Precision Simulation of Low-Frequency Radio Interferometers

by

Adam Eudene Lanman

B. S., Physics and Astronomy, University of Rochester, 2013

B. A., Mathematics, University of Rochester, 2013

Sc. M., Brown University, 2015

A dissertation submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in the Department of Physics at Brown University

Providence, Rhode Island

February 2020

This dissertation by Adam Eudene Lanman is accepted in its present form by
the Department of Physics as satisfying the dissertation requirement
for the degree of Doctor of Philosophy.

Date ______________ ________________________________
Jonathan Pober, Director

Recommended to the Graduate Council

Date ______________ ________________________________
Ian Dell'Antonio, Reader

Date ______________ ________________________________
Gregory Tucker, Reader

Approved by the Graduate Council

Date ______________ ________________________________
Andrew Campbell
Dean of the Graduate School

Abstract

From the release of the CMB at redshift $z \sim 1100$, when the CMB was released, to around $z \sim 6$, where the most distant quasars are observed, lies an unobserved stretch of cosmic history. During this time, the first large scale structures grew from small perturbations, the first stars ignited, and early galaxies emitted UV radiation which re-ionized the intergalactic medium (IGM) into the fully ionized universe we see today. This last stage, when these luminous sources heated and ionized the universe, is known as the Epoch of Reionization (EoR). The 21 cm hyperfine transition of neutral hydrogen (HI) offers a window into this period, as it can be used to directly map the three-dimensional structure of the IGM throughout cosmic history. Experiments to measure this signal must contend with foreground contaminants 5 orders of magnitude brighter than the expected signal. Though bright, these foregrounds have very smooth spectra, which allows their power to be isolated from the EoR signal in Fourier space provided the observing instrument does not introduce any spectral structure. Understanding and mitigating these sources of spectral structure has been an ongoing effort.

Precise and detailed simulations of the instrument response have revealed several effects that were unknown in radio astronomy, but relevant to the challenge of 21 cm detection. Previous generations of simulators have lacked sufficient standards of verification, and made it impossible to determine whether some observed artifacts were due to the simulation architecture or realistic to the instrument. In this work, I present two new instrument simulation tools. The simulator pyuvsim has been collaboratively written within a robust test framework to provide the most precise and accurate simulated visibilities available, and has come to serve as a reference standard within the HERA collaboration. In addition, I have written a new diffuse simulator called healvis which has been validated against pyuvsim and analytic calculations. With healvis simulations, I have studied the time-correlation properties of drift-scan observations of EoR-like signals and measured the sample variance of realistic observations of such signals, as a fundamental lower limit on the precision that can be obtained from current experiments. In addition, I have looked at the impact of various instrumental effects on the spread of diffuse foreground power in Fourier space, testing the limits of what Fourier modes can be detected.

Curriculum Vitae

Education

B.S. Physics and Astronomy, University of Rochester, 2013

B.A. Mathematics, University of Rochester, 2013

M.S. Physics, Brown University, 2016

Research

My current focus is on applying numerical simulation to validate analysis and calibration tools for low-frequency radio telescope arrays. This work is divided between making instrument simulators and making sky models as input to them.

With Jonathan Pober and the HERA and MWA Collaborations, Brown University (January 2016 – present)

Explored techniques of mapping a realistic 21 cm signal from a cubic simulation volume onto a set of sky maps while preserving 2-point statistics.

Contributed to the development of the `Fast Holographic Deconvolution` (FHD) and `pyuvdata` software packages.

Used instrument simulations to assess sample variance and visibility correlation properties of drift-scanning interferometers.

Developer of a state of the art radio interferometry simulation tool `pyuvsim`.

Publications

Lead Author

1. **Lanman** , A.E., Pober, J.C., Kern, N.S., Acedo, E. d.L., DeBoer, D.R., and Fagnoni, N. (2019b). Quantifying EoR delay spectrum contamination from diffuse radio emission. *arXiv:1910.10573 [astro-ph]*.

2. **Lanman** , A.E. and Pober, J.C. (2019). Fundamental uncertainty levels of 21 cm power spectra from a delay analysis. *Monthly Notices of the Royal Astronomical Society* , 487(4):5840–5853.

3. **Lanman** , A., Hazelton, B., Jacobs, D., Kolopanis, M., Pober, J., Aguirre, J., and Thyagarajan, N. (2019a). pyuvsim: A comprehensive simulation package for radio interferometers in python. *Journal of Open Source Software* , 4(37):1234.

Co-author

1. Li, W., Pober, J.C., Barry, N., Hazelton, B.J., Morales, M.F., Trott, C.M., and **Lanman** , A. [et.al.] (2019). First Season MWA Phase II EoR Power Spectrum Results at Redshift 7. *arXiv:1911.10216 [astro-ph]*.

2. Kern, N.S., Parsons, A.R., Dillon, J.S., and **Lanman** , A.E. (2019b). Mitigating Internal Instrument Coupling I: Temporal and Spectral Modeling. *arXiv:1909.11732 [astro-ph]*.

3. Kern, N.S., Parsons, A.R., Dillon, J.S., and **Lanman** , A. E. . [et.al.] (2019c). Mitigating Internal Instrument Coupling II: A Method Demonstration with the Hydrogen Epoch of Reionization Array. *arXiv:1909.11733 [astro-ph]*.

Software Products

1. **Lanman** , A.E. and Kern, N. (2019). healvis: Radio interferometric visibility simulator based on HEALpix maps. *Astrophysics Source Code Library* , page ascl:1907.002.

Collaboration Papers

1. Kern, N.S., Dillon, J.S., Parsons, A.R., Carilli, C.L., Bernardi, G., Abdurashidova, Z., Aguirre, J.E., and Alexander, Paul [et al. including **Lanman** , A.] (2019a). Absolute Calibration for the Hydrogen Epoch of Reionization Array and Its Impact on the 21 cm Power Spectrum. *arXiv:1910.12943 [astro-ph]*.

2. Barry, N., Wilensky, M., Trott, C.M., Pindor, B., Beardsley, A.P., Hazelton, B.J., Sullivan, I.S., Morales, M.F., Pober, J.C., Line, J., Greig, B., Byrne, R., and **Lanman** , A. [et.al.] (2019). Improving the Epoch of Reionization Power Spectrum Results from Murchison Widefield Array Season 1 Observations. *The Astrophysical Journal* , 884(1):1.

3. Fagnoni, N., Acedo, E. d.L., DeBoer, D.R., Abdurashidova, Z., Aguirre, J.E., and Alexander, P. [et al. including **Lanman**, A. E.] (2019). Electrical and electromagnetic co-simulations of the HERA Phase I receiver system including the effects of mutual coupling, and impact on the EoR window. *arXiv:1908.02383 [astro-ph].*

4. Kerrigan, J., LaPlante, P., Kohn, S., Pober, J.C., Aguirre, J., Abdurashidova, Z., Alexander, P., and [et al.including **Lanman** , A. (2019). Optimizing Sparse RFI Prediction using Deep Learning. *arXiv:1902.08244 [astro-ph].*

5. Li, W., Pober, J.C., Hazelton, B.J., Barry, N., Morales, M.F., Sullivan, I., Parsons, A.R., and [et al, i.t. (2018). Comparing Redundant and Sky-model-based Interferometric Calibration: A First Look with Phase II of the MWA. *The Astrophysical Journal* , 863(2):170.

6. Kerrigan, J.R., Pober, J.C., and [et al.including **Lanman** , A. (2018). Improved 21 cm Epoch of Reionization Power Spectrum Measurements with a Hybrid Foreground Subtraction and Avoidance Technique. *The Astrophysical Journal* , 864(2):131.

Conference Talks

Science at Low Frequencies

2016 Precision Simulations of Cosmic Dawn Experiments

2019 pyuvsim: A comprehensive radio interferometry simulator in Python

URSI National Radio Science Meeting

2017 Precision Simulations of Cosmic Dawn Experiments

2018 Realistic EoR Instrument Simulations with FHD

American Astronomical Society Meetings

Jan. 2019 Recovering 21cm Power Spectra in Simulated Delay Spectrum Pipelines

Teaching

Instructor:

CEPI 0905, *From the Solar System to the Universe: An Introduction to Astrophysics and Cosmology*, Summer 2015 and 2016.

This Summer@Brown course surveyed the major topics of astrophysics and cosmology for a group of advanced high-school students.

Teaching Assistant (TA) at Brown University:

Phys 0270, *Introduction to Astronomy*, Fall 2013, Instructor: Greg Tucker

Phys 0220, *Astronomy*, Spring 2014, Instructor: Savvas Koushiappas

Phys 2010, *Techniques in Experimental Physics*, Fall 2014, Instructor: Gang Xiao

Phys 0220, *Astronomy*, Spring 2015, Instructor: Greg Tucker

Administration and Service

(2016-2018) Organized and hosted the *Brown Astrophysics Seminar Series*, bringing in speakers from around New England for bi-weekly talks on astrophysics.

(2017-2018) Co-founded the *Physics Organization* and organized its community development activities, including:

A peer-mentoring program offering supplemental guidance to first-year graduate students.

Career skills workshops offered to undergraduates.

(2018) Coordinated and ran the *Sidewalk Astronomy* outreach program, bringing telescopes to Providence WaterFire events to promote astronomy to the public.

(2019) Peer-review services for the Astrophysical Journal.

Past Research

With Ulrich Heintz and the CMS Collaboration, Brown University

Tests of neutron-irradiated silicon strip sensors for properties of charge sharing after long beta-radiation exposure. (March 2015 – December 2015)

Analysis of CMS data for tri-lepton events. (August 2015 – December 2015)

With Savvas Koushiappas, Brown University:

Implementing a new technique for detecting gamma-ray pulsar contributions to the diffuse gamma-ray background in Fermi LAT data, (unpublished), June – September 2014.

Acknowledgements

The work presented here was made by possible by the generous effort of many wonderful people, and I am forever grateful to all of them for their guidance, expertise, and support. My adviser, Jonathan Pober, is an exemplary mentor and teacher, and the list of ways he has helped me grow as a scientist and as a professional is too long to recount – *please* give him tenure as soon as possible! Thank you as well to Ian Dell'Antonio and Greg Tucker for being on my thesis and prelim committees, and for various recommendation letters and other support over the years, and to all the faculty for their help with finding speakers for BASS. Thank you also to Bob Horton for supporting the sidewalk astronomy program and keeping me involved with outreach opportunities.

I must also give particular thanks to Bryna Hazelton for teaching me so much about collaborative software development, and of course for reviewing all of my pull requests. Thanks as well to Danny Jacobs, Matt Kolopanis, Nithya Thyagarajan, James Aguirre, and all other RASG developers for their efforts in making `pyuvsim` happen. Thank you to Nichole Barry and Ian Sullivan for their help with FHD and eppsilon. Overall, working with the HERA and MWA collaborations has been a rewarding experience that I can only hope to find again.

In the lab, I'd like to say thank you to Josh Kerrigan, Wenyang Li, Jacob Burba, Theodora Kunicki, Daniya Seitova, Zhang Zheng, and Peter Sims for being great coworkers and occasional distractions. Best of luck to all of you! Outside the lab, the Brown Physics community has been a consistently welcoming and supportive environment – Thanks to Ben Weiner, Liz Baesler, and all the other FP folks for hosting art and movie nights.

Finally, I have to thank my family for their love and support. This includes new – Lisa, Peter, Tyler, Sandy; old – Ben, Nate, and Papa; and four-legged – Callie, Zoe, Boo, and Lola.

Above all, I am grateful to Jessica, who I am honored to call my wife, for all she has done for me through this journey.

