradient along the x-direction that is solved for by the absolute calibration step. We calculate the derivative of the phase gradient with respect to frequency, and present the results as a function of frequency for the vertical (left), horizontal (middle), and tilting (right) motions. The derivative for the fiducial model is nearly constant. However, the phase gradient for the perturbed case exhibits high-frequency structure, which contributes to the spectral leakage especially at high delay modes.

Figure 3.12 demonstrates a frequency Fourier transform of the calibrated antenna gains, after the redundant calibration (dashed lines) and after the full calibration, including the post absolute calibration (solid lines). The Fourier transform of the gain shows the frequency structure in the gain that will be mapped to power spectrum estimates. The delay in the x-axis is a Fourier dual to the frequency. To derive the y-axis, we divide the perturbed gains by the fiducial one, and multiply them by a 7th-order Blackman-Harris tapering window function, which helps to suppress the side lobes and achieves a large dynamic range (Lanman et al., 2020). We then perform the frequency Fourier transform, normalize the results to the peaks, and average the amplitudes over all antennas.

We found that the Fourier transform of the fiducial gain, with the perfect calibration, reaches a floor of about 10^{-9} at high delays, which is consistent with the true gain. This means the floor above about 10^{-7} shown in Figure 3.12 may arise from the calibration error

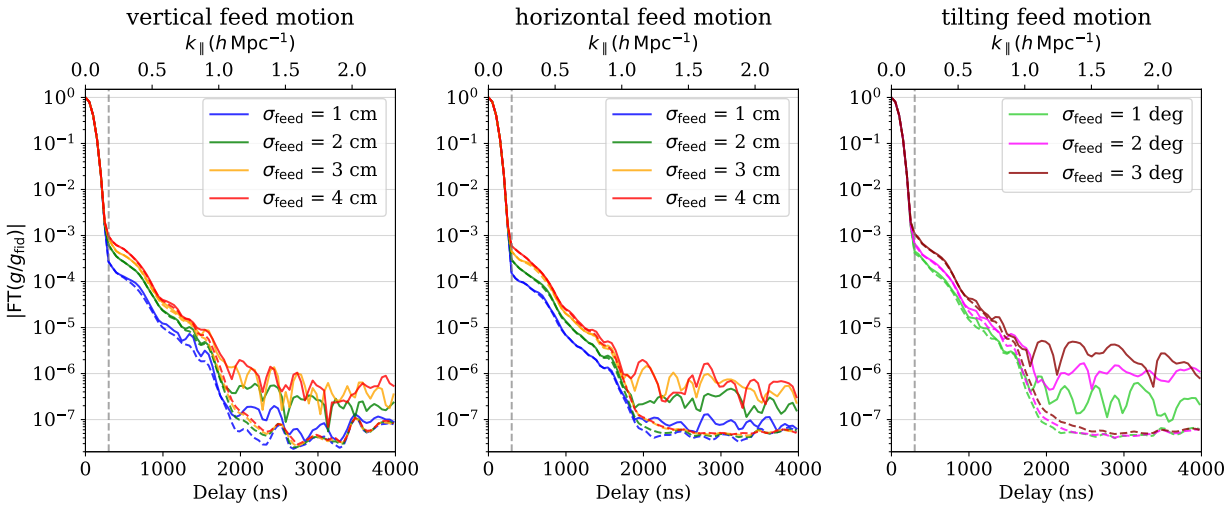

Figure 3.12: Averaged Fourier transformed antenna gain over all antennas for the axial (left), the lateral (middle), and the tilting (right) feed motions. The perturbed gain is normalized to the fiducial one. The dashed and solid lines indicate the results after the redundant-baseline and after the full calibration including the absolute calibration, respectively. Different colored lines denote different levels of perturbation in feed motions. Even in the case when the feed is least perturbed in our choice, a broad wing $\gtrsim$ 300 ns or 0.17 $h\,\mathrm{Mpc}^{-1}$ (vertical line) is present. This is due to chromatic gain errors caused by the nonredundancy in visibilities emerging from redundant-baseline calibration. Existence of the wing is critical since it affects detection of cosmological signals at low $k_\parallel$ where high signal-to-ratios are expected. The chromatic gain errors above 2000 ns mainly come from the absolute calibration.

due to the perturbed beams. In addition, unwanted broad wings start to appear at delay larger than 300 ns, which are primarily caused by chromatic errors in the gain solutions from the redundant-baseline calibration as shown in dashed lines. Larger perturbation in feed motion results in larger nonredundancy errors, and thus larger amplitude of the wing. In principle, the antenna gain from the perfect calibration that is convolved with smooth spectrum of foregrounds makes the foreground power isolated in the wedge of a power spectrum. When the gain includes chromatic errors, however, convolution of the gain and the foregrounds can lead to foreground leakage outside the wedge owing to the broad wing in the Fourier transform. Additional chromatic gain errors at delay larger than $\sim$2000 ns are observed mainly due to the absolute calibration step.

3.5 Power Spectrum Estimation with Feed Motion

An infinite frequency Fourier transform of a visibility for spectrally flat foreground emission from a single source forms a Dirac delta function in the delay domain. Realistic consideration with smoothly varying foreground and instrumental responses with frequency as well as the

finite band width turns the Dirac delta function into a broadened delay-spectrum, but the width remains small as long as the variation is smooth enough. The delay-spectrum is centered at the geometric delay of the source, $\tau_{\rm g} = |\mathbf{b}| \sin\theta/c$ where $|\mathbf{b}|$ is the norm of the baseline vector, θ is the angle from zenith to the source, and c is the speed of light. The frequency Fourier transform of a visibility for all foreground sources is then the superposition of such delay-spectra across the field-of-view of the sky, and the maximum delay mode is defined by the largest geometric delay between two antennas forming the visibility. With a zenith-pointed array, the maximum geometric delay is set by the horizon limit and thus the delay modes are bound to $\tau \lesssim \tau_{\rm hor} = |\mathbf{b}|/c$. The complex-spectrum cosmological signal due to line-of-sight fluctuations, however, is widely distributed across all delay modes even beyond the horizon limit, which leaves a room for detecting the cosmological signals at $\tau > \tau_{\rm hor}$.

3.5.1 The Power Spectrum for Cosmological Signals

The observable cosmological quantity measured from redshifted 21-cm observation is the brighteness temperature of the neutral hydrogen that is defined by the spin temperature $T_{\rm S}$ relative to the background radiation temperature T_γ (e.g., Furlanetto et al., 2006),

$$\delta T_{\rm b}(z) = \frac{T_{\rm S} - T_\gamma(z)}{1+z}(1 - e^{-\tau_{\nu_0}}) \tag{3.13}$$

$$\approx 27 x_{\rm HI}(1+\delta_{\rm m})\left(\frac{H(z)}{{\rm d}v_\parallel/{\rm d}r_\parallel + H(z)}\right)\left(1 - \frac{T_\gamma}{T_{\rm S}}\right) \times \left(\frac{1+z}{10}\frac{0.15}{\Omega_{\rm m}h^2}\right)^{\frac{1}{2}}\left(\frac{\Omega_{\rm b}h^2}{0.023}\right) \quad [{\rm mK}], \tag{3.14}$$

where τ_{ν_0} is the optical depth at the rest 21-cm frequency $\nu_0 = 1420$ MHz, $x_{\rm HI}$ is the fraction of the neutral hydrogen, $\delta_{\rm m}$ is the matter density fluctuation, $H(z)$ is the Hubble parameter, and ${\rm d}v_\parallel/{\rm d}r_\parallel$ is the gradient of the line-of-sight velocity.

We adopt a cosmological signal from Mesinger et al. (2016) generated with `21cmFAST` (Mesinger et al., 2011) using the Faint Galaxies model that is consistent with Ly α forest observations. The size of the simulation box is 1600^3 cMpc^3 with 1024 pixels on each side. We chose the coeval cube at $z \sim 7.4$ corresponding to 170 MHz. The cosmological model at this redshift predicts the ionization fraction of hydrogen gas $\lesssim 0.5$.

Figure 3.13 shows the power spectrum of EoR signals defined in $(k_\perp,\ k_\parallel)$ space where $k_\perp$ and $k_\parallel$ are Fourier modes perpendicular to and parallel to the line-of-sight, respectively.

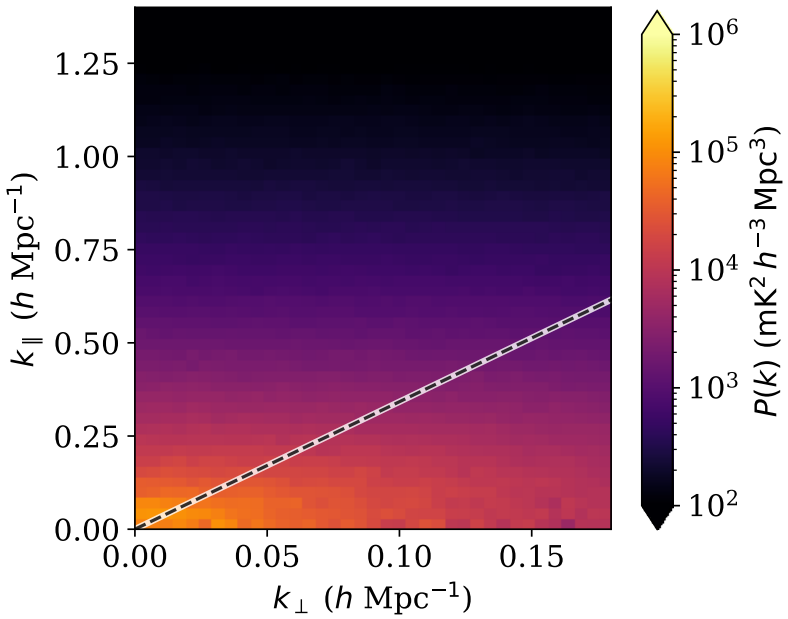

Figure 3.13: EoR power spectrum estimated at $z \sim 7.4$. Spectral structure in the cosmological signal allow the power spectrum to spread beyond the horizon limit (black dashed line) defined in Equation (3.20), forming the isotropic power spectrum.

The cosmological power spectrum is estimated with

$$\hat{P}_{21}(\mathbf{k}, z) = \frac{\langle |\delta\tilde{T}_\mathrm{b}(\mathbf{k}, z)|^2 \rangle}{V} \quad [\mathrm{mK}^2\, h^{-3}\, \mathrm{Mpc}^3], \tag{3.15}$$

where $\delta\tilde{T}_\mathrm{b}(\mathbf{k}, z)$ is a 3D spatial Fourier transform of $\delta T_\mathrm{b}(\mathbf{x}, z)$ and V is the simulation volume. The effective frequency bandwidth with the 7th-order Blackman-Harris tapering function is about 5.4 MHz or $\Delta z = 0.27$ in which the universe can be regarded as coeval and the light-cone effect on the power spectrum can be ignored (Datta et al., 2012). As expected, the EoR signal forms an isotropic power in the 2D power spectrum that is extended to high $k_\parallel$, providing a region for detecting the EoR above the horizon limit. This power spectrum will be used as a reference to quantify the foreground leakage beyond the horizon limit.

3.5.2 The Foreground Leakage due to Non-uniform Primary Beam Models

In principle, the foregrounds and the cosmological signal are separable in the 2D power spectrum thanks to the different behaviors of the sources in the Fourier domain. The power

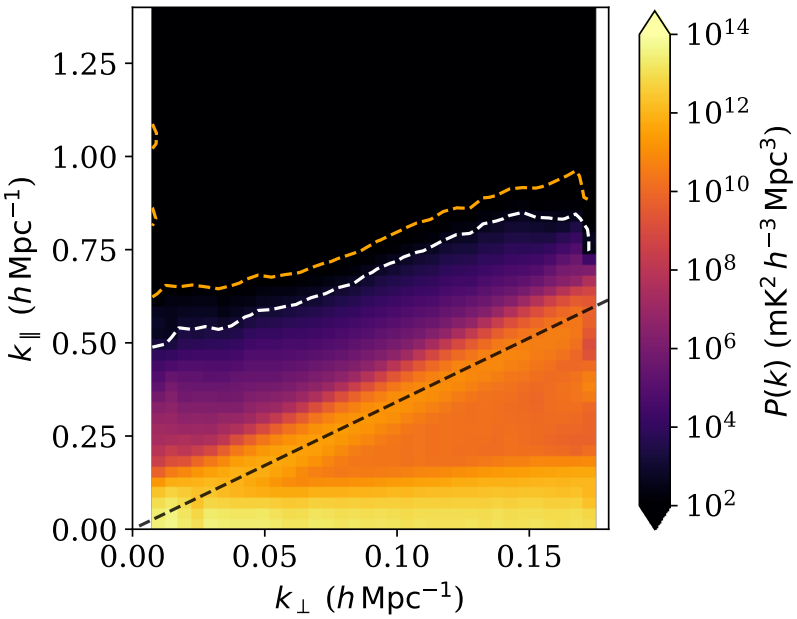

Figure 3.14: Foreground power spectrum for the fiducial beam model with the combined sky model. The power spectrum is estimated with the calibrated visibility at LST = 2.25 hours for the 160–180 MHz band. The black dashed line indicates the geometric horizon limit. The strong emission at low $k_\parallel$ is due to the point sources located at zenith while the emission aligned with the horizon limit is due to the diffuse source near the horizon. The white and orange contours are where the foreground spillovers are the same as and 10% of the EoR power, respectively. The foreground spillover beyond the horizon limit is not due to calibration errors but due to the intrinsic chromaticity of the beam in the far side lobes convolved with the diffuse power originating from the pitchfork effect. Without the fringe-rate filter, the cosmological information below $k_\parallel \sim 0.5\,h\,\mathrm{Mpc}^{-1}$ is obscured by the foreground power, even with the perfect calibration.

spectrum can be obtained based on the Fourier transform of a visibility along frequency,

$$\tilde{V}(\mathbf{u},\tau) = \int w(\nu) V(\mathbf{u},\nu) e^{2\pi i \nu \tau} d\nu. \tag{3.16}$$

Here, $\mathbf{u} = \mathbf{b}/\lambda$ and $w(\nu)$ is a tapering function applied along the frequency axis to down-weight the edge effect of the bandpass. As described in Section 4.3, we chose the 7th-order Blackman-Harris window function.

We coherently average complex visibilities with the same redundant baselines, which helps reduce the spectral structure arising from the chromatic gain errors. The coherently averaged visibilites, V_{coh}, are converted to delay spectra and the square of them yields the power spectrum estimate,

$$\hat{P}(k_\perp, k_\parallel) = \frac{X^2 Y}{B_{\mathrm{pp}} \Omega_{\mathrm{pp}}} |\tilde{V}_{\mathrm{coh}}(\mathbf{u},\tau)|^2, \tag{3.17}$$

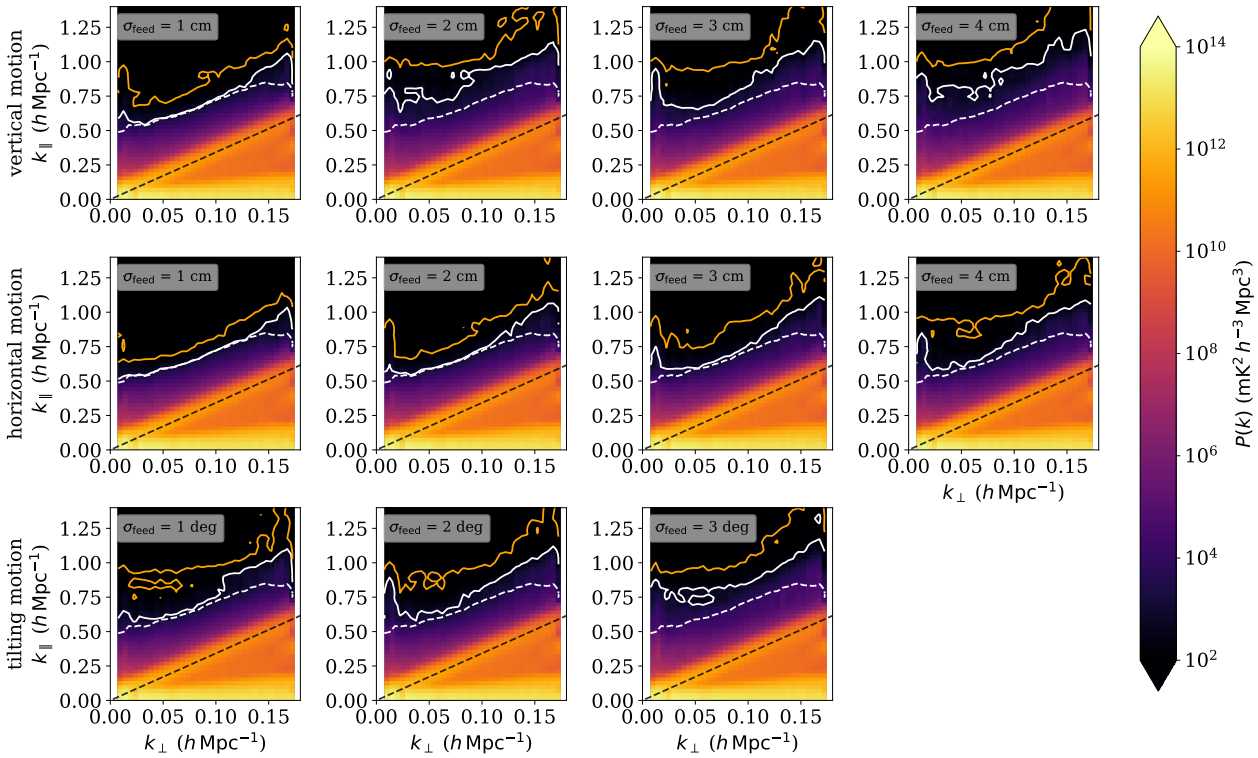

Figure 3.15: Foreground power spectra for vertical (top row), horizontal (middle row), and tilting (bottom row) feed motions with the combined sky model. From left, power spectra of $\sigma_{\rm feed} = 1$, 2, 3, and 4 cm for the translation motions and $\sigma_{\rm feed} = 1^\circ$, 2°, and 3° for the tilting motion are presented. The white and orange solid contours indicate where the foreground is equal to and 10% of the EoR power, respectively. The white dashed contour representing the fiducial model is shown in all panels for comparison. As expected, there is contamination when the feed positions are perturbed. We do not apply any mitigation technique to reduce the chromatic gain errors such as down-weighting the effect of long baselines in calibration or smoothing the gain solutions along the frequency axis, which will be discussed in the subsequent paper.

where $B_{\rm pp}$ is the effective band width defined as $B_{\rm pp} = \int |w(\nu)|^2 d\nu$ and $\Omega_{\rm pp}$ is the spatial integral of the squared primary beam (Parsons et al., 2014). X and Y are scaling factors relating $|\mathbf{u}|$ and τ to $k_\perp$ and $k_\parallel$ in cosmological units,

$$k_\perp = \frac{2\pi|\mathbf{u}|}{X} \tag{3.18}$$

$$k_\parallel = \frac{2\pi\tau}{Y}, \tag{3.19}$$

where $X = D(z)$, $Y = c(1+z)^2/(\nu_0 H(z))$, and $D(z)$ is the comoving distance. The horizon limit, $\tau_{\rm hor} = |\mathbf{b}|/c$, is then expressed in cosmological units as

$$k_{\parallel,\rm hor} = \frac{H(z)D(z)}{c(1+z)} k_\perp. \tag{3.20}$$

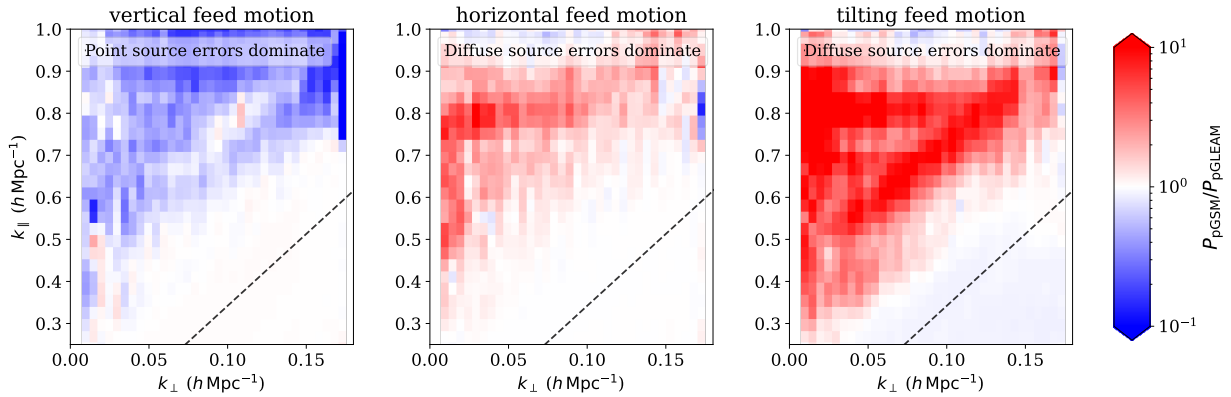

Figure 3.16: Relative foreground power spectra, $P_{\rm pGSM}/P_{\rm pGLEAM}$, for vertical (left), horizontal (middle), and tilting (right) feed motions. $P_{\rm pGSM}$ is from the visibility that is simulated using the GSM with the perturbed beam and the GLEAM with the fiducial beam. $P_{\rm pGLEAM}$ indicates the opposite case when the GLEAM is with the perturbed beam and the GSM is with the fiducial beam. The black dashed line indicates the horizon limit. Base on the analyses with $\sigma_{\rm red}$ in Figure 3.7, $P_{\rm pGLEAM}$ is expected to have larger foreground leakage than $P_{\rm pGSM}$ for the vertical motion while $P_{\rm pGSM}$ is in charge of leaking the foreground into the EoR window for the horizontal and tilting motions.

We utilized `hera_pspec`[8] that is a publicly available software to construct the power spectrum estimation.

Figure 4.13 presents the estimated power spectrum of the combined sky model for the fiducial feed position. For the fiducial model, the foreground power is expected to fall off rapidly with $k_\parallel$ beyond the horizon limit (black dashed line). However, significant amount of foreground spillovers are observed beyond the horizon limit despite the perfect calibration. This is an intrinsic spillover associated with the powerful emission near the horizon known as the "pitchfork effect" along the black dashed line (Thyagarajan et al., 2015a,b). The pitchfork effect can be understood as a result of the response of short projected baselines to strong diffuse sky emissions from the horizon lined up with the horizon limit in the ($k_\perp$, $k_\parallel$) space.

At LST $\sim$ 2 hours which is relatively foreground free in the direction of the zenith-pointed sky, the diffuse galactic plane lies near the horizon and forms the pitchfork. Because the galactic plane emanates strong synchrotron emissions, convolving the primary beam response with the sky emission in the Fourier domain results in an excess of power above the horizon limit as shown in Figure 4.13. There is a trend that low $k_\perp$ corresponding to short baselines has more excess of power than high $k_\perp$ and this is because short baselines less resolve out the diffuse source and thus capture more power on the horizon leading to more power at high $k_\parallel$ than long baselines. The overplotted white and orange contours correspond to the power

[8] https://github.com/HERA-Team/hera_pspec

level where the foreground power is equal to and 10% of the cosmological signal, respectively. These contours reveal the foreground spillovers extended above the horizon limit clearly.

This result is somewhat different from the result of Orosz et al. (2019) who showed smaller spillovers beyond the horizon limit. One major difference between their and our simulations is that they included point sources only while we consider diffuse sources as well as point sources. We found if we only consider point sources throughout the analysis, we came into a similar result to Orosz et al. (2019) with considerably suppressed spillover beyond the horizon limit. We include the GSM because it represents a more realistic sky model. The pitchfork effect can be mitigated by filtering out small fringe-rate regions on the sky (Parsons et al., 2016), and the effects of applying this technique will be addressed in Kim et al. (in prep).

In Figure 3.15, from left to right, power spectra perturbed by $\sigma_{\rm feed}$ = 1, 2, 3, and 4 cm for the vertical (top row) and the horizontal (middle row) feed motion and $\sigma_{\rm feed}$ = 1°, 2°, and 3° for the tilting motion (bottom row) are shown. Across all feed motions, as the feed moves away from the fiducial point, the solid contours representing the perturbed models depart from the dashed contour which is the fiducial one shown in Figure 4.13. This lift of the foreground leakage is associated with the antenna gains corrupted by the chromatic errors as discussed in Figure 3.12. At short baselines (i.e., low $k_\perp$) which have a number of redundant baselines, coherent averaging in the visibility space can reduce high-frequency structure introduced by the chromatic gain errors. As a result, for $\sigma_{\rm feed}$ = 1 cm or $\sigma_{\rm feed}$ = 1° when the feed displacement is relatively small, the averaged visibility yields a similar foreground power to that of the fiducial model at low $k_\perp$. Longer baselines, however, have relatively smaller number of redundant baselines and the averaging is not sufficient to smooth out the high-frequency features imprinted by the chromatic gain errors on top of the intrinsic spectral structure. This makes the long baselines suffer from unsmoothed frequency structure, displaying the leakage along $k_\parallel$ in the power spectrum analogous to the broad wing of the perturbed gains. For a larger size of perturbation, the effect of the chromatic gain errors become more significant and the chromatic errors along with the strong foreground power observed at the short baselines lead to leakage at very low $k_\perp$ modes. This is consistent with the results of Orosz et al. (2019) in the sense that there is a trend of the foreground leakage at very low and high $k_\perp$.

The amount of foreground leakage is different for different type of feed motion. For the vertical and tilting feed motions, the power spectra exhibit stronger evidence of leakage compared to the horizontal motions. We expect this behavior from the large nonredundancy error in visibilities of the vertical and tilting motions as discussed in Section 3.3.2. Especially when $\sigma_{\rm feed}$ = 4 cm and $\sigma_{\rm feed}$ = 3°, we see the low $k_\perp$ modes are heavily contaminated by the foregrounds and are not usable for EoR detection. For the horizontal feed motion which shows smaller $\sigma_{\rm red}$, the contours are least pushed away from the fiducial one but noticeable

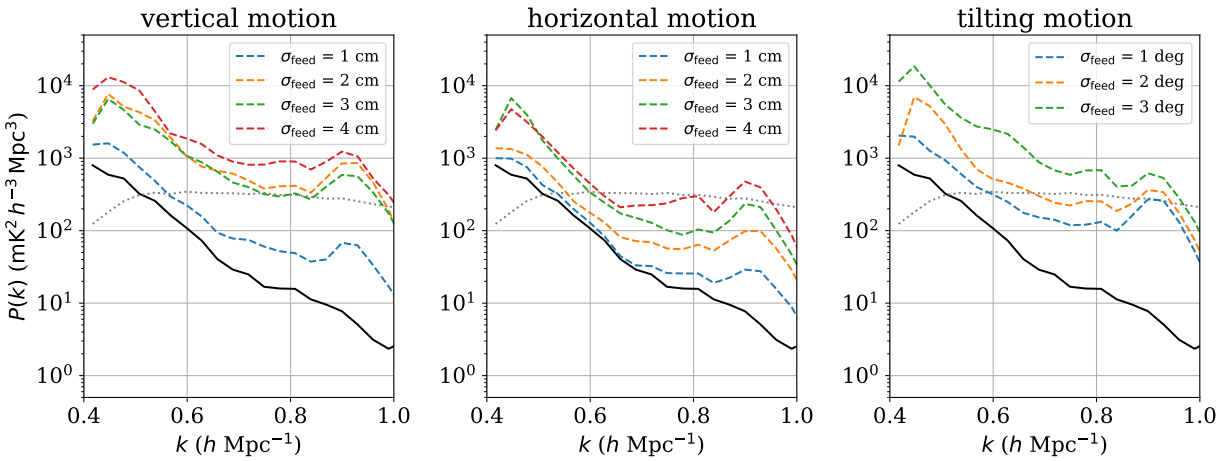

Figure 3.17: Spherically averaged 1D foreground power spectrum estimates for vertical (left), horizontal (middle), and tilting (right) feed motion. The power spectra are constructed from the 2D power spectra (Figure 3.15) with k modes satisfying $k_\parallel \geq k_{\parallel,\rm hor} + 0.4\,h\,{\rm Mpc}^{-1}$ to minimize the intrinsic foreground spillovers due to the spectral structure in the primary beam along with the pitchfork effect. The black solid curve is the power spectrum estimate of the fiducial model and colored dashed curves are for the perturbed beams labelled in the legend. The dotted grey curve denotes the EoR power spectrum with the same cut applied.

contamination is still allowed at high $k_\perp$. The foreground contamination is expected to be reduced by applying the baseline cut-off in calibration and smooth calibration along with the fringe-rate filter, and this will be discussed in the subsequent study.

As shown previously in Figure 3.7, the nonredundancy in visibilities depends not only on the feed motion but also on the sky model. In order to relate the result to the foreground leakage in the power spectrum, we simulated visibilities by perturbing the beam model for one sky model but keeping the fiducial beam model for the other. We then fed the visibility into the calibration pipeline and estimated the power spectrum. The ratio between power spectra, $P_{\rm pGSM}$ (i.e., fiducial GLEAM + perturbed GSM) and $P_{\rm pGLEAM}$ (i.e., perturbed GLEAM + fiducial GSM) is shown in Figure 3.16.

We consider $\sigma_{\rm feed} = 3$ cm for the vertical (left panel) and horizontal (middle) motions and $\sigma_{\rm feed} = 3^\circ$ for the tilting motion (right). For the vertical motion, the small $P_{\rm pGSM}/P_{\rm pGLEAM}$ is predicted because of predominant $\sigma_{\rm red}$ of the GLEAM over $\sigma_{\rm red}$ of the GSM. This may demonstrate the variation in the width of the main lobe with the vertical feed motion is the key feature inducing the foreground leaking into the EoR window. On the contrary, the larger $P_{\rm pGSM}/P_{\rm pGLEAM}$ for the tilting motion can be explained by the fact that $\sigma_{\rm red}$ of GSM is predominant over $\sigma_{\rm red}$ of GLEAM at short baselines, which may arise from the variation of the side lobes covering the bright galactic plane. The horizontal motion, which has relatively small $\sigma_{\rm red}$ for both GLEAM and GSM compared to other motions, reveals the impact of GSM is greater than that of GLEAM in forming the foreground leakage. This can be understood

in the context that the beam-weighted flux density of the GSM can be significantly larger than that of the GLEAM. For example, the flux density of the point sources inside the main lobe (e.g., inside 20×20 degree sq around zenith) is about 1.4×10^3 Jy, while that of GSM inside the side lobe ($60^\circ <$ zenith angle $< 90^\circ$) is about 8×10^5 Jy. For the feed displacement by 5 cm along the x-axis, the change to the power beam area for the main lobe is about 1.4×10^{-4}, whereas that for the side lobe is about 2.6×10^{-5}. This means the change to the beam-weighted flux density due to the feed motion is ∼0.2 Jy for the point sources and is ∼20 Jy for the GSM. This implies the contribution of the GSM with the feed motion would be 100 times larger than that of the point source and that's why we see a larger impact of the GSM for the horizontal motion.