Contents

List of Tables

List of Figures

Chapter 1

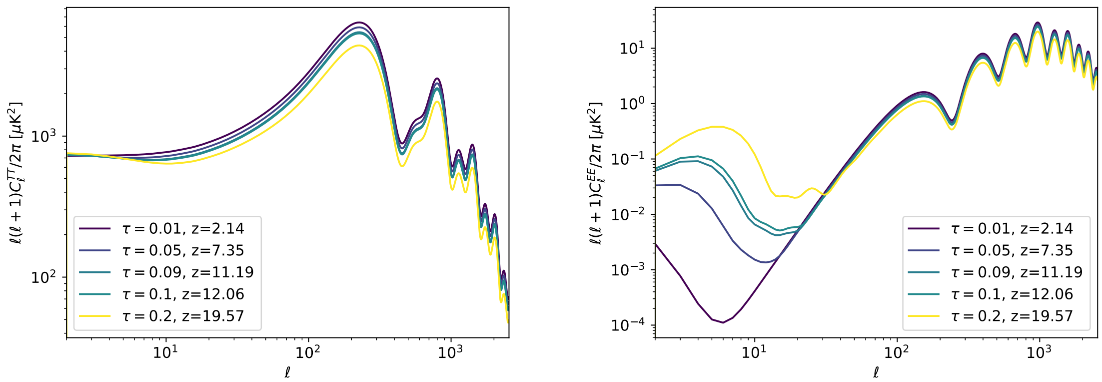
The Cosmic Dawn

The history of the universe may be succinctly described as a gradual change from a simple, uniform mixture of materials, in complete thermal equilibrium, into a web of complex hierarchical structures. This structure arose from small perturbations to the homogeneous intial state of the universe, as precision cosmology research has revealed over the last few decades. Observations of the Cosmic Microwave Background (CMB) have revolutionized our understanding of the early universe, because the CMB remains imprinted with the structures present when it decoupled from baryonic matter at $z \sim 1100$ (Dodelson, 2003). Surveys of galaxies and weak gravitational lensing trace the structure matter at later times, forming a complementary picture of the evolution of structure for $z \lesssim 6$ (Becker et al., 2015).

These existing cosmological probes leaves an enormous gap of cosmological time largely unexplored: the *Dark Ages* after the CMB decoupling epoch, when the baryonic content of the universe consisted of almost entirely neutral gas, and the *Epoch of Reionization* (EoR), when the first galaxies formed and emitted UV radiation that ionized the intergalactic medium (IGM). Together, these eras comprise the *Cosmic Dawn*, a period that spans roughly 900 Myrs of cosmological time yet lacks detailed observation. This is largely because the cosmic dawn lacks sources bright enough to be detected at such distance, and the intervening gas is optically thick, quickly extinguishing the light of whatever sources there may be.

Information about the overall extent of reionization can be inferred from the cumulative effect of the post-EoR (ionized) universe on the CMB. Free electrons in the ionized IGM following reionization scatter CMB photons via Thomson scattering, introducing an optical depth to the CMB related to the integrated column density of free electrons along the line of sight (Reichardt, 2016),

$$\tau = \int_0^{z_r} n_e(z)\sigma_T \frac{c}{(1+z)H(z)} dz \tag{1.1}$$



(a) C_ℓ^{TT} (b) C_ℓ^{EE}

Figure 1.1: Angular power spectra of the CMB for varied optical depth τ, as generated by CAMB. (a) shows the temperature-temperature angular spectra, while (b) shows the E-E power. Each curve corresponds with a different optical depth, and mean redshift of reionization. These were made using the Code for Anisotropies in the Microwave Background (CAMB) (Lewis et al., 2000; Howlett et al., 2012).

where n_e is the density of free electrons, σ_T is the Thomson scattering cross section, and the remaining term in the integrand is the differential proper distance along the line of sight. The parameter that can be determined from this is z_r, the redshift at which reionization ended. Increasing this redshift implies that reionization finished earlier and leads to a larger optical depth. In this form, it is assumed that reionization occurred abruptly, such that any partially ionized IGM from $z > z_r$ is negligible. There is some degeneracy between the duration of reionization and the redshift that fits the definition in eq. (1.1). In addition, the effects of the optical depth on the TT power spectrum are degenerate with variations in the amplitude of the primordial power spectrum, and so an additional probe is needed to break the degeneracy. The E-mode polarization of the CMB provides this.

The net effect of isotropic scattering such, as Thomson scattering, is to smooth anisotropies (Hu & Sugiyama, 1995). Angular scales that entered the horizon before reionization are subject to this smoothing, such that the effect of this optical depth is to suppress power on scales smaller than the horizon at z_r. This is demonstrated in Fig. 1.1a, which shows the impact of varying the z_r and τ on the temperature-temperature CMB angular power spectrum. Overall, the earlier reionization took place, the longer the ionized universe has had to suppress power and the larger the ℓ mode where the suppression starts to take effect.

Reionization also affects the E-mode polarization of the CMB. Thomson scattering imparts a net linear polarization on light, provided the radiation field has a nonzero quadrupole anisotropy (Reichardt, 2016). This anisotropy can exist on scales larger than the horizon, so the ionized universe post-reionization sees a quadrupole on scales larger than the horizon at reionization, and can introduce a net polarization. This amplifies the E-E angular power spectrum of the CMB on small ℓ scales. This can be seen in Fig. 1.1b, which shows strong variation in the power spectrum on low-ℓ scales in response to changing optical depth.

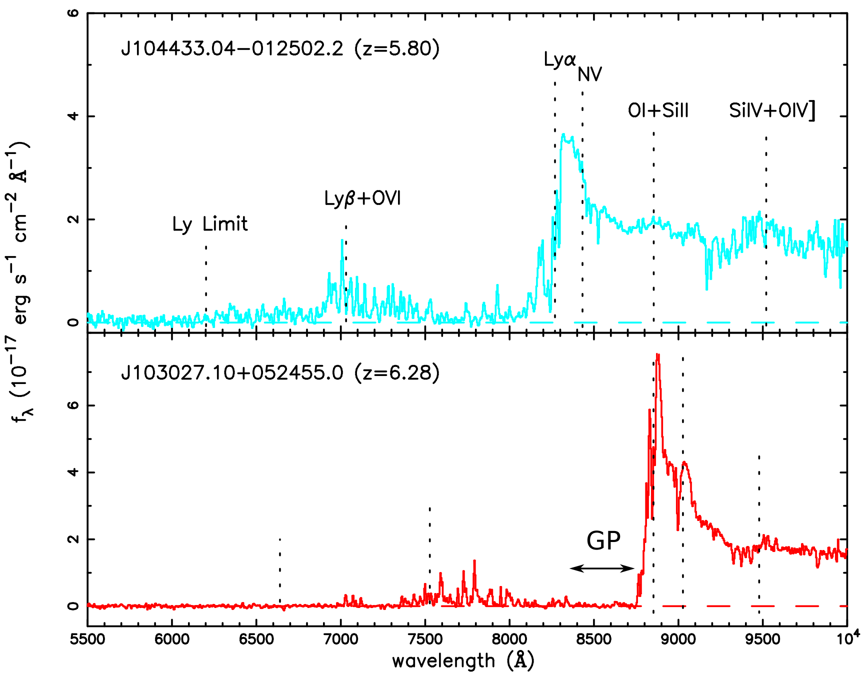


Figure 1.2: Adapted from Figure 1 of (Becker et al., 2001), these are spectra of two SDSS quasars. The vertical dashed lines mark the positions of significant emission lines. The lower redshift quasar spectrum on top shows a more gradual dropoff in flux to the left of the Ly-α peak, compared with the $z \sim 6.28$ quasar which has a complete lack of flux left of the Ly-α peak. This is characteristic of a Gunn-Peterson trough.

The late stages of reionization can be observed via Lyman-α absorption in the spectra of quasars and distant galaxies. The frequency at which Ly-α absorption occurs directly maps to the distance to the neutral gas that absorbed it, under a given cosmological model, so the absorption features in a broad specrum bright source like a quasar provides details on the distribution of opaque clouds along the line of sight. In the immediate vicinity of a quasar, gas is ionized and hence does not absorb Ly-α radiation, leaving a spike in the spectrum. If the quasar is embedded in a mostly neutral IGM, however, this ionized pocket has a finite extent. The neutral gas beyond this ionized region strongly absorbs at the Ly-α transition, creating a strong suppression in the spectrum called

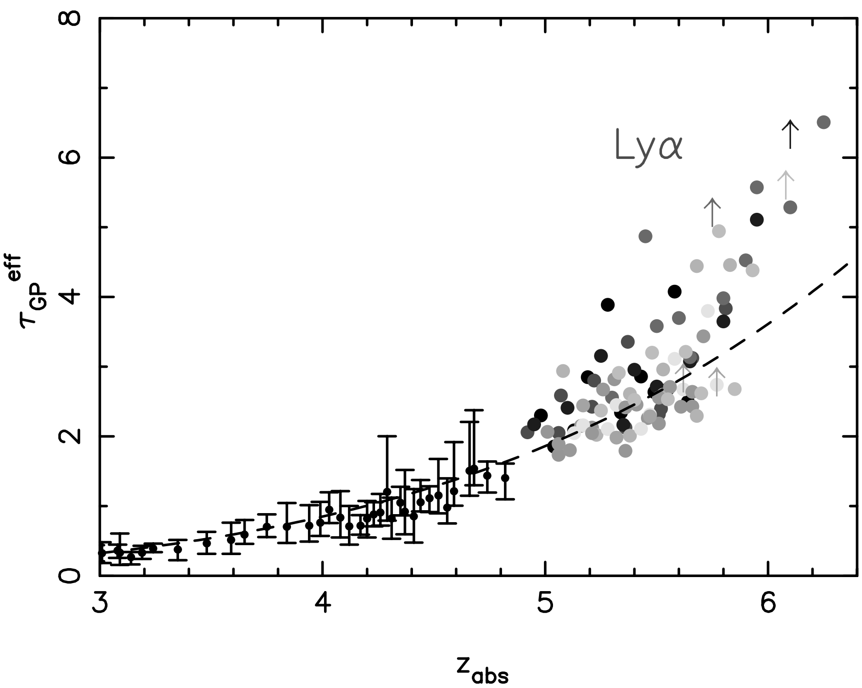


Figure 1.3: Mean effective Ly-α optical depth vs. redshift from Ly-α forest measurements of a sample of 19 $z > 6$ quasars, reproduced from Fan et al. (2006). At $z > 5.7$, the optical depth diverges from the trend at lower redshift, demonstrating the sudden presence of an optically-thick IGM.

a Gunn-Peterson trough (Gunn & Peterson, 1965).

The first detection of a Gunn-Peterson trough was claimed by Becker et al. (2001), observed in a high-z quasar spectrum from the Sloan Digital Sky Survey (SDSS). This spectrum is shown in Fig. 1.2, in comparison with a $z = 5.8$ quasar spectrum. The strong damping at shorter wavelengths (hence, smaller redshift) of the Ly-α peak implies the presence of a neutral IGM between the quasar and the observer. Subsequent study with more samples, probing different lines of sight, estimated the mean neutral fraction at the end of the EoR. Fig. 1.3, reproduced from Fan et al. (2006), shows the effective Ly-α optical depth at a variety of redshifts. This is obtained from measuring the depths of Ly-α absorption lines from a sample of 19 quasars. Absorption above $z \sim 5.7$ consistently saturates, indicating that the intervening IGM is at least partially neutral.

Quasar absorption spectra and CMB measurements have provided some of the best constraints on reionization so far. Planck Collaboration et al. (2016) measured $\tau = 0.058 \pm 0.012$, assuming a rapid reionization model, and estimated the redshift of reionization $z_r = 7.8 \pm 0.9$, with a maximum duration $\Delta z < 2.8$ when combined with data from the South Pole Telescope (George et al., 2015).