A 1D power spectrum is formed based on the 2D power spectrum shown in Figure 3.15. We first selected the region least affected by the intrinsic foreground spillovers, even shown in the fiducial model, arising from the chromatic primary beam convolved with the pitchfork effect by setting a constant buffer above the horizon limit, $k_\parallel \geq k_{\parallel,\mathrm{hor}} + k_{\parallel,\mathrm{buffer}}$. For a conservative choice to minimize the effect of the spillovers, $k_{\parallel,\mathrm{buffer}} = 0.4\,h\,\mathrm{Mpc}^{-1}$ is taken into account. Though this buffer size is aggressive in terms of removing significant amount of k modes with relatively stronger EoR signals, it enables us to explore the behavior of the foreground leakage due to the feed motions.

We then turned the filtered power spectrum into the spherically averaged 1D power spectrum by averaging cylindrical bins into corresponding spherical bins. Figure 3.17 shows the results for the vertical (left), horizontal (middle), and tilting (right) feed motions. The fiducial model (black solid line) crosses the EoR power spectrum (grey dotted line) at $k \sim 0.5\,h\,\mathrm{Mpc}^{-1}$ where $P(k) \sim 300\,\mathrm{mK}^2\,h^{-3}\,\mathrm{Mpc}^3$. Since the EoR power spectrum declines with k, we focus on $0.5 < k < 1.0\,h\,\mathrm{Mpc}^{-1}$ where we can achieve relatively high sensitivity compared to $k > 1.0\,h\,\mathrm{Mpc}^{-1}$ if thermal noise is included. For the given EoR model, our noiseless simulation shows the foreground power spectrum is smaller than the EoR power spectrum at $k \gtrsim 0.6\,h\,\mathrm{Mpc}^{-1}$ for the lateral motions except for $\sigma_{\mathrm{feed}} = 4$ cm. For the vertical feed motion, the foreground power spectrum smaller than the EoR power when $\sigma_{\mathrm{feed}} = 1$ cm but the EoR power is buried under the foreground power spectrum for other cases. Our choice of tilts also makes the foreground power spectrum similar to or greater than the EoR power spectrum for most cases of the perturbation. We expect the foreground power can be dropped below the EoR level and set a requirement of feed positioning if appropriate mitigation for the chromatic gain errors is applied, which will be explored in Kim et al. (in prep).

3.6 Conclusions

In this study, we have characterized the effects of feed positional perturbations on the electromagnetic properties of the HERA antenna's beam pattern by using the CST's full-wave electromagnetic time-domain solver. Previous studies have so far been limited to studying analytic models of per-antenna beam perturbations for HERA. We significantly extend upon these works by using realistic CST simulations in a full forward-model context, and apply a realistic calibration pipeline to explore its impact on modern data analysis pipelines.

We separate the feed motions into three classes, including vertical, horizontal, and tilting motions. The vertical feed motion was found to mainly drive changes to the width of the main lobe, whereas the horizontal and tilting motions mainly perturb the pointing angle of the main lobe. Relatively larger perturbation in side lobes is observed for the tilting motion than the translation motions.

With the perturbed beams for 320 antennas, we simulated visibility measurements using the GLEAM sky survey and the GSM sky model for point sources and diffuse sky, respectively. Figure 3.7 shows different feed motions are responsible for different patterns of nonredundancy in cross-correlation visibilities. The uniform nonredundancy across the antenna separation shown in the vertical feed motion is primarily caused by the response of the point sources to variations in the main lobe, while the high nonredundancy concentrated at short baselines shown in the horizontal and tilting motions arises from the response of the diffuse source to changes of the side lobes.

The introduction of non-redundancies in the visibility due to per-antenna feed motions breaks the assumptions of redundant calibration and thus introduces chromatic errors into the gains. It also imparts gain errors into the absolute calibration step performed after redundant calibration. Fractional gain errors and χ^2 statistics reveal that the chromatic gain errors increase with the size of perturbation in the feed motion. The Fourier transform of the gain solutions shows excess power at delay $\gtrsim 300$ ns compared to the fiducial case, which results in foreground leakage outside the wedge in the power spectrum.

In the 2D power spectrum, the chromatic gain errors cause the foreground to leak from the wedge, which can significantly reduce the accessible size of the EoR window. If the perturbation is small ($\sigma_{\rm feed} \lesssim 1$ cm or $\sigma_{\rm feed} \lesssim 1^\circ$), we found that the foreground power can be suppressed to a level similar to the fiducial case at low $k_\perp$ (i.e., short baseline) if we coherently average visibilities over redundant baselines. Since high $k_\perp$ modes have less redundant baselines, more foreground leakage is therefore observed relative to low $k_\perp$ for all cases of the feed motions. When the feed positions are perturbed more than 1 cm in vertical displacement or 1° in tilt, there are considerable foreground contamination at both low and high $k_\perp$ modes due to the chromatic gain errors.

Figure 3.16 demonstrates the foreground leakage is mainly caused by GLEAM point sources in the main lobe of the perturbed beam from the vertical feed motions. We also see that the diffuse GSM component in the side lobes are the main contributors of the leakage in the case of horizontal and tilting feed motions, which is corroborated by Figure 3.7.

Based on the spherically averaged 1D power spectrum analysis, $\sigma_{\text{feed}} = 1$ cm for the horizontal motion may allow us to retain the EoR window with the least foreground bias. Unlike the lateral feed motion, the vertical and tilting motions introduce more foreground power leakage and the EoR signal is barely above the foreground power spectrum when the feeds are perturbed more than $\sigma_{\text{feed}} = 1$ cm or $\sigma_{\text{feed}} = 1°$. We expect the stringent feed positioning requirement can be loosened once fine spectral structure in gain solutions can be suppressed using some of the mitigation techniques proposed below.

So far, we have illustrated that antenna feed position offsets can introduce unwanted foreground leakage beyond the horizon limit in the power spectrum. The level of the foreground leakage is expected to be minimized with several different mitigation methods. One major intrinsic foreground spillover, even in the ideal fiducial model, is the pitchfork effect. This pitchfork power leakage level can be reduced by excluding the sky around the horizon, which can be accomplished by filtering out small fringe rates (Parsons et al., 2016). Regarding the leakage associated with chromatic gain errors, including only certain baselines in calibration is a potential solution. For example, Ewall-Wice et al. (2017) and Orosz et al. (2019) found down-weighting long baselines in calibration can reduce the chromatic gain errors and the leakage of foreground in the power spectrum. Another approach is to smooth out high-frequency structure in antenna gain solutions (e.g., Kern et al., 2020a). The effects of these mitigation methods on the power spectrum estimate will be discussed in Kim et al. (in prep).

Although this study focuses on the middle range of the HERA band, it provides a framework to quantify the instrument configuration requirements and systematic errors. A similar approach can be extended for the rest of the HERA band in future work.

3.7 Appendix

3.7.1 Beam interpolation along the feed motion direction

The primary beams simulated by CST are sampled at regular grid positions, which means the beams interpolated at random feed positions of interest may contain errors arising from the interpolation along the feed motion direction. It is important to evaluate and minimize the potential errors in the interpolation to study the effect of perturbed beams driven by the

feed motion.

The interpolation is performed on far-field electric fields expressed in Equation (3.1). Because the far-field electric field consists of two components, E_θ and E_ϕ, and each component is complex, the interpolation is carried out 4 times at a given frequency, zenith angle, and azimuthal angle along the feed motion direction using `scipy` python package (Virtanen et al., 2020).

To test the accuracy of the interpolation, we ran CST simulations at randomly chosen feed positions or tilts as noted in the title of each panel of Figure 3.18 and compare the CST simulated and interpolated beam. Each pair plot consists of 1D profiles of power beams for the CST simulated and interpolated beams with zenith angle in the EW direction (top) and their fractional difference (bottom). For the pairs in the top row representing the vertical feed offset, the simulated beams and interpolated ones present a good agreement. The fractional difference indicates the interpolation error is less than 0.05%, which means the error is insignificant compared to the perturbation in the beams due to feed motions. For the horizontal feed motion (pairs in the middle row), the error is about 1% or less. The errors can be as large as 5% for the tilts shown in the pair plots in the bottom row but the interpolation error is still less significant than the error of the perturbed beam as shown in Figure 3.1.

Figure 3.19 shows comparison between the CST simulated and interpolated beams with frequency at the zenith point. From top to bottom, each pair plot indicates the vertical, horizontal, and tilting feed motions, respectively. Across all panels, the simulated and interpolated beams are well lined up and we found the fractional difference is less than 1%. In the bottom panel of each pair plot, we show the frequency Fourier transform of the interpolated beam which is consistent with that of the simulated beam, forming the numerical noise floor at around 10^{-7}. This means there is no additional high-frequency structure introduced by the interpolation. This leaves us with the conclusion that chromatic gain errors and leakage in the power spectrum are mainly caused by the beam error induced by the feed motion rather than the interpolation error.

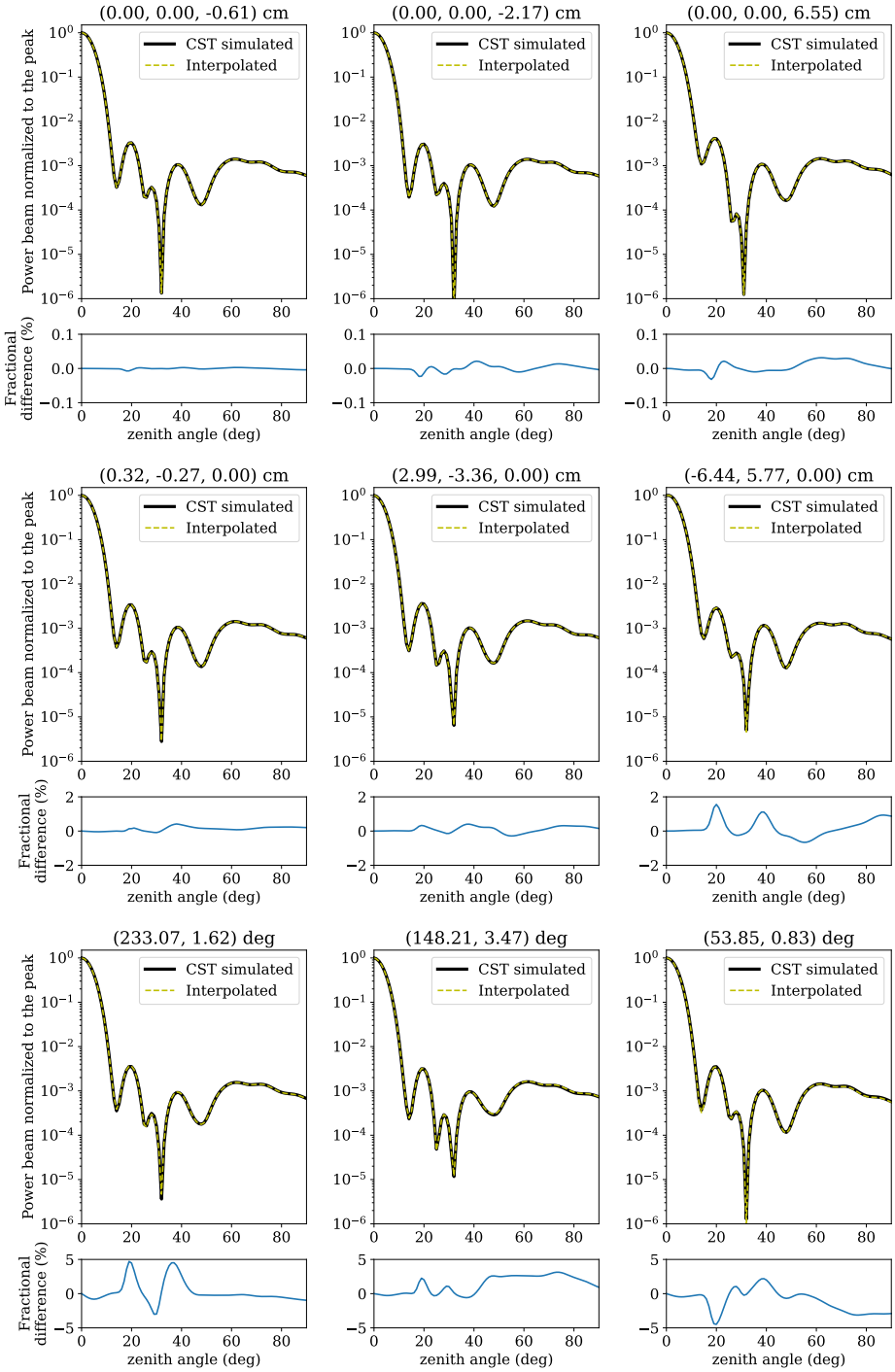

Figure 3.18: Comparison of CST simulated beams and interpolated ones at off-grid feed positions or tilts labeled in the title of each panel at 165 MHz. The title indicates (x, y, z) feed positions for vertical and horizontal feed offsets, and (ϕ, θ) for tilts. Each pair consists of the line profiles of the power beams with zenith angle in the EW direction (top) and their fractional difference (bottom). The top three are for the vertical feed offset, the middle three for the horizontal displacement and the bottom three for the tilting motion.

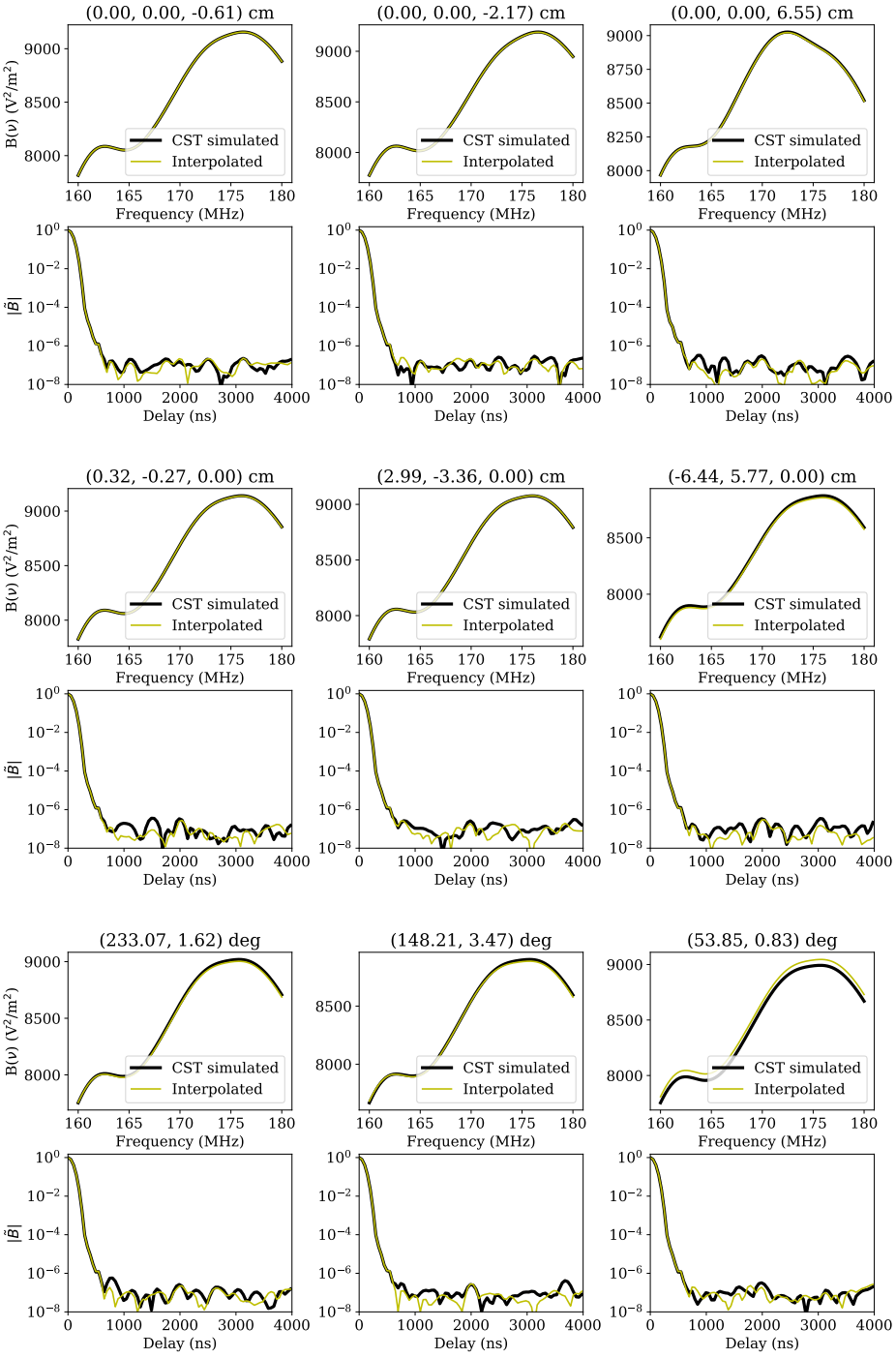

Figure 3.19: Comparison of CST simulated beams and interpolated ones at off-grid feed positions along frequency. The zenith point is chosen for the validation. From top to bottom, each pair plot represents the feed offset in the z-axis, in the xy plane, and in tilt. Overall, two power beams in each panel agree better than ∼1%, and the Fourier transform of the interpolated beam does not show extra high-frequency structure compared to that of the CST simulated one.

Chapter 4

The Impact of Beam Variations on Power Spectrum Estimation for 21 cm Cosmology II: Mitigation of Foreground Systematics for HERA

This work is published in The Astrophysical Journal. Author lists include Honggeun Kim, Nicholas S. Kern, Jacqueline N. Hewitt, Bang D. Nhan, Joshua S. Dillon, Eloy de Lera Acedo, Scott B. C. Dynes, Nivedita Mahesh, Nicolas Fagnoni, and David R. DeBoer.

Abstract

One key challenge in detecting 21 cm cosmological signal at $z > 6$ is to separate the cosmological signal from foreground emission. This can be studied in a power spectrum space where the foreground is confined to low delay (or equivalently, $k_{\parallel}$) modes whereas the cosmological signal can spread out to high delay modes. When there is a calibration error, however, chromaticity of gain errors propagates to the power spectrum estimate and contaminates the modes for cosmological detection. The Hydrogen Epoch of Reionization Array (HERA) employs a high-precision calibration scheme using redundancy in measurements. In this study, we focus on the gain errors induced by nonredundancies arising from feed offset relative to the HERA's 14-meter parabolic dish element, and investigate how to mitigate the chromatic gain errors using three different methods: restricting baseline lengths for calibration, smoothing the antenna gains, and applying a temporal filter prior to calibration. The combination of the baseline cut and temporal filtering indicates that the spurious gain feature due to nonredundancies is significantly reduced, and the power spectrum recovers the clean foreground-free region. We found that the mitigation technique works even for large feed motions but in order to keep a stable calibration process, the feed positions need to be constrained to 2 cm for translation motions and 2° for tilting offset relative to the dish's

vertex.

4.1 Introduction

Understanding the formation of the first stars and galaxies in the universe is key to our broader picture of structure formation over cosmic time. This era, known as Cosmic Dawn, and the subsequent Epoch of Reionization (EoR), when these sources injected ionizing photons into the intergalactic medium (IGM), set the stage for the emergence of modern galaxies as we see them in the present universe; however, we currently have a poor understanding of the astrophysics of these eras. A powerful probe of Cosmic Dawn astrophysics is the 21 cm hyperfine signal emanating from neutral hydrogen in the IGM at redshifts $z > 6$. With the efforts to detect the cosmological signal, there are past, current and upcoming instruments including the Giant Metre Wave Radio Telescope (GMRT; Paciga et al., 2013), the Murchison Widefield Array (MWA; Barry et al., 2019, Beardsley et al., 2016, Dillon et al., 2014, Ewall-Wice et al., 2016a, Tingay et al., 2013, Trott et al., 2020), the Donald C. Backer Precision Array for Probing the Epoch of Reionization (PAPER; Cheng et al., 2018, Kolopanis et al., 2019, Parsons et al., 2010), the Low Frequency Array (LOFAR; Gehlot et al., 2019, Mertens et al., 2020, Patil et al., 2017, van Haarlem et al., 2013), the Hydrogen Epoch of Reionization Array (HERA; DeBoer et al., 2017, Dillon and Parsons, 2016, The HERA Collaboration et al., 2022a,b, 2023), and the Square Kilometre Array (SKA; Koopmans et al., 2015).

A key challenge in making robust measurements of 21 cm emission at the EoR are the bright synchrotron foreground emission coming from the foreground galaxy and from extragalactic sources, which are at least 4-5 orders of magnitude brighter. One of the ways to circumvent the foreground obstacle is to employ a 2D power spectrum approach that has a potential to separate the cosmological signals from the foregrounds. The spectrally smooth foreground emission, in principle, is confined to a certain range of the delay (τ) or the line-of-sight cosmological modes ($k_{\parallel}$), a Fourier dual to the frequency. In contrast, the cosmological signals trace the evolution of the neutral hydrogen with fluctuations in a brightness temperature field along the redshift or frequency, and hence are distributed across a wide range of the delay modes. This provides foreground-free delay modes for the cosmological detection (Datta et al., 2010, Liu et al., 2014, Morales et al., 2012, Morales et al., 2018, Parsons et al., 2012, Pober et al., 2014, Thyagarajan et al., 2013, Trott et al., 2012, Vedantham et al., 2012).

In order to adopt the strategy described above, high-precision antenna gain calibration is required. Especially if there are frequency-dependent (chromatic) calibration gain errors larger than in a fraction of 10^{-5}, the spectral artifacts from the gain solutions can significantly reduce the detectability of the 21 cm cosmological signals. Redundant baseline

calibration proposes a way of precise calibration that uses redundancy in measurements of baselines that sample the same Fourier mode (Dillon et al., 2020, Liu et al., 2010, Wieringa, 1992). If all antennas are identical, the calibration process should return exact gain solutions independent of our knowledge of the sky. However, if there are nonredundancies in the data, for example due to per-antenna perturbations, the redundant calibration can give rise to spurious chromaticity in gain solutions (Byrne et al., 2019, Choudhuri et al., 2021, Orosz et al., 2019). Furthermore, redundant calibration schemes have degenerate degrees of freedom that must be solved using an external sky model (Byrne et al., 2019, Dillon et al., 2018, Kern et al., 2020a, Liu et al., 2010, Zheng et al., 2017).

Kim et al. (2022, hereafter, Paper I) numerically model the deformation of the primary beams for HERA according to different types of per-antenna feed perturbations, including vertical, horizontal, and tilting offsets. Such a beam deformation can introduce nonredundancies into measurements and cause chromatic gain errors from redundant calibration. As a result, the chromatic gain errors will give rise to the leakage of foreground modes into the cosmological modes, which are supposed to be foreground-free in the ideal redundant calibration.

Ewall-Wice et al. (2017) and Orosz et al. (2019) look primarily at the nonredundancy effect on gain errors with simplistic per-antenna primary beam deformations. They show that down-weighting the contribution of long baselines to the calibration can reduce the calibration bias and alleviate the foreground contamination in the 2D power spectrum. Motivated by these studies, we try to mitigate the spectral gain errors found in Paper I with various mitigation methods.

In this study, similar to Paper I, we focus on the Phase II system of HERA with zenith-pointing 14-m dish elements. We explore three different techniques of mitigating the chromatic gain errors. These are

- *Baseline cut*: we restrict the baseline lengths used for calibration, similar to **?** and Orosz et al. (2019).
- *Gain smoothing*: we apply a smoothing filter to the recovered gain solutions across frequency to reject spurious features caused by the inherent nonredundancies in the data.
- *Temporal visibility filtering*: we apply the temporal filter technique (Charles et al., 2023, Parsons and Backer, 2009), or fringe-rate filtering, which filters the data prior to calibration to down-weight strong nonredundant features in the data originating from galactic diffuse emission entering the side lobes of the primary beam.

Note that Charles et al. (2023 in prep) also discuss the use of temporal filtering to suppress

nonredundant errors in calibration, but they focus on its application in mitigating mutual coupling systematics, whereas this work focuses solely on nonredundancies caused by per-antenna feed displacements.

In Section 4.2, we explain about the visibility simulation and the input gain model. Section 4.3 describes the redundant baseline calibration scheme, followed by the absolute calibration, and presents the calibration gain errors before mitigation. Section 4.4 discusses the effects of three different mitigation methods and their combinations on reducing the chromatic gain errors. The power spectrum analysis without and with mitigation is performed in Section 4.5. Section 4.6 summarizes the main results of this study. Throughout the paper, we adopted the cosmological parameters from Planck Collaboration et al. (2016) which are consistent with what we used in Paper I: $\Omega_\Lambda = 0.6844$, $\Omega_\mathrm{m} = 0.3156$, $\Omega_\mathrm{b} = 0.04911$, and $H_0 = 67.27\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

4.2 Visibility Simulations with Perturbed Primary Beams

As described in Paper I, we simplified the feed motions by separating them into vertical, horizontal, and tilting offsets. We utilized a three-dimensional Vivaldi feed model for the calculation of the primary beam at each feed position using the CST electromagnetic simulation software. The fiducial model locates the feed at the place, 5 m above the vertex of the 14-m parabolic dish (Fagnoni et al., 2021). For the perturbed feed models, we displaced the feed away from the fiducial position and simulated far-field electric fields. The far-field electric fields were calculated for the 160–180 MHz band, sampled at 0.125 MHz. The behavior of the primary beam responses alters for different feed perturbations. The vertical feed offset mainly causes a change to the width of the main lobe of the primary beam, while the horizontal or tilting ones mainly results in a change to the pointing angle of the main lobe.

We utilized the HERA core array configuration consisting of 320 antenna elements for visibility simulations (HERA-320; Dillon and Parsons, 2016). The HERA array is designed to be compact with a large number of redundant baselines, so that it can achieve high-precision calibration and high sensitivity to the 21 cm power spectrum. The array configuration requires 51,360 visibility calculations per frequency bin and per time integration.

The visibility measurement for a given antenna pair, frequency, and time is calculated as

follows[1] (e.g., Smirnov, 2011).

$$V_{ij}(\nu, t) = \sum_{n=1}^{N_{\rm src}} B_{ij}(\hat{s}_n(t), \nu) S(\hat{s}_n(t), \nu) e^{-2\pi i \nu \mathbf{b}_{ij} \cdot \hat{s}_n(t)/c}, \tag{4.1}$$

where V_{ij} is the complex visibility at time t and frequency bin ν, $\mathbf{b}_{ij}$ is a baseline vector between antennas i and j, and $\hat{s}_n(t)$ is a pointing vector to a source varying with time in the local observer frame. We consider the middle band of HERA, 160–180 MHz, corresponding to the redshift of $\sim$7 at a cadence of 0.25 MHz. There are three components involved with the calculation. B_{ij} is the power beam computed from far-field electric fields of the antenna pair i and j. S is the flux density of a sky source in Jy that corresponds to either a compact source or a discretized pixel of a diffuse source. The last exponential term is the baseline interferometric fringe response.

The power beam term is defined by $B_{ij}(\theta, \phi, \nu) = E^p_{i,\theta} {E^p_{j,\theta}}^* + E^p_{i,\phi} {E^p_{j,\phi}}^*$ where p is the feed polarization (e.g., Kohn et al., 2019). We use an East-oriented feed polarization throughout the study. Each far-field electric field of antenna i and j is perturbed by the feed motions. We consider vertical, horizontal, and tilting motions separately. Those motions follow a Gaussian distribution with zero mean and standard deviation $\sigma_{\rm feed} = 2$, 3, and 4 cm for the vertical and horizontal feed motions and $\sigma_{\rm feed} = 2^\circ$ and 3° feed tilting at the fiducial feed position, providing 8 different simulation sets. The perturbed power beam responses were then fed into the sky visibility calculation. Descriptions of the beam simulations and visibility simulations are detailed in Paper I.

For the foreground emission, we consider two different sky components, a point source model and a diffuse sky model. The former comes from the GaLactic and Extragalactic All-sky MWA (GLEAM) survey. We combined GLEAM I (Hurley-Walker et al., 2017) and GLEAM II (Hurley-Walker et al., 2019) point source catalogs as well as peeled bright sources listed in Table 2 of Hurley-Walker et al. (2017) and Fornax A (e.g., Bernardi et al., 2013). The diffuse sky model is constructed with Hierarchical Equal Area isoLatitude Pixelization (`HEALPix`; Gorski et al., 2005) using the Global Sky Model (GSM; Zheng et al., 2017) that removed the bright compact sources from the sky model. This makes us minimize double counting bright point sources when we combine the point source model and the diffuse sky model. We chose `HEALPix Nside` = 256, containing 786,432 pixels with the pixel area of 1.60×10^{-5} sr. When the pixel area is larger than the angular resolution of a particular baseline ($\theta \sim \lambda/|\mathbf{b}|$), an error may be introduced into the power spectrum due to under-sampling of the discretized map (Lanman et al., 2020). For the array configuration and the frequency band we adopted, the angular resolution corresponding to the longest baseline, 292 m, is $\sim 3 \times 10^{-5}$ sr that satisfies the Nyquist limit. This means for most of baselines,

[1]We developed a visibility simulator for fast computation of visibilities using multiple beam models that are randomly perturbed. It is publicly available in https://github.com/vispb/vispb.

especially for short baselines that could achieve high sensitivity in power spectrum estimate, the adopted pixel area of the GSM does not have a critical impact on our power spectrum analysis.

The cosmological signals are simulated as a reference for the detection of the EoR. To simulate the EoR signal, we employed the `21cmFAST` public software (Mesinger et al., 2011) with the fiducial model studied by Park et al. (2019). The cube size and the number of pixels per side are 250 Mpc and 128, respectively. The cube is simulated at the redshifts of interest spanning 160–180 MHz, and each cube is converted to the `HEALPix` map to feed into the visibility simulation by following the method described in Kittiwisit et al. (2017). For the conversion, we considered `Nside` = 8192 so that at least two `HEALPix` pixels are sampled at each cube cell, and degraded the resolution of the `HEALPix` to `Nside` = 256 that is computationally feasible for the visibility simulation. The EoR visibility is simulated with the fiducial primary beam. Because the EoR visibility is only for assessing the detectability of the EoR in the power spectrum analysis, no gains and calibration are applied.