The 2011 detection of the quasar ULAS J1120+0641 at $z = 7.1$ (Mortlock et al., 2011), the farthest yet found, shows a Ly-α damping wing consistent with a partially-ionized intervening IGM. Bayesian inference from simulations showed that the observed damping is consistent with a neutral fraction of $\bar{x}_{\rm HI} = 0.40^{+0.21}_{-0.19}$ (Greig et al., 2016).

Though useful information can be obtained from quasar and CMB measurements, these are fundamentally limited probes in several ways. CMB measurements can only look at quantities integrated along the line of sight, providing very little information on the evolution of structure and timing of key processes in reionization. Quasar absorption can probe along the line of sight well, but is limited to the handful of sightlines of high-z quasars that have been detected. Further, since the Ly-α absorption saturates so quickly, it can only really probe the latest stages of reionization.

1.1 HI Intensity Mapping

The most promising tool for studying these eras is to map the emission of the neutral IGM itself through the 21 cm hyperfine transition of neutral hydrogen (HI). This is a transition between the singlet and triplet states of the hydrogen $n = 1$ ground state, which is split into two energy levels by the interactions of the spin magnetic moments of the proton and electron. The transition between the two hyperfine states emits a photon with frequency 1420.4 MHz ($\lambda = 21.1$ cm) in its rest frame. The rate of the triplet to singlet transition is given by the Einstein-A coefficient $A_{10} = 2.85 \times 10^{-15}\ \mathrm{s}^{-1}$, meaning a single triplet state has a lifetime on the scale of 10^7 years. Nonetheless, the enormous quantity of hydrogen in the universe allows this to produce a detectable signal. Further, since it is a line transition, the redshift of 21 cm emission/absorption directly corresponds to distance when cosmic expansion is the dominant factor. In this way, mapping the distribution of redshifted 21 cm brightness can map the three-dimensional structure of neutral hydrogen during the Cosmic Dawn.

Constraints from CMB and quasar probes imply that the EoR 21 cm signal will be redshifted to wavelengths of $\sim 1.2 - 6.0$ m, or frequencies of $50 - 250$ MHz. Thus, from Earth, 21 cm emission and absorption from the EoR is detectable by radio telescopes. To better understand the signals that may be observed, this section will derive the expected radio brightness, in connection with the physics of reionization, and discuss the challenges to its detection.

1.1.1 The 21 cm Line

To see how the observable signal connects to the properties of the gas, I will follow the steps of Pritchard & Loeb (2012) to derive the brightness of the 21 cm line from an HI cloud, and relate that to the detectable contrast in brightness against a background source as seen on Earth from cosmological distances. The source of 21 cm emission within the cloud is set by the relative occupation of the triplet and singlet states, characterized by the spin temperature T_S,

$$\frac{n_1}{n_0} = 3e^{T_{21}/T_S} \tag{1.2}$$

where $T_{21} = h\nu_{21}/k_B \approx 0.0681$ K is the temperature associated with the 21 cm line (for Boltzmann constant k_B and Planck constant h) and n_1/n_0 are the number densities of triplet/singlet state hydrogen, respectively. The factor of 3 accounts for the degeneracy of the triplet state.

The radiative transfer equation is written in terms of specific intensity of light I_ν and a source S_ν as

$$\frac{dI_\nu}{d\tau_\nu} = -I_\nu + S_\nu, \tag{1.3}$$

where τ_ν is the optical depth. For this case, the source of radiation is the CMB, so $S_\nu = B_\nu(T)$, the Planck function for blackbody emission. I_ν has units of power per unit area per solid angle per frequency (e.g., W m^{-2} sr^{-1} Hz^{-1}). The Planck function describes the background radiation field, which is just the CMB. In the Rayleigh-Jeans limit, there is a linear relationship between temperature and specific intensity,

$$I_\nu = T_B \frac{2k_B\nu^2}{c^2},$$

where c is the speed of light. The *brightness temperature* T_B is thus equivalent to the specific intensity, but directly comparable to other physical properties. Expressing I_ν in terms of T_B and replacing $B_\nu(T)$ with the Rayleigh-Jeans law, eq. (1.3) simplifies to

$$\frac{dT_B}{d\tau_\nu} = -T_B + T_\gamma, \tag{1.4}$$

where T_γ is the CMB temperature. This has the solution

$$T_B(\nu) = T_S\left(1 - e^{-\tau_\nu}\right) + T_\gamma e^{-\tau_\nu}, \tag{1.5}$$

where $T_S = T_B(\tau_\nu = 0)$. To summarize, the relative occupations of the two hyperfine states within the gas will remain constant, described by T_S, in the absence of any emission or absorption into the cloud. The factor of $(1 - e^{-\tau_\nu})$ describes the probability of 21 cm photons being emitted from the cloud, and $e^{-\tau_\nu}$ gives the probability of CMB photons being absorbed at the 21 cm line.

The optical depth is given by the 21 cm absorption cross section, σ_{01}, and a line width function $\phi(\nu)$ defined such that $\int \phi(\nu)d\nu = 1$. The cross section constant σ_{01} is related to the spontaneous emission Einstein coefficient A_{10} via (Hilborn, 1982):

$$\sigma_{01} = \frac{3c^2 A_{10}}{8\pi\nu^2}$$

which can be integrated through the cloud to obtain the optical depth. From here, most authors evaluate the optical depth by assuming the line width is dominated by velocity broadening, where the velocity is due to Hubble expansion. $\phi(\nu) \sim c/(LH(z)\nu)$ for a region of linear extent L (Furlanetto et al., 2006b; Pritchard & Loeb, 2012; Zaroubi, 2013). The result, inserting the full form of σ_{10}, is

$$\begin{aligned}\tau_\nu &= \frac{3}{32\pi}\frac{hc^3 A_{10}}{k_B T_S \nu^2}\frac{x_{HI} n_H}{(1+z)(dv_\parallel/dr_\parallel)} \\ &\approx 0.0092(1+\delta_B)(1+z)^{3/2}\frac{x_{\rm HI}}{T_S}\left[\frac{H(z)/(1+z)}{dv_\parallel/dr_\parallel}\right]\end{aligned} \tag{1.6}$$

Here, $x_{\rm HI}$ is the fraction of neutral hydrogen, n_H is the total number density of hydrogen, $(1+\delta_B)$ is the fractional overdensity of baryons, and $dv_\parallel/dr_\parallel$ is the gradient of proper velocity along the line of sight (accounting for peculiar velocities). For the second line, the spin temperature must be expressed in Kelvin.

The quantity of interest is the *differential brightness temperature* $\delta T_B = T_B - T_{\rm bg}$, which is the observed contrast between the brightness temperature and a radiating background $T_{\rm bg}$. For the Cosmic Dawn, the background is just the CMB. From eq. (1.5), this is

$$\delta T_B = \frac{T_S(z) - T_\gamma(z)}{1+z}\left(1 - e^{-\tau_\nu}\right) \approx \frac{T_S(z) - T_\gamma(z)}{1+z}\tau_\nu, \tag{1.7}$$

where the $(1+z)$ factor accounts for the effect of cosmological expansion on the observed temperature, and we assume that τ_ν is small in the second step. Substituting in the optical depth expression and a full expression for $H(z)$, this becomes:

$$\delta T_B \approx (9\text{ mK})\, x_{\rm HI}\, (1+\delta_B)(1+z)^{1/2}\left[1 - \frac{T_\gamma(z)}{T_S}\right]\left[\frac{H(z)/(1+z)}{dv_\parallel/dr_\parallel}\right] \tag{1.8}$$

This gives the observable brightness contrast between the 21 cm spin temperature and the CMB as a function of redshift, depending on the baryon density, ionization fraction, and velocity gradients. In the following subsection, I will discuss the various factors that set the spin temperature over cosmological time, and show how this gives affects the *global*, or sky-averaged, brightness temperature. In the subsequent subsection, I will discuss the spatial statistics of the 21 cm signal and the power spectrum.

1.1.2 Global Signal

Averaging the differential brightness temperature over the whole sky gives the global signal, which depends on the spin temperature and neutral fraction $x_{\rm HI}$. The global signal is nonzero whenever the spin temperature T_S deviates from the CMB temperature T_γ and the neutral fraction is nonzero. The existence of a differential brightness signal depends on whether the spin temperature is coupled to the gas kinetic temperature T_K or the CMB T_γ. At early times, the gas expands adiabatically

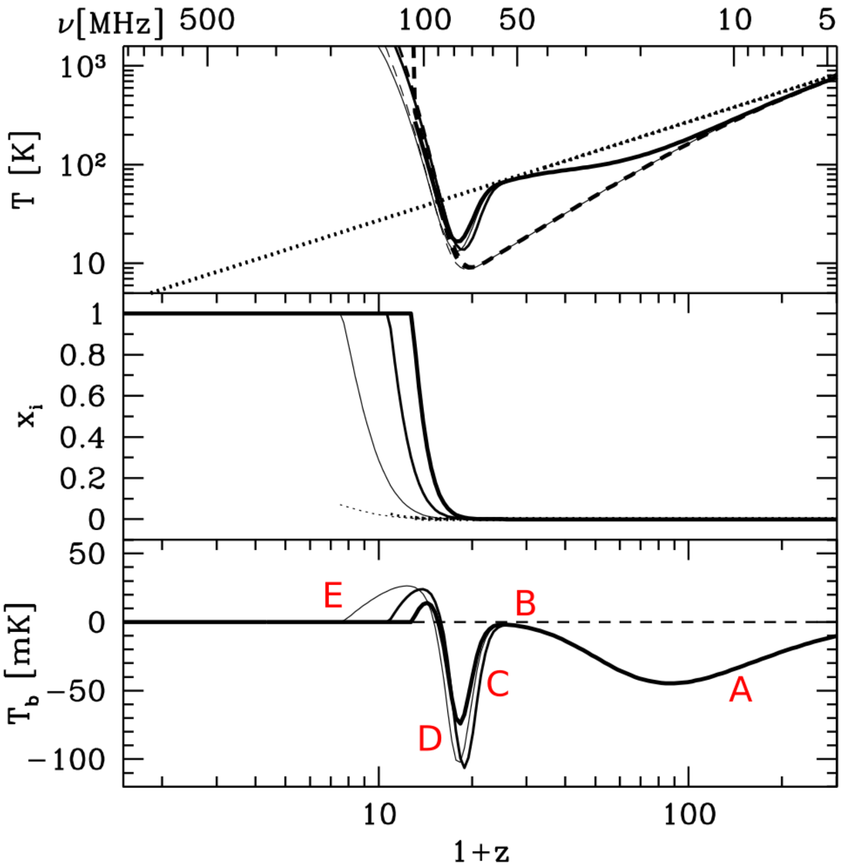


Figure 1.4: Top: CMB temperature (dotted), gas kinetic temperature (dashed), and spin temperature (solid) for three EoR models. Middle: Corresponding ionized fraction of $x_i = (1 - x_{\rm HI})$ for the three models. Bottom: Differential brightness temperature. Reproduced from Pritchard & Loeb (2008). I have added letters indicating the various key turning points listed in the text.

such that $T_K \propto (1+z)^2$, while $T_\gamma \propto (1+z)^1$, so coupling to the kinetic temperature generally produces an absorption signal.

There are a few coupling mechanisms governing the evolution of T_S. When the gas density is high, frequent collisions among hydrogen atoms shift singlet hydrogens into triplet states. This collisional coupling sets $T_S \sim T_K$. Weak radiative coupling, due to the interactions between CMB photons and hydrogen atoms, drives $T_S \sim T_\gamma$. When strong Ly-α radiation is present, it pumps ground state hydrogen into excited states, which can end up in a different hyperfine states when it decays back down. This is known as the Wouthuysen-Field effect (Wouthuysen, 1952; Field, 1959), and it causes the spin temperature to couple to the Ly-α color temperature (see, e.g., section 2.3

of Furlanetto et al. (2006b)).

The evolution of relevant quantities is shown in Fig. 1.4, including the spin, CMB, and kinetic temperatures and the ionized fraction. Several turning points are also marked, at which certain physical processes take effect:

A IGM gas is dense, so collisional coupling keeps $T_S \sim T_K$. An absorption signal is seen.

B Collisional coupling becomes weaker than radiative coupling, and $T_S \sim T_\gamma$. The signal vanishes.

C The first stars ignite, producing a strong Ly-α background that couples $T_S \sim T_K$ via the Wouthuysen-Field effect. An absorption feature is seen.

D X-ray emission heats the IGM rapidly. Wouthuysen-Field coupling keeps $T_S \sim T_K$ as T_K grows, passing T_γ and producing an emission signal.

E As the IGM is ionized, $x_{\mathrm{HI}} \to 0$ and 21 cm emission diminishes.

1.1.3 Power Spectrum

One of the main difficulties in observing the global signal is that it evolves relatively slowly over frequency, making it difficult to distinguish its smooth turning points from the smooth spectral signatures of foreground emission and other instrumental effects (I will say more about foregrounds in section 1.2). In contrast, the spatial variations in the 21 cm brightness show relatively sharp features – discrete regions with distinct heating and Ly-α emission brightnesses impart fluctuations in the 21 cm brightness. Mapping these fluctuations would provide a clear image of the IGM during reionization, but requires rather large instruments to achieve (see the following subsection). A more feasible first step is to measure the statistics of the 21 cm fluctuations, such as the power spectrum. Under the cosmological principle, these quantities can be averaged into one dimensional expressions, increasing signal to noise by combining multiple samples of each Fourier mode.

The fluctuation in the 21 cm temperature may be defined as the deviation of $\delta T_B(z)$ from its sky average, $\langle \delta T_B(z) \rangle$[1],

$$\delta_{21}(\boldsymbol{x}) = \delta T_B(\boldsymbol{x}, z) - \langle \delta T_B(z) \rangle \,, \tag{1.9}$$

where I've written in the explicit dependence on position $\boldsymbol{x}$. As shown in eq. (1.8), δT_B depends on several spatially-varying quantities, such as the spin temperature T_S, the neutral fraction x_H, the baryon overdensity δ_B, and the peculiar velocities $dv_\parallel / dr_\parallel$.

[1] This is usually divided through by the mean temperature fluctuation as well, leaving a dimensionless quantity that can be easily compared with other dimensional overdensity quantities. In measurement, however, the fluctuation is defined in temperature units

The spatial Fourier transform is defined,

$$\widetilde{\delta}_{21}(\boldsymbol{k}) = \int \delta_{21}(\boldsymbol{x}) e^{-i\boldsymbol{x}\cdot\boldsymbol{k}} d^3x \tag{1.10}$$

$$\delta_{21}(\boldsymbol{x}) = \frac{1}{(2\pi)^3} \int \widetilde{\delta}_{21}(\boldsymbol{k}) e^{i\boldsymbol{x}\cdot\boldsymbol{k}} d^3k \tag{1.11}$$

where $\boldsymbol{k}$ is a wave vector, and the integrals are taken over all of space or Fourier space, respectively. The power spectrum is defined as the ensemble average of the product of the Fourier-transformed fluctuations:

$$\left\langle \widetilde{\delta}_{21}(\boldsymbol{k}) \widetilde{\delta}^*_{21}(\boldsymbol{k}') \right\rangle = (2\pi)^3 \delta_D(\boldsymbol{k} + \boldsymbol{k}') P(\boldsymbol{k}) \tag{1.12}$$

Note that under the Fourier convention used here, $P(\boldsymbol{k})$ has units of temperature squared times volume, e.g. mK^2 Mpc^3. It is common to rescale this by $k^3/2\pi^2$ for $k = |\boldsymbol{k}|$, producing the *dimensionless* power spectrum $\Delta^2(k)$.[2] In our convention, this has units of temperature squared:

$$\Delta^2(k) = \frac{k^3}{2\pi^2} P(k)$$

Cosmological isotropy implies that $P(\boldsymbol{k}) = P(k)$, allowing one to average over directions of the wave vector. In reality, redshift space distortions, included in eq. (1.8) as $dv_\parallel / dr_\parallel$, break this isotropy. For instance – an overdense region will have gas infalling into it, increasing the redshift along the line of sight slightly, and making the region appear denser along the line of sight than transverse to it. It turns out that these anisotropies can be used to separate some of the astrophysical factors that determine P_{21} (see e.g., Barkana & Loeb (2005)).

Fig. 1.5 shows the redshift evolution of power spectra for a typical EoR model across the later stages of reionization (between points D and E in Fig. 1.4). At earlier times (i.e., higher redshift), power shifts slightly from large scales to smaller as structure collapses, causing the small k side of the curves to drift downward with decreasing z. As ionization takes over, power on small scales is lost, causing a rapid drop on the large k end.

1.2 Challenges

The ionized bubbles of the EoR are expected to appear on scales of tens of Mpc at the later stages of reionization. Taking the characteristic bubble size to be 20 Mpc at $z = 8$ (see Fig. 1 of Furlanetto et al. (2006a)), this corresponds with an angular order of about 6 arcmin and requires an interferometer baseline about 76 km to resolve. Thus, imaging EoR features will likely require an instrument of the scale of the upcoming Square Kilometre Array (SKA; Mellema et al. 2013; Furlanetto et al. 2006b).

[2] The dimensionless power spectrum is usually only used after spherically-binning in k.

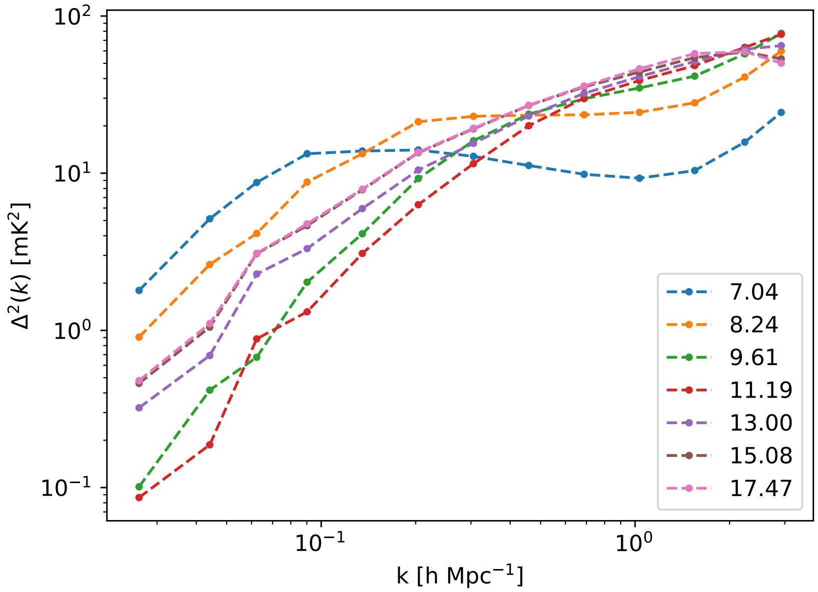


Figure 1.5: Power spectra at several redshifts approaching the end of reionization, produced from a default-parameter simulation with the seminumerical code `21cmFAST` (Mesinger et al., 2011). As redshift decreases, ionized bubbles grow, reducing power on small scales (large k).

Current generation experiments seek to measure the power spectrum and global signal. Radio interferometers are well suited to power spectrum measurements, since they directly sample the angular Fourier transform of the sky, and can achieve high spectral resolution, which provides good sampling of line-of-sight Fourier modes. Many of these instruments feature highly redundant, compact layouts that maximize sensitivity to large angular scales, aiming to detect the EoR signal through their fine spectral resolution first (Parsons et al., 2012a). Though an actual detection of the power spectrum has yet to be made, experiments like the GMRT (Paciga et al., 2013, 2011), MWA (Barry et al., 2019; Beardsley et al., 2016; Bowman et al., 2013), PAPER (Kolopanis et al., 2019; Cheng et al., 2018; Parsons et al., 2010), and LOFAR (Patil et al., 2017), have placed upper limits on the power spectrum amplitude at several redshifts. The best upper limits on power spectrum amplitude are shown in Fig. 1.6, which shows amplitude vs k at different redshifts.

The primary obstacle to these experiments is contamination of bright foreground sources, such as Galactic synchrotron emission, free-free emission, radio galaxies, pulsars, and supernova remnants, which are typically five orders of magnitude brighter than the expected EoR signal (Furlanetto et al., 2006b; Oh & Mack, 2003; Matteo et al., 2002). Fortunately, the smooth spectra exhibited by these sources distinguish them from the relatively complex structure of EoR signals, making it

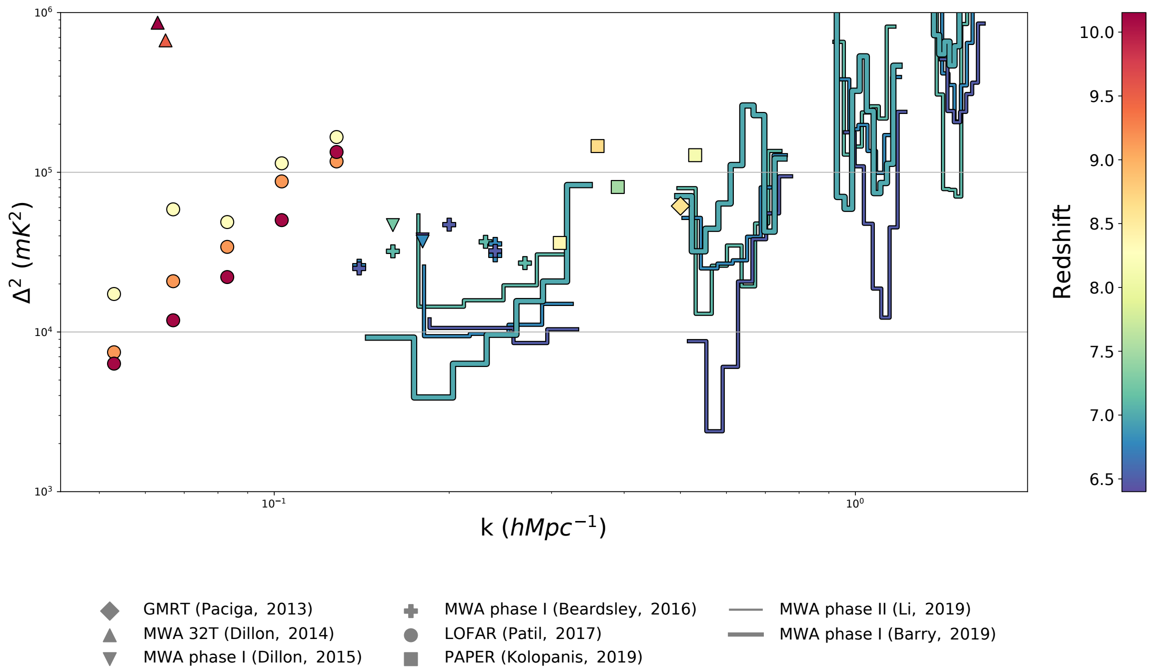


Figure 1.6: Upper limits on the power spectrum amplitude at different redshfits (color) and k modes (x axis). Plotted using `https://github.com/EoRImaging/eor_limits`.

possible – in principle – to model and subtract them from data (Oh & Mack, 2003; Liu & Tegmark, 2012; Chapman et al., 2015; Sims et al., 2016; Carroll et al., 2016) or avoid them in Fourier space (Parsons et al., 2012b,a; Pober et al., 2013b; Chapman et al., 2016). Since frequency maps directly to redshift and distance, this separation occurs in Fourier space parallel to the line of sight ($k_{\|}$). Smooth-spectrum power will cluster around $k_{\|} = 0$, while EoR signal power will spread farther out.