As the sky rotates with time, the nonredundancy effect can alter at a different local sidereal time (LST). For example, Paper I found that when strong diffuse emission, such as the galactic plane, is located near the horizon, the deformation of side lobes due to feed motions can induce significant nonredundancies in the visibility data. Additionally, another type of nonredundancy can arise when strong compact radio sources come in and out the main lobe with time. This effect is expected to be stronger for the vertical feed motion since it predominantly affects the beam width of the main lobe comparing to the horizontal and tilting perturbations.

In this work, we simulated the visibility over LST of 0–3 hours at a cadence of 50 seconds. The range and sampling rate along the LST is chosen for the temporal filter defined in the fringe rate (f), carrying units of millihertz (mHz), the Fourier dual of time defined as $\tilde{V}(f) = \int dt\ e^{2\pi i t f} V(t)$. That is, we have enough resolution (δf $\sim$ 0.046 mHz) and coverage (Δf $\sim$ 10 mHz) in the fringe-rate domain to perform the fringe-rate filter which will filter out signals $<$ 0.3 mHz. Throughout the analysis, the feed positions are fixed, and we do not take into account time-dependent feed motions.

The simulated raw visibilities are corrupted by multiplying complex antenna gains to yield the "observed" visibilities, $V_{ij}^{\rm obs} = g_i g_j^* V_{ij}^{\rm raw}$. The antenna gains g_i are modeled by following the calculation of the voltage response described in Equation of (5) of Fagnoni et al. (2021). The calculation is based on the impedance of the antenna, the electric-field pattern, and the scattering parameter. The gain model is interpolated by using Gaussian process regression (GPR) with a fixed kernel size of 1 MHz, which effectively smooths out frequency structure smaller than 1 MHz (e.g. Kern et al., 2018, Lanman et al., 2020). Then the input gains are randomly perturbed in amplitudes (η_i) and cable delays (τ_i), by $\exp(\eta_i + 2\pi i \tau_i \nu)$, according to a Gaussian distribution of zero mean and a standard deviation of 0.2 and 20 ns, respectively.

Details of the full antenna gain derivation are described in Paper I. We ignore thermal noise in generating the measured visibilities to focus on the foreground contamination leaking into the EoR window that are supposed to be foreground-free.

4.3 Redundant-baseline Calibration and Observed Chromatic Gain Errors

HERA benefits from the highly redundant baseline configuration for precise calibration. We follow the standard HERA direction-independent calibration scheme consisting of a redundant baseline calibration step, followed by an absolute calibration step that solves for the inherent degeneracies in redundant calibration. Instead of using prior information of the sky, the redundant baseline calibration uses the redundancy of measurements of the same baselines. For antenna pairs with the same baseline separation, they are assumed to share the same unique visibility solution, $V^{\rm sol}_{i-j}$. Namely, the redundant calibration is carried out by solving $V^{\rm obs}_{ij} = g_i g_j^* V^{\rm sol}_{i-j}$ to obtain g_i, g_j, and $V^{\rm sol}_{i-j}$ by minimizing the χ^2 statistic,

$$\chi^2 = \sum_{i<j} \frac{|V^{\rm obs}_{ij} - g_i g_j^* V^{\rm sol}_{i-j}|^2}{\sigma^2_{ij}}, \tag{4.2}$$

where σ^2_{ij} is the thermal noise corresponding to each $V^{\rm obs}_{ij}$.

The antenna gains and unique visibility solutions derived from the redundant baseline calibration can be degenerate. There are four well-known degenerate parameters, including the overall amplitude ($g_i \rightarrow \hat{A} g_i$, $V^{\rm sol}_{i-j} \rightarrow V^{\rm sol}_{i-j}/\hat{A}^2$), the overall phase ($g_i \rightarrow e^{i\psi} g_i$, $g_j^* \rightarrow e^{-i\psi} g_j^*$), and two phase gradients associated with tip-tilt slopes across the array along the east-west ($g_i \rightarrow g_i e^{i\Phi_x x_i}$ and $V^{\rm sol}_{i-j} \rightarrow V^{\rm sol}_{i-j} e^{-i\Phi_x (x_i - x_j)}$) and the north-south direction ($g_i \rightarrow g_i e^{i\Phi_y y_i}$ and $V^{\rm sol}_{i-j} \rightarrow V^{\rm sol}_{i-j} e^{-i\Phi_y (y_i - y_j)}$) where $\hat{A}$, ψ, Φ_x, and Φ_y are scalar free parameters and x_i and y_i are antenna positions (Byrne et al., 2019, Dillon et al., 2018, Dillon et al., 2020, Kern et al., 2020a, Li et al., 2018, Liu et al., 2010, Zheng et al., 2014). To solve for these degeneracies, an absolute calibration step using sky information is needed. In this second calibration step, the raw visibility simulated with the unperturbed beam is chosen for the model visibility. As discussed in Paper I, when the feed positions are perturbed, majority of the chromatic gain errors come from the redundant baseline calibration step, and the absolute calibration mainly causes the errors at high delays.

During the calibration, we exclude bad antennas with relatively larger feed displacement that may degrade calibration solutions. The process is implemented by ruling out antennas with a modified Z-score > 4 in the redundant-baseline calibration. The modified Z-score for antenna i, described in The HERA Collaboration et al. (2022a), is defined as $Z_i^{\rm mod} =$

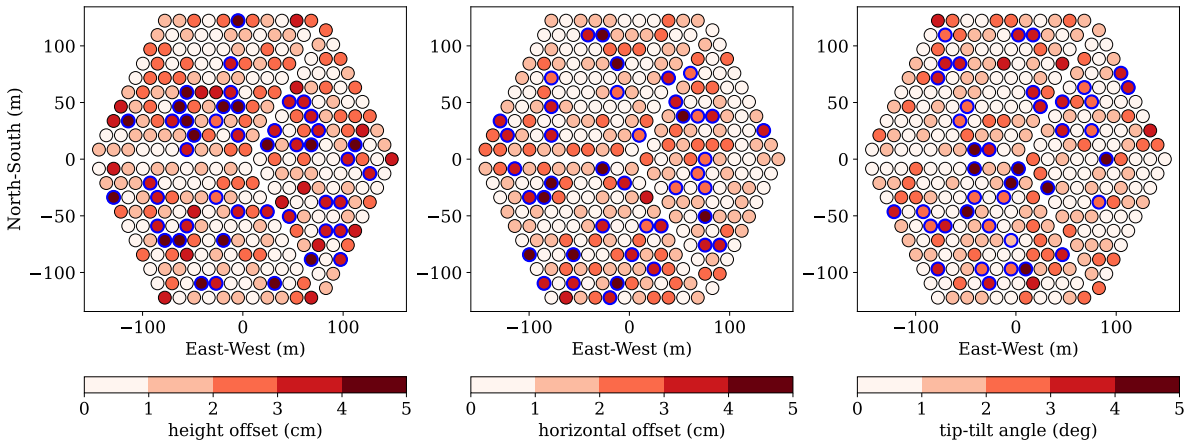

Figure 4.1: Antenna layout with antennas colored according to the feed offset. This is an example when the feed positions are perturbed by following a Gaussian distribution with zero mean and $\sigma_{\rm feed} = 2$ cm for the vertical and horizontal feed motions and $\sigma_{\rm feed} = 2^\circ$ for the tilting feed motion. The color bar indicates absolute values of the offsets that are sampled to 5 bins for illustrative purposes. The flagged antennas based on the modified Z-score metric in the redundant-baseline calibration are marked with blue circular boundaries. Overall, the metric works well enough to exclude bad antennas with large offset.

$(x_i - \mathrm{med}(x))/\sigma^{\rm mad}$ where $\sigma^{\rm mad} = 1.482 \times \mathrm{med}|x - \mathrm{med}(x)|$ which is the median absolute deviation. x_i is the mean visibility amplitude computed by $x_i = \sum_{j\neq i,\nu,t} |V_{ij}(\nu,t)|/(N-1)$ where N is the number of antennas and the sum is performed over $j \neq i$, ν, and t. This metric takes an iterative process that is repeated three times with the removal of bad antennas (i.e., modified Z-score > 4) in every iteration.

Figure 4.1 illustrates how the modified Z-score works to exclude bad antennas during the calibration process. The feed is displaced according to a Gaussian distribution of the zero mean and $\sigma_{\rm feed} = 2$ cm for the translation feed motion and $\sigma_{\rm feed} = 2^\circ$ for the tilting feed motion. The flagged antennas with the modified Z-score > 4 are highlighted with bold edges of blue circles. This result indicates that many antennas with large feed offset around 3–5 cm or 3–5° are ruled out during the calibration and the modified Z-score metric is an effective tool to identify bad antennas that may harm the calibration. We found that this reduces the overall χ^2 defined in Equation (4.2) but has little impact on reducing spurious spectral gain structure and thus no advance in terms of reducing the foreground power leakage in the power spectrum.

In practice, antennas that contribute large errors to the calibration solution are flagged. To be consistent with actual HERA calibration techniques, calibration solutions in this analysis are calculated with flagged antenna omitted. We only consider $\sigma_{\rm feed} = 2$ cm for the axial and lateral feed motions and $\sigma_{\rm feed} = 2^\circ$ for the tip-tilt motion throughout the paper unless it is explicitly stated otherwise.

4.3.1 Chromatic Gain Errors

For the fiducial model where all antenna elements have identical beam responses, the calibration is expected to be perfect without introducing additional frequency structure to the antenna gain solutions. When the feeds are perturbed, however, the key assumption of redundancy is violated, and the mismatch between visibilities in the redundant baseline group may introduce artifacts to the gain solution to account for the nonredundancy in the data. In Figure 4.2, we demonstrate how the artifacts look like for the vertical (second column), horizontal (third column), and tilting (forth column) feed motions compared to the fiducial one (first column). We select three antennas to illustrate the effects of different types of feed motions on the recovered gain solutions: one antenna displaced by -4.3 cm in height from the vertical feed motion simulation, another off by $+3.4$ cm in east-direction from the horizontal motion simulation, and the other perturbed by $+3.1^\circ$ in east-direction from the tip-tilt simulation.

As the input gain does not change in time, we expect calibration gains from the fiducial beam model should also be time-constant as shown in the first panel of Figure 4.2. When feeds are perturbed, regardless of the type of feed motions, calibration solutions show additional spectral and temporal structure. The artificial features are evident in the difference between the calibration gains and the input gains (bottom row).

There are two different scales of artifacts. One is the change in the bulk shape of the gain that can be understood based on the primary beam patterns changing with the feed motions. For example, in the bottom panel of the vertical motion, there are negative and positive features along frequency at LST$\sim$3 hr that is close to Fornax A located at 3.3 hr, one of the brightest radio sources in the HERA stripe (e.g., Dillon et al., 2020). Paper I shows that the change in the feed height can drive the change to the beam width. With the height offset of -4.3 cm, the beam width is narrower than the fiducial beam at 160 MHz and gets broader as the frequency increases, and reaches a similar beam width level as the fiducial one at $\sim$165 MHz (see Figure 3 of Paper I). Such a behavior indicates the flux density of Fornax A is more/less attenuated by the primary beam than the fiducial one at frequency below/above 165 MHz, which is consistent with the pattern shown in Figure 4.2.

Similarly, when the feed moves in the horizontal direction or tilts, the primary beam's center position shifts. According to the results of Paper I, when the feed moves in the east-direction, the beam shifts in the opposite direction, and hence the beam-weighted flux density of Fornax A becomes dimmer than the fiducial case at LST$\sim$3 hr, resulting in smaller gain amplitudes. Meanwhile, when the feed tilts in the east-direction, the beam shifts in the same direction, and the beam-weighted flux density of Fornax A is brighter than the fiducial one at LST$\sim$3 hr, leading to larger gain amplitudes than the fiducial beam model. Additionally, there are also horizontal patterns between 1.5 and 2.5 hr in the bottom panels

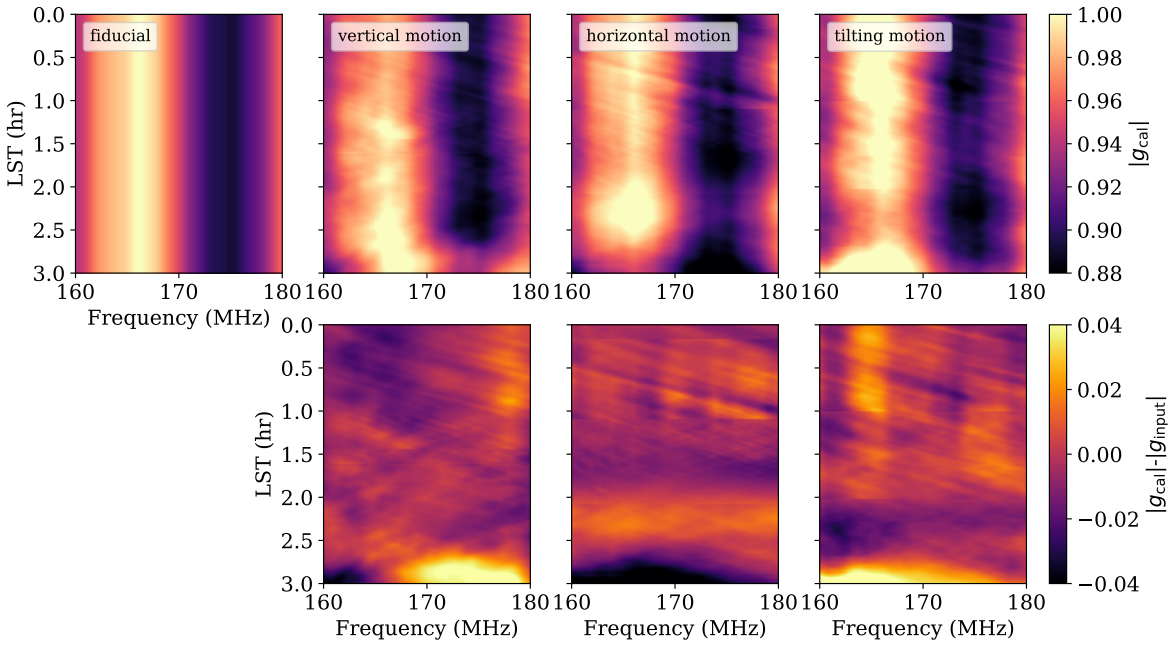

Figure 4.2: Antenna gain solutions derived from the perturbed beam simulations with LST compared to the fiducial one. Three randomly perturbed antennas are picked from the simulation sets for vertical, horizontal, and tilting motions, respectively. Each antenna is perturbed with −4.3 cm of height offset (center left), +3.4 cm of east-direction displacement (center right), and +3.1° of east-direction tip-tilt (far right). Top panels show the gain amplitudes and the bottom panels show the difference between the calibrated gains and the input gains. For the fiducial model, the gain solutions have constant temporal structure as the input gains without any artificial features. When the feed is perturbed, fine-scale artifacts show up and result in corrupting the cosmological modes in the EoR window.

of horizontal and tilting motions. Those are caused by two point sources transiting zenith at LST∼2 hr, and can be understood in the same way as the LST∼3 hr feature.

Another noticeable artificial features are fine-scale structure in frequency. Because we include long baselines in the calibration, artificial gain features arising from the baselines are expected to form fine structure in frequency, resulting in complex patterns as observed in Figure 4.2. These chromatic gain errors corrupting high delay modes are ones responsible for the foreground leakage observed in Paper I. Figure 4.3 presents how the chromatic gain errors of the perturbed beam models are shaped in the delay domain compared to the fiducial model. To transform the gain to the delay spectrum, we divided the gain by the input gain, multiplied it by a seven-term Blackman-Harris tapering function along the frequency axis, and performed the frequency Fourier transform. Then we averaged the amplitudes of the peak-normalized delay spectra over unflagged antennas at each LST bin. In the first column, we show results of the fiducial model with the time-constant gain solutions (top) and the LST-averaged gain solution consistent with the input gain as expected (bottom).

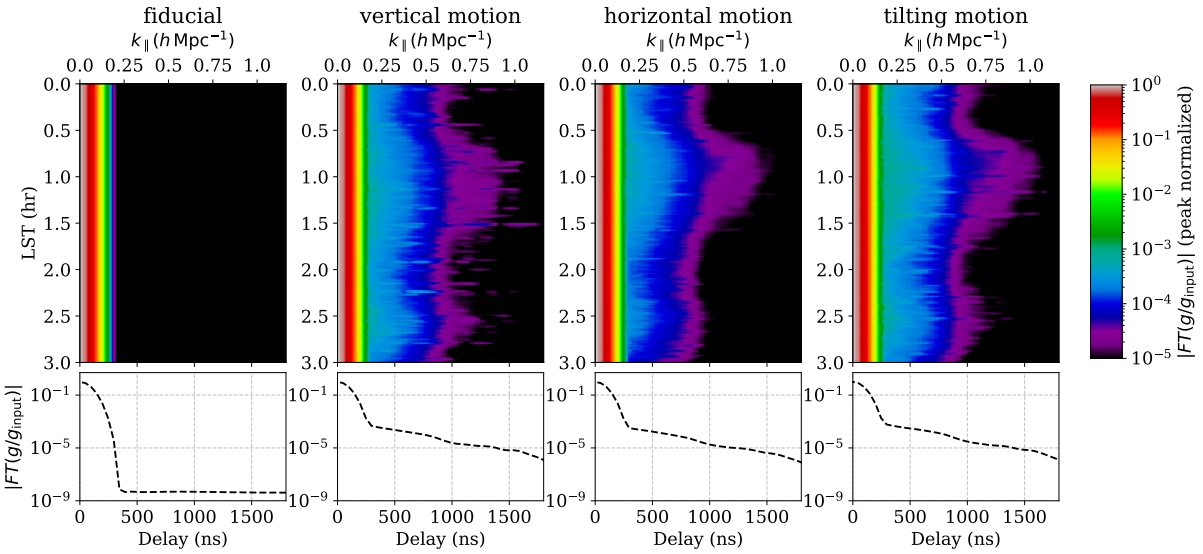

Figure 4.3: Delay spectrum of the calibrated gains without mitigation. Top panels show the gain delay spectrum varying with LST. Bottom panels present the LST-averaged delay spectrum. The fiducial beam model (far left panel) recovers the input gain, providing the clean window at high delay modes > 400 ns. g/g_{input} is constant with frequency and the bell-shaped curve shown in the first bottom panel comes from the seven-term Blackman-Harris tapering function. For the perturbed cases, the window $k_\parallel > 0.25\,h\,\mathrm{Mpc}^{-1}$ is corrupted by chromatic gain errors arising from the nonredundancy in measurements. The bulge occurring at 0.5–1.5 hr is due to the galactic plane located near the horizon.

As shown in Figure 4.3 from the second left to right columns, when the feeds are displaced from the fiducial position, gain errors bleed into high delay modes. The overall patterns are similar regardless of the type of feed motions, showing more power leakage than the fiducial model across all LST. The strongest bleeding regions appear to be at LST range of 0.5–1.5 hr, when the galactic plane lies near the horizon and sets as the Earth rotates as depicted in Figure 4.4. Since the perturbed primary beams induced by the feed motions are responsive to the strong diffuse emission from horizon, it can be understood that the bulge produced at LST$\sim$1 hr is most likely due to the galactic plane near the horizon. In addition, because the feed offsets, especially vertical ones, are sensitive to the stochastic distribution of point sources, they introduce irregularly narrow structures across each LST bin.

Bottom panels of Figure 4.3 show the averaged gain delay spectra over LST. For the perturbed feed models, power in the delay spectrum is 4–5 orders of magnitudes stronger than the input gain at 500–1000 ns. Because the power spectrum of cosmological signals is expected to be 10^{10} times smaller than the foregrounds (e.g., Thyagarajan et al., 2016), the delay spectrum should be smaller than 10^{-5} at around 700 ns or $k_\parallel \sim 0.4\ h\,\mathrm{Mpc}^{-1}$, which indicates we need to mitigate the chromatic gain errors to achieve robust detection of the EoR[2]. We chose $k_\parallel \sim 0.4\ h\,\mathrm{Mpc}^{-1}$ since it provides a foreground-free region for the fiducial

[2]Though averaging uncorrelated gain errors in forming a power spectrum may lessen this stringent re-

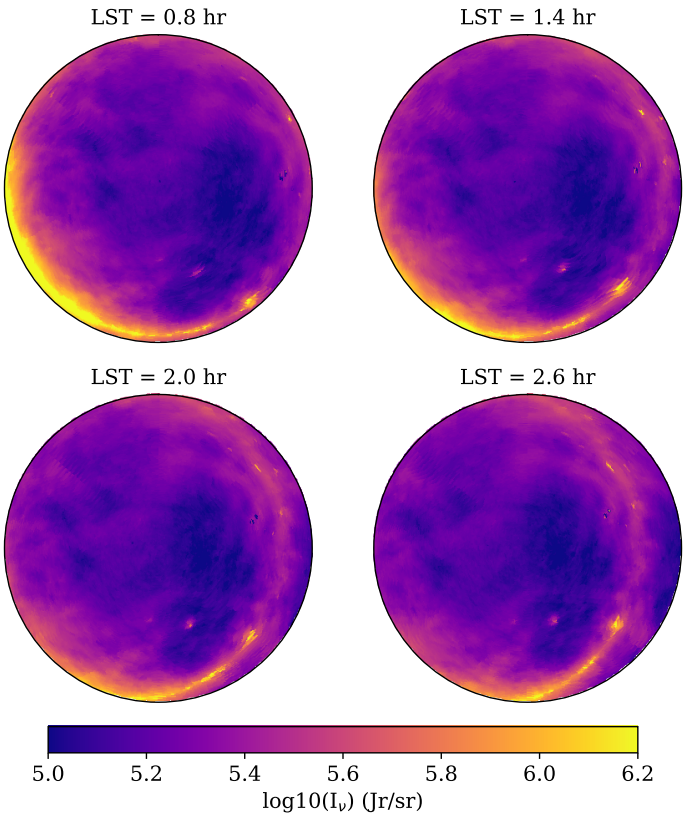

Figure 4.4: GSM maps viewed in horizon coordinates from the HERA observing site at (−30.72°S, 21.43°E), plotted at different local sidereal times. At LST∼0.8 hr, the galactic plane lies near the horizon, and the bright emission disappears as the Earth rotates.

power spectrum (see Figure 4.13) with high sensitivity when thermal noise is included. In the following section, we introduce three different ways of mitigating these chromatic gain errors.

4.4 Mitigation of the spectral structure in the gain solutions

In this section, we explore three different ways of mitigating spurious fine-scale frequency structure in the derived antenna gains due to nonredundancy created by random feed displacements. Two of the methods, baseline cut and temporal filtering, treat the visibilities before calibration, thus mitigating potential errors introduced by the calibration process itself. Meanwhile, the other method, gain smoothing, treats the derived gains after performing

quirement of 10^{-5} in gain powers, we use the requirement as it guarantees the recovery of a clean EoR window as shown in Figure 4.14.

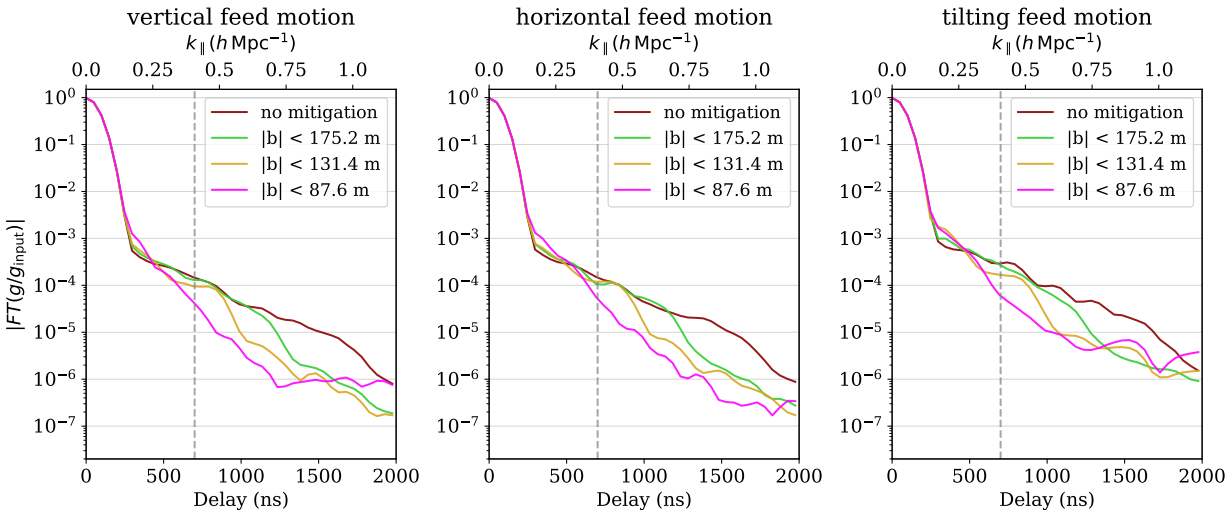

Figure 4.5: The gain delay spectra with different maximum baseline cut-offs for vertical (left), horizontal (middle), and tilting feed motions (right) at 1 hr of LST. The Fourier transform of the antenna gain divided by the true input gain is averaged over the unflagged antennas. There is a trend that the broad wing appearing at high delays gets reduced as we consider tighter baseline limits for the calibration, indicating the chromatic gain errors are mitigated with the restriction of the baseline lengths. There is an order of magnitude improvement at 700 ns or $k = 0.4\,h\,\mathrm{Mpc}^{-1}$ (vertical dashed line) for $|\mathbf{b}| < 87.6$ m case compared to no mitigation case.

calibration. Baseline cut is to down-weight long baselines in the data, which contribute to gain errors at fine spectral scales during calibration. Gain smoothing is to suppress the fine-scale frequency structure in the derived gain solutions using a low-pass frequency filter after the calibration, thus smoothing the gains. Temporal filtering is to apply the fringe-rate notch filter that excludes the horizon emission before performing the redundant calibration. This is effective to reduce the nonredundancy in the visibility when strong emission such as the galactic plane is lying on the horizon. All different strategies are discussed separately and then combined to illustrate the composite effect of the mitigation on the power spectrum in Section 4.5.

4.4.1 Down-weighting Long Baselines

High-frequency errors in antenna gains caused by nonredundancy associated with compact radio sources can predominantly arise from long baselines with high delays ($|\mathbf{b}|/c$). For example, visibilities suffering from nonredundancy in a 200-m baseline group, which has a delay of ∼700 ns, can be affected by nonredundancy errors at the scale of ∼1 MHz or smaller, which will propagate into the same scale of gain errors during the calibration process. Ewall-Wice et al. (2017) and Orosz et al. (2019) found that chromatic gain errors can be reduced by placing a restriction on baseline lengths used in the redundant calibration. Following the

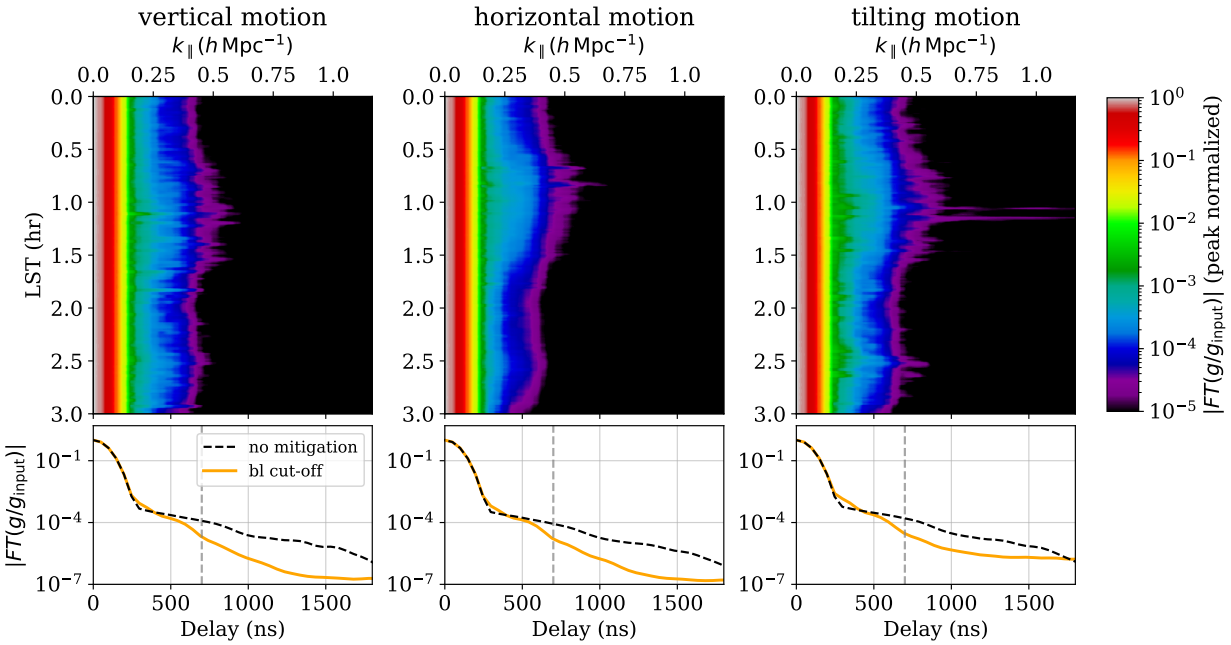

Figure 4.6: Delay spectrum of the gain solutions with down-weighting long baselines. We consider 87.6 m baseline cut-off for the mitigation. The top row shows the delay spectrum for the mitigated gain with LST for the vertical (left), horizontal (middle), and tilting (right) feed motions. Overall, the delay spectra fall down to $< 5 \times 10^{-5}$ at 700 ns, which is also manifested in the bottom row for the LST-averaged delay spectra.

approach of Orosz et al. (2019), we considered a binary weight to down-weight baselines that are longer than certain threshold (i.e., 1 for shorter baselines and 0 otherwise).

In order to look for an optimal maximum baseline cut-off, we examined the improvement of calibration gains by changing the baseline cut-off used for the redundant calibration and absolute calibration. For the baseline cut-off values, we chose 87.6, 131.4, and 175.2 m which correspond to 6×, 9×, and 12× the shortest baseline (14.6 m), respectively.[3] Figure 4.5 presents the gain delay spectra with different baseline cut-offs for the vertical (left), horizontal (middle), and tilting feed motion (right) at LST = 1 hr. Each curve is computed as the Fourier transform of the ratio of antenna gains and input gains, weighted by a seven-term Blackman-Harris window function, then averaged over all unflagged antennas. Regardless of the types of feed motions, it is apparent that the power leakage at high delays is suppressed more as the baseline limit becomes tighter, which supports the finding of Orosz et al. (2019).