For global signal experiments, this spectral smoothness implies that the foreground spectrum can be fit by a low order polynomial and subtracted. The Experiment to Detect the Global EoR Signal (EDGES) claimed to have detected an absorption signature at 78 MHz (corresponding with point D in Fig. 1.4), having subtracted a five term polynomial foreground model from their measured spectrum. They verified the detection by comparing two independent processing pipelines, comparing results from different hardware configurations, and checking subsets of the total observing time to look for variations (as an isotropic signal, it should be present at all times). The measured absorption feature is much deeper than would be expected from theory, leading to a flurry of speculation about why that might be, and a handful of responses arguing that the absorption could be due to a systematic error such as an unmodeled spectral ripple (Hills et al., 2018) or the effects of soil conductivity under the antenna (Bradley et al., 2019). The EDGES detection cannot be truly confirmed until another global experiment detects a similar feature. Since this experiment is the cutting-edge of global experiments, it will take time for others to catch up to the level of precision EDGES has achieved.

Current power spectrum experiments are being developed in tandem with complementary analysis methods, which allow for quicker confirmation of a claimed detection. In the meantime, each published measurement sets a deeper upper limit on the amplitude of the power spectrum at different k-modes and redshifts, as shown in Fig. 1.6. Improved upper limits come with more effective foreground removal strategies, better calibration, and more data. With each limit, however, we must be certain that the analysis procedure itself has not introduced unexpected artifacts or, in the worst case, potentially subtracted some of the signal power away. To claim to have set an upper limit on the EoR, one must show that the foreground subtraction technique used could not have acccidentally removed the EoR signal itself. This signal loss problem led to the retraction of an upper limit from the PAPER experiment (Ali et al., 2015), with subsequent (improved) analysis providing the new limits presented in (Kolopanis et al., 2019) (and shown in Fig. 1.6).

Since then, the HERA collaboration has put a more concentrated effort into systematically verifying analysis and power spectrum estimation software against simulations of foreground sources, EoR signal models, and systematic effects such as cable reflects, cross-talk, and thermal noise. In this thesis, I present my work in developing new simulation tools and applying them to studying foreground contamination and EoR signals. In the following chapter, I introduce the theory behind radio interferometry and present two common forms of estimating power spectra from intererometric data. In chapter 3, I describe two new instrument simulators, `pyuvsim` and `healvis`, which have come to play important roles within the HERA collaboration. In chapter 4, I derive several analytic models for testing simulators of diffuse maps, and use them to test the accuracy and precision of `healvis`. In chapters 5 and 6, I use `healvis` to study the behavior of diffuse foreground sources and an EoR-like signal, respectively, as a means of understanding what can be measured and how it may be used for precision cosmology.

Chapter 2

Radio Interferometry

The simulation software described in chapter 3 evaluates the Radio Interferometry Measurement Equation (RIME) in full, including the effects of polarization and direction-dependent gain. In this chapter, I will follow the steps in Clark (1999) to derive the RIME from basic optics, extending the derivation by including the effects of polarization. This derivation is also similarly presented in Liu & Shaw (2019). From there, I will show how this fully polarized form reduces to the more commonly seen scalar expressions under some basic assumptions, and how those scalar expressions may be used to estimate the power spectrum. These power spectrum estimators will be used in chapters 5 and 6 with simulated data.

2.1 Measurement Equation

The electromagnetic radiation emitted from an astronomical object has the form of a time-varying electric field, $\boldsymbol{E}(\boldsymbol{\rho}, t)$, where $\boldsymbol{\rho}$ is a vector describing the position of the source. Equivalently, this may be expressed as a set of *quasi-monochromatic* components $\boldsymbol{E}_\nu(\boldsymbol{\rho})$, for frequencies ν, by taking the Fourier transform of the E-field in time. Assuming that the intervening space is empty vacuum, the propagation of this electric field from the source at $\boldsymbol{\rho}$ to a receiver at $\boldsymbol{r}$ is given by Huygens' principle, which says that the electric field propagates as a spherical wave:

$$\boldsymbol{E}_\nu(\boldsymbol{r}) = \int \boldsymbol{E}_\nu(\boldsymbol{\rho}) \frac{e^{2\pi i \nu |\boldsymbol{\rho} - \boldsymbol{r}|/c}}{|\boldsymbol{\rho} - \boldsymbol{r}|} d^3\rho \tag{2.1}$$

The integral is taken over all of space. To an observer on the Earth, the sky appears as a spherical surface with a varying "surface brightness." We may describe the sky as a "celestial sphere" at a large distance R, such that $\boldsymbol{\rho} = R\hat{s}$ where $\hat{s}$ is a unit vector. Letting $\boldsymbol{\mathcal{E}}_\nu(\hat{s})$ represent the surface distribution of the electric field, eq. (2.1) becomes

$$\boldsymbol{E}_\nu(\boldsymbol{r}) = \int \boldsymbol{\mathcal{E}}_\nu(\hat{s}) \frac{e^{2\pi i\nu|R\hat{s}-\boldsymbol{r}|/c}}{|R\hat{s}-\boldsymbol{r}|} R^2 d\Omega, \tag{2.2}$$

where $d\Omega$ is a differential surface area. The integral is taken over the whole sphere. However, before simplifying further we need to account for the directional sensitivity of the antenna receiving the emission.

Each antenna in an interferometer consists of a set of *feeds* that are sensitive to orthogonal polarizations of the incoming light. In reality, these feeds may not be perfectly orthogonal and each may have some sensitivity to the other polarization state, which will cause an unpolarized signal to appear polarized. In this derivation, we assume ideally orthogonal feeds. A time-varying voltage is induced in each feed by incident electric fields. Letting p and q denote two orthogonal feeds, the voltages induced in each feed may be written as a vector $\boldsymbol{v} = (v_p, v_q)$. The contribution to the voltage from a single source is

$$\boldsymbol{v} = \mathbf{J}(\hat{s}, \nu)\, \boldsymbol{\mathcal{E}}_\nu(\hat{s}), \tag{2.3}$$

where $\mathbf{J}$ is a 2×2 *Jones* matrix, which gives the relative sensitivity of the antenna feeds to each electric field component in the direction $\hat{s}$. Specifically, this Jones matrix is made up of the components of the *E-field primary beam* of the antenna (or just the primary beam), and just represents the effect of summing incident waves from across the antenna onto a focal point at position $\boldsymbol{r}$. As its name suggests, the E-field beam is related to an electric field pattern in the far field. I will say more about this in section 2.1.2.

Additional Jones matrices may be introduced to account for other linear effects on the incoming signal, or within the signal chain of the instrument. For example,

$$\mathbf{G}_i = \begin{pmatrix} g_{ip} & 0 \\ 0 & g_{iq} \end{pmatrix}$$

describes the gains in each feed of antenna i, and

$$\mathbf{D}_i = \begin{pmatrix} 0 & d_{ip} \\ -d_{iq} & 0 \end{pmatrix}$$

describes the errors of feeds that are not ideally sensitive to orthogonal polarizations (Hamaker et al., 1996). Faraday rotation due to the ionosphere can also be describe by a direction-dependent rotation matrix. For now, we will focus on the primary beam Jones matrix, with the understanding that additional effects can be included within this formalism.

Since the beam Jones matrix is direction-dependent, it must be included in the integral when summing contributions from multiple sources. Hence, eq. (2.2) becomes

$$\boldsymbol{v} = \int \mathbf{J}(\hat{s}, \nu)\, \boldsymbol{\mathcal{E}}_\nu(\hat{s}) \frac{e^{2\pi i\nu|R\hat{s}-\boldsymbol{r}|/c}}{|R\hat{s}-\boldsymbol{r}|} R^2 d\Omega. \tag{2.4}$$

This expression gives the voltage response in each feed due to a distribution of electric fields across the sky. This is measured by a single antenna.

An interferometer measures the correlations among the feeds for a pair of antennas in a baseline. In terms of the quasi-monochromatic E-field components described before, this amounts to cross-multiplying the voltages and taking an expectation. For a pair of antennas labelled 1 and 2, with feeds p and q, the correlation may be written as a 2×2 matrix:

$$\mathbf{V}_{12} = \begin{pmatrix} V_{12}^{pp} & V_{12}^{pq} \\ V_{12}^{qp} & V_{12}^{qq} \end{pmatrix} \equiv \begin{pmatrix} \left\langle v_{1p} v_{2p}^* \right\rangle & \left\langle v_{1p} v_{2q}^* \right\rangle \\ \left\langle v_{1q} v_{2p}^* \right\rangle & \left\langle v_{1q} v_{2q}^* \right\rangle \end{pmatrix} = \left\langle \boldsymbol{v}_a \times \boldsymbol{v}_b^\dagger \right\rangle \tag{2.5}$$

Asterisks denote complex conjugation. In terms of eq. (2.4), we get the expectation of a product of integrals:

$$\mathbf{V}_{12} = \int\int \mathbf{J}_1(\hat{s}, \nu) \left\langle \boldsymbol{\mathcal{E}}_\nu(\hat{s}) \boldsymbol{\mathcal{E}}_\nu^\dagger(\hat{s}') \right\rangle \mathbf{J}_2^\dagger(\hat{s}', \nu) \frac{e^{2\pi i \nu |R\hat{s} - \boldsymbol{r_1}|/c}}{|R\hat{s} - \boldsymbol{r_1}|} \frac{e^{-2\pi i \nu |R\hat{s}' - \boldsymbol{r_2}|/c}}{|R\hat{s}' - \boldsymbol{r_2}|} R^4 d\Omega d\Omega' \tag{2.6}$$

Note the subscripts on the Jones matrices – antennas in principle do not have identical beams. We can simplify this further with two well-motivated assumptions. First, we assume that the sky emission is not coherent, such that

$$\left\langle \boldsymbol{\mathcal{E}}_\nu(\hat{s}) \boldsymbol{\mathcal{E}}_\nu^\dagger(\hat{s}') \right\rangle = \boldsymbol{\mathcal{C}}_\nu(\hat{s})\, \delta_D(\hat{s} - \hat{s}')/R^2, \tag{2.7}$$

with Dirac delta-function δ_D. $\boldsymbol{\mathcal{C}}$ is a 2×2 *coherency matrix*, describing the expectation of the products of electric field components at the position $\hat{s}$. If we write $\boldsymbol{E}_\nu = (E_x, E_y)$, such that x and y are in the tangent plane of the sky, this may be written as

$$\boldsymbol{\mathcal{C}}(\hat{s}) = \begin{pmatrix} \langle E_x(\hat{s}) E_x^*(\hat{s}) \rangle & \left\langle E_x(\hat{s}) E_y^*(\hat{s}) \right\rangle \\ \langle E_y(\hat{s}) E_x^*(\hat{s}) \rangle & \left\langle E_y(\hat{s}) E_y^*(\hat{s}) \right\rangle \end{pmatrix}, \tag{2.8}$$

with asterisks denoting complex conjugation. This can be expressed in terms of the Stokes parameters, as we will see. For now, we note that the delta-function eliminates one of the integrals over the sphere and sets $\hat{s}'$ to $\hat{s}$.