With the cut-off of 87.6 m, the power leakage is reduced by an order of magnitude at $k_\parallel \sim 0.4\,h\,\mathrm{Mpc}^{-1}$ (vertical dashed line). For all cases, the mitigation brings the power leakage down to 5×10^{-5}, which is still a bit higher than desired. There is an additional bump at

[3]In practice, we also impose a minimum baseline cut-off to reduce the effect of modeling uncertainty of diffuse emission in the absolute calibration. Thus, choosing maximum baseline cut-off smaller than 87.6 m may cause lack of redundant baselines in the system.

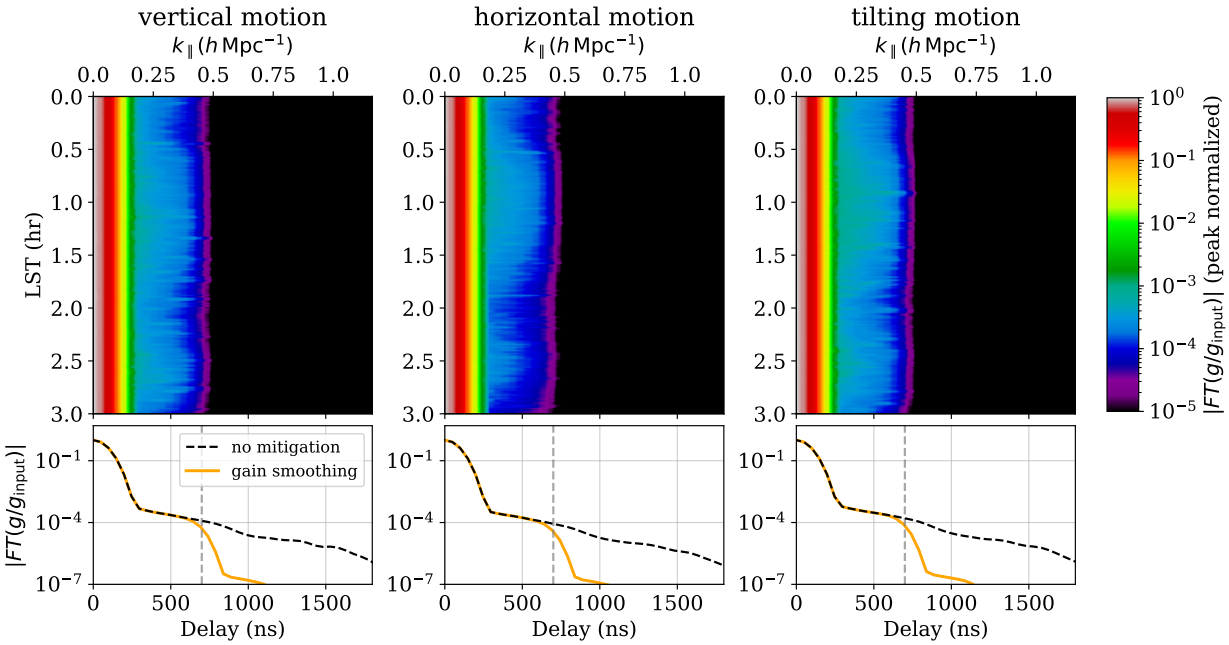

Figure 4.7: Delay spectrum of the gain solutions with gain smoothing along the frequency. The format is the same as that of Figure 4.6. As the Fourier filter rules out the spectral scales smaller than 1 MHz, or larger than 500 ns, the top panels show the clean window beyond 700 ns for all types of the feed motions over all LSTs as expected.

$\sim$300 ns (magenta curve) compared to no mitigation case since redundant calibration is most sensitive to the nonredundancy errors arsing from the cut-off baseline corresponding to the delay of 300 ns (i.e., $87.6/c \sim 300$ ns).

In Figure 4.6, we illustrate how the mitigation with the 87.6-m cut-off is effective over the LST range. Across all LSTs, the mitigated gain has reduced spectral artifacts compared to the case without mitigation. The mitigation is relatively less effective at LST between 0.5 and 1.5 hr, where significant chromatic gain errors result from strong horizon emission captured by the perturbed beams. This is even worse for the tilting motion as there are more variations in the side lobes owing to that motion than other motions. The LST-averaged delay spectra in the bottom row of Figure 4.6 indicate that the chromatic gain errors are sufficiently mitigated at high delays but need to be suppressed further at $k_\parallel = 0.4\,h\,\mathrm{Mpc}^{-1}$. We adopt 87.6 m as the optimal baseline cut-off for the remaining sections.

4.4.2 Gain Smoothing

The mitigation of fine-scale spurious features in the derived antenna gain solutions can also be achieved by smoothing the gains in some manner. Barry et al. (2016) proposed doing this with low-order polynomials, whereas Kern et al. (2020a), The HERA Collaboration et al.

(2022a) propose doing this with Fourier filters, while Yatawatta (2015) enforce spectral smoothness in the process of calibration itself. In this work, we use a formalism similar to the one used in The HERA Collaboration et al. (2022a) for smoothing the derived gains in frequency with a low-pass filter. Our input gain is designed to have features at a scale larger than 1 MHz. Thus, we find that a 1-MHz smoothing scale to be the most effective in rejecting spectral gain errors while retaining major components of the true gain structure that need to be calibrated out.

The delay spectrum results after applying the low-pass filter with the filter size of 500 ns, corresponding to the 1 MHz smoothing scale, are shown in Figure 4.7. As expected, power leakage beyond 700 ns (vertical dashed line) is well suppressed below the level of 10^{-5} across all LST bins for all three types of feed motions. However, unlike other mitigation methods used in this study, this method is essentially a low-pass filtering thus, by design, ineffective in reducing gain error power leakage at delays < 500 ns, which corrupts low k-modes in the power spectrum. The LST-averaged delay spectra for all feed offset types (bottom panels) also display a fast drop-off of the power beyond 700 ns, and no mitigation of the chromatic errors inside 500 ns.

4.4.3 Fringe-rate Filtering

For interferometric drift-scan observations, which do not track any particular field on the sky and simply observe the sky as it passes through the baseline fringes, there is a useful mapping from time-based Fourier modes in the visibility (i.e., fringe-rates) to the location on the sky where the signal originated from. Parsons and Backer (2009) first studied this relationship for the purposes of low-frequency interferometric calibration. This was later studied as a means for understanding both foregrounds and instrumental features seen in drift-scan observations (e.g. Parsons et al., 2016, Shaw et al., 2014). For HERA, these kinds of filters have been used to effectively mitigate instrumental systematics (Josaitis et al., 2022, Kern et al., 2019, 2020b, The HERA Collaboration et al., 2022a).

Recently, Charles et al. (2023) proposed using fringe-rate filters to improve the fidelity of gain calibration for low-frequency 21 cm interferometers. Their analysis focused solely on mitigating the impact of poorly modeled diffuse emission on absolute calibration, while Charles et al. (2023 in prep) investigates how this technique can mitigate nonredundancies arising from mutual coupling. Here, we adopt this fringe-rate filtering technique to study its efficacy in improving redundant calibration in the presence of nonredundancies caused by antenna beam variations. Specifically, we use a "fringe-rate notch" filter that rejects $f \sim 0$ mHz modes in the data similar to the one described in Charles et al. (2023). In our paper I, we showed that diffuse emission from the observer's horizon, which has $f \sim 0$ mHz, is one of the main sources of the nonredundancy error, especially for the feed tilting offset (see

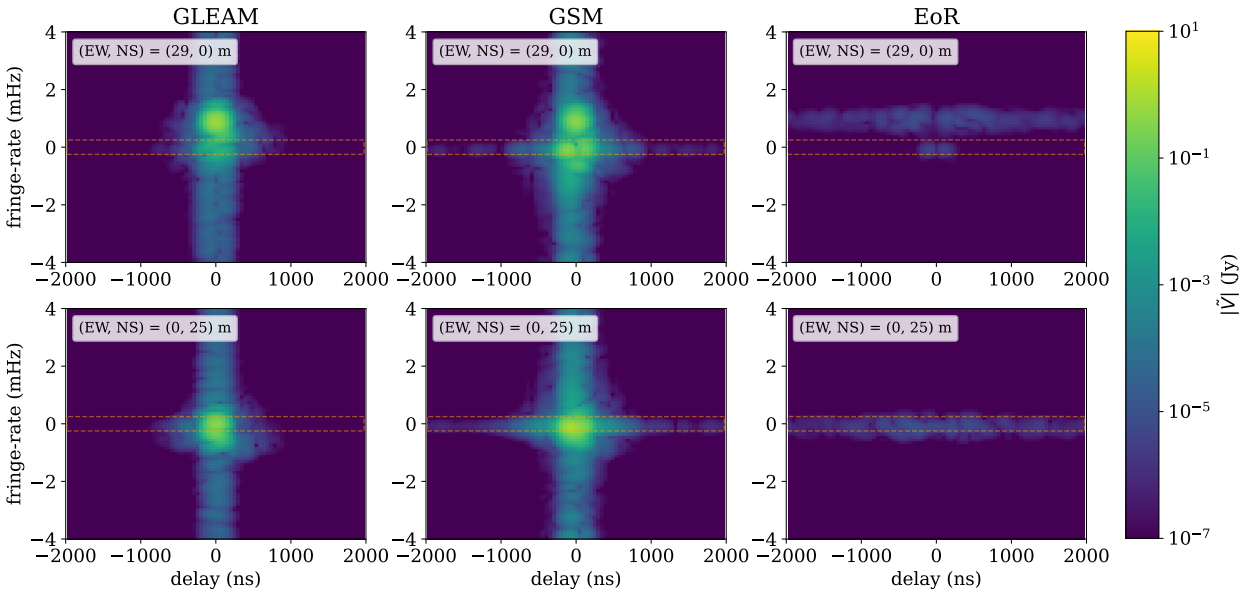

Figure 4.8: Fourier transform of simulated HERA visibilities for GLEAM (left), GSM (middle), and EoR (right) sky models in the delay and fringe-rate domain. Two antenna pairs with similar separations but with different orientations are presented as labeled in the plot. For the north-south only baseline (bottom row), all types of sky models have their power located near zero fringe rates, which drift away from the center as the baseline length in the east-west direction increases. On the top row, the two central bloblike features in the GSM panel that originate from the pitchfork effect of the horizon emission. The fringe-rate notch filter excludes the horizon emission within the region outlined by the orange dashed box.

Figure 7 of Paper I). This implies that applying the fringe-rate notch filter should reduce nonredundancies in the data, and improve the calibration process.

Figure 4.8 describes where different kinds of sky sources are located in the delay and fringe-rate space. We considered raw visibilities (i.e., no gains and calibration applied) for GLEAM (left), GSM (middle), and EoR (right) signals, separately to present Fourier transform as functions of the frequency and time. Two short baselines were chosen: the east-west direction (top panels) and the north-south direction (bottom panels). If we look at the delay axis first, most power of the GLEAM sources resides in the low delay modes spanning -200 to 200 ns since the point sources are spectrally smooth. In the case of diffuse emission based on the GSM, there are additional features such as two central bright blobs from the pitchfork effect (Thyagarajan et al., 2015a,b) and extended power in high delay modes. As for the EoR signal, the spectrum is fluctuating along frequency, and thus the power is uniformly distributed in the delay domain.

Now let's consider the fringe-rate axis. The north-south baseline maps the sky inside the main lobe to near zero fringe rates and thus contains majority of the power at $f \sim 0$ mHz as shown in the bottom panels. For the east-west baseline, sky sources in the main lobe of

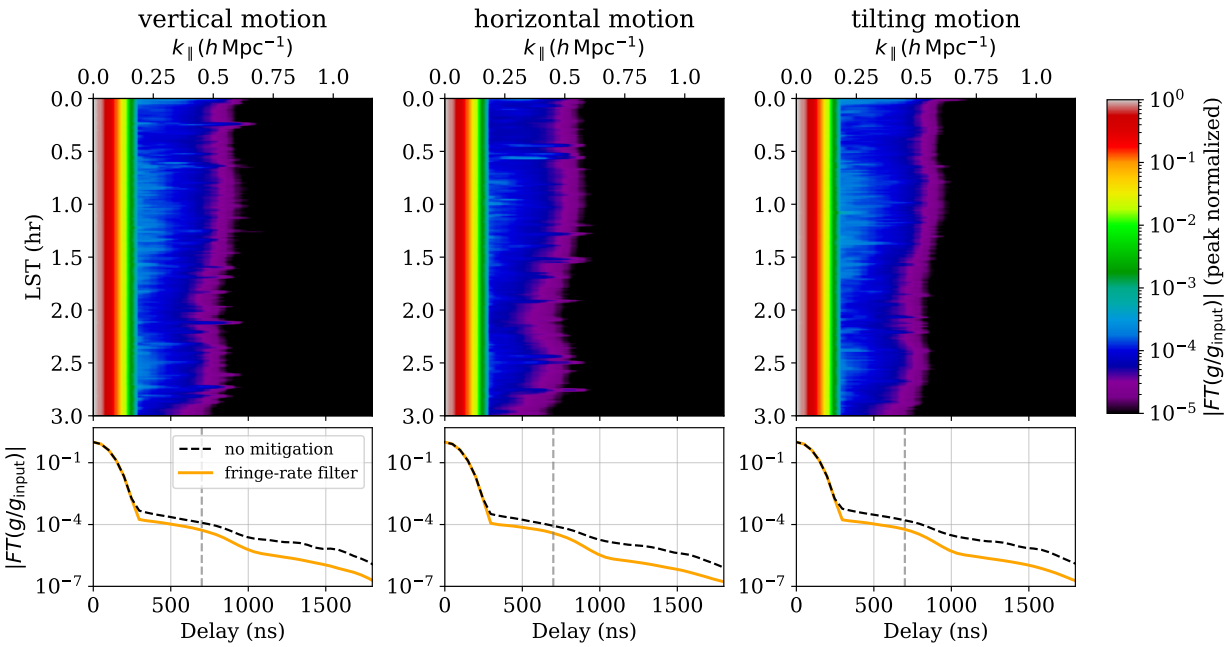

Figure 4.9: Delay spectrum of the gain solutions with fringe-rate filtering. The format is the same as Figure 4.6. We considered the half filter width $f_w = 0.3$ mHz corresponding to the orange dashed box in Figure 4.8. The mitigation takes effect across all delays. The improvement at around 300–500 ns is notable compared to previous mitigation methods.

the primary beam, especially at zenith, have positive fringe rates and the main lobe drifts in the positive direction as the east-west baseline length increases. However, for the GSM, there is still a significant fraction of power remaining at zero fringe rate that originates from the horizon galactic emission. Namely, the fringe-rate notch filtering can suppress nonredundancies dominated by the diffuse horizon emission by removing signals at around zero fringe rates within the orange dashed box. Further details about the baseline-dependent relation between the sky positions and fringe rates are described in Parsons et al. (2016).

For fringe-rate notch filtering (analogous to a high-pass temporal filtering), we adopted the `DAYENU`[4] filter (Ewall-Wice et al., 2020) which is a linear filter defined in a basis of Discrete Prolate Spheroidal Sequences (DPSS; Slepian, 1978). This filter takes an analytic covariance model to suppress signal inside a half-filter width (f_w) with a suppression factor (ϵ), where $\epsilon = 10^{-8}$ was chosen for this study.

The high-pass DPSS-based filter is applied to the observed visibilities before the calibration along the time axis. The filtered visibilities are then calibrated through the redundant and absolute calibration steps with the model visibilities that are also fringe-rate filtered. An optimal half filter width of $f_w = 0.3$ mHz was determined by accessing degrees of mitigation

[4]The full name is DPSS Approximate lazY filtEriNg of foregroUnds. The software library is publicly available in https://github.com/HERA-Team/uvtools.

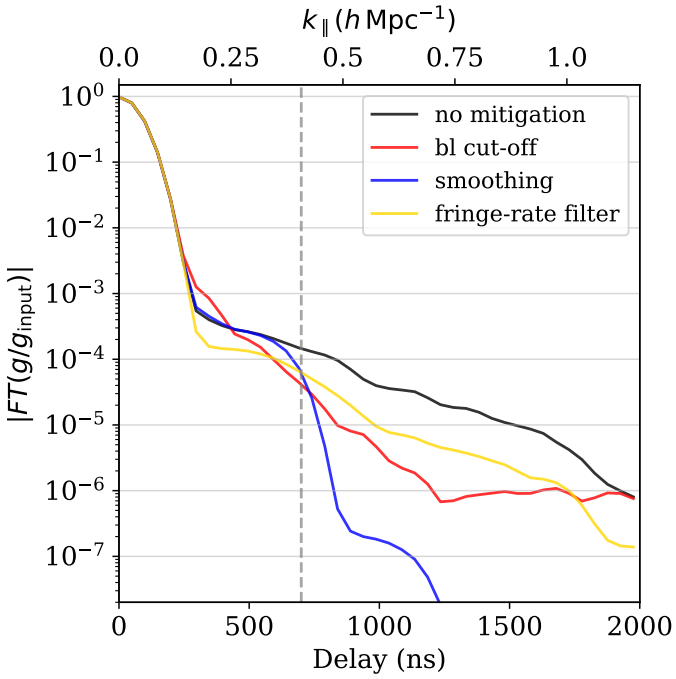

Figure 4.10: Delay spectra with the three gain error mitigation techniques are contrasted to the delay spectrum before mitigation is applied (black curve). Without loss of generality, only the results for vertical feed perturbations are shown. The delay spectra at LST = 1 hr are averaged over unflagged antennas. All different strategies provide mitigated delay spectrum results but in different shapes. See the main context for more details.

on chromatic gain errors with various filter sizes. A smaller or larger than the optimal filter size returns a larger calibration bias.

Figure 4.9 summarizes results of the fringe-rate filtering mitigation. The strong bulge observed at around LST of 1 hr in Figure 4.3 has almost vanished, which indicates the effectiveness of the fringe-rate filter in suppressing the emission from the galactic plane lying on the horizon. Compared to the baseline cut-off mitigation, the frequency structure at ∼700 ns is less effectively reduced. However, the delay spectrum at low delay modes, ∼300–500 ns, is a few factor lower than the case without any mitigation, which is not accomplished by other mitigation methods.

As different mitigation techniques result in different mitigated gain results, we anticipate combining the mitigation methods would improve the overall gain solution, as discussed in the following section.

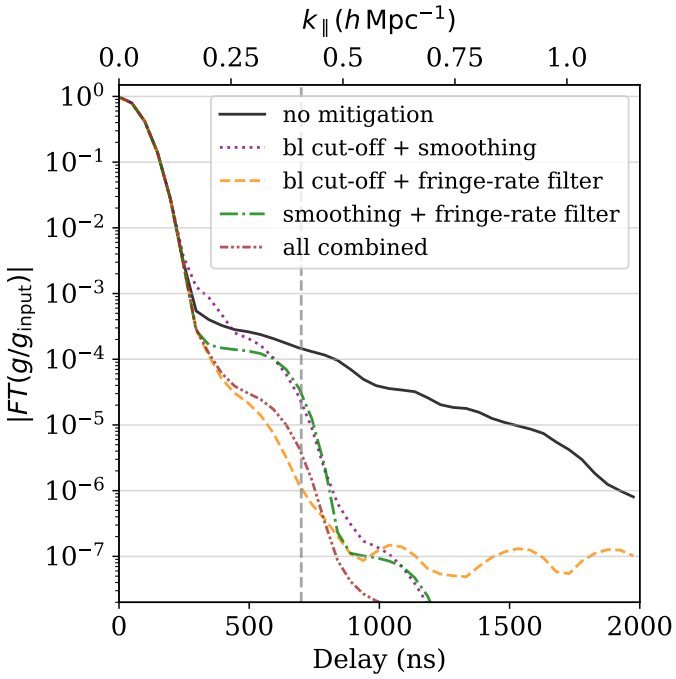

Figure 4.11: Delay spectra with different combinations of mitigation strategies. The delay spectra are shown for the vertical feed motion with a sky at LST = 1 hr. The most effective method (orange curve) is the combination of baseline cut-off and fringe-rate filtering. It brings the suppression to 10^{-6} at $k_\parallel \sim 0.4\,h\,\mathrm{Mpc}^{-1}$ to satisfy the EoR detection requirements.

4.4.4 Summary of Mitigation

In this section, we summarize the three mitigation methods and discuss their hybrid applications. Figure 4.10 characterizes the result of each mitigation at LST = 1 hr, showing the delay spectra with suppressed power at high delay modes. Without loss of generality, only the vertical feed perturbation cases are represented, as all feed motions share a similar trend. All mitigation techniques alleviate the chromatic gain errors but in different ways. The baseline cut-off mitigation that down-weights long baselines (> 87.6 m) in calibration mainly reduces the fine-frequency structure induced by the baselines, and thus suppresses the delay spectrum at high delays. Meanwhile, the fringe-rate notch filtering, excluding diffuse galactic emission located at nearly zero fringe rates across all delays to 2000 ns, drops the delay spectrum by at least a few factors beyond 300 ns. The smoothing that rules out fine spectral structure with the low-pass frequency filter brings a rapid fall of the delay spectrum above 700 ns, but has little improvement below 500 ns.

Since different mitigation approaches alleviate the chromatic gain errors in different ways, a combination of the mitigation techniques can provide a more effective way to control the leakage. In Figure 4.11, four different combinations of the methods are presented: baseline cut-off + smoothing (purple), baseline cut-off + fringe-rate filtering (orange), smoothing +

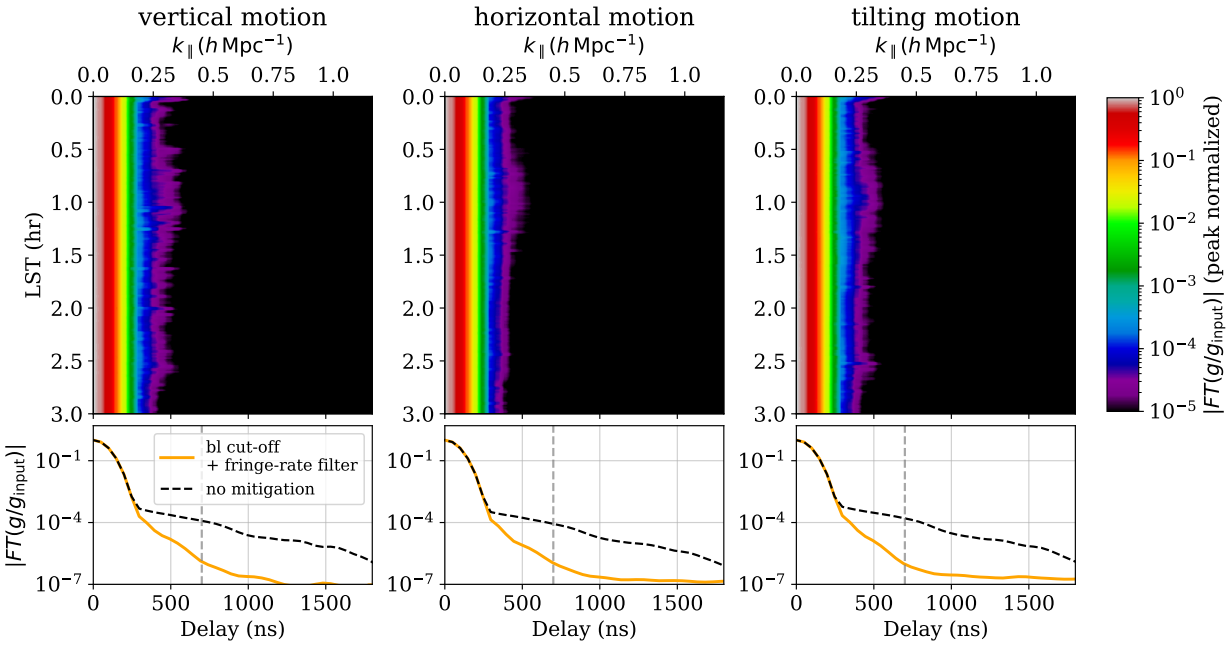

Figure 4.12: Delay spectrum of the gain solutions with both baseline cut-off and fringe-rate filtering mitigation applied. The leakage due to chromatic gain errors shown in Figure 4.3 retreats to $k_\parallel \sim 0.3\,h\,\mathrm{Mpc}^{-1}$ over all LSTs, which indicates that gain errors are suppressed enough to recover the EoR window on the power spectrum.

fringe-rate filtering (green), and all three combined method (brown). Either the baseline cut-off or the fringe-rate filtering method together with smoothing results in a fast drop at delays > 700 ns, but there is little improvement at delays < 500 ns compared to the method alone.

The combination of baseline cut-off and fringe-rate filtering is distinct from the smoothing joint cases. Limiting the baselines to short ones leads to less nonredundancy errors from point sources and applying the fringe-rate notch filter reduces the nonredundancy caused by the diffuse emission. As a result, this combination accounts for the nonredundancy effects for both types of sources, and substantially suppresses the spectral artifacts arising from calibration. The orange line in Figure 4.11, with the baseline cut-off $= 87.6$ m and $f_w = 0.3$ mHz for the fringe-rate filter, shows that the delay spectrum is significantly reduced across all delay modes, and can reach down to 10^{-6} at $k_\parallel \sim 0.4\,h\,\mathrm{Mpc}^{-1}$, which in return helps recover the clean EoR window similarly to the fiducial one. The fully combined case also reduces the delay spectrum considerably but there is a bump artifact created by the low-pass filter for the smoothing at 500 ns.

Figure 4.12 presents the delay spectrum along LST when the baseline cut-off and fringe-rate notch filtering are both applied. Across all LSTs, the bleeding gain errors shown in Figure 4.3 are reduced significantly, restrained to $k_\parallel \sim 0.3\,h\,\mathrm{Mpc}^{-1}$ for all three types of

feed perturbations, which offers at least two orders of magnitude in improvement at $k_\parallel \sim 0.4\,h\,\mathrm{Mpc}^{-1}$.

So far, we have mainly focused on the mitigation effect in terms of the gain solutions. The gain solutions will be applied to the observed visibility, and calibration errors will propagate to the calibrated visibility. One way to assess the calibration errors imprinted in the visibility is to compute the power spectrum. In addition, our test is limited to a certain size of feed perturbations: $\sigma_{\mathrm{feed}} = 2$ cm for axial and lateral feed motions and $\sigma_{\mathrm{feed}} = 2^\circ$ for feed tilting. For a comprehensive understanding on the effectiveness of the mitigation, we test various sizes of feed perturbations and discuss the mitigation effect on the power spectrum in Section 4.5.

4.5 Power Spectrum Estimation

In Section 4.3 and 4.4, we introduced the delay spectrum of gain solutions. Equivalently, the visibility delay spectrum can be produced by performing the frequency Fourier transform on the calibrated visibilities,

$$\tilde{V}(\mathbf{u},\tau) = \int d\nu\, w(\nu) V(\mathbf{u},\nu) e^{2\pi i \nu \tau}, \tag{4.3}$$

where $\mathbf{u} = \mathbf{b}/\lambda$ and $w(\nu)$ is the seven-term Blackman-Harris tapering window function. According to Equation (4.3), for a given baseline, the smooth frequency spectrum of the foreground confines its delay spectrum to near zero delay modes with a width limited by the horizon limit, $\tau_{\mathrm{hor}} = |\mathbf{b}|/c$ (Datta et al., 2010, Liu et al., 2014, Morales et al., 2012, Morales et al., 2018, Parsons et al., 2012, Pober et al., 2014, Thyagarajan et al., 2013, Trott et al., 2012, Vedantham et al., 2012). In contrast, the 21-cm cosmological signal with fluctuating spectral structure allows its delay spectrum to spread beyond the horizon limit.

4.5.1 2D Power Spectrum Estimation

The 2D power spectrum, or cylindrically averaged power spectrum, is a tool to separate the cosmological signals from the foregrounds. It is defined by cosmological Fourier modes perpendicular and parallel to the line-of-sight, $k_\perp$ and $k_\parallel$, which are linearly scaled from $|\mathbf{u}|$ and τ. More specifically, the relationship is

$$k_\perp = \frac{2\pi|\mathbf{u}|}{X},\ k_\parallel = \frac{2\pi\tau}{Y}, \tag{4.4}$$

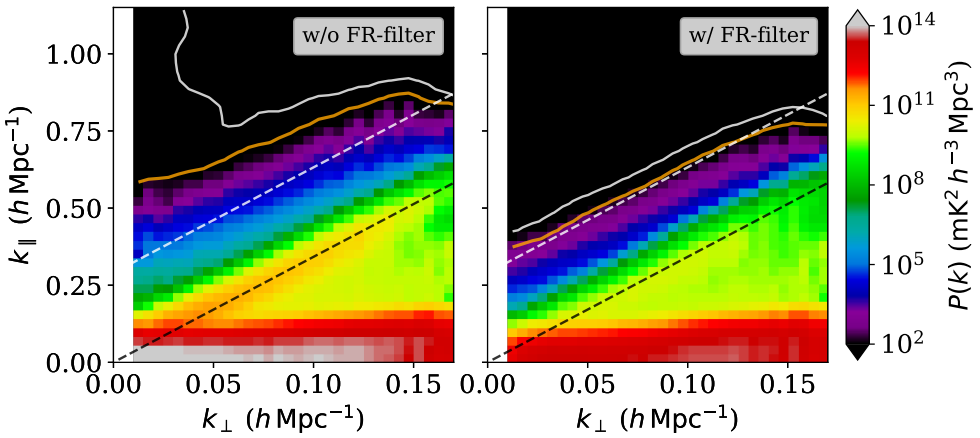

Figure 4.13: Foreground power spectrum estimation with the fiducial beam model before (left) and after (right) the fringe-rate notch filter. Orange and white contours represent the EoR power equal to and 5 times larger than the foreground power, respectively. The pitchfork along the horizon limit (black dashed line) shown in the left panel is suppressed in the right panel after the fringe-rate notch filtering. As a result, the spillover of the foregrounds is constrained inside the buffer line (white dashed line), offering a room for the EoR detection above the white dashed line. Note that the fringe-rate filter here is not being applied to visiblities prior to calibration but applied to visibilities after calibration prior to forming the power spectrum merely to suppress the pitchfork.

where $X = D(z)$, $Y = c(1+z)^2/(\nu_0 H(z))$, $D(z)$ is the comoving distance, and $H(z)$ is the Hubble parameter. With increasing $k_\perp$, $k_{\parallel,\mathrm{hor}}$ corresponding to τ_{hor} also increases, forming the foreground wedge below the horizon limit $k_{\parallel,\mathrm{hor}} = \frac{H(z)D(z)}{c(1+z)} k_\perp$.