The second assumption is that the lengths of the antenna position vectors are much smaller than the radius R – $|\boldsymbol{r}_1|, |\boldsymbol{r}_2| \ll R$. This lets us eliminate terms of order $\mathcal{O}(r/R)$. To this order,

$$|R\hat{s} - \boldsymbol{r}| \simeq R - \hat{s} \cdot \boldsymbol{r}$$
$$|R\hat{s} - \boldsymbol{r}_1||R\hat{s} - \boldsymbol{r}_2| \simeq R^2 \left(1 - \frac{\hat{s} \cdot (\boldsymbol{r}_1 + \boldsymbol{r}_2)}{R} + \frac{(\hat{s} \cdot \boldsymbol{r}_1)(\hat{s} \cdot \boldsymbol{r}_2)}{R^2} \right) \approx R^2$$

With these approximations, the final expression is independent of R:

$$\mathbf{V}_{12} = \int \mathbf{J}_1(\hat{s}, \nu)\, \boldsymbol{\mathcal{C}}_\nu(\hat{s})\, \mathbf{J}_2^\dagger(\hat{s}, \nu) e^{2\pi i \hat{s} \cdot (\boldsymbol{r}_2 - \boldsymbol{r}_1)\nu/c} d\Omega \tag{2.9}$$

This is the full form of the RIME which in the Jones matrix formalism. The difference in the exponential is usually denoted as a *baseline vector* $\boldsymbol{b} = \boldsymbol{r}_2 - \boldsymbol{r}_1$.

To reiterate, the assumptions made here are:

1. Sources are very far away, compared to the dimensions of the instrument.
2. Light propagates from the sources through vacuum.
3. Emission is spatially incoherent.

Among these, assumption (2) is the least consistently true, as the light from deep space must pass through the ionized IGM and the Earth's ionosphere to reach the telescope, which causes phase delays and Faraday rotation. Such effects can be accounted for in the Jones matrix formalism (Smirnov, 2011; Martinot et al., 2018). Assumption (1) is generally true except for very-long baseline interferometry (VLBI) observations of solar system objects, and is certaintly the case for cosmological observations. Assumption (3) is violated in some limited cases, such as by pulsar and maser emission, but in those cases the spatial coherence is over very small angular sizes (the size of the emitter). Scattering in space can also impart coherence, if light from the same point is scattered off two separated locations (Clark, 1999; Thompson et al., 2017).

2.1.1 Coherency

The coherency matrix described in eq. (2.8) may be expressed in terms of the Stokes parameters (I, Q, U, V) (Smirnov, 2011; Born et al., 1999). In the Cartesian basis defined before,

$$\begin{aligned} I &= 0.5 \left\langle |E_x|^2 \right\rangle + \left\langle |E_y|^2 \right\rangle \\ Q &= 0.5 \left\langle |E_x|^2 \right\rangle - \left\langle |E_y|^2 \right\rangle \\ U &= \langle |E_x||E_y| \cos\delta \rangle \\ V &= \langle |E_x||E_y| \sin\delta \rangle , \end{aligned}$$

where δ is the phase difference between E_x and E_y. Inverting these relations,

$$\boldsymbol{\mathcal{C}} = \begin{pmatrix} I + Q & U + iV \\ U - iV & I - Q \end{pmatrix}.$$

2.1.2 Jones Matrix Elements

The Jones matrices $\mathbf{J}$ in eq. (2.9) describe the relative weighting of electric field components across the sky in their impact on the feeds of the antenna. This section follows the steps in Appendix B of Martinot et al. (2018) and Section 5 of Liu & Shaw (2019). Equation (2.3) may be written out in full:

$$\begin{pmatrix} v_p \\ v_q \end{pmatrix} = \begin{pmatrix} A_p^x(\hat{s},\nu) & A_p^y(\hat{s},\nu) \\ A_q^x(\hat{s},\nu) & A_q^y(\hat{s},\nu) \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix} \tag{2.10}$$

The rows of the Jones matrix may be viewed as two vectors, one for each feed:

$$v_p = \int \boldsymbol{A}_p(\hat{s}) \cdot \boldsymbol{\mathcal{E}}(\hat{s}) d\Omega$$
$$v_q = \int \boldsymbol{A}_q(\hat{s}) \cdot \boldsymbol{\mathcal{E}}(\hat{s}) d\Omega$$

The meaning of the vectors $\boldsymbol{A}_p$ and $\boldsymbol{A}_q$ becomes clear if we think of the antenna feed as a transmitter, instead of a receiver. Some signal v_p in the feed will produce a set of plane waves $\boldsymbol{\mathcal{E}}_p(\hat{s})$ in the far field that satisfy $\boldsymbol{\mathcal{E}}_p \cdot \hat{s} = 0$ (i.e., are perpendicular to the line of sight). Note that the subscript here indicates that this is the far-field electric field pattern transmitted by the feed p.

By the reciprocity theorem, this electric field pattern gives the response v in the feed, so we can determine the response of the feed to each direction based on the far-field electric field pattern it radiates. In short, the E-field beam is proportional to the electric field radiated by the feed as a transmitter:

$$\boldsymbol{A}_p(\hat{s}) = \frac{1}{M}\boldsymbol{\mathcal{E}}_p(\hat{s}) \tag{2.11}$$

where M is some fixed normalization factor. Typically, the beam is normalized so that the power at some point is 1. For example, to normalize such that the power at zenith is 1, then

$$1 = |\boldsymbol{A}_p(\hat{s}_z)|^2 \qquad \Longrightarrow \qquad M = \sqrt{|\boldsymbol{\mathcal{E}}_p(\hat{s}_z)|},$$

where $\hat{s}_z$ is the unit vector pointing to the zenith.

2.1.3 Simplified Case

We'll make two simplifying assumptions, which reduce the RIME for each antenna feed to a single scalar expression.

1. The emission is unpolarized, so $Q = U = V = 0$.
2. Both antennas in the baseline have the same beam, so $\mathsf{J}_1 = \mathsf{J}_2 \equiv \mathsf{J}$.

With the first assumption, the coherency matrix reducese to $\boldsymbol{\mathcal{C}} = I_\nu\,\mathbf{1}$ for 2×2 identity matrix $\mathbf{1}$. The integrand in eq. (2.9) is then just I_ν times a the exponential fringe term and the product of the Jones matrix with its complex conjugate,

$$\mathsf{JJ}^\dagger = \begin{pmatrix} |A_p^x|^2 + |A_p^y|^2 & A_p^x A_q^{*x} + A_p^y A_q^{*y} \\ A_q^x A_p^{*x} + A_q^y A_p^{*y} & |A_q^x|^2 + |A_q^y|^2 \end{pmatrix} \equiv \begin{pmatrix} A_{pp} & A_{pq} \\ A_{qp} & A_{qq} \end{pmatrix}$$

The components of this matrix are often called the *power beams*, because they are analogous to the power in an electric field. This leaves us with,

$$\begin{pmatrix} V_{pp} & V_{pq} \\ V_{qp} & V_{qq} \end{pmatrix} = \int I_\nu(\hat{s}) \begin{pmatrix} A_{pp}(\hat{s},\nu) & A_{pq}(\hat{s},\nu) \\ A_{qp}(\hat{s},\nu) & A_{qq}(\hat{s},\nu) \end{pmatrix} e^{-2\pi i \hat{s}\cdot \boldsymbol{b}\nu/c} d\Omega, \tag{2.12}$$

where the spatial and frequency dependence is now made explicit and $\boldsymbol{b}$ is the baseline vector. If we now consider just the visibility formed from the two p feeds – V_{pp}, we can drop the feed subscripts. When not concerned with polarimetry, this is usually the form of the visibility equation that is most commonly used:

$$V(\boldsymbol{b},\nu) = \int I_\nu(\hat{s}) A_\nu(\hat{s}) e^{-2\pi i \hat{s}\cdot \boldsymbol{b}\nu/c} d\Omega \tag{2.13}$$

To recap, the form shown in eq. (2.13) gives the cross correlation of two identical antenna feeds separated by a vector $\boldsymbol{b}$, receiving an unpolarized sky signal. This form is evaluted by the simulator `healvis` (see section 3.2), and is generally acceptable for simulations of the EoR signal, which is theoretically unpolarized. The full form in eq. (2.9) is evaluated by the simulator `pyuvsim` (section 3.1).

2.1.4 Imaging

Producing an image from interferometric observations involves inverting eq. (2.13) to obtain the Stokes I surface brightness $I_\nu(\hat{s})$. By changing coordinates of the integral and making a simplifying assumption, the measurement equation has the form of a 2D Fourier transform sampled at the baseline positions, and can be inverted to get an image.

Fig. 2.1 shows the coordinate system. The position $\hat{s}_0$, called the *phase center*, defines a tangent plane to the celestial sphere with axes (l, m). The m direction points toward to the north celestial pole (N). The baseline (blue arrow) vector has components (u, v, w), expressed in wavelengths:

$$\boldsymbol{u} = u\hat{u} + v\hat{v} + w\hat{s}_0 \equiv \boldsymbol{b}\frac{\nu}{c} \tag{2.14}$$

$$\text{Such that } \boldsymbol{u}\cdot\hat{s}_0 = w$$

We can now write the unit vector $\hat{s}$ in this Cartesian frame as

$$\hat{s} = l\hat{u} + m\hat{v} + (\sqrt{1 - l^2 - m^2})\hat{s_0} \tag{2.15}$$

Note that the $\hat{u}$ and $\hat{v}$ directions point eastward and northward, respectively, but are not necessarily tangent to the Earth's surface. The (u, v) plane is orthogonal to the phase center direction, which only coincides with the tangent plane of the Earth's surface if the phase center is at zenith. This is a common choice for *drift scanning* instruments. Furthermore, the phase center does not necessarily coincide with the *pointing center*, which is the direction of maximum sensitivity of an antenna. When antennas are steerable, typically they will be made to track a particular pointing center as the Earth rotates. Instruments like HERA, however, have a fixed zenith pointing and are drift scanning.

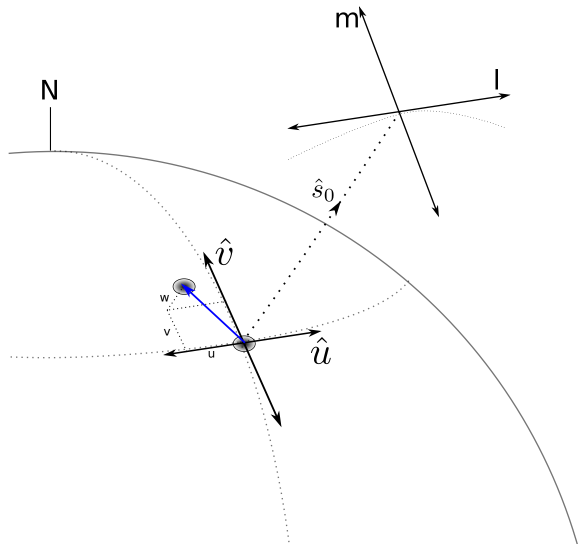


Figure 2.1: A diagram of relevant coordinates used in this subsection. The $\hat{u}$ and $\hat{v}$ directions are parallel to the l and m directions, which define a plane tangent to the celestial sphere at the phase center $\hat{s}_0$.

In this new coordinate system, the RIME has the following form:

$$V(u,v,w) = \int I_\nu(l,m)A_\nu(l,m)e^{-2\pi i(ul+vm+w\sqrt{1-l^2-m^2})}\frac{dldm}{\sqrt{1-l^2-m^2}}, \tag{2.16}$$

where the denominator comes from the Jacobian of the transformation for angle to (l,m,n) coordinates.

From here, we can take the *flat sky approximation* by assuming that $w \approx 0$. Defining $I'_\nu(l,m) = I_\nu(l,m)/\sqrt{1-l^2-m^2}$, we get

$$V(u,v) = \int I'_\nu(l,m)A_\nu(l,m)e^{-2\pi i(ul+vm)}dldm \tag{2.17}$$

Alternatively, we can assume that the beam is very weak outside of some small region around $\hat{s}_0$. In this case, we can rewrite the integral in terms of $\boldsymbol{\sigma} = \hat{s} - \hat{s}_0$, and drop terms of order σ^2.

Under this approximation, $\boldsymbol{\sigma}$ is perpendicular to $\hat{s}_0$, which means we can write $\boldsymbol{\sigma} = l\hat{x} + m\hat{y}$.

$$\begin{aligned} V(u,v) &= \int I_\nu(l,m)A_\nu(l,m)e^{-2\pi i\boldsymbol{u}\cdot(\boldsymbol{\sigma}+\hat{s}_0)}dldm \\ &= e^{-2\pi iw}\int I_\nu(l,m)A_\nu(l,m)e^{-2\pi i(ul+vm)}dldm \end{aligned} \tag{2.18}$$

The square root in the denominator goes away because $l^2 + m^2 = \sigma^2 \approx 0$.

In both of these forms, the measurement equation has the form of a 2D Fourier transform, with (l,m) dual to (u,v). As such, this can be easily inverted:

$$I_\nu(l,m)A_\nu(l,m) = \int V(u,v)e^{2\pi iul+vm}dudv \tag{2.19}$$

2.2 Power Spectrum Measurement

Methods to estimate the power spectrum of HI emission from interferometric data generally fall into two broad categories (Morales et al., 2019). The traditional method to measuring a power spectrum is to map the pattern of emission to comoving coordinates in space and then take a 3D Fourier transform. Because this method, and others like it, involve recovering an image of the sky at different frequencies, these may be broadly categorized as "imaging" methods. The other category covers techniques that take advantage of the fact that visibilities are already measured in a quasi-Fourier basis – since the UV coordinates are already the Fourier dual of image space – and so the visibilities are Fourier-transformed in frequency (from frequency ν to "delay" τ), to produce something equivalent to a 3D Fourier-transformed volume. With appropriate scaling, this approximately corresponds to the Fourier transform of a comoving cosmological volume. These methods are categorized as "delay-transform methods".

In this section, I will briefly go over the key features of each method. I used delay spectrum estimators for the remainder of results in this dissertation, so those will be discussed in more detail. With both methods, it is convenient to think about the power spectrum wavevectors as comprising one component parallel to the line of sight, $k_\parallel$, and two components perpendicular $\boldsymbol{k}_\perp = (k_x, k_y)$. The $\boldsymbol{k}_\perp$ modes are detected through the angular sensitivity of the instrument, and in general long baselines $|\boldsymbol{u}|$ add sensitivity to large $|\boldsymbol{k}_\perp|$ modes. Spatial structure along the line of sight corresponds with structure in the spectrum, due to the direct mapping between distance and redshift, so measurements of $k_\parallel$ modes are achieved through frequency spectrum measurements.

2.2.1 Sky-Reconstruction Estimator

The power spectrum is defined, as in eq. (1.12), via

$$\left\langle \widetilde{T}(\boldsymbol{k})\widetilde{T}^*(\boldsymbol{k}')\right\rangle = (2\pi)^3\delta_D(\boldsymbol{k}+\boldsymbol{k}')P(\boldsymbol{k}),$$

where I've let T stand for the brightness temperature fluctuation. We can form an estimator of the power spectrum from a measured volume $\mathbb{V}$,

$$\hat{P}(\boldsymbol{k}) \equiv \frac{\left\langle |\widetilde{T}(\boldsymbol{k})|^2 \right\rangle}{\mathbb{V}}, \tag{2.20}$$

where the average is taken over multiple samples of the Fourier mode $\boldsymbol{k}$, and $\mathbb{V}$ is the volume surveyed.

In sky-reconstruction estimators, the visibility data is used to make an estimate of the spatial brightness temperature distribution $T(\hat{s}, \chi_\nu)$, where χ_ν is the comoving distance corresponding with the frequency ν and $\hat{s}$ is a direction. This is done by constructing an image from the visibility data.

Recall that the flat-sky approximation visibility equation has the form,

$$V_\nu(u, v) = \int A_\nu(l, m) I_\nu(l, m) e^{-2\pi i(ul+vm)} dl dm, \tag{2.21}$$

which may be inverted to get

$$I_\nu(l, m) A_\nu(l, m) = \frac{1}{(2\pi)^2} \int V_\nu(u, v) e^{2\pi i(ul+vm)} du dv. \tag{2.22}$$

However, the instrument only measures the visibility at specific positions, and so we must introduce a sampling to eq. (2.21). This sampling may be described by a,

$$S(u, v) = \sum_{p \in \mathbb{B}} \delta_D(u - u_p) \delta_D(v - v_p),$$

where $\mathbb{B}$ is the set of baselines in the array and (u_p, v_p) are the UV coordinates of the p$^{\mathrm{th}}$ baseline. This is convolved with the visibility function to give the sampled data. Applying the convolution theorem to eq. (2.21), this looks like,

$$I_\nu(l, m) A_\nu(l, m) = \frac{1}{(2\pi)^2} \sum_{p \in \mathbb{B}} V_\nu^p e^{2\pi i(u_p l + v_p m)}, \tag{2.23}$$

where V_ν^p is the visibility measured by baseline p. This is, essentially, a sum of plane waves weighted by the visibility amplitudes. In principle, the more positions sampled in (u, v) space, the better the interference of these plane waves will describe the true brightness distribution of the sky. Often, these sampled visibilities are *gridded* to a regular 2D grid of (u,v) points, such that the data may be transformed into an image using a Fast Fourier Transform (FFT). This requires that the visibilities be convolved with a chosen *gridding kernel*. In *holographic* imaging, this is chosen to be the Fourier-transform of the primary beam itself, which is chosen because the sampled data are known to be a weighted sum of the beam-convolved UV plane (Morales & Matejek, 2009). Holographic mapmaking is optimal, unlike mapmaking with other gridding kernels, such that it preserves the information available in the measurement. The effects of the sampling and beam weighting can also be reduced by applying a form of deconvolution to the data.

As a particular example, one analysis pipeline of the MWA experiment is made up of the software Fast Holographic Deconvolution (FHD), which does calibration, forward-modeling, and imaging, and a power spectrum estimator "error propagated power spectrum with interleaved observed noise" (εppsilon), which performs DFTs along each axis to estimate power spectra and errors. FHD grids visibilities using a primary beam model interpolated from saved data (i.e., a data-defined beam), and simultaneously produces "model" data from a given point source model of the sky using the same response as is used for processing the data. The images are saved, along with weights reflecting the various gridding kernels applied, and read into εppsilon, which does full DFTs (not FFTs) in the image plane, divides out the weights, and estimates the line of sight Fourier transform using a Lomb-Scargle periodogram. Barry et al. (2019) summarizes the FHD$\rightarrow$ εppsilon pipeline well.

Once an image is formed, there is a direct correspondence between the data and the spatial distribution of brightness temperature. Letting $\theta_x = \arcsin(l)$ and $\theta_y = \arcsin m$,

$$\Delta r_x = \chi(z)\theta_x \qquad k_x = \frac{2\pi u}{\chi(z)} \tag{2.24}$$

$$\Delta r_y = \chi(z)\theta_y \qquad k_y = \frac{2\pi v}{\chi(z)} \tag{2.25}$$

$$\Delta r_z = \frac{c(1+z)^2}{H(z)\nu_{21}}\Delta\nu \qquad k_z = \frac{2\pi H(z)\nu_{21}}{c(1+z)^2}\eta, \tag{2.26}$$

where $\nu_{21} = 1420$ MHz is the 21 cm frequency, $\chi(z)$ is the comoving transverse distance,[1] and $k_{x,y,z}$ are the Fourier duals to $\Delta r_{x,y,z}$ (Morales & Hewitt, 2004). η is the Fourier dual to the frequency ν.

2.2.2 Delay Transform Estimator

As we've seen, visibilities are defined in a hybrid of Fourier and physical space, such that the baseline vector corresponds with a Fourier mode perpendicular to the line of sight, while the frequency axis directly maps to distance along the line of sightfor the 21 cm signal. Fourier-transforming along the frequency axis puts the visibilities into a 3D Fourier space, which can be used to estimate the power spectrum. The delay transform, introduced in Parsons & Backer (2009); Parsons et al. (2012b, 2014), does exactly this – it is an inverse Fourier transform along the frequency axis, transforming frequency into "delay" τ.

$$\widetilde{V}(\boldsymbol{b},\tau) = \int \phi(\nu)V(\boldsymbol{b},\nu)e^{2\pi i\nu\tau}d\nu \tag{2.27}$$

$$= \int\int \phi(\nu)A(\hat{s})\,T(\hat{s},\nu)e^{-2\pi i(\boldsymbol{b}\cdot\hat{s}\nu/c-\nu\tau)}d\Omega d\nu \tag{2.28}$$

[1]The comoving transverse distance is equivalent to the comoving radial distance for flat cosmological models.

We've taken the transform while applying a *spectral window function* $\phi(\nu)$, which accounts for the finiteness of the bandpass and allows us to weight frequencies. The quantity $\widetilde{V}$ corresponds with the Fourier-transform of the brightness temperature field, $\widetilde{T}(\boldsymbol{k})$, and so we can estimate the power spectrum through some quadratic form. Note that tildes here are used to denote quantities that are in a 3D Fourier space, which has a slightly different meaning for I and V. $\widetilde{I}(\boldsymbol{k})$ is obtained by doing a 3D spatial Fourier transform, but V, which is already in a hybrid Fourier basis, is turned to $\widetilde{V}$ by a 1D transform over frequency. The results correspond with each other.

If we consider a set of baselines G with the same length and orientation (called "redundant"), they form independent samples of a set of $k_\parallel$ modes for a given $\boldsymbol{k}_\perp$ mode corresponding with the baseline vector. We cross-multiply the delay-transformed visibilities among this redundant group to form an estimate of the power spectrum:

$$\begin{aligned} \widehat{P}(k_{b,\tau}) &\equiv \left(\frac{\lambda^2}{2k_B}\right)^2 \frac{X^2 Y}{B_{pp}\Omega_{pp}} \left\langle \widetilde{V}_\alpha(\tau)\widetilde{V}^*_\beta(\tau)\right\rangle_{\alpha,\beta\in G_b} \\ k_{b,\tau} &= 2\pi\sqrt{\left(\frac{\tau}{Y}\right)^2 + \left(\frac{b}{\lambda X}\right)^2} \end{aligned} \tag{2.29}$$

The Fourier mode $k_{b,\tau}$ probed by this estimator depends on the baseline length u (in wavelengths) as well as the delay mode τ. Note that cross-multiplying data for a single baseline will lead to a bias in the presence of thermal noise. Cross-multiplying redundant baselines, which measure the same signal but have independent noise realizations, avoids this bias. We can also avoid this bias by cross-multiplying measurements of the same LSTs across multiple nights.

The multiplicative factors in front approximate the inverse volume factor of eq. (2.20). X and Y are redshift-dependent cosmological scaling factors, with units of length per angle and length per frequency, respectively. B_{pp} is an effective bandwidth, given by $B_{pp} = \int |\phi(\nu)|^2 d\nu$, which relates to the total line of sight extent observed. The primary beam squared integral $\Omega_{pp} = \int |A(\hat{s})|^2 d\Omega$ gives the angular area. The exact derivation of this factor is given by Appendix B of (Parsons et al., 2014).[2]

The cosmological factors are

$$X(z) = \chi(z) \qquad Y(z) = \frac{c(1+z)^2}{H(z)\nu_{21}}, \tag{2.30}$$

where z is redshift, χ is the comoving distance, $H(z)$ is the Hubble parameter, ν_{21} is the rest-frame 21 cm frequency, and c is the speed of light. The $\boldsymbol{k}_\perp$ and $k_\parallel$ modes corresponding to a given baseline $\boldsymbol{u}$ and delay τ are given by

$$\boldsymbol{k}_\perp = \frac{2\pi\boldsymbol{u}}{X} \qquad k_\parallel = \frac{2\pi\tau}{Y} \tag{2.31}$$

[2] See also Memos #27 and #43 at reionization.org.