The 2D power spectrum estimate is calculated from the delay spectrum as follows,

$$\hat{P}(k_\perp, k_\parallel) = \frac{X^2 Y}{B_{\mathrm{pp}} \Omega_{\mathrm{pp}}} |\tilde{V}(\mathbf{u}, \tau)|^2, \tag{4.5}$$

where $B_{\mathrm{pp}} = \int d\nu \, |w(\nu)|^2$ and Ω_{pp} is the spatial integral of the squared primary beam (Parsons et al., 2014). The complex visibility delay spectra were averaged over redundant baselines and fed into Equation (4.5). We implemented the power spectrum calculation via the publicly available software, `hera_pspec`[5].

Figure 4.13 shows the 2D foreground power spectrum with the fiducial beam model. The power spectrum is averaged over LST of 0–3 hr. The foregrounds are indicated in the color bar. We also over-plotted the cosmological signal in contours, the orange and white solid contours corresponding to the region that the EoR power is equal to and 5 times larger than the foregrounds, respectively. The black dashed line indicates the horizon limit, and the white dashed line is for a 500-ns buffer above the horizon limit.

[5] https://github.com/HERA-Team/hera_pspec

Because the fiducial case does not suffer from calibration errors, in principle the foreground should not leak into high $k_\parallel$ modes beyond the horizon limit. In the left panel, we see the foreground power still spills into the EoR window above the horizon limit, which is different from the leakage due to calibration errors. This is because of the pitchfork effect, along the horizon limit line, that is convolved with the primary beam. The pitchfork and associated spillover can be reduced by the fringe-rate notch filter (Kern et al., 2020a,b, Thyagarajan et al., 2016) and thus we applied the fringe-rate notch filter with $f_w = 0.3$ mHz to calibrated visibilities.

In addition, as illustrated in Figure 4.8, the EoR signals with short east-west baselines are influenced by a significant signal loss. Therefore, in forming the power spectrum we included only baselines longer than 15 m in the east-west direction for both foreground and EoR power spectra. See Kern et al. 2021 (HERA memo #96[6]) for more discussion on the signal loss due to the fringe-rate notch filter. In the right panel of Figure 4.13, after the fringe-rate filtering, the bright foreground along the black dashed line, the pitchfork, is suppressed and as a result, the intrinsic spillover above the white dashed line is also reduced. In the following section, the power spectra with perturbed beams are also results of the fringe-rate filtering to suppress the pitchfork.

4.5.2 Power Spectrum Estimation with Mitigation

As discussed above, in the ideal case when there is no calibration bias, the foreground is confined to the foreground wedge. However, when there are chromatic gain errors due to per-antenna perturbations, the convolution of gain errors with the observed visibility in Fourier space can result in foreground leakage on the power spectrum, analogous to Figure 4.3. In the top panels of Figure 4.14, the resultant power spectra without applying any mitigation strategy, averaged across 0–3 hr, are shown. All feed motions display the corrupted EoR window with the detection border line (orange contour) pushed beyond $k_\parallel \sim 0.7\,h\,\mathrm{Mpc}^{-1}$, losing potential detectability at low k modes where relatively higher sensitivity can be obtained when the thermal noise is included. The foreground leakage is predominantly derived by the 0.5–1.5 hr gain corruption shown in Figure 4.3.

The middle panels show the case for the mitigation with baseline cut-off that has been explored by previous studies (Ewall-Wice et al., 2017, Orosz et al., 2019). We set 87.6 m cut-off for the mitigation. The foreground corruption in the EoR window is largely restrained compared to no mitigation case, which is consistent with the finding of Orosz et al. (2019). One notable thing is that the low $k_\perp$ modes are still corrupted by the foreground leakage, making the EoR detection difficult at low k modes without further suppression of the foreground contamination. For the tilting motion, the mitigation is less effective than other

[6] http://reionization.org/science/memos/

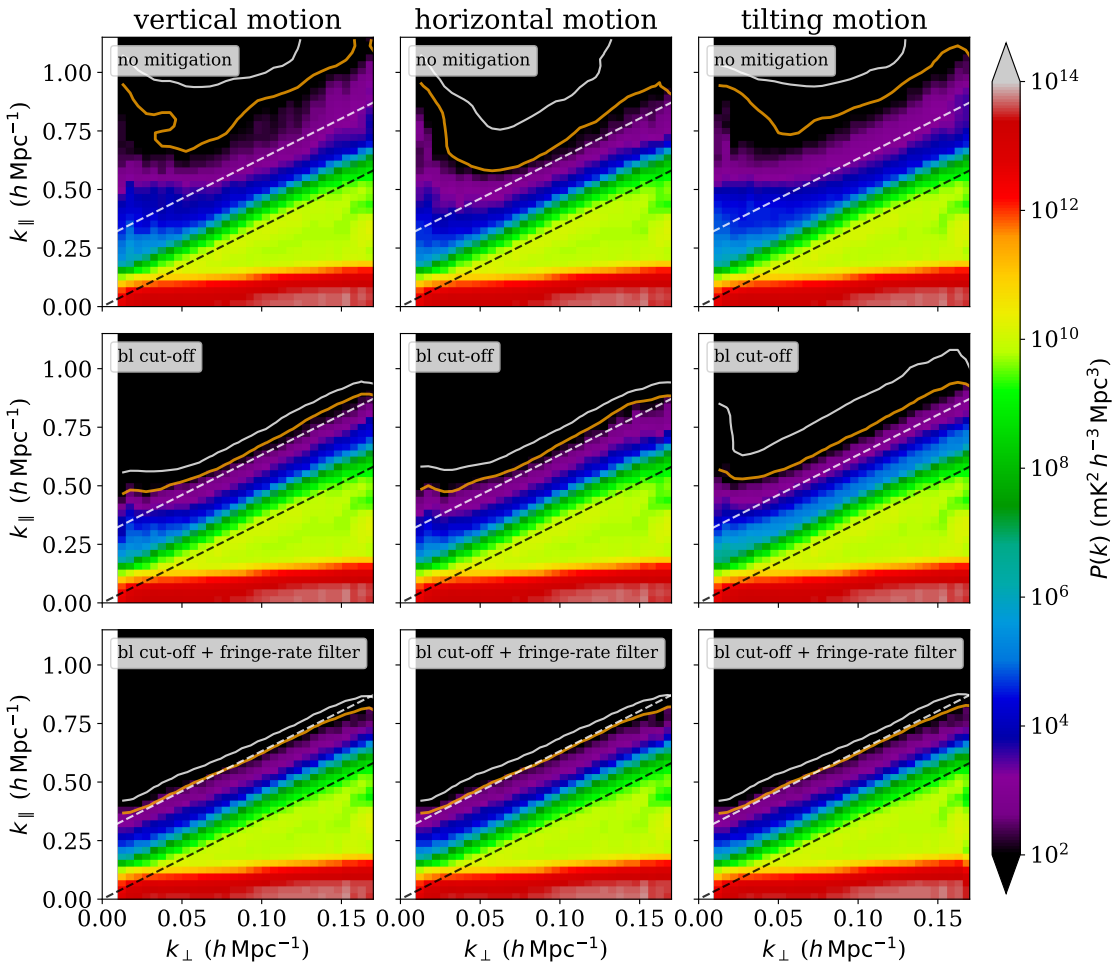

Figure 4.14: Foreground power spectrum estimation for perturbed beams without and with the mitigation strategies. Top row shows the foreground power spectrum for the vertical (left column), horizontal (middle column) and tilting (right column) motions before the mitigation on the chromatic gain errors. The foreground leakage contaminates the cosmological modes that are available for the EoR detection above the buffer line (white dashed line). The middle row presents the results after the baseline cut-off mitigation with 87.6 m cut-off. The mitigation reduces the leakage to some extent, but the low Fourier modes are still corrupted by the foreground contamination. In the bottom row, the combined method of the baseline cut-off and fringe-rate filtering leads to a considerable reduction of the foreground leakage for all types of feed perturbations and recovers the clean EoR window similar to the fiducial power spectrum shown in Figure 4.13.

motions as the gain errors due to the tip-tilt primarily derive from diffuse emission.

With the combined strategy of the baseline cut-off and fringe-rate filtering, we showed that chromatic gain errors are significantly improved. In the bottom panels of Figure 4.14, the results with the mitigated gains are given. Regardless of the feed motions, the mitigation method removes the foreground contamination and achieve as clean a EoR window as the fiducial model, enabling EoR detectability at $k_\parallel = 0.3$–$0.5\,h\,\mathrm{Mpc}^{-1}$ where the cosmological

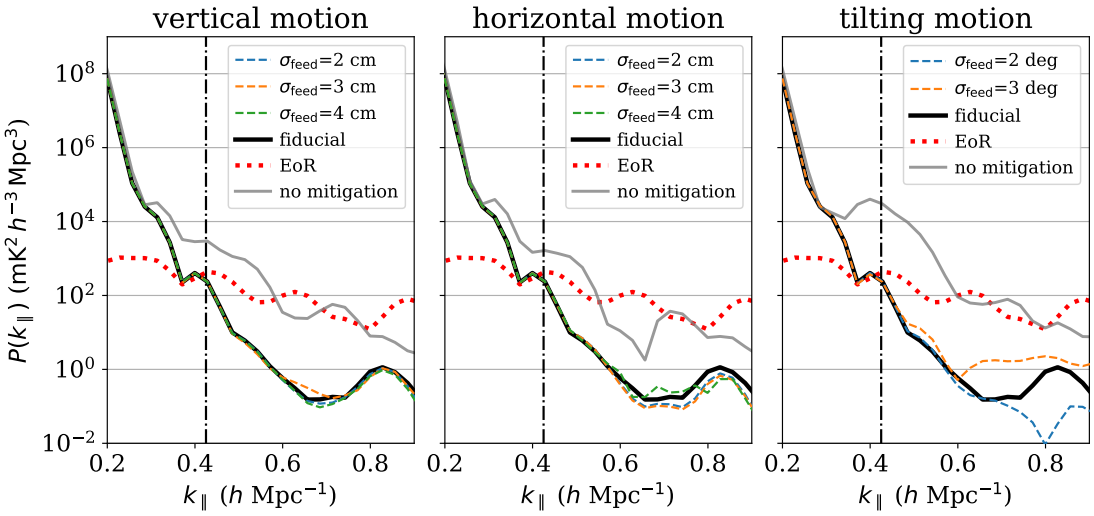

Figure 4.15: 1D power spectra along $k_\parallel$ at $k_\perp = 0.04\,h\,\mathrm{Mpc}^{-1}$ for the foregrounds (dashed and solid lines) and the EoR (red dotted line). The vertical line indicates the 500-ns buffer line. We present the case without mitigation for $\sigma_{\rm feed} = 2$ cm or 2° (gray solid line) that corrupts k modes in the EoR window by the foreground residual. The combined mitigation method with the baseline cut-off and fringe-rate filtering makes the perturbed power spectra comparable to the fiducial one, and leave the EoR detection feasible beyond the buffer line, regardless of the sizes of perturbations.

signal detection may be feasible with a high sensitivity.

Paper I shows that larger feed motions cause larger calibration errors, and thus larger amounts of foreground leakage. With the mitigation technique, we examined how the foreground contamination can be alleviated for the larger perturbation cases. Figure 4.15 presents 1D power spectra along $k_\parallel$ at $k_\perp = 0.04\,h\,\mathrm{Mpc}^{-1}$, averaged over 0–3 hr. From left to right, results for vertical, horizontal, and tilting motions are given. For the vertical and horizontal feed motions, $\sigma_{\rm feed} = 2$, 3, and 4 cm models are considered and for the tip-tilt, $\sigma_{\rm feed} = 2°$ and 3° models are used. For comparison, we also show the case without mitigation with $\sigma_{\rm feed} = 2$ cm and 2° for the translation and tilting motion, respectively. The vertical line indicates the 500-ns buffer, $k_{\parallel,\rm buffer} = k_{\parallel,\rm hor} + 0.29\ h\,\mathrm{Mpc}^{-1}$.

The foreground power spectrum with no mitigation (gray solid line) shrinks available k modes for the EoR detection compared to the fiducial case (black solid line). We found that the joint mitigation method with the baseline cut-off and the fringe-rate filtering (dashed lines) makes the foreground power in the EoR window comparable to the fiducial model, regardless of the types of feed motions and the perturbation sizes. The excess EoR power spectrum beyond the buffer line relative to the foregrounds means the EoR detection is free from the foreground bias. However, we noticed that as the perturbation size increases, the calibration pipeline tends to fail to converge to the global minimum because of large nonredundancies, ending up flagging more data to keep well-calibrated results. The models with $\sigma_{\rm feed} > 2$ cm or 2° result in the unstable calibration solutions, which may indicate the

feed motion should be constrained to $\sigma_{\rm feed} = 2$ cm or 2° perturbations for better calibration performance.

4.6 Conclusions

Feed motions cause a deviation of the primary beam shape from the fiducial one. The per-antenna perturbations from the feed motions cause nonredundancies in measured visibilities and introduce artificial spectral structure in the calibration gain solutions from the redundant and absolute calibration, employed by HERA, propagating to the power spectrum estimation. The perturbation in the beam shape and the resultant chromatic gain errors are discussed in Paper I.

To extend the work of Paper I, we simulated longer LST mock observations covering 0–3 hr including the GLEAM sky model and the GSM that represent point sources and diffuse emission, respectively. Analogous to Paper I, we separated the feed motions into the vertical, horizontal, and tilting motions. All feed motions introduce artificial structure in the calibration gains compared to the fiducial model, but with different overall aspects for different feed motions as shown in Figure 4.2. It can be understood in terms of the different behaviors of the primary beam induced by each feed motion. In addition, fine spectral and temporal gain structure is produced for all feed motions. The gain delay spectrum shows that the chromatic gain errors corrupt the delay modes $k_\parallel > 0.2\,h\,{\rm Mpc}^{-1}$, forming a bulge at LST = 0.5–1.5 hr where the galactic plane are located near the horizon (Figure 4.3). Mitigating the chromatic gain errors is the main goal of this study: attaining $\sim 10^{-5}$ suppression of the gain delay spectrum at ≤ 700 ns (or $k_\parallel \leq 0.4\,h\,{\rm Mpc}^{-1}$).

We have explored three different mitigation techniques: placing a limit on the maximum baseline length used for the calibration, smoothing the gain structure along the frequency axis after the calibration, and applying the fringe-rate notch filter on the measured visibilities prior to the calibration. All mitigation methods reduce the fine spectral artifacts to some extent but the aspects of the improvement are different. The baseline cut-off method constrains the nonredundancy arising from long baselines and results in a reduction in the fine-frequency artifacts at > 400 ns. The gain delay spectrum is suppressed to $\sim 5 \times 10^{-5}$ at 700 ns, which is still a bit higher than desired. The smoothing with the low-pass frequency filter effectively smooths out the small scale frequency features, presenting the gain delay spectrum with a rapid drop after $\sim$700 ns. Nonetheless, it has little effect at delays < 500 ns. The fringe-rate notch filtering improves the calibration process by eliminating the diffuse horizon emission, one of the main nonredundancy sources, from the visibility measurements. It leads to the delay spectrum mitigated by a few factor across all delay modes > 300 ns, which is still not sufficient to acheive the required suppression.

Because each method alone is not enough to accomplish the required suppression of the chromatic gain errors, joint mitigation strategies are also examined. The combinations between the smoothing and other methods only improve the delay spectrum beyond 700 ns and low delay modes have little improvement. When the baseline cut-off method is joint with the fringe-rate filtering, there is substantial improvement. Because the baseline cut-off and the fringe-rate filtering strategies are mainly responsible for reducing the nonredundancies from the point sources and diffuse horizon emission, respectively, the joint method addresses both nonredundancy effects. It brings the delay spectrum down to $\sim 10^{-6}$ at 700 ns that satisfies the need for suppression of gain errors for EoR detection.

As the chromatic gain errors propagate to the power spectrum estimate, we see that the EoR window is corrupted by the foreground leakage in the 2D power spectra when there is no mitigation for all feed motions. If the baseline cut-off mitigation is applied as Ewall-Wice et al. (2017) and Orosz et al. (2019) did, the foreground contamination is suppressed to some degree but some contamination still remains in the low k modes. The combined method of the baseline cut-off and the fringe-rate filtering clears away the contamination in the EoR window and arrives at a power spectrum similar to the fiducial one with perfect calibration.

In Section 4.3–4.5, we focused on the modest feed motions by $\sigma_{\text{feed}} = 2$ cm and 2° for the translation motions and the tip-tilts, respectively. We also investigated the effectiveness of the mitigation technique with larger perturbation sizes. We found that the 1D power spectra with the combined mitigation of the baseline cut-off and fringe-rate filtering present the results similar to the fiducial one even for $\sigma_{\text{feed}} = 4$ cm or 3°, demonstrating that the mitigation technique is effective enough to suppress the chromatic gain errors for larger perturbations. However, we also noticed that the larger nonredundancies due to the feed motion may lead to failure in finding appropriate gain solutions, which implies we need to constrain the feed motions to $\sigma_{\text{feed}} = 2$ cm and 2°.

In this study, we only focused on the middle band corresponding to the redshift of ~ 7. HERA is also designed to study the Cosmic Dawn at lower frequencies. The low band with a wider primary beam may bring about a larger per-antenna perturbation than the middle band for the same feed position deviations. Hence, it is important to test the mitigation scheme for the low band as well to set a requirement for feed positioning with a more comprehensive analysis. We defer this to future work.

Chapter 5

Conclusion

In this chapter, I provide a brief summary of the thesis work and discuss the novel contributions to the research field.

5.1 Summary

This thesis presents my contributions to observational 21 cm cosmology using the HERA radio interferometer. The primary focus of this work is on investigating the impact of instrument systematics on our statistical measurements, specifically the power spectrum and one-point statistics. The one-point statistics measured from the HERA Phase I observations examine how systematics affect the Gaussian and non-Gaussian features in the images. Using future Phase II mock observations, I investigate how systematics, stemming from per-antenna perturbations, impact data calibration and power spectrum estimation. I also devise methods to mitigate calibration errors and reduce foreground contamination in the power spectrum analysis.

Chapter 2 describes the results of one-point statistics measured from the HERA Phase I observational data. First, in order to remove the bright foreground emission, I evaluate the effectiveness of three different foreground removal methods using simulation data: `DAYENU`-filtering, standard wedge-masking, and PCA-based foreground subtraction. The first two techniques employ foreground avoidance, while the last one uses foreground subtraction in image space. I find that the `DAYENU`-filtering method surpasses the other methods in minimizing the amount of foreground residuals in the map.

I then apply the wedge-filtering (`DAYENU`) foreground removal technique to the observational data and produce images by using the direct optimal mapping, which optimally

down-weights the data with large noise variance, suitable in detecting faint cosmological signals. To infer statistical properties from the data, I use forward modeling to simulate mock observations of thermal noise and the 21 cm signal. This approach takes into account pixel correlations introduced by the instrument response and frequency correlations induced by the wedge-filtering. One-point statistics, such as the second moment (m_2), third moment (m_3) and standardized third moment (S_3), representing variance and skewness, are computed for all datasets and compared. The instrument response substantially diminishes the amplitude of the one-point statistics measurements. This reduction is further compounded, particularly in m_3 and S_3, by wedge-filtering, making it difficult to probe non-Gaussianity. Nevertheless, there may remain an increasing trend in S_3 toward lower redshifts.

The comparison of the observational data to the theoretical expectation shows that the foreground-removed maps are limited by systematics rather than by noise, despite efforts in systematics mitigation during the observational data processing. The likelihood analysis, given the observational data, is likely to reject the cold reionization model characterized by the inefficient X-ray heating, which is consistent with the earlier power spectrum studies (The HERA Collaboration et al., 2022b, 2023).

In the chapter, I also include forecasts on the detectability of non-Gaussianity for future observations with the full HERA core array. I conclude that the S/N for the perfect foreground removal, assuming no signal loss of 21 cm signals, reach about 6 for m_2 and 2 for m_3. This S/N is limited by cosmic variance rather than thermal noise. However, once wedge-filtering is applied, the S/N drops significantly to below 1, indicating potential difficulties in measuring non-Gaussianity. Despite this, the Fisher information matrix, which combines information from all frequency bins, suggests that m_3, along with m_2, may provide a better constraint on the EoR model than m_2 alone.

In Chapter 3, I investigate the calibration errors induced by per-antenna perturbations, specifically feed misalignment, a potential systematic issue that HERA may encounter. Previous studies (e.g., Ewall-Wice et al., 2017, Orosz et al., 2019) have explored the effects of per-antenna perturbations using airy beams. However, in practice, the primary beams are more complex than the simple airy beam models, making them insufficient to capture the detailed behavior of the HERA system. To address this, I use an electromagnetic simulator (CST) to create more realistic beam models by perturbing the feed positions. These models are then used to generate mock Phase II observations for foregrounds and 21 cm signals. I consider three different directions of the feed misalignment: vertical, horizontal, and tip-tiling displacement. Additionally, I separately account for point-source and diffuse foregrounds, as they each contribute differently to calibration errors. By varying the size of the perturbations, I examine their impact on calibration errors and how these errors propagate to the power spectrum calculation. I conclude that stringent constraints on feed misalignment are necessary ($\sigma_{\rm feed} \leq 1$ cm or $\sigma_{\rm feed} \leq 1^\circ$) in the absence of mitigation efforts.

Finally, Chapter 4 addresses various strategies to minimize the calibration errors. I employ three different techniques including restricting long baselines during calibration, smoothing gain solutions post-calibration, and applying a temporal filter (fringe-rate notch filter) prior to calibration. Through a combination of these techniques, calibration errors are substantially reduced, allowing recovery of a clean EoR window. The findings indicate that the acceptable tolerance for feed positioning can be relaxed to $\sigma_{\text{feed}} \sim 2$ cm or 2°.

5.2 Novel Contributions

In Chapter 2, I present the results for one-point statistics. This study is unique compared to previous ones because it: 1) employs a robust foreground-filtering method, 2) considers the real instrument response or PSF when forming mock observations of noise and 21 cm simulations, and 3) forward models each data set to account for correlation effects in reconstructed maps. Additionally, this work is a milestone as it is the first to measure m_2 and m_3 from HERA Phase I observations, which represent the deepest available data targeting high redshift. Although the results were limited by systematics and thermal noise, this work offers a direction for future observations to study the non-Gaussianity of 21 cm signals.

The systematics we observe in the one-point statistics may arise from calibration errors, inducing foreground leakage into the EoR window. In Chapter 3, I showcase the possible scenarios of per-antenna perturbations by using physically motivated feed models, alerting the stringent requirements of feed positioning to minimize corruption of the EoR window.

In Chapter 4, I demonstrate the effectiveness of the fringe-rate filtering in minimizing systematics. While the fringe-rate notch filter has often been used for crosstalk mitigation post-calibration, this is the first time it has been used to address nonredundancies from per-antenna perturbations prior to calibration. If visibility data contain nonredundancies due to feed displacement or other factors, the fringe-rate notch filter, combined with the baseline cut-off method, is expected to significantly improve redundant calibration performance and reduce the effects of systematics in future observations. Recently, Charles et al. (2024, in prep) show that nonredundancies caused by mutual coupling between antennas can be mitigated by applying the fringe-rate notch filter prior to calibration.

In the study of feed displacement, I conducted a large number of visibility simulations across different frequencies, times, and baselines for each set of perturbed feed models. Due to the high computational cost, using public visibility simulators designed for general use was not feasible. Consequently, I developed a new visibility simulator called `VisPB`[1]. This simulator is optimized for rapid visibility calculations, particularly for arrays with redundant

[1] https://github.com/vispb/vispb

layouts and perturbed antenna primary beams. For instance, to streamline the process, I avoided redundant calculations for baselines with identical configurations by computing the unchanging parts of the visibility calculation for one baseline and reusing them for the others. In addition, I vectorized the calculations as much as possible and parallelized computationally intensive tasks, such as beam interpolations at sky positions and fringe term calculations over baselines. The HERA collaboration is actively investigating the mutual coupling effects between antennas in the Phase II system. Since mutual coupling causes perturbations in beam shapes, `VisPB` may contribute to facilitating the calculations of these effects on the visibility within a reasonable timeframe.

References

David Alonso, Philip Bull, Pedro G. Ferreira, and Mário G. Santos. Blind foreground subtraction for intensity mapping experiments. *Monthly Notices of the Royal Astronomical Society*, 447(1):400–416, 12 2014. ISSN 0035-8711. doi:10.1093/mnras/stu2474. URL https://doi.org/10.1093/mnras/stu2474.

Jacob Baars. *The Paraboloidal Reflector Antenna in Radio Astronomy and Communication: Theory and Practice*. 01 2007. ISBN 978-0-387-69733-8. doi:10.1007/978-0-387-69734-5.

N. Barry, B. Hazelton, I. Sullivan, M. F. Morales, and J. C. Pober. Calibration requirements for detecting the 21 cm epoch of reionization power spectrum and implications for the SKA. *Monthly Notices of the Royal Astronomical Society*, 461(3):3135–3144, September 2016. doi:10.1093/mnras/stw1380.

N. Barry, M. Wilensky, C. M. Trott, B. Pindor, A. P. Beardsley, B. J. Hazelton, I. S. Sullivan, M. F. Morales, J. C. Pober, J. Line, B. Greig, R. Byrne, A. Lanman, W. Li, C. H. Jordan, R. C. Joseph, B. McKinley, M. Rahimi, S. Yoshiura, J. D. Bowman, B. M. Gaensler, J. N. Hewitt, D. C. Jacobs, D. A. Mitchell, N. Udaya Shankar, S. K. Sethi, R. Subrahmanyan, S. J. Tingay, R. L. Webster, and J. S. B. Wyithe. Improving the epoch of reionization power spectrum results from murchison widefield array season 1 observations. *The Astrophysical Journal*, 884(1):1, oct 2019. doi:10.3847/1538-4357/ab40a8. URL https://doi.org/10.3847/1538-4357/ab40a8.

Nichole Barry and Aman Chokshi. The role of the instrumental response in 21 cm epoch of reionization power spectrum gridding analyses. *Astrophysical Journal*, 929(1):64, apr 2022. doi:10.3847/1538-4357/ac5903. URL https://doi.org/10.3847/1538-4357/ac5903.

Neil Bassett, David Rapetti, Keith Tauscher, Bang D. Nhan, David D. Bordenave, Joshua J. Hibbard, and Jack O. Burns. Lost horizon: Quantifying the effect of local topography on global 21 cm cosmology data analysis. *Astrophysical Journal*, 923(1):33, dec 2021. doi:10.3847/1538-4357/ac1cde. URL https://doi.org/10.3847/1538-4357/ac1cde.

A. P. Beardsley, B. J. Hazelton, I. S. Sullivan, P. Carroll, N. Barry, M. Rahimi, B. Pindor, C. M. Trott, J. Line, Daniel C. Jacobs, M. F. Morales, J. C. Pober, G. Bernardi, Judd D. Bowman, M. P. Busch, F. Briggs, R. J. Cappallo, B. E. Corey, A. de Oliveira-Costa, Joshua S. Dillon, D. Emrich, A. Ewall-Wice, L. Feng, B. M. Gaensler, R. Goeke, L. J. Greenhill, J. N. Hewitt, N. Hurley-Walker, M. Johnston-Hollitt, D. L. Kaplan, J. C.

Kasper, H. S. Kim, E. Kratzenberg, E. Lenc, A. Loeb, C. J. Lonsdale, M. J. Lynch, B. McKinley, S. R. McWhirter, D. A. Mitchell, E. Morgan, A. R. Neben, Nithyanandan Thyagarajan, D. Oberoi, A. R. Offringa, S. M. Ord, S. Paul, T. Prabu, P. Procopio, J. Riding, A. E. E. Rogers, A. Roshi, N. Udaya Shankar, Shiv K. Sethi, K. S. Srivani, R. Subrahmanyan, M. Tegmark, S. J. Tingay, M. Waterson, R. B. Wayth, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, C. Wu, and J. S. B. Wyithe. FIRST SEASON MWA EOR POWER SPECTRUM RESULTS AT REDSHIFT 7. *The Astrophysical Journal*, 833(1):102, dec 2016. doi:10.3847/1538-4357/833/1/102. URL https://doi.org/10.3847/1538-4357/833/1/102.

G. Bernardi, L. J. Greenhill, D. A. Mitchell, S. M. Ord, B. J. Hazelton, B. M. Gaensler, A. de Oliveira-Costa, M. F. Morales, N. Udaya Shankar, R. Subrahmanyan, R. B. Wayth, E. Lenc, C. L. Williams, W. Arcus, B. S. Arora, D. G. Barnes, J. D. Bowman, F. H. Briggs, J. D. Bunton, R. J. Cappallo, B. E. Corey, A. Deshpande, L. deSouza, D. Emrich, R. Goeke, D. Herne, J. N. Hewitt, M. Johnston-Hollitt, D. Kaplan, J. C. Kasper, B. B. Kincaid, R. Koenig, E. Kratzenberg, C. J. Lonsdale, M. J. Lynch, S. R. McWhirter, E. Morgan, D. Oberoi, J. Pathikulangara, T. Prabu, R. A. Remillard, A. E. E. Rogers, A. Roshi, J. E. Salah, R. J. Sault, K. S. Srivani, J. Stevens, S. J. Tingay, M. Waterson, R. L. Webster, A. R. Whitney, A. Williams, and J. S. B. Wyithe. A 189 MHz, 2400 deg^2 Polarization Survey with the Murchison Widefield Array 32-element Prototype. *Astrophysical Journal*, 771(2):105, July 2013. doi:10.1088/0004-637X/771/2/105.

Judd D. Bowman, Alan E. E. Rogers, Raul A. Monsalve, Thomas J. Mozdzen, and Nivedita Mahesh. An absorption profile centred at 78 megahertz in the sky-averaged spectrum. *Nature*, 555(4):67–70, 03 2018. ISSN 1476-4687. doi:10.1038/nature25792. URL https://doi.org/10.1038/nature25792.

Ruby Byrne, Miguel F. Morales, Bryna Hazelton, Wenyang Li, Nichole Barry, Adam P. Beardsley, Ronniy Joseph, Jonathan Pober, Ian Sullivan, and Cathryn Trott. Fundamental Limitations on the Calibration of Redundant 21 cm Cosmology Instruments and Implications for HERA and the SKA. *Astrophysical Journal*, 875(1):70, April 2019. doi:10.3847/1538-4357/ab107d.