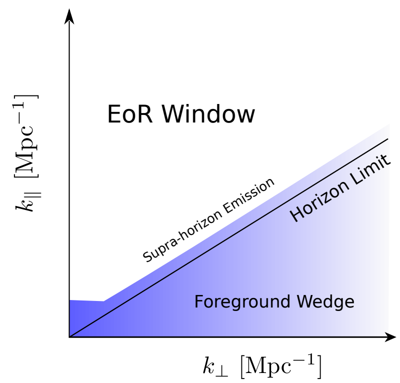
For the above relations, X and Y are evaluated at the center of the bandpass and assumed not to evolve much over the chosen bandpass. The correspondence between $k_{\parallel}$ and τ is only approximate, and gets worse for longer baselines.

The delay spectrum estimator has the advantage that it can be used to form power spectrum estimates directly from measured visbilities, without going through the steps of imaging. This is especially beneficial to highly redundant arrays like HERA, which lack the dense UV coverage for good imaging (Morales et al., 2019). The main downside to the delay spectrum approach comes from the approximate correspondence between τ and $k_{\parallel}$ on long baselines. Since baseline length is expressed in wavelengths, it changes with frequency, and changes more rapidly for long baselines. This causes some coupling between the $k_{\perp}$ and $k_{\parallel}$ modes measured by long baselines. The delay mode τ is approximately equivalent to the Fourier dual η mentioned in the previous subsection. The key difference is that the delay transform is done on a per-baseline basis, while the Fourier transform over frequency in image-based estimators is done on an "image cube", which has accumulated information from multiple baselines.

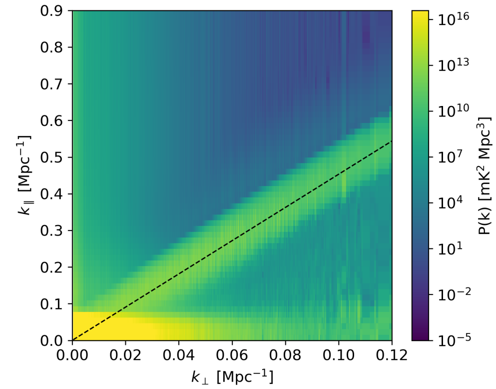
2.2.3 Foreground Separation

With both power spectrum estimation approaches, line of sight Fourier modes are detected through the frequency response of the instrument. The spectrum of the 21 cm signal from cosmological distances directly corresponds with the spatial distribution of the IGM absorbing and emitting it. In contrast, bright foreground sources have very smooth spectra. In Fourier space, this will cluster around $k_{\parallel} \approx 0$, leaving space at larger $k_{\parallel}$ that should, in principle, be free of foreground contamination. This uncontaminated region of Fourier space is called the *EoR Window*, while the region dominated by foregrounds is called the *Foreground Wedge* (for reasons that will be clear in a moment).

Spectral structure in the sky and in the instrument will cause this foreground power to spread toward higher $k_{\parallel}$. In particular, there is an implicit chromaticity to an interferometer, due to the change in baseline length with frequency. The antennas are at a fixed physical separation, but the distance between them enters the measurement equation in units of *wavelength*, meaning that the effective length of the baseline changes with frequency. This is equivalent to the baseline drifting through Fourier space with frequency, which creates a frequency dependence in the visibility measurements. This couples $k_{\perp}$ and $k_{\parallel}$ and spreads power to higher $k_{\parallel}$ modes (Datta et al., 2010; Vedantham et al., 2012; Morales et al., 2012; Parsons et al., 2012b; Trott et al., 2012; Hazelton et al., 2013; Thyagarajan et al., 2013; Liu et al., 2014; Pober et al., 2014). The longer the baseline, the more important this chromaticity becomes, and the farther power spreads. This leads to a characteristic wedge shape.



(a) Foreground Wedge diagram



(b) Simulated foregrounds.

Figure 2.2: 2.2a shows a diagram of the 2D Fourier space, with the foreground wedge, EoR window, and horizon limit labelled. Power that extends beyond the horizon limit is fittingly called "suprahorizon power". The distance into the EoR window in which it extends is the "spillover". 2.2b shows the calculated power spectrum from a simulation of a diffuse foreground model (more about this simulator in section 3.2). Foreground power is largely contained within the wedge, but extends beyond the horizon to a limited extent.

The effects of baseline chromaticity are, in principle, still confined to a maximum $k_{\|}$ set by the angular extent of the field of view, called the "horizon limit." Fig. 2.2 shows the various regions in 2D Fourier space, as well as the power spectrum of a simulated diffuse sky model. Some power leaks beyond the horizon limit, called "suprahorizon power", due to a variety of effects. This suprahorizon leakage will be discussed more in chapter 5.

Chapter 3

Instrument Simulation

To achieve the level of precision necessary to detect a 21 cm signal, inteferometry experiments have had to uncover previously unknown or otherwise neglected sources of spectral structure and miscalibration that can contaminate the EoR window. Detailed simulations of the interferometer response have been essential to this effort. Often these simulations uncover effects that had not yet been seen in data, but are later observed. Datta et al. (2010) performed the most advanced simulations at the time and uncovered the foreground wedge before it was found in data. Thyagarajan et al. (2015a) simulated sources down to the horizon for the first time, and uncovered a previously-unobserved "pitchfork" foreground signature. Barry et al. (2016) used instrument simulations to show that unmodeled emission would corrupt calibration.

The industry standard for simulation was the `simobserve` utility in *Common Astronomy Software Applications* package (`casa`), which is optimized for simulating observations of small fields of view with tracking antennas.[1] This was the tool used by Datta et al. (2010) for modeling EoR foregrounds and predicting the wedge. To address the need for sensitivity to wide fields of view, the Precision Radio Interferometry Simulator (`PRISim`) was written, which is parallelized over the source axis to optimize support for large numbers of sources within the field of view.[2] With sensitivity to the horizon, this simulator was able to detect the foreground pitchfork effect (discussed in more detail in chapter 5). The simulator built into Fast Holographic Deconvolution (`FHD`), the software workhorse of the US side of MWA EoR science, is designed to work with data-defined beam models and efficiently handle large numbers of sources. FHD simulations use the same infrastructure as the data reduction pipeline, allowing them to directly track the effects of changes to the analysis. It has been largely successful in verifying foreground subtraction in MWA analyses

[1] `https://casa.nrao.edu/casadocs/casa-5.6.0/simulation`

[2] `https://github.com/nithyanandan/PRISim`

(Barry et al., 2019).

As of October of 2017, instrument simulation within the HERA collaboration was done using this handful of software packages, each with specialized design choices and varied feature coverage. Simulated data could be compared with real data, but it was difficult to determine whether differences were due to some artifact of the simulator itself or due to an unmodeled instrument effect. Further, no simulator then available could cover all of the factors known to be significant to EoR science, and no simulator had successfully simulated data for an EoR-like signal on the sky and recovered the correct power spectrum from the simulated visibilities. For these reasons, it was decided that a new simulator would be developed from the ground up to serve as a reference standard for others, incorporating lessons from previous generations of simulation tools. This simulator, called `pyuvsim`, would be written collaboratively and with rigorous standards of testing and review (Lanman et al., 2019b).

3.1 `pyuvsim`

`pyuvsim` aims to evaluate the full polarized radio interferometric measurement equation eq. (2.9) to high precision, making as few approximations as possible along the way. As it was first written, it was designed to simulate visibilities for unresolved *point* sources, which provides a natural discretization of the measurement equation, allowing for numerical evaluation. The coherency of a collection of unpolarized point sources may be described by a sum of delta functions:

$$\boldsymbol{\mathcal{C}}(\hat{s}) = \sum_p \boldsymbol{\mathcal{C}}_p \delta_D(\hat{s} - \hat{s}_p) \tag{3.1}$$

where $\boldsymbol{\mathcal{C}}_p$ is the coherency of the p^{th} source. Inserting this into eq. (2.9), the integral becomes a sum:

$$\mathbf{V}_{12} = \sum_p \mathbf{J}_1(\hat{s}_p, \nu)\, \boldsymbol{\mathcal{C}}_p(\hat{s}_p, \nu)\, \mathbf{J}_2^\dagger(\hat{s}_p, \nu) e^{2\pi i\, \hat{s}_p \cdot \boldsymbol{b}_{12} \nu / c} \tag{3.2}$$

In the above, the source positions $\hat{s}_p$ are time-dependent due to the rotation of the Earth.

3.1.1 Design Philosophy

The impetus to develop `pyuvsim` came from an uncertainty in the validity of simulation data produced by other software packages. The best software available at the time was optimized to specific use cases, and it was decided that generalizing them would take more effort and be fraught with more bugs than to start from scratch.

Several key goals of `pyuvsim` include:

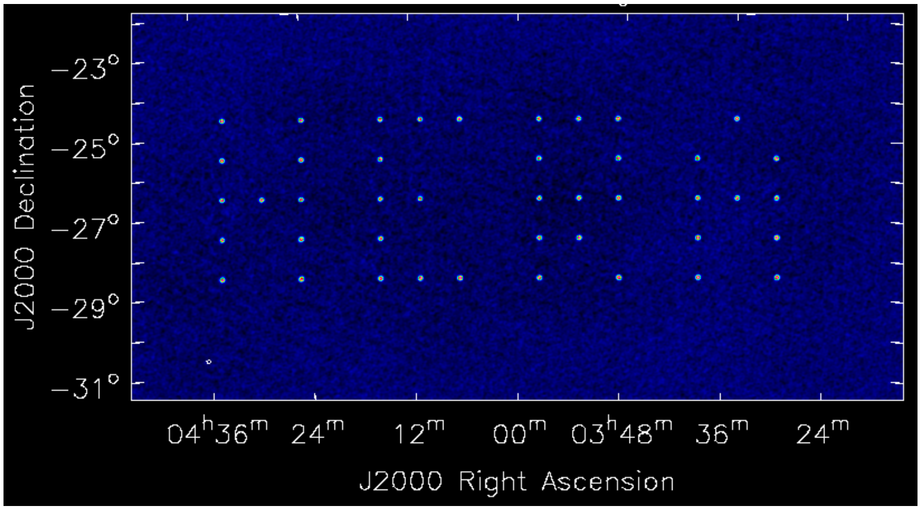


Figure 3.1: An image made in CASA using the CLEAN algorithm from `pyuvsim`-simulated observations of a mock catalog spelling out the word "HERA" in point sources. Test patterns such as this are used to verify that sources are correctly positioned. This particular simulation is one of the *reference simulations*, which are used as a point of comparison with other simulators and as a test for `pyuvsim` updates.

1. High level of test coverage including accuracy (design goal is 97%).

   "Test coverage" here refers to the fraction of code that is run through an automated unit testing framework, checking acceptability of results whenever a change is made. This is done with the `py.test` toolset. In addition to tests of the user interface and utility functions, there are several tests that verify simulated visibilities in controlled cases against analytic calculations for the same cases. For instance, one test confirms that the visibilities measured for a 1 Jy unpolarized point source at zenith are 0.5 Jy, for V_{pp} and V_{qq}. Another test confirms the calculation for an off-zenith source.

2. Comparison with external simulations and reference cases.

   In addition to unit tests, `pyuvsim` is verified through a variety of external comparisons