Ntsikelelo Charles, Nicholas Kern, Gianni Bernardi, Landman Bester, Oleg Smirnov, Nicolas Fagnoni, and Eloy de Lera Acedo. On the use of temporal filtering for mitigating galactic synchrotron calibration bias in 21 cm reionization observations. *arXiv e-prints*, art. arXiv:2302.00269, February 2023. doi:10.48550/arXiv.2302.00269.

Carina Cheng, Aaron R. Parsons, Matthew Kolopanis, Daniel C. Jacobs, Adrian Liu, Saul A. Kohn, James E. Aguirre, Jonathan C. Pober, Zaki S. Ali, Gianni Bernardi, Richard F. Bradley, Chris L. Carilli, David R. DeBoer, Matthew R. Dexter, Joshua S. Dillon, Pat Klima, David H. E. MacMahon, David F. Moore, Chuneeta D. Nunhokee, William P. Walbrugh, and Andre Walker. Characterizing signal loss in the 21 cm reionization power spectrum: A revised study of PAPER-64. *The Astrophysical Journal*, 868(1):26, nov 2018. doi:10.3847/1538-4357/aae833. URL https://doi.org/10.3847/1538-4357/aae833.

Samir Choudhuri, Philip Bull, and Hugh Garsden. Patterns of primary beam non-redundancy in close-packed 21 cm array observations. *Monthly Notices of the Royal Astronomical Society*, 506(2):2066–2088, 06 2021. ISSN 0035-8711. doi:10.1093/mnras/stab1795. URL https://doi.org/10.1093/mnras/stab1795.

J. J. Condon. Confusion and Flux-Density Error Distributions. *Astrophysical Journal*, 188: 279–286, March 1974. doi:10.1086/152714.

Tim J. Cornwell. Multiscale clean deconvolution of radio synthesis images. *IEEE Journal of Selected Topics in Signal Processing*, 2(5):793–801, 2008. doi:10.1109/JSTSP.2008.2006388.

Tim J. Cornwell, Kumar Golap, and Sanjay Bhatnagar. The noncoplanar baselines effect in radio interferometry: The w-projection algorithm. *IEEE Journal of Selected Topics in Signal Processing*, 2(5):647–657, 2008. doi:10.1109/JSTSP.2008.2005290.

Steven Cunnington, Yichao Li, Mario G Santos, Jingying Wang, Isabella P Carucci, Melis O Irfan, Alkistis Pourtsidou, Marta Spinelli, Laura Wolz, Paula S Soares, Chris Blake, Philip Bull, Brandon Engelbrecht, José Fonseca, Keith Grainge, and Yin-Zhe Ma. Hi intensity mapping with MeerKAT: power spectrum detection in cross-correlation with WiggleZ galaxies. *Monthly Notices of the Royal Astronomical Society*, 518(4):6262–6272, 10 2022. ISSN 0035-8711. doi:10.1093/mnras/stac3060. URL https://doi.org/10.1093/mnras/stac3060.

A. Datta, J. D. Bowman, and C. L. Carilli. Bright Source Subtraction Requirements for Redshifted 21 cm Measurements. *Astrophysical Journal*, 724(1):526–538, November 2010. doi:10.1088/0004-637X/724/1/526.

Kanan K. Datta, Garrelt Mellema, Yi Mao, Ilian T. Iliev, Paul R. Shapiro, and Kyungjin Ahn. Light-cone effect on the reionization 21-cm power spectrum. *Monthly Notices of the Royal Astronomical Society*, 424(3):1877–1891, 08 2012. ISSN 0035-8711. doi:10.1111/j.1365-2966.2012.21293.x. URL https://doi.org/10.1111/j.1365-2966.2012.21293.x.

David R. DeBoer, Aaron R. Parsons, James E. Aguirre, Paul Alexander, Zaki S. Ali, Adam P. Beardsley, Gianni Bernardi, Judd D. Bowman, Richard F. Bradley, Chris L. Carilli, Carina Cheng, Eloy de Lera Acedo, Joshua S. Dillon, Aaron Ewall-Wice, Gcobisa Fadana, Nicolas Fagnoni, Randall Fritz, Steve R. Furlanetto, Brian Glendenning, Bradley Greig, Jasper Grobbelaar, Bryna J. Hazelton, Jacqueline N. Hewitt, Jack Hickish, Daniel C. Jacobs, Austin Julius, MacCalvin Kariseb, Saul A. Kohn, Telalo Lekalake, Adrian Liu, Anita Loots, David MacMahon, Lourence Malan, Cresshim Malgas, Matthys Maree, Zachary Martinot, Nathan Mathison, Eunice Matsetela, Andrei Mesinger, Miguel F. Morales, Abraham R. Neben, Nipanjana Patra, Samantha Pieterse, Jonathan C. Pober, Nima Razavi-Ghods, Jon Ringuette, James Robnett, Kathryn Rosie, Raddwine Sell, Craig Smith, Angelo Syce, Max Tegmark, Nithyanandan Thyagarajan, Peter K. G. Williams, and Haoxuan Zheng. Hydrogen epoch of reionization array (HERA). *Publications of*

the Astronomical Society of the Pacific, 129(974):045001, mar 2017. doi:10.1088/1538-3873/129/974/045001. URL https://doi.org/10.1088/1538-3873/129/974/045001.

Joshua S. Dillon and Aaron R. Parsons. REDUNDANT ARRAY CONFIGURATIONS FOR 21 cm COSMOLOGY. *The Astrophysical Journal*, 826(2):181, jul 2016. doi:10.3847/0004-637x/826/2/181. URL https://doi.org/10.3847/0004-637x/826/2/181.

Joshua S. Dillon, Adrian Liu, Christopher L. Williams, Jacqueline N. Hewitt, Max Tegmark, Edward H. Morgan, Alan M. Levine, Miguel F. Morales, Steven J. Tingay, Gianni Bernardi, Judd D. Bowman, Frank H. Briggs, Roger C. Cappallo, David Emrich, Daniel A. Mitchell, Divya Oberoi, Thiagaraj Prabu, Randall Wayth, and Rachel L. Webster. Overcoming real-world obstacles in 21 cm power spectrum estimation: A method demonstration and results from early murchison widefield array data. *Phys. Rev. D*, 89:023002, Jan 2014. doi:10.1103/PhysRevD.89.023002. URL https://link.aps.org/doi/10.1103/PhysRevD.89.023002.

Joshua S. Dillon, Max Tegmark, Adrian Liu, Aaron Ewall-Wice, Jacqueline N. Hewitt, Miguel F. Morales, Abraham R. Neben, Aaron R. Parsons, and Haoxuan Zheng. Mapmaking for precision 21 cm cosmology. *Phys. Rev. D*, 91:023002, Jan 2015. doi:10.1103/PhysRevD.91.023002. URL https://link.aps.org/doi/10.1103/PhysRevD.91.023002.

Joshua S Dillon, Saul A Kohn, Aaron R Parsons, James E Aguirre, Zaki S Ali, Gianni Bernardi, Nicholas S Kern, Wenyang Li, Adrian Liu, Chuneeta D Nunhokee, and Jonathan C Pober. Polarized redundant-baseline calibration for 21 cm cosmology without adding spectral structure. *Monthly Notices of the Royal Astronomical Society*, 477(4):5670–5681, 04 2018. ISSN 0035-8711. doi:10.1093/mnras/sty1060. URL https://doi.org/10.1093/mnras/sty1060.

Joshua S. Dillon, Max Lee, Zaki S. Ali, Aaron R. Parsons, Naomi Orosz, Chuneeta Devi Nunhokee, Paul La Plante, Adam P. Beardsley, Nicholas S. Kern, Zara Abdurashidova, James E. Aguirre, Paul Alexander, Yanga Balfour, Gianni Bernardi, Tashalee S. Billings, Judd D. Bowman, Richard F. Bradley, Phil Bull, Jacob Burba, Steve Carey, Chris L. Carilli, Carina Cheng, David R. DeBoer, Matt Dexter, Eloy de Lera Acedo, John Ely, Aaron Ewall-Wice, Nicolas Fagnoni, Randall Fritz, Steven R. Furlanetto, Kingsley Gale-Sides, Brian Glendenning, Deepthi Gorthi, Bradley Greig, Jasper Grobbelaar, Ziyaad Halday, Bryna J. Hazelton, Jacqueline N. Hewitt, Jack Hickish, Daniel C. Jacobs, Austin Julius, Joshua Kerrigan, Piyanat Kittiwisit, Saul A. Kohn, Matthew Kolopanis, Adam Lanman, Telalo Lekalake, David Lewis, Adrian Liu, Yin-Zhe Ma, David MacMahon, Lourence Malan, Cresshim Malgas, Matthys Maree, Zachary E. Martinot, Eunice Matsetela, Andrei Mesinger, Mathakane Molewa, Miguel F. Morales, Tshegofalang Mosiane, Steven Murray, Abraham R. Neben, Bojan Nikolic, Robert Pascua, Nipanjana Patra, Samantha Pieterse, Jonathan C. Pober, Nima Razavi-Ghods, Jon Ringuette, James Robnett, Kathryn Rosie, Mario G. Santos, Peter Sims, Craig Smith, Angelo Syce, Max Tegmark, Nithyanandan Thyagarajan, Peter K. G. Williams, and Haoxuan Zheng. Redundant-baseline calibration

of the hydrogen epoch of reionization array. *Monthly Notices of the Royal Astronomical Society*, 499(4):5840–5861, December 2020. doi:10.1093/mnras/staa3001.

Douspis, M., Aghanim, N., Ili´c, S., and Langer, M. A new parameterization of the reionisation history. *A&A*, 580:L4, 2015. doi:10.1051/0004-6361/201526543. URL https://doi.org/10.1051/0004-6361/201526543.

Michael W. Eastwood, Marin M. Anderson, Ryan M. Monroe, Gregg Hallinan, Benjamin R. Barsdell, Stephen A. Bourke, M. A. Clark, Steven W. Ellingson, Jayce Dowell, Hugh Garsden, Lincoln J. Greenhill, Jacob M. Hartman, Jonathon Kocz, T. Joseph W. Lazio, Danny C. Price, Frank K. Schinzel, Gregory B. Taylor, Harish K. Vedantham, Yuankun Wang, and David P. Woody. The radio sky at meter wavelengths: m-mode analysis imaging with the ovro-lwa. *The Astronomical Journal*, 156(1):32, jun 2018. doi:10.3847/1538-3881/aac721. URL https://dx.doi.org/10.3847/1538-3881/aac721.

A. Ewall-Wice, Joshua S. Dillon, J. N. Hewitt, A. Loeb, A. Mesinger, A. R. Neben, A. R. Offringa, M. Tegmark, N. Barry, A. P. Beardsley, G. Bernardi, Judd D. Bowman, F. Briggs, R. J. Cappallo, P. Carroll, B. E. Corey, A. de Oliveira-Costa, D. Emrich, L. Feng, B. M. Gaensler, R. Goeke, L. J. Greenhill, B. J. Hazelton, N. Hurley-Walker, M. Johnston-Hollitt, Daniel C. Jacobs, D. L. Kaplan, J. C. Kasper, HS Kim, E. Kratzenberg, E. Lenc, J. Line, C. J. Lonsdale, M. J. Lynch, B. McKinley, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, Nithyanandan Thyagarajan, D. Oberoi, S. M. Ord, S. Paul, B. Pindor, J. C. Pober, T. Prabu, P. Procopio, J. Riding, A. E. E. Rogers, A. Roshi, N. Udaya Shankar, Shiv K. Sethi, K. S. Srivani, R. Subrahmanyan, I. S. Sullivan, S. J. Tingay, C. M. Trott, M. Waterson, R. B. Wayth, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, C. Wu, and J. S. B. Wyithe. First limits on the 21 cm power spectrum during the Epoch of X-ray heating. *Monthly Notices of the Royal Astronomical Society*, 460(4):4320–4347, 05 2016a. ISSN 0035-8711. doi:10.1093/mnras/stw1022. URL https://doi.org/10.1093/mnras/stw1022.

Aaron Ewall-Wice, Jacqueline Hewitt, Andrei Mesinger, Joshua S. Dillon, Adrian Liu, and Jonathan Pober. Constraining high-redshift X-ray sources with next generation 21-cm power spectrum measurements. *Monthly Notices of the Royal Astronomical Society*, 458 (3):2710–2724, 02 2016b. ISSN 0035-8711. doi:10.1093/mnras/stw452. URL https://doi.org/10.1093/mnras/stw452.

Aaron Ewall-Wice, Joshua S. Dillon, Adrian Liu, and Jacqueline Hewitt. The impact of modelling errors on interferometer calibration for 21 cm power spectra. *Monthly Notices of the Royal Astronomical Society*, 470(2):1849–1870, 05 2017. ISSN 0035-8711. doi:10.1093/mnras/stx1221. URL https://doi.org/10.1093/mnras/stx1221.

Aaron Ewall-Wice, Nicholas Kern, Joshua S Dillon, Adrian Liu, Aaron Parsons, Saurabh Singh, Adam Lanman, Paul La Plante, Nicolas Fagnoni, Eloy de Lera Acedo, David R DeBoer, Chuneeta Nunhokee, Philip Bull, Tzu-Ching Chang, T Joseph W Lazio, James Aguirre, and Sean Weinberg. DAYENU: a simple filter of smooth foregrounds for intensity

mapping power spectra. *Monthly Notices of the Royal Astronomical Society*, 500(4):5195–5213, 10 2020. ISSN 0035-8711. doi:10.1093/mnras/staa3293. URL https://doi.org/10.1093/mnras/staa3293.

Nicolas Fagnoni, Eloy de Lera Acedo, David R DeBoer, Zara Abdurashidova, James E Aguirre, Paul Alexander, Zaki S Ali, Yanga Balfour, Adam P Beardsley, Gianni Bernardi, Tashalee S Billings, Judd D Bowman, Richard F Bradley, Phil Bull, Jacob Burba, Chris L Carilli, Carina Cheng, Matt Dexter, Joshua S Dillon, Aaron Ewall-Wice, Randall Fritz, Steve R Furlanetto, Kingsley Gale-Sides, Brian Glendenning, Deepthi Gorthi, Bradley Greig, Jasper Grobbelaar, Ziyaad Halday, Bryna J Hazelton, Jacqueline N Hewitt, Jack Hickish, Daniel C Jacobs, Alec Josaitis, Austin Julius, Nicholas S Kern, Joshua Kerrigan, Honggeun Kim, Piyanat Kittiwisit, Saul A Kohn, Matthew Kolopanis, Adam Lanman, Paul La Plante, Telalo Lekalake, Adrian Liu, David MacMahon, Lourence Malan, Cresshim Malgas, Matthys Maree, Zachary E Martinot, Eunice Matsetela, Juan Mena Parra, Andrei Mesinger, Mathakane Molewa, Miguel F Morales, Tshegofalang Mosiane, Abraham R Neben, Bojan Nikolic, Aaron R Parsons, Nipanjana Patra, Samantha Pieterse, Jonathan C Pober, Nima Razavi-Ghods, James Robnett, Kathryn Rosie, Peter Sims, Craig Smith, Angelo Syce, Nithyanandan Thyagarajan, Peter K G Williams, and Haoxuan Zheng. Understanding the HERA Phase I receiver system with simulations and its impact on the detectability of the EoR delay power spectrum. *Monthly Notices of the Royal Astronomical Society*, 500(1):1232–1242, 10 2020. ISSN 0035-8711. doi:10.1093/mnras/staa3268. URL https://doi.org/10.1093/mnras/staa3268.

Nicolas Fagnoni, Eloy de Lera Acedo, Nick Drought, David R. DeBoer, Daniel Riley, Nima Razavi-Ghods, Steven Carey, and Aaron R. Parsons. Design of the new wideband vivaldi feed for the hera radio-telescope phase ii. *IEEE Transactions on Antennas and Propagation*, 69(12):8143–8157, 2021. doi:10.1109/TAP.2021.3083788.

George B. Field. Excitation of the hydrogen 21-cm line. *Proceedings of the IRE*, 46(1): 240–250, 1958. doi:10.1109/JRPROC.1958.286741.

R. A. Fisher. The logic of inductive inference. *Journal of the Royal Statistical Society*, 98 (1):39–82, 1935. ISSN 09528385. URL http://www.jstor.org/stable/2342435.

Steven R. Furlanetto, S. Peng Oh, and Frank H. Briggs. Cosmology at low frequencies: The 21cm transition and the high-redshift universe. *Physics Reports*, 433(4):181–301, 2006. ISSN 0370-1573. doi:https://doi.org/10.1016/j.physrep.2006.08.002. URL https://www.sciencedirect.com/science/article/pii/S0370157306002730.

Samuel Gagnon-Hartman, Yue Cui, Adrian Liu, and Siamak Ravanbakhsh. Recovering the wedge modes lost to 21-cm foregrounds. *Monthly Notices of the Royal Astronomical Society*, 504(4):4716–4729, 04 2021. ISSN 0035-8711. doi:10.1093/mnras/stab1158. URL https://doi.org/10.1093/mnras/stab1158.

B K Gehlot, F G Mertens, L V E Koopmans, M A Brentjens, S Zaroubi, B Ciardi, A Ghosh, M Hatef, I T Iliev, V Jelić, R Kooistra, F Krause, G Mellema, M Mevius, M Mitra,

A R Offringa, V N Pandey, A M Sardarabadi, J Schaye, M B Silva, H K Vedantham, and S Yatawatta. The first power spectrum limit on the 21-cm signal of neutral hydrogen during the Cosmic Dawn at z = 20–25 from LOFAR. *Monthly Notices of the Royal Astronomical Society*, 488(3):4271–4287, 07 2019. ISSN 0035-8711. doi:10.1093/mnras/stz1937. URL https://doi.org/10.1093/mnras/stz1937.

Bharat K. Gehlot, Daniel C. Jacobs, Judd D. Bowman, Nivedita Mahesh, Steven G. Murray, Matthew Kolopanis, Adam P. Beardsley, Zara Abdurashidova, James E. Aguirre, Paul Alexander, Zaki S. Ali, Yanga Balfour, Gianni Bernardi, Tashalee S. Billings, Richard F. Bradley, Phil Bull, Jacob Burba, Steve Carey, Chris L. Carilli, Carina Cheng, David R. DeBoer, Matt Dexter, Eloy de Lera Acedo, Joshua S. Dillon, John Ely, Aaron Ewall-Wice, Nicolas Fagnoni, Randall Fritz, Steven R. Furlanetto, Kingsley Gale-Sides, Brian Glendenning, Deepthi Gorthi, Bradley Greig, Jasper Grobbelaar, Ziyaad Halday, Bryna J. Hazelton, Jacqueline N. Hewitt, Jack Hickish, Austin Julius, Nicholas S. Kern, Joshua Kerrigan, Piyanat Kittiwisit, Saul A. Kohn, Adam Lanman, Paul La Plante, Telalo Lekalake, David Lewis, Adrian Liu, Yin-Zhe Ma, David MacMahon, Lourence Malan, Cresshim Malgas, Matthys Maree, Zachary E. Martinot, Eunice Matsetela, Andrei Mesinger, Mathakane Molewa, Raul A. Monsalve, Miguel F. Morales, Tshegofalang Mosiane, Abraham R. Neben, Bojan Nikolic, Aaron R. Parsons, Robert Pascua, Nipanjana Patra, Samantha Pieterse, Jonathan C. Pober, Nima Razavi-Ghods, Jon Ringuette, James Robnett, Kathryn Rosie, Mario G. Santos, Peter Sims, Craig Smith, Angelo Syce, Max Tegmark, Nithyanandan Thyagarajan, Peter K. G. Williams, and Haoxuan Zheng. Effects of model incompleteness on the drift-scan calibration of radio telescopes. *Monthly Notices of the Royal Astronomical Society*, 506(3):4578–4592, September 2021. doi:10.1093/mnras/stab2072.

Vera Gluscevic and Rennan Barkana. Statistics of 21-cm fluctuations in cosmic reionization simulations: PDFs and difference PDFs. *Monthly Notices of the Royal Astronomical Society*, 408(4):2373–2380, 10 2010. ISSN 0035-8711. doi:10.1111/j.1365-2966.2010.17293.x. URL https://doi.org/10.1111/j.1365-2966.2010.17293.x.

K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. HEALPix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere. *Astrophysical Journal*, 622(2):759–771, apr 2005. doi:10.1086/427976. URL https://doi.org/10.1086/427976.

Bradley Greig and Andrei Mesinger. The global history of reionization. *Monthly Notices of the Royal Astronomical Society*, 465(4):4838–4852, 11 2016. ISSN 0035-8711. doi:10.1093/mnras/stw3026. URL https://doi.org/10.1093/mnras/stw3026.

Bradley Greig, Yuan-Sen Ting, and Alexander A Kaurov. Exploring the cosmic 21-cm signal from the epoch of reionization using the wavelet scattering transform. *Monthly Notices of the Royal Astronomical Society*, 513(2):1719–1741, 04 2022. ISSN 0035-8711. doi:10.1093/mnras/stac977. URL https://doi.org/10.1093/mnras/stac977.

Bradley Greig, Yuan-Sen Ting, and Alexander A Kaurov. Detecting the non-Gaussianity

of the 21-cm signal during reionization with the wavelet scattering transform. *Monthly Notices of the Royal Astronomical Society*, 519(4):5288–5303, 03 2023. ISSN 0035-8711. doi:10.1093/mnras/stac3822. URL https://doi.org/10.1093/mnras/stac3822.

Geraint J. A. Harker, Saleem Zaroubi, Rajat M. Thomas, Vibor Jelić, Panagiotis Labropoulos, Garrelt Mellema, Ilian T. Iliev, Gianni Bernardi, Michiel A. Brentjens, A. G. De Bruyn, Benedetta Ciardi, Leon V. E. Koopmans, V. N. Pandey, Andreas H. Pawlik, Joop Schaye, and Sarod Yatawatta. Detection and extraction of signals from the epoch of reionization using higher-order one-point statistics. *Monthly Notices of the Royal Astronomical Society*, 393(4):1449–1458, 02 2009. ISSN 0035-8711. doi:10.1111/j.1365-2966.2008.14209.x. URL https://doi.org/10.1111/j.1365-2966.2008.14209.x.

C. G. T. Haslam, C. J. Salter, H. Stoffel, and W. E. Wilson. A 408-MHZ All-Sky Continuum Survey. II. The Atlas of Contour Maps. , 47:1, January 1982.

J. A. Högbom. Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines. , 15:417, June 1974.

N. Hurley-Walker, J. R. Callingham, P. J. Hancock, T. M. O. Franzen, L. Hindson, A. D. Kapińska, J. Morgan, A. R. Offringa, R. B. Wayth, C. Wu, Q. Zheng, T. Murphy, M. E. Bell, K. S. Dwarakanath, B. For, B. M. Gaensler, M. Johnston-Hollitt, E. Lenc, P. Procopio, L. Staveley-Smith, R. Ekers, J. D. Bowman, F. Briggs, R. J. Cappallo, A. A. Deshpande, L. Greenhill, B. J. Hazelton, D. L. Kaplan, C. J. Lonsdale, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, D. Oberoi, S. M. Ord, T. Prabu, N. Udaya Shankar, K. S. Srivani, R. Subrahmanyan, S. J. Tingay, R. L. Webster, A. Williams, and C. L. Williams. GaLactic and Extragalactic All-sky Murchison Widefield Array (GLEAM) survey - I. A low-frequency extragalactic catalogue. *Monthly Notices of the Royal Astronomical Society*, 464(1):1146–1167, January 2017. doi:10.1093/mnras/stw2337.

N. Hurley-Walker, P. J. Hancock, T. M. O. Franzen, J. R. Callingham, A. R. Offringa, L. Hindson, C. Wu, M. E. Bell, B. Q. For, B. M. Gaensler, M. Johnston-Hollitt, A. D. Kapińska, J. Morgan, T. Murphy, B. McKinley, P. Procopio, L. Staveley-Smith, R. B. Wayth, and Q. Zheng. GaLactic and Extragalactic All-sky Murchison Widefield Array (GLEAM) survey II: Galactic plane $345^\circ < l < 67^\circ$, $180^\circ < l < 240^\circ$. *Publications of the Astronomical Society of Australia*, 36:e047, November 2019. doi:10.1017/pasa.2019.37.

Anne Hutter, Catherine A Watkinson, Jacob Seiler, Pratika Dayal, Manodeep Sinha, and Darren J Croton. The 21cm bispectrum during reionization: a tracer of the ionization topology. *Monthly Notices of the Royal Astronomical Society*, 492(1):653–667, 11 2019. ISSN 0035-8711. doi:10.1093/mnras/stz3139. URL https://doi.org/10.1093/mnras/stz3139.

Alec T Josaitis, Aaron Ewall-Wice, Nicolas Fagnoni, and Eloy de Lera Acedo. Array element coupling in radio interferometry I: a semi-analytic approach. *Monthly Notices of the Royal Astronomical Society*, 514(2):1804–1827, 04 2022. ISSN 0035-8711. doi:10.1093/mnras/stac916. URL https://doi.org/10.1093/mnras/stac916.

Mohd Kamran, Raghunath Ghara, Suman Majumdar, Rajesh Mondal, Garrelt Mellema, Somnath Bharadwaj, Jonathan R Pritchard, and Ilian T Iliev. Redshifted 21-cm bispectrum – II. Impact of the spin temperature fluctuations and redshift space distortions on the signal from the Cosmic Dawn. *Monthly Notices of the Royal Astronomical Society*, 502(3):3800–3813, 02 2021. ISSN 0035-8711. doi:10.1093/mnras/stab216. URL https://doi.org/10.1093/mnras/stab216.

N. Kern, G. Bernardi, and Carilli C. HERA MEMORANDUM #42: ABSOLUTE CALIBRATION OF HERA-47 IDR1. Technical report, Department of Astronomy, University of California, Berkeley, CA, January 2018. HERA Memo #42.

Nicholas S Kern and Adrian Liu. Gaussian process foreground subtraction and power spectrum estimation for 21cm cosmology. *Monthly Notices of the Royal Astronomical Society*, 501(1):1463–1480, 12 2020. ISSN 0035-8711. doi:10.1093/mnras/staa3736. URL https://doi.org/10.1093/mnras/staa3736.

Nicholas S. Kern, Aaron R. Parsons, Joshua S. Dillon, Adam E. Lanman, Nicolas Fagnoni, and Eloy de Lera Acedo. Mitigating internal instrument coupling for 21 cm cosmology. i. temporal and spectral modeling in simulations. *The Astrophysical Journal*, 884(2):105, oct 2019. doi:10.3847/1538-4357/ab3e73. URL https://dx.doi.org/10.3847/1538-4357/ab3e73.

Nicholas S. Kern, Joshua S. Dillon, Aaron R. Parsons, Christopher L. Carilli, Gianni Bernardi, Zara Abdurashidova, James E. Aguirre, Paul Alexander, Zaki S. Ali, Yanga Balfour, Adam P. Beardsley, Tashalee S. Billings, Judd D. Bowman, Richard F. Bradley, Philip Bull, Jacob Burba, Steven Carey, Carina Cheng, David R. DeBoer, Matt Dexter, Eloy de Lera Acedo, John Ely, Aaron Ewall-Wice, Nicolas Fagnoni, Randall Fritz, Steve R. Furlanetto, Kingsley Gale-Sides, Brian Glendenning, Deepthi Gorthi, Bradley Greig, Jasper Grobbelaar, Ziyaad Halday, Bryna J. Hazelton, Jacqueline N. Hewitt, Jack Hickish, Daniel C. Jacobs, Austin Julius, Joshua Kerrigan, Piyanat Kittiwisit, Saul A. Kohn, Matthew Kolopanis, Adam Lanman, Paul La Plante, Telalo Lekalake, Adrian Liu, David MacMahon, Lourence Malan, Cresshim Malgas, Matthys Maree, Zachary E. Martinot, Eunice Matsetela, Andrei Mesinger, Mathakane Molewa, Miguel F. Morales, Tshegofalang Mosiane, Steven G. Murray, Abraham R. Neben, Bojan Nikolic, Chuneeta D. Nunhokee, Nipanjana Patra, Samantha Pieterse, Jonathan C. Pober, Nima Razavi-Ghods, Jon Ringuette, James Robnett, Kathryn Rosie, Peter Sims, Craig Smith, Angelo Syce, Nithyanandan Thyagarajan, Peter K. G. Williams, and Haoxuan Zheng. Absolute calibration strategies for the hydrogen epoch of reionization array and their impact on the 21 cm power spectrum. *The Astrophysical Journal*, 890(2):122, feb 2020a. doi:10.3847/1538-4357/ab67bc. URL https://doi.org/10.3847/1538-4357/ab67bc.

Nicholas S. Kern, Aaron R. Parsons, Joshua S. Dillon, Adam E. Lanman, Adrian Liu, Philip Bull, Aaron Ewall-Wice, Zara Abdurashidova, James E. Aguirre, Paul Alexander, Zaki S. Ali, Yanga Balfour, Adam P. Beardsley, Gianni Bernardi, Judd D. Bowman, Richard F. Bradley, Jacob Burba, Chris L. Carilli, Carina Cheng, David R. DeBoer, Matt Dexter, Eloy de Lera Acedo, Nicolas Fagnoni, Randall Fritz, Steve R. Furlanetto, Brian Glendenning,

Deepthi Gorthi, Bradley Greig, Jasper Grobbelaar, Ziyaad Halday, Bryna J. Hazelton, Jacqueline N. Hewitt, Jack Hickish, Daniel C. Jacobs, Austin Julius, Joshua Kerrigan, Piyanat Kittiwisit, Saul A. Kohn, Matthew Kolopanis, Paul La Plante, Telalo Lekalake, David MacMahon, Lourence Malan, Cresshim Malgas, Matthys Maree, Zachary E. Martinot, Eunice Matsetela, Andrei Mesinger, Mathakane Molewa, Miguel F. Morales, Tshegofalang Mosiane, Steven G. Murray, Abraham R. Neben, Aaron R. Parsons, Nipanjana Patra, Samantha Pieterse, Jonathan C. Pober, Nima Razavi-Ghods, Jon Ringuette, James Robnett, Kathryn Rosie, Peter Sims, Craig Smith, Angelo Syce, Nithyanandan Thyagarajan, Peter K. G. Williams, and Haoxuan Zheng. Mitigating internal instrument coupling for 21 cm cosmology. II. a method demonstration with the hydrogen epoch of reionization array. *The Astrophysical Journal*, 888(2):70, jan 2020b. doi:10.3847/1538-4357/ab5e8a. URL https://doi.org/10.3847/1538-4357/ab5e8a.

Honggeun Kim, Bang D. Nhan, Jacqueline N. Hewitt, Nicholas S. Kern, Joshua S. Dillon, Eloy de Lera Acedo, Scott B. C. Dynes, Nivedita Mahesh, Nicolas Fagnoni, and David R. DeBoer. The impact of beam variations on power spectrum estimation for 21 cm cosmology. i. simulations of foreground contamination for hera. *The Astrophysical Journal*, 941(2):207, dec 2022. doi:10.3847/1538-4357/ac9eaf. URL https://dx.doi.org/10.3847/1538-4357/ac9eaf.

Honggeun Kim, Nicholas S. Kern, Jacqueline N. Hewitt, Bang D. Nhan, Joshua S. Dillon, Eloy de Lera Acedo, Scott B. C. Dynes, Nivedita Mahesh, Nicolas Fagnoni, and David R. DeBoer. The impact of beam variations on power spectrum estimation for 21 cm cosmology. ii. mitigation of foreground systematics for hera. *The Astrophysical Journal*, 953(2):136, aug 2023. doi:10.3847/1538-4357/ace35e. URL https://dx.doi.org/10.3847/1538-4357/ace35e.

Piyanat Kittiwisit, Judd D Bowman, Daniel C Jacobs, Adam P Beardsley, and Nithyanandan Thyagarajan. Sensitivity of the Hydrogen Epoch of Reionization Array and its build-out stages to one-point statistics from redshifted 21cm observations. *Monthly Notices of the Royal Astronomical Society*, 474(4):4487–4499, 12 2017. ISSN 0035-8711. doi:10.1093/mnras/stx3099. URL https://doi.org/10.1093/mnras/stx3099.

Piyanat Kittiwisit, Judd D Bowman, Steven G Murray, Bharat K Gehlot, Daniel C Jacobs, and Adam P Beardsley. Measurements of one-point statistics in 21-cm intensity maps via foreground avoidance strategy. *Monthly Notices of the Royal Astronomical Society*, 517(2):2138–2150, 10 2022. ISSN 0035-8711. doi:10.1093/mnras/stac2826. URL https://doi.org/10.1093/mnras/stac2826.

Saul A. Kohn, James E. Aguirre, Paul La Plante, Tashalee S. Billings, Paul M. Chichura, Austin F. Fortino, Amy S. Igarashi, Roshan K. Benefo, Samavarti Gallardo, Zachary E. Martinot, Chuneeta D. Nunhokee, Nicholas S. Kern, Philip Bull, Adrian Liu, Paul Alexander, Zaki S. Ali, Adam P. Beardsley, Gianni Bernardi, Judd D. Bowman, Richard F. Bradley, Chris L. Carilli, Carina Cheng, David R. DeBoer, Eloy de Lera Acedo, Joshua S. Dillon, Aaron Ewall-Wice, Gcobisa Fadana, Nicolas Fagnoni, Randall Fritz, Steven R.

Furlanetto, Brian Glendenning, Bradley Greig, Jasper Grobbelaar, Bryna J. Hazelton, Jacqueline N. Hewitt, Jack Hickish, Daniel C. Jacobs, Austin Julius, MacCalvin Kariseb, Matthew Kolopanis, Telalo Lekalake, Anita Loots, David MacMahon, Lourence Malan, Cresshim Malgas, Matthys Maree, Nathan Mathison, Eunice Matsetela, Andrei Mesinger, Miguel F. Morales, Abraham R. Neben, Bojan Nikolic, Aaron R. Parsons, Nipanjana Patra, Samantha Pieterse, Jonathan C. Pober, Nima Razavi-Ghods, Jon Ringuette, James Robnett, Kathryn Rosie, Raddwine Sell, Craig Smith, Angelo Syce, Max Tegmark, Nithyanandan Thyagarajan, Peter K. G. Williams, and Haoxuan Zheng. The HERA-19 commissioning array: Direction-dependent effects. *The Astrophysical Journal*, 882(1):58, sep 2019. doi:10.3847/1538-4357/ab2f72. URL https://doi.org/10.3847/1538-4357/ab2f72.

Matthew Kolopanis, Daniel C. Jacobs, Carina Cheng, Aaron R. Parsons, Saul A. Kohn, Jonathan C. Pober, James E. Aguirre, Zaki S. Ali, Gianni Bernardi, Richard F. Bradley, Chris L. Carilli, David R. DeBoer, Matthew R. Dexter, Joshua S. Dillon, Joshua Kerrigan, Pat Klima, Adrian Liu, David H. E. MacMahon, David F. Moore, Nithyanandan Thyagarajan, Chuneeta D. Nunhokee, William P. Walbrugh, and Andre Walker. A simplified, lossless reanalysis of PAPER-64. *The Astrophysical Journal*, 883(2):133, sep 2019. doi:10.3847/1538-4357/ab3e3a. URL https://doi.org/10.3847/1538-4357/ab3e3a.

Leon Koopmans, J Pritchard, G Mellema, J Aguirre, K Ahn, R Barkana, I van Bemmel, G Bernardi, A Bonaldi, F Briggs, A. G. de Bruyn, T. C. Chang, E Chapman, X Chen, B Courty, P. Dayal, A. Ferrara, A. Fialkov, F Fiore, K. Ichiki, I. T. Illiev, S Inoue, V Jelic, M Jones, J Lazio, Umberto Maio, S Majumdar, K. J. Mack, A. Mesinger, M F. Morales, A. Parsons, U.L. Pen, M Santos, R Schneider, B Semelin, R S de Souza, R Subrahmanyan, T Takeuchi, H Vedantham, J Wagg, R Webster, S Wyithe, Kanan Kumar Datta, and C. Trott. The Cosmic Dawn and Epoch of Reionisation with SKA. *PoS*, AASKA14:001, 2015. doi:10.22323/1.215.0001.

Kenji Kubota, Shintaro Yoshiura, Hayato Shimabukuro, and Keitaro Takahashi. Expected constraints on models of the epoch of reionization with the variance and skewness in redshifted 21cm-line fluctuations. *Publications of the Astronomical Society of Japan*, 68 (4):61, 06 2016. ISSN 0004-6264. doi:10.1093/pasj/psw059. URL https://doi.org/10.1093/pasj/psw059.

Adam E. Lanman, Jonathan C. Pober, Nicholas S. Kern, Eloy de Lera Acedo, David R. DeBoer, and Nicolas Fagnoni. Quantifying EoR delay spectrum contamination from diffuse radio emission. *Monthly Notices of the Royal Astronomical Society*, 494(3):3712–3727, May 2020. doi:10.1093/mnras/staa987.

W. Li, J. C. Pober, B. J. Hazelton, N. Barry, M. F. Morales, I. Sullivan, A. R. Parsons, Z. S. Ali, J. S. Dillon, A. P. Beardsley, J. D. Bowman, F. Briggs, R. Byrne, P. Carroll, B. Crosse, D. Emrich, A. Ewall-Wice, L. Feng, T. M. O. Franzen, J. N. Hewitt, L. Horsley, D. C. Jacobs, M. Johnston-Hollitt, C. Jordan, R. C. Joseph, D. L. Kaplan, D. Kenney, H. Kim, P. Kittiwisit, A. Lanman, J. Line, B. McKinley, D. A. Mitchell, S. Murray, A. Neben, A. R. Offringa, D. Pallot, S. Paul, B. Pindor, P. Procopio, M. Rahimi, J. Riding, S. K.

Sethi, N. Udaya Shankar, K. Steele, R. Subrahmanian, M. Tegmark, N. Thyagarajan, S. J. Tingay, C. Trott, M. Walker, R. B. Wayth, R. L. Webster, A. Williams, C. Wu, and S. Wyithe. Comparing redundant and sky-model-based interferometric calibration: A first look with phase II of the MWA. *The Astrophysical Journal*, 863(2):170, aug 2018. doi:10.3847/1538-4357/aad3c3. URL https://doi.org/10.3847/1538-4357/aad3c3.

A Liu, M Tegmark, S Morrison, A Lutomirski, and M Zaldarriaga. Precision calibration of radio interferometers using redundant baselines. *Monthly Notices of the Royal Astronomical Society*, 408:1029–1050, oct 2010. doi:10.1111/j.1365-2966.2010.17174.x.

Adrian Liu and Aaron R. Parsons. Constraining cosmology and ionization history with combined 21cm power spectrum and global signal measurements. *Monthly Notices of the Royal Astronomical Society*, 457(2):1864–1877, 02 2016. ISSN 0035-8711. doi:10.1093/mnras/stw071. URL https://doi.org/10.1093/mnras/stw071.

Adrian Liu and J. Richard Shaw. Data analysis for precision 21 cm cosmology. *Publications of the Astronomical Society of the Pacific*, 132(1012):062001, apr 2020. doi:10.1088/1538-3873/ab5bfd. URL https://doi.org/10.1088/1538-3873/ab5bfd.

Adrian Liu, Aaron R. Parsons, and Cathryn M. Trott. Epoch of reionization window. i. mathematical formalism. *Phys. Rev. D*, 90:023018, Jul 2014. doi:10.1103/PhysRevD.90.023018. URL https://link.aps.org/doi/10.1103/PhysRevD.90.023018.

Suman Majumdar, Jonathan R Pritchard, Rajesh Mondal, Catherine A Watkinson, Somnath Bharadwaj, and Garrelt Mellema. Quantifying the non-Gaussianity in the EoR 21-cm signal through bispectrum. *Monthly Notices of the Royal Astronomical Society*, 476(3): 4007–4024, 02 2018. ISSN 0035-8711. doi:10.1093/mnras/sty535. URL https://doi.org/10.1093/mnras/sty535.

Suman Majumdar, Mohd Kamran, Jonathan R Pritchard, Rajesh Mondal, Arindam Mazumdar, Somnath Bharadwaj, and Garrelt Mellema. Redshifted 21-cm bispectrum – I. Impact of the redshift space distortions on the signal from the Epoch of Reionization. *Monthly Notices of the Royal Astronomical Society*, 499(4):5090–5106, 10 2020. ISSN 0035-8711. doi:10.1093/mnras/staa3168. URL https://doi.org/10.1093/mnras/staa3168.

Matthew McQuinn. The evolution of the intergalactic medium. *Annual Review of Astronomy and Astrophysics*, 54(1):313–362, 2016. doi:10.1146/annurev-astro-082214-122355. URL https://doi.org/10.1146/annurev-astro-082214-122355.

F. G. Mertens, A. Ghosh, and L. V. E. Koopmans. Statistical 21-cm signal separation via Gaussian Process Regression analysis. *Monthly Notices of the Royal Astronomical Society*, 478(3):3640–3652, Aug 2018. doi:10.1093/mnras/sty1207.

F G Mertens, M Mevius, L V E Koopmans, A R Offringa, G Mellema, S Zaroubi, M A Brentjens, H Gan, B K Gehlot, V N Pandey, A M Sardarabadi, H K Vedantham, S Yatawatta, K M B Asad, B Ciardi, E Chapman, S Gazagnes, R Ghara, A Ghosh, S K Giri, I T Iliev, V Jelić, R Kooistra, R Mondal, J Schaye, and M B Silva. Improved upper limits on the 21cm signal power spectrum of neutral hydrogen at z 9.1 from LOFAR. *Monthly*

Notices of the Royal Astronomical Society, 493(2):1662–1685, 02 2020. ISSN 0035-8711. doi:10.1093/mnras/staa327. URL https://doi.org/10.1093/mnras/staa327.

Andrei Mesinger, Steven Furlanetto, and Renyue Cen. 21cmfast: a fast, seminumerical simulation of the high-redshift 21-cm signal. *Monthly Notices of the Royal Astronomical Society*, 411(2):955–972, 02 2011. ISSN 0035-8711. doi:10.1111/j.1365-2966.2010.17731.x. URL https://doi.org/10.1111/j.1365-2966.2010.17731.x.

Andrei Mesinger, Aaron Ewall-Wice, and Jacqueline Hewitt. Reionization and beyond: detecting the peaks of the cosmological 21cm signal. *Monthly Notices of the Royal Astronomical Society*, 439(4):3262–3274, 02 2014. ISSN 0035-8711. doi:10.1093/mnras/stu125. URL https://doi.org/10.1093/mnras/stu125.

Andrei Mesinger, Bradley Greig, and Emanuele Sobacchi. The Evolution Of 21 cm Structure (EOS): public, large-scale simulations of Cosmic Dawn and reionization. *Monthly Notices of the Royal Astronomical Society*, 459(3):2342–2353, 04 2016. ISSN 0035-8711. doi:10.1093/mnras/stw831. URL https://doi.org/10.1093/mnras/stw831.

Marius Millea and François Bouchet. Cosmic microwave background constraints in light of priors over reionization histories. *A&A*, 617:A96, 2018. doi:10.1051/0004-6361/201833288. URL https://doi.org/10.1051/0004-6361/201833288.

Sourav Mitra, T. Roy Choudhury, and Andrea Ferrara. Cosmic reionization after Planck. *Monthly Notices of the Royal Astronomical Society: Letters*, 454(1):L76–L80, 09 2015. ISSN 1745-3925. doi:10.1093/mnrasl/slv134. URL https://doi.org/10.1093/mnrasl/slv134.

Rajesh Mondal, Garrelt Mellema, Abinash Kumar Shaw, Mohd Kamran, and Suman Majumdar. The Epoch of Reionization 21-cm bispectrum: the impact of light-cone effects and detectability. *Monthly Notices of the Royal Astronomical Society*, 508(3):3848–3859, 10 2021. ISSN 0035-8711. doi:10.1093/mnras/stab2900. URL https://doi.org/10.1093/mnras/stab2900.

Miguel F. Morales, Bryna Hazelton, Ian Sullivan, and Adam Beardsley. Four Fundamental Foreground Power Spectrum Shapes for 21 cm Cosmology Observations. *Astrophysical Journal*, 752(2):137, June 2012. doi:10.1088/0004-637X/752/2/137.

Miguel F Morales, Adam Beardsley, Jonathan Pober, Nichole Barry, Bryna Hazelton, Daniel Jacobs, and Ian Sullivan. Understanding the diversity of 21 cm cosmology analyses. *Monthly Notices of the Royal Astronomical Society*, 483(2):2207–2216, 10 2018. ISSN 0035-8711. doi:10.1093/mnras/sty2844. URL https://doi.org/10.1093/mnras/sty2844.

A Mouri Sardarabadi and L V E Koopmans. Quantifying suppression of the cosmological 21-cm signal due to direction-dependent gain calibration in radio interferometers. *Monthly Notices of the Royal Astronomical Society*, 483(4):5480–5490, 12 2018. ISSN 0035-8711. doi:10.1093/mnras/sty3444. URL https://doi.org/10.1093/mnras/sty3444.

Steven G. Murray and C. M. Trott. The effect of baseline layouts on the epoch of reionization foreground wedge: A semianalytical approach. *The Astrophysical Journal*, 869(1):25, dec 2018. doi:10.3847/1538-4357/aaebfa. URL https://dx.doi.org/10.3847/1538-4357/aaebfa.

Steven G. Murray, Bradley Greig, Andrei Mesinger, Julian B. Muñoz, Yuxiang Qin, Jaehong Park, and Catherine A. Watkinson. 21cmfast v3: A python-integrated c code for generating 3d realizations of the cosmic 21cm signal. *Journal of Open Source Software*, 5(54):2582, 2020. doi:10.21105/joss.02582. URL https://doi.org/10.21105/joss.02582.

Naomi Orosz, Joshua S. Dillon, Aaron Ewall-Wice, Aaron R. Parsons, and Nithyanandan Thyagarajan. Mitigating the effects of antenna-to-antenna variation on redundant-baseline calibration for 21 cm cosmology. *Monthly Notices of the Royal Astronomical Society*, 487 (1):537–549, July 2019. doi:10.1093/mnras/stz1287.

Gregory Paciga, Joshua G. Albert, Kevin Bandura, Tzu-Ching Chang, Yashwant Gupta, Christopher Hirata, Julia Odegova, Ue-Li Pen, Jeffrey B. Peterson, Jayanta Roy, J. Richard Shaw, Kris Sigurdson, and Tabitha Voytek. A simulation-calibrated limit on the Hi power spectrum from the GMRT Epoch of Reionization experiment. *Monthly Notices of the Royal Astronomical Society*, 433(1):639–647, 05 2013. ISSN 0035-8711. doi:10.1093/mnras/stt753. URL https://doi.org/10.1093/mnras/stt753.

Jaehong Park, Andrei Mesinger, Bradley Greig, and Nicolas Gillet. Inferring the astrophysics of reionization and cosmic dawn from galaxy luminosity functions and the 21-cm signal. *Monthly Notices of the Royal Astronomical Society*, 484(1):933–949, 01 2019. ISSN 0035-8711. doi:10.1093/mnras/stz032. URL https://doi.org/10.1093/mnras/stz032.

Aaron R. Parsons and Donald C. Backer. Calibration of Low-Frequency, Wide-Field Radio Interferometers Using Delay/Delay-Rate Filtering. *The Astronomical Journal*, 138(1): 219–226, July 2009. doi:10.1088/0004-6256/138/1/219.

Aaron R. Parsons, Donald C. Backer, Griffin S. Foster, Melvyn C. H. Wright, Richard F. Bradley, Nicole E. Gugliucci, Chaitali R. Parashare, Erin E. Benoit, James E. Aguirre, Daniel C. Jacobs, Chris L. Carilli, David Herne, Mervyn J. Lynch, Jason R. Manley, and Daniel J. Werthimer. THE PRECISION ARRAY FOR PROBING THE EPOCH OF RE-IONIZATION: EIGHT STATION RESULTS. *The Astronomical Journal*, 139(4): 1468–1480, mar 2010. doi:10.1088/0004-6256/139/4/1468. URL https://doi.org/10.1088/0004-6256/139/4/1468.

Aaron R. Parsons, Jonathan C. Pober, James E. Aguirre, Christopher L. Carilli, Daniel C. Jacobs, and David F. Moore. A PER-BASELINE, DELAY-SPECTRUM TECHNIQUE FOR ACCESSING THE 21 cm COSMIC REIONIZATION SIGNATURE. *The Astrophysical Journal*, 756(2):165, aug 2012. doi:10.1088/0004-637x/756/2/165. URL https://doi.org/10.1088/0004-637x/756/2/165.

Aaron R. Parsons, Adrian Liu, James E. Aguirre, Zaki S. Ali, Richard F. Bradley, Chris L. Carilli, David R. DeBoer, Matthew R. Dexter, Nicole E. Gugliucci, Daniel C. Jacobs, Pat Klima, David H. E. MacMahon, Jason R. Manley, David F. Moore, Jonathan C. Pober, Irina I. Stefan, and William P. Walbrugh. NEW LIMITS ON 21 cm EPOCH OF REIONIZATION FROM PAPER-32 CONSISTENT WITH AN x-RAY HEATED INTERGALACTIC MEDIUM ATz= 7.7. *The Astrophysical Journal*, 788(2):106, may

2014. doi:10.1088/0004-637x/788/2/106. URL https://doi.org/10.1088/0004-637x/788/2/106.

Aaron R. Parsons, Adrian Liu, Zaki S. Ali, and Carina Cheng. OPTIMIZED BEAM SCULPTING WITH GENERALIZED FRINGE-RATE FILTERS. *The Astrophysical Journal*, 820(1):51, mar 2016. doi:10.3847/0004-637x/820/1/51. URL https://doi.org/10.3847/0004-637x/820/1/51.

A. H. Patil, S. Yatawatta, L. V. E. Koopmans, A. G. de Bruyn, M. A. Brentjens, S. Zaroubi, K. M. B. Asad, M. Hatef, V. Jelić, M. Mevius, A. R. Offringa, V. N. Pandey, H. Vedantham, F. B. Abdalla, W. N. Brouw, E. Chapman, B. Ciardi, B. K. Gehlot, A. Ghosh, G. Harker, I. T. Iliev, K. Kakiichi, S. Majumdar, G. Mellema, M. B. Silva, J. Schaye, D. Vrbanec, and S. J. Wijnholds. Upper limits on the 21 cm epoch of reionization power spectrum from one night with LOFAR. *The Astrophysical Journal*, 838(1):65, mar 2017. doi:10.3847/1538-4357/aa63e7. URL https://doi.org/10.3847/1538-4357/aa63e7.

T. J. Pearson and A. C. S. Readhead. Image formation by self-calibration in radio astronomy. *Annual Review of Astronomy and Astrophysics*, 22(1):97–130, 1984. doi:10.1146/annurev.aa.22.090184.000525. URL https://doi.org/10.1146/annurev.aa.22.090184.000525.

Planck Collaboration, Ade, P. A. R., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., Banday, A. J., Barreiro, R. B., Bartlett, J. G., Bartolo, N., Battaner, E., Battye, R., Benabed, K., Benoît, A., Benoit-Lévy, A., Bernard, J.-P., Bersanelli, M., Bielewicz, P., Bock, J. J., Bonaldi, A., Bonavera, L., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Bucher, M., Burigana, C., Butler, R. C., Calabrese, E., Cardoso, J.-F., Catalano, A., Challinor, A., Chamballu, A., Chary, R.-R., Chiang, H. C., Chluba, J., Christensen, P. R., Church, S., Clements, D. L., Colombi, S., Colombo, L. P. L., Combet, C., Coulais, A., Crill, B. P., Curto, A., Cuttaia, F., Danese, L., Davies, R. D., Davis, R. J., de Bernardis, P., de Rosa, A., de Zotti, G., Delabrouille, J., Désert, F.-X., Di Valentino, E., Dickinson, C., Diego, J. M., Dolag, K., Dole, H., Donzelli, S., Doré, O., Douspis, M., Ducout, A., Dunkley, J., Dupac, X., Efstathiou, G., Elsner, F., Enßlin, T. A., Eriksen, H. K., Farhang, M., Fergusson, J., Finelli, F., Forni, O., Frailis, M., Fraisse, A. A., Franceschi, E., Frejsel, A., Galeotta, S., Galli, S., Ganga, K., Gauthier, C., Gerbino, M., Ghosh, T., Giard, M., Giraud-Héraud, Y., Giusarma, E., Gjerløw, E., González-Nuevo, J., Górski, K. M., Gratton, S., Gregorio, A., Gruppuso, A., Gudmundsson, J. E., Hamann, J., Hansen, F. K., Hanson, D., Harrison, D. L., Helou, G., Henrot-Versillé, S., Hernández-Monteagudo, C., Herranz, D., Hildebrandt, S. R., Hivon, E., Hobson, M., Holmes, W. A., Hornstrup, A., Hovest, W., Huang, Z., Huffenberger, K. M., Hurier, G., Jaffe, A. H., Jaffe, T. R., Jones, W. C., Juvela, M., Keihänen, E., Keskitalo, R., Kisner, T. S., Kneissl, R., Knoche, J., Knox, L., Kunz, M., Kurki-Suonio, H., Lagache, G., Lähteenmäki, A., Lamarre, J.-M., Lasenby, A., Lattanzi, M., Lawrence, C. R., Leahy, J. P., Leonardi, R., Lesgourgues, J., Levrier, F., Lewis, A., Liguori, M., Lilje, P. B., Linden-Vørnle, M., López-Caniego, M., Lubin, P. M., Macías-Pérez, J. F., Maggio, G., Maino,

D., Mandolesi, N., Mangilli, A., Marchini, A., Maris, M., Martin, P. G., Martinelli, M., Martínez-González, E., Masi, S., Matarrese, S., McGehee, P., Meinhold, P. R., Melchiorri, A., Melin, J.-B., Mendes, L., Mennella, A., Migliaccio, M., Millea, M., Mitra, S., Miville-Deschênes, M.-A., Moneti, A., Montier, L., Morgante, G., Mortlock, D., Moss, A., Munshi, D., Murphy, J. A., Naselsky, P., Nati, F., Natoli, P., Netterfield, C. B., Nørgaard-Nielsen, H. U., Noviello, F., Novikov, D., Novikov, I., Oxborrow, C. A., Paci, F., Pagano, L., Pajot, F., Paladini, R., Paoletti, D., Partridge, B., Pasian, F., Patanchon, G., Pearson, T. J., Perdereau, O., Perotto, L., Perrotta, F., Pettorino, V., Piacentini, F., Piat, M., Pierpaoli, E., Pietrobon, D., Plaszczynski, S., Pointecouteau, E., Polenta, G., Popa, L., Pratt, G. W., Prézeau, G., Prunet, S., Puget, J.-L., Rachen, J. P., Reach, W. T., Rebolo, R., Reinecke, M., Remazeilles, M., Renault, C., Renzi, A., Ristorcelli, I., Rocha, G., Rosset, C., Rossetti, M., Roudier, G., Rouillé d´Orfeuil, B., Rowan-Robinson, M., Rubiño-Martín, J. A., Rusholme, B., Said, N., Salvatelli, V., Salvati, L., Sandri, M., Santos, D., Savelainen, M., Savini, G., Scott, D., Seiffert, M. D., Serra, P., Shellard, E. P. S., Spencer, L. D., Spinelli, M., Stolyarov, V., Stompor, R., Sudiwala, R., Sunyaev, R., Sutton, D., Suur-Uski, A.-S., Sygnet, J.-F., Tauber, J. A., Terenzi, L., Toffolatti, L., Tomasi, M., Tristram, M., Trombetti, T., Tucci, M., Tuovinen, J., Türler, M., Umana, G., Valenziano, L., Valiviita, J., Van Tent, F., Vielva, P., Villa, F., Wade, L. A., Wandelt, B. D., Wehus, I. K., White, M., White, S. D. M., Wilkinson, A., Yvon, D., Zacchei, A., and Zonca, A. Planck 2015 results - xiii. cosmological parameters. *A&A*, 594:A13, 2016. doi:10.1051/0004-6361/201525830. URL https://doi.org/10.1051/0004-6361/201525830.

Planck Collaboration, Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Battye, R., Benabed, K., Bernard, J.-P., Bersanelli, M., Bielewicz, P., Bock, J. J., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Bucher, M., Burigana, C., Butler, R. C., Calabrese, E., Cardoso, J.-F., Carron, J., Challinor, A., Chiang, H. C., Chluba, J., Colombo, L. P. L., Combet, C., Contreras, D., Crill, B. P., Cuttaia, F., de Bernardis, P., de Zotti, G., Delabrouille, J., Delouis, J.-M., Di Valentino, E., Diego, J. M., Doré, O., Douspis, M., Ducout, A., Dupac, X., Dusini, S., Efstathiou, G., Elsner, F., Enßlin, T. A., Eriksen, H. K., Fantaye, Y., Farhang, M., Fergusson, J., Fernandez-Cobos, R., Finelli, F., Forastieri, F., Frailis, M., Fraisse, A. A., Franceschi, E., Frolov, A., Galeotta, S., Galli, S., Ganga, K., Génova-Santos, R. T., Gerbino, M., Ghosh, T., González-Nuevo, J., Górski, K. M., Gratton, S., Gruppuso, A., Gudmundsson, J. E., Hamann, J., Handley, W., Hansen, F. K., Herranz, D., Hildebrandt, S. R., Hivon, E., Huang, Z., Jaffe, A. H., Jones, W. C., Karakci, A., Keihänen, E., Keskitalo, R., Kiiveri, K., Kim, J., Kisner, T. S., Knox, L., Krachmalnicoff, N., Kunz, M., Kurki-Suonio, H., Lagache, G., Lamarre, J.-M., Lasenby, A., Lattanzi, M., Lawrence, C. R., Le Jeune, M., Lemos, P., Lesgourgues, J., Levrier, F., Lewis, A., Liguori, M., Lilje, P. B., Lilley, M., Lindholm, V., López-Caniego, M., Lubin, P. M., Ma, Y.-Z., Macías-Pérez, J. F., Maggio, G., Maino, D., Mandolesi, N., Mangilli, A., Marcos-Caballero, A., Maris, M., Martin, P. G., Martinelli, M., Martínez-

González, E., Matarrese, S., Mauri, N., McEwen, J. D., Meinhold, P. R., Melchiorri, A., Mennella, A., Migliaccio, M., Millea, M., Mitra, S., Miville-Deschênes, M.-A., Molinari, D., Montier, L., Morgante, G., Moss, A., Natoli, P., Nørgaard-Nielsen, H. U., Pagano, L., Paoletti, D., Partridge, B., Patanchon, G., Peiris, H. V., Perrotta, F., Pettorino, V., Piacentini, F., Polastri, L., Polenta, G., Puget, J.-L., Rachen, J. P., Reinecke, M., Remazeilles, M., Renzi, A., Rocha, G., Rosset, C., Roudier, G., Rubiño-Martín, J. A., Ruiz-Granados, B., Salvati, L., Sandri, M., Savelainen, M., Scott, D., Shellard, E. P. S., Sirignano, C., Sirri, G., Spencer, L. D., Sunyaev, R., Suur-Uski, A.-S., Tauber, J. A., Tavagnacco, D., Tenti, M., Toffolatti, L., Tomasi, M., Trombetti, T., Valenziano, L., Valiviita, J., Van Tent, B., Vibert, L., Vielva, P., Villa, F., Vittorio, N., Wandelt, B. D., Wehus, I. K., White, M., White, S. D. M., Zacchei, A., and Zonca, A. Planck 2018 results - vi. cosmological parameters. *A&A*, 641:A6, 2020. doi:10.1051/0004-6361/201833910. URL https://doi.org/10.1051/0004-6361/201833910.

Jonathan C. Pober, Adrian Liu, Joshua S. Dillon, James E. Aguirre, Judd D. Bowman, Richard F. Bradley, Chris L. Carilli, David R. DeBoer, Jacqueline N. Hewitt, Daniel C. Jacobs, Matthew McQuinn, Miguel F. Morales, Aaron R. Parsons, Max Tegmark, and Dan J. Werthimer. WHAT NEXT-GENERATION 21 cm POWER SPECTRUM MEASUREMENTS CAN TEACH US ABOUT THE EPOCH OF REIONIZATION. *The Astrophysical Journal*, 782(2):66, jan 2014. doi:10.1088/0004-637x/782/2/66. URL https://doi.org/10.1088/0004-637x/782/2/66.

David Prelogović, Andrei Mesinger, Steven Murray, Giuseppe Fiameni, and Nicolas Gillet. Machine learning astrophysics from 21 cm lightcones: impact of network architectures and signal contamination. *Monthly Notices of the Royal Astronomical Society*, 509(3):3852–3867, 11 2021. ISSN 0035-8711. doi:10.1093/mnras/stab3215. URL https://doi.org/10.1093/mnras/stab3215.

Yuxiang Qin, Vivian Poulin, Andrei Mesinger, Bradley Greig, Steven Murray, and Jaehong Park. Reionization inference from the CMB optical depth and E-mode polarization power spectra. *Monthly Notices of the Royal Astronomical Society*, 499(1):550–558, 07 2020. ISSN 0035-8711. doi:10.1093/mnras/staa2797. URL https://doi.org/10.1093/mnras/staa2797.

Janakee Raste, Girish Kulkarni, Catherine A Watkinson, Laura C Keating, and Martin G Haehnelt. The 21-cm bispectrum from neutral hydrogen islands at z
lt; 6. *Monthly Notices of the Royal Astronomical Society*, page stae492, 02 2024. ISSN 0035-8711. doi:10.1093/mnras/stae492. URL https://doi.org/10.1093/mnras/stae492.

Urvashi Rau, Sanjay Bhatnagar, Maxim A. Voronkov, and Tim J. Cornwell. Advances in calibration and imaging techniques in radio interferometry. *Proceedings of the IEEE*, 97 (8):1472–1481, 2009. doi:10.1109/JPROC.2009.2014853.

N. Razavi-Ghods, S. H. Carey, J. A. Ely, and P. F Scott. Hera rf and calibration system design. In *2017 Progress In Electromagnetics Research Symposium - Spring (PIERS)*, pages 1544–1548, 2017. doi:10.1109/PIERS.2017.8261992.

P. A. G. Scheuer. A statistical method for analysing observations of faint radio stars. *Mathematical Proceedings of the Cambridge Philosophical Society*, 53(3):764–773, 1957. doi:10.1017/S0305004100032825.

J. Richard Shaw, Kris Sigurdson, Ue-Li Pen, Albert Stebbins, and Michael Sitwell. ALL-SKY INTERFEROMETRY WITH SPHERICAL HARMONIC TRANSIT TELESCOPES. *The Astrophysical Journal*, 781(2):57, jan 2014. doi:10.1088/0004-637x/781/2/57. URL https://doi.org/10.1088/0004-637x/781/2/57.

Hayato Shimabukuro, Shintaro Yoshiura, Keitaro Takahashi, Shuichiro Yokoyama, and Kiyotomo Ichiki. Studying 21cm power spectrum with one-point statistics. *Monthly Notices of the Royal Astronomical Society*, 451(1):467–474, 05 2015. ISSN 0035-8711. doi:10.1093/mnras/stv965. URL https://doi.org/10.1093/mnras/stv965.

Hayato Shimabukuro, Shintaro Yoshiura, Keitaro Takahashi, Shuichiro Yokoyama, and Kiyotomo Ichiki. 21 cm line bispectrum as a method to probe cosmic dawn and epoch of reionization. *Monthly Notices of the Royal Astronomical Society*, 458(3):3003–3011, 03 2016. ISSN 0035-8711. doi:10.1093/mnras/stw482. URL https://doi.org/10.1093/mnras/stw482.

D. Slepian. Prolate spheroidal wave functions, fourier analysis, and uncertainty — v: the discrete case. *The Bell System Technical Journal*, 57(5):1371–1430, 1978. doi:10.1002/j.1538-7305.1978.tb02104.x.

O. M. Smirnov. Revisiting the radio interferometer measurement equation - ii. calibration and direction-dependent effects. *A&A*, 527:A107, 2011. doi:10.1051/0004-6361/201116434. URL https://doi.org/10.1051/0004-6361/201116434.

Paula S Soares, Catherine A Watkinson, Steven Cunnington, and Alkistis Pourtsidou. Gaussian Process Regression for foreground removal in Hi Intensity Mapping experiments. *Monthly Notices of the Royal Astronomical Society*, 510(4):5872–5890, 09 2021. ISSN 0035-8711. doi:10.1093/mnras/stab2594. URL https://doi.org/10.1093/mnras/stab2594.

Marta Spinelli, Isabella P Carucci, Steven Cunnington, Stuart E Harper, Melis O Irfan, José Fonseca, Alkistis Pourtsidou, and Laura Wolz. SKAO Hi intensity mapping: blind foreground subtraction challenge. *Monthly Notices of the Royal Astronomical Society*, 509(2):2048–2074, 10 2021. ISSN 0035-8711. doi:10.1093/mnras/stab3064. URL https://doi.org/10.1093/mnras/stab3064.

I. S. Sullivan, M. F. Morales, B. J. Hazelton, W. Arcus, D. Barnes, G. Bernardi, F. H. Briggs, J. D. Bowman, J. D. Bunton, R. J. Cappallo, B. E. Corey, A. Deshpande, L. deSouza, D. Emrich, B. M. Gaensler, R. Goeke, L. J. Greenhill, D. Herne, J. N. Hewitt, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, B. B. Kincaid, R. Koenig, E. Kratzenberg, C. J. Lonsdale, M. J. Lynch, S. R. McWhirter, D. A. Mitchell, E. Morgan, D. Oberoi, S. M. Ord, J. Pathikulangara, T. Prabu, R. A. Remillard, A. E. E. Rogers, A. Roshi, J. E. Salah, R. J. Sault, N. Udaya Shankar, K. S. Srivani, J. Stevens, R. Subrahmanyan, S. J. Tingay, R. B. Wayth, M. Waterson, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, and J. S. B. Wyithe. Fast holographic deconvolution: A new technique for precision

radio interferometry. *The Astrophysical Journal*, 759(1):17, oct 2012. doi:10.1088/0004-637X/759/1/17. URL https://dx.doi.org/10.1088/0004-637X/759/1/17.

Max Tegmark. How to make maps from cosmic microwave background data without losing information. *The Astrophysical Journal*, 480(2):L87, may 1997. doi:10.1086/310631. URL https://dx.doi.org/10.1086/310631.

The HERA Collaboration, Zara Abdurashidova, James E. Aguirre, Paul Alexander, Zaki S. Ali, Yanga Balfour, Adam P. Beardsley, Gianni Bernardi, Tashalee S. Billings, Judd D. Bowman, Richard F. Bradley, Philip Bull, Jacob Burba, Steve Carey, Chris L. Carilli, Carina Cheng, David R. DeBoer, Matt Dexter, Eloy de Lera Acedo, Taylor Dibblee-Barkman, Joshua S. Dillon, John Ely, Aaron Ewall-Wice, Nicolas Fagnoni, Randall Fritz, Steven R. Furlanetto, Kingsley Gale-Sides, Brian Glendenning, Deepthi Gorthi, Bradley Greig, Jasper Grobbelaar, Ziyaad Halday, Bryna J. Hazelton, Jacqueline N. Hewitt, Jack Hickish, Daniel C. Jacobs, Austin Julius, Nicholas S. Kern, Joshua Kerrigan, Piyanat Kittiwisit, Saul A. Kohn, Matthew Kolopanis, Adam Lanman, Paul La Plante, Telalo Lekalake, David Lewis, Adrian Liu, David MacMahon, Lourence Malan, Cresshim Malgas, Matthys Maree, Zachary E. Martinot, Eunice Matsetela, Andrei Mesinger, Mathakane Molewa, Miguel F. Morales, Tshegofalang Mosiane, Steven G. Murray, Abraham R. Neben, Bojan Nikolic, Chuneeta D. Nunhokee, Aaron R. Parsons, Nipanjana Patra, Robert Pascua, Samantha Pieterse, Jonathan C. Pober, Nima Razavi-Ghods, Jon Ringuette, James Robnett, Kathryn Rosie, Peter Sims, Saurabh Singh, Craig Smith, Angelo Syce, Nithyanandan Thyagarajan, Peter K. G. Williams, and Haoxuan Zheng. First results from HERA phase i: Upper limits on the epoch of reionization 21 cm power spectrum. *The Astrophysical Journal*, 925(2):221, feb 2022a. doi:10.3847/1538-4357/ac1c78. URL https://doi.org/10.3847/1538-4357/ac1c78.

The HERA Collaboration, Zara Abdurashidova, James E. Aguirre, Paul Alexander, Zaki S. Ali, Yanga Balfour, Rennan Barkana, Adam P. Beardsley, Gianni Bernardi, Tashalee S. Billings, Judd D. Bowman, Richard F. Bradley, Philip Bull, Jacob Burba, Steve Carey, Chris L. Carilli, Carina Cheng, David R. DeBoer, Matt Dexter, Eloy de Lera Acedo, Joshua S. Dillon, John Ely, Aaron Ewall-Wice, Nicolas Fagnoni, Anastasia Fialkov, Randall Fritz, Steven R. Furlanetto, Kingsley Gale-Sides, Brian Glendenning, Deepthi Gorthi, Bradley Greig, Jasper Grobbelaar, Ziyaad Halday, Bryna J. Hazelton, Stefan Heimersheim, Jacqueline N. Hewitt, Jack Hickish, Daniel C. Jacobs, Austin Julius, Nicholas S. Kern, Joshua Kerrigan, Piyanat Kittiwisit, Saul A. Kohn, Matthew Kolopanis, Adam Lanman, Paul La Plante, Telalo Lekalake, David Lewis, Adrian Liu, Yin-Zhe Ma, David MacMahon, Lourence Malan, Cresshim Malgas, Matthys Maree, Zachary E. Martinot, Eunice Matsetela, Andrei Mesinger, Jordan Mirocha, Mathakane Molewa, Miguel F. Morales, Tshegofalang Mosiane, Julian B. Muñoz, Steven G. Murray, Abraham R. Neben, Bojan Nikolic, Chuneeta D. Nunhokee, Aaron R. Parsons, Nipanjana Patra, Samantha Pieterse, Jonathan C. Pober, Yuxiang Qin, Nima Razavi-Ghods, Itamar Reis, Jon Ringuette, James Robnett, Kathryn Rosie, Mario G. Santos, Sudipta Sikder, Peter Sims, Craig Smith,

Angelo Syce, Nithyanandan Thyagarajan, Peter K. G. Williams, and Haoxuan Zheng. Hera phase i limits on the cosmic 21 cm signal: Constraints on astrophysics and cosmology during the epoch of reionization. *The Astrophysical Journal*, 924(2):51, jan 2022b. doi:10.3847/1538-4357/ac2ffc. URL https://dx.doi.org/10.3847/1538-4357/ac2ffc.

The HERA Collaboration, Zara Abdurashidova, Tyrone Adams, James E. Aguirre, Paul Alexander, Zaki S. Ali, Rushelle Baartman, Yanga Balfour, Rennan Barkana, Adam P. Beardsley, Gianni Bernardi, Tashalee S. Billings, Judd D. Bowman, Richard F. Bradley, Daniela Breitman, Philip Bull, Jacob Burba, Steve Carey, Chris L. Carilli, Carina Cheng, Samir Choudhuri, David R. DeBoer, Eloy de Lera Acedo, Matt Dexter, Joshua S. Dillon, John Ely, Aaron Ewall-Wice, Nicolas Fagnoni, Anastasia Fialkov, Randall Fritz, Steven R. Furlanetto, Kingsley Gale-Sides, Hugh Garsden, Brian Glendenning, Adélie Gorce, Deepthi Gorthi, Bradley Greig, Jasper Grobbelaar, Ziyaad Halday, Bryna J. Hazelton, Stefan Heimersheim, Jacqueline N. Hewitt, Jack Hickish, Daniel C. Jacobs, Austin Julius, Nicholas S. Kern, Joshua Kerrigan, Piyanat Kittiwisit, Saul A. Kohn, Matthew Kolopanis, Adam Lanman, Paul La Plante, David Lewis, Adrian Liu, Anita Loots, Yin-Zhe Ma, David H. E. MacMahon, Lourence Malan, Keith Malgas, Cresshim Malgas, Matthys Maree, Bradley Marero, Zachary E. Martinot, Lisa McBride, Andrei Mesinger, Jordan Mirocha, Mathakane Molewa, Miguel F. Morales, Tshegofalang Mosiane, Julian B. Muñoz, Steven G. Murray, Vighnesh Nagpal, Abraham R. Neben, Bojan Nikolic, Chuneeta D. Nunhokee, Hans Nuwegeld, Aaron R. Parsons, Robert Pascua, Nipanjana Patra, Samantha Pieterse, Yuxiang Qin, Nima Razavi-Ghods, James Robnett, Kathryn Rosie, Mario G. Santos, Peter Sims, Saurabh Singh, Craig Smith, Hilton Swarts, Jianrong Tan, Nithyanandan Thyagarajan, Michael J. Wilensky, Peter K. G. Williams, Pieter van Wyngaarden, and Haoxuan Zheng. Improved constraints on the 21 cm eor power spectrum and the x-ray heating of the igm with hera phase i observations. *The Astrophysical Journal*, 945(2):124, mar 2023. doi:10.3847/1538-4357/acaf50. URL https://dx.doi.org/10.3847/1538-4357/acaf50.

Nithyanandan Thyagarajan, N. Udaya Shankar, Ravi Subrahmanyan, Wayne Arcus, Gianni Bernardi, Judd D. Bowman, Frank Briggs, John D. Bunton, Roger J. Cappallo, Brian E. Corey, Ludi deSouza, David Emrich, Bryan M. Gaensler, Robert F. Goeke, Lincoln J. Greenhill, Bryna J. Hazelton, David Herne, Jacqueline N. Hewitt, Melanie Johnston-Hollitt, David L. Kaplan, Justin C. Kasper, Barton B. Kincaid, Ronald Koenig, Eric Kratzenberg, Colin J. Lonsdale, Mervyn J. Lynch, S. Russell McWhirter, Daniel A. Mitchell, Miguel F. Morales, Edward H. Morgan, Divya Oberoi, Stephen M. Ord, Joseph Pathikulangara, Ronald A. Remillard, Alan E. E. Rogers, D. Anish Roshi, Joseph E. Salah, Robert J. Sault, K. S. Srivani, Jamie B. Stevens, Prabu Thiagaraj, Steven J. Tingay, Randall B. Wayth, Mark Waterson, Rachel L. Webster, Alan R. Whitney, Andrew J. Williams, Christopher L. Williams, and J. Stuart B. Wyithe. A STUDY OF FUNDAMENTAL LIMITATIONS TO STATISTICAL DETECTION OF REDSHIFTED h i FROM THE EPOCH OF REIONIZATION. *The Astrophysical Journal*, 776(1):6, sep 2013. doi:10.1088/0004-637x/776/1/6. URL https://doi.org/10.1088/0004-637x/776/1/6.

Nithyanandan Thyagarajan, Daniel C. Jacobs, Judd D. Bowman, N. Barry, A. P. Beardsley, G. Bernardi, F. Briggs, R. J. Cappallo, P. Carroll, B. E. Corey, A. de Oliveira-Costa, Joshua S. Dillon, D. Emrich, A. Ewall-Wice, L. Feng, R. Goeke, L. J. Greenhill, B. J. Hazelton, J. N. Hewitt, N. Hurley-Walker, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, Han-Seek Kim, P. Kittiwisit, E. Kratzenberg, E. Lenc, J. Line, A. Loeb, C. J. Lonsdale, M. J. Lynch, B. McKinley, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, A. R. Neben, D. Oberoi, A. R. Offringa, S. M. Ord, Sourabh Paul, B. Pindor, J. C. Pober, T. Prabu, P. Procopio, J. Riding, A. E. E. Rogers, A. Roshi, N. Udaya Shankar, Shiv K. Sethi, K. S. Srivani, R. Subrahmanyan, I. S. Sullivan, M. Tegmark, S. J. Tingay, C. M. Trott, M. Waterson, R. B. Wayth, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, C. Wu, and J. S. B. Wyithe. FOREGROUNDS IN WIDE-FIELD REDSHIFTED 21 cm POWER SPECTRA. *The Astrophysical Journal*, 804(1):14, apr 2015a. doi:10.1088/0004-637x/804/1/14. URL https://doi.org/10.1088/0004-637x/804/1/14.

Nithyanandan Thyagarajan, Daniel C. Jacobs, Judd D. Bowman, N. Barry, A. P. Beardsley, G. Bernardi, F. Briggs, R. J. Cappallo, P. Carroll, A. A. Deshpande, A. de Oliveira-Costa, Joshua S. Dillon, A. Ewall-Wice, L. Feng, L. J. Greenhill, B. J. Hazelton, L. Hernquist, J. N. Hewitt, N. Hurley-Walker, M. Johnston-Hollitt, D. L. Kaplan, Han-Seek Kim, P. Kittiwisit, E. Lenc, J. Line, A. Loeb, C. J. Lonsdale, B. McKinley, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, A. R. Neben, D. Oberoi, A. R. Offringa, S. M. Ord, Sourabh Paul, B. Pindor, J. C. Pober, T. Prabu, P. Procopio, J. Riding, N. Udaya Shankar, Shiv K. Sethi, K. S. Srivani, R. Subrahmanyan, I. S. Sullivan, M. Tegmark, S. J. Tingay, C. M. Trott, R. B. Wayth, R. L. Webster, A. Williams, C. L. Williams, and J. S. B. Wyithe. CONFIRMATION OF WIDE-FIELD SIGNATURES IN REDSHIFTED 21 cm POWER SPECTRA. *The Astrophysical Journal*, 807(2):L28, jul 2015b. doi:10.1088/2041-8205/807/2/l28. URL https://doi.org/10.1088/2041-8205/807/2/l28.

Nithyanandan Thyagarajan, Aaron R. Parsons, David R. DeBoer, Judd D. Bowman, Aaron M. Ewall-Wice, Abraham R. Neben, and Nipanjana Patra. Effects of antenna beam chromaticity on redshifted 21 cm power spectrum and implications for hydrogen epoch of reionization array. *The Astrophysical Journal*, 825(1):9, jun 2016. doi:10.3847/0004-637X/825/1/9. URL https://dx.doi.org/10.3847/0004-637X/825/1/9.

S. J. Tingay, R. Goeke, J. D. Bowman, D. Emrich, S. M. Ord, D. A. Mitchell, M. F. Morales, T. Booler, B. Crosse, R. B. Wayth, and et al. The murchison widefield array: The square kilometre array precursor at low radio frequencies. *Publications of the Astronomical Society of Australia*, 30:e007, 2013. doi:10.1017/pasa.2012.007.

Cathryn M. Trott, Randall B. Wayth, and Steven J. Tingay. THE IMPACT OF POINT-SOURCE SUBTRACTION RESIDUALS ON 21 cm EPOCH OF REIONIZATION ESTIMATION. *The Astrophysical Journal*, 757(1):101, sep 2012. doi:10.1088/0004-637x/757/1/101. URL https://doi.org/10.1088/0004-637x/757/1/101.

Cathryn M Trott, C H Jordan, S Midgley, N Barry, B Greig, B Pindor, J H Cook, G Sleap, S J Tingay, D Ung, P Hancock, A Williams, J Bowman, R Byrne, A Chokshi, B J

Hazelton, K Hasegawa, D Jacobs, R C Joseph, W Li, J L B Line, C Lynch, B McKinley, D A Mitchell, M F Morales, M Ouchi, J C Pober, M Rahimi, K Takahashi, R B Wayth, R L Webster, M Wilensky, J S B Wyithe, S Yoshiura, Z Zhang, and Q Zheng. Deep multiredshift limits on Epoch of Reionization 21 cm power spectra from four seasons of Murchison Widefield Array observations. *Monthly Notices of the Royal Astronomical Society*, 493(4):4711–4727, 02 2020. ISSN 0035-8711. doi:10.1093/mnras/staa414. URL https://doi.org/10.1093/mnras/staa414.

van Haarlem, Wise, M. W., Gunst, A. W., Heald, G., McKean, J. P., Hessels, J. W. T., de Bruyn, A. G., Nijboer, R., Swinbank, J., Fallows, R., Brentjens, M., Nelles, A., Beck, R., Falcke, H., Fender, R., Hörandel, J., Koopmans, L. V. E., Mann, G., Miley, G., Röttgering, H., Stappers, B. W., Wijers, R. A. M. J., Zaroubi, S., van den Akker, M., Alexov, A., Anderson, J., Anderson, K., van Ardenne, A., Arts, M., Asgekar, A., Avruch, I. M., Batejat, F., Bähren, L., Bell, M. E., Bell, M. R., van Bemmel, I., Bennema, P., Bentum, M. J., Bernardi, G., Best, P., Bîrzan, L., Bonafede, A., Boonstra, A.-J., Braun, R., Bregman, J., Breitling, F., van de Brink, R. H., Broderick, J., Broekema, P. C., Brouw, W. N., Brüggen, M., Butcher, H. R., van Cappellen, W., Ciardi, B., Coenen, T., Conway, J., Coolen, A., Corstanje, A., Damstra, S., Davies, O., Deller, A. T., Dettmar, R.-J., van Diepen, G., Dijkstra, K., Donker, P., Doorduin, A., Dromer, J., Drost, M., van Duin, A., Eislöffel, J., van Enst, J., Ferrari, C., Frieswijk, W., Gankema, H., Garrett, M. A., de Gasperin, F., Gerbers, M., de Geus, E., Grießmeier, J.-M., Grit, T., Gruppen, P., Hamaker, J. P., Hassall, T., Hoeft, M., Holties, H. A., Horneffer, A., van der Horst, A., van Houwelingen, A., Huijgen, A., Iacobelli, M., Intema, H., Jackson, N., Jelic, V., de Jong, A., Juette, E., Kant, D., Karastergiou, A., Koers, A., Kollen, H., Kondratiev, V. I., Kooistra, E., Koopman, Y., Koster, A., Kuniyoshi, M., Kramer, M., Kuper, G., Lambropoulos, P., Law, C., van Leeuwen, J., Lemaitre, J., Loose, M., Maat, P., Macario, G., Markoff, S., Masters, J., McFadden, R. A., McKay-Bukowski, D., Meijering, H., Meulman, H., Mevius, M., Middelberg, E., Millenaar, R., Miller-Jones, J. C. A., Mohan, R. N., Mol, J. D., Morawietz, J., Morganti, R., Mulcahy, D. D., Mulder, E., Munk, H., Nieuwenhuis, L., van Nieuwpoort, R., Noordam, J. E., Norden, M., Noutsos, A., Offringa, A. R., Olofsson, H., Omar, A., Orrú, E., Overeem, R., Paas, H., Pandey-Pommier, M., Pandey, V. N., Pizzo, R., Polatidis, A., Rafferty, D., Rawlings, S., Reich, W., de Reijer, J.-P., Reitsma, J., Renting, G. A., Riemers, P., Rol, E., Romein, J. W., Roosjen, J., Ruiter, M., Scaife, A., van der Schaaf, K., Scheers, B., Schellart, P., Schoenmakers, A., Schoonderbeek, G., Serylak, M., Shulevski, A., Sluman, J., Smirnov, O., Sobey, C., Spreeuw, H., Steinmetz, M., Sterks, C. G. M., Stiepel, H.-J., Stuurwold, K., Tagger, M., Tang, Y., Tasse, C., Thomas, I., Thoudam, S., Toribio, M. C., van der Tol, B., Usov, O., van Veelen, M., van der Veen, A.-J., ter Veen, S., Verbiest, J. P. W., Vermeulen, R., Vermaas, N., Vocks, C., Vogt, C., de Vos, M., van der Wal, E., van Weeren, R., Weggemans, H., Weltevrede, P., White, S., Wijnholds, S. J., Wilhelmsson, T., Wucknitz, O., Yatawatta, S., Zarka, P., Zensus, A., and van Zwieten, J. Lofar: The low-frequency array. *A&A*, 556:A2, 2013.

doi:10.1051/0004-6361/201220873. URL https://doi.org/10.1051/0004-6361/201220873.

Harish Vedantham, N. Udaya Shankar, and Ravi Subrahmanyan. IMAGING THE EPOCH OF REIONIZATION: LIMITATIONS FROM FOREGROUND CONFUSION AND IMAGING ALGORITHMS. *The Astrophysical Journal*, 745(2):176, jan 2012. doi:10.1088/0004-637x/745/2/176. URL https://doi.org/10.1088/0004-637x/745/2/176.

Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi:10.1038/s41592-019-0686-2.

B P Wakker and UJ Schwarz. The multi-resolution clean and its application to the short-spacing problem in interferometry. *Astronomy and Astrophysics (ISSN 0004-6361), vol. 200, no. 1-2, July 1988, p. 312-322. Research supported by ZWO.*, 200:312–322, 1988.

C. A. Watkinson and J. R. Pritchard. Distinguishing models of reionization using future radio observations of 21-cm 1-point statistics. *Monthly Notices of the Royal Astronomical Society*, 443(4):3090–3106, 08 2014. ISSN 0035-8711. doi:10.1093/mnras/stu1384. URL https://doi.org/10.1093/mnras/stu1384.

C. A. Watkinson and J. R. Pritchard. The impact of spin-temperature fluctuations on the 21-cm moments. *Monthly Notices of the Royal Astronomical Society*, 454(2):1416–1431, 10 2015. ISSN 0035-8711. doi:10.1093/mnras/stv2010. URL https://doi.org/10.1093/mnras/stv2010.

Catherine A Watkinson, Sambit K Giri, Hannah E Ross, Keri L Dixon, Ilian T Iliev, Garrelt Mellema, and Jonathan R Pritchard. The 21-cm bispectrum as a probe of non-Gaussianities due to X-ray heating. *Monthly Notices of the Royal Astronomical Society*, 482(2):2653–2669, 10 2018. ISSN 0035-8711. doi:10.1093/mnras/sty2740. URL https://doi.org/10.1093/mnras/sty2740.

Catherine A Watkinson, Bradley Greig, and Andrei Mesinger. Epoch of reionization parameter estimation with the 21-cm bispectrum. *Monthly Notices of the Royal Astronomical Society*, 510(3):3838–3848, 01 2022. ISSN 0035-8711. doi:10.1093/mnras/stab3706. URL https://doi.org/10.1093/mnras/stab3706.

M. H. Wieringa. An investigation of the telescope based calibration methods 'redundancy' and 'self-cal'. *Experimental Astronomy*, 2:203–225, 1992. doi:10.1007/BF00420576.

J. Stuart B. Wyithe and Miguel F. Morales. Biased reionization and non-Gaussianity in redshifted 21-cm intensity maps of the reionization epoch. *Monthly Notices of the Royal*

Astronomical Society, 379(4):1647–1657, 07 2007. ISSN 0035-8711. doi:10.1111/j.1365-2966.2007.12048.x. URL https://doi.org/10.1111/j.1365-2966.2007.12048.x.

Zhilei Xu, Jacqueline N. Hewitt, Kai-Feng Chen, Honggeun Kim, Joshua S. Dillon, Nicholas S. Kern, Miguel F. Morales, Bryna J. Hazelton, Ruby Byrne, Nicolas Fagnoni, Eloy de Lera Acedo, Zara Abdurashidova, Tyrone Adams, James E. Aguirre, Paul Alexander, Zaki S. Ali, Rushelle Baartman, Yanga Balfour, Adam P. Beardsley, Gianni Bernardi, Tashalee S. Billings, Judd D. Bowman, Richard F. Bradley, Philip Bull, Jacob Burba, Steven Carey, Chris L. Carilli, Carina Cheng, David R. DeBoer, Matt Dexter, Nico Eksteen, John Ely, Aaron Ewall-Wice, Randall Fritz, Steven R. Furlanetto, Kingsley Gale-Sides, Brian Glendenning, Deepthi Gorthi, Bradley Greig, Jasper Grobbelaar, Ziyaad Halday, Jack Hickish, Daniel C. Jacobs, Austin Julius, MacCalvin Kariseb, Joshua Kerrigan, Piyanat Kittiwisit, Saul A. Kohn, Matthew Kolopanis, Adam Lanman, Paul La Plante, Adrian Liu, Anita Loots, Yin-Zhe Ma, David Harold Edward MacMahon, Lourence Malan, Cresshim Malgas, Keith Malgas, Bradley Marero, Zachary E. Martinot, Andrei Mesinger, Mathakane Molewa, Tshegofalang Mosiane, Steven G. Murray, Abraham R. Neben, Bojan Nikolic, Hans Nuwegeld, Aaron R. Parsons, Nipanjana Patra, Samantha Pieterse, Jonathan C. Pober, Nima Razavi-Ghods, James Robnett, Kathryn Rosie, Peter Sims, Craig Smith, Hilton Swarts, Nithyanandan Thyagarajan, Pieter Van Van Wyngaarden, Peter K. G. Williams, Haoxuan Zheng, and (HERA Collaboration). Direct optimal mapping for 21 cm cosmology: A demonstration with the hydrogen epoch of reionization array. *The Astrophysical Journal*, 938(2):128, oct 2022. doi:10.3847/1538-4357/ac9053. URL https://dx.doi.org/10.3847/1538-4357/ac9053.

Sarod Yatawatta. Distributed radio interferometric calibration. *Monthly Notices of the Royal Astronomical Society*, 449(4):4506–4514, June 2015. doi:10.1093/mnras/stv596.

Shintaro Yoshiura, Hayato Shimabukuro, Keitaro Takahashi, Rieko Momose, Hiroyuki Nakanishi, and Hiroshi Imai. Sensitivity for 21 cm bispectrum from Epoch of Reionization. *Monthly Notices of the Royal Astronomical Society*, 451(1):266–274, 05 2015. ISSN 0035-8711. doi:10.1093/mnras/stv855. URL https://doi.org/10.1093/mnras/stv855.

H. Zheng, M. Tegmark, V. Buza, J. S. Dillon, H. Gharibyan, J. Hickish, E. Kunz, A. Liu, J. Losh, A. Lutomirski, S. Morrison, S. Narayanan, A. Perko, D. Rosner, N. Sanchez, K. Schutz, S. M. Tribiano, M. Valdez, H. Yang, K. Zarb Adami, I. Zelko, K. Zheng, R. P. Armstrong, R. F. Bradley, M. R. Dexter, A. Ewall-Wice, A. Magro, M. Matejek, E. Morgan, A. R. Neben, Q. Pan, R. F. Penna, C. M. Peterson, M. Su, J. Villasenor, C. L. Williams, and Y. Zhu. MITEoR: a scalable interferometer for precision 21 cm cosmology. *Monthly Notices of the Royal Astronomical Society*, 445(2):1084–1103, 10 2014. ISSN 0035-8711. doi:10.1093/mnras/stu1773. URL https://doi.org/10.1093/mnras/stu1773.

H. Zheng, M. Tegmark, J. S. Dillon, D. A. Kim, A. Liu, A. R. Neben, J. Jonas, P. Reich, and W. Reich. An improved model of diffuse galactic radio emission from 10 MHz to 5 THz. *Monthly Notices of the Royal Astronomical Society*, 464(3):3486–3497, January 2017. doi:10.1093/mnras/stw2525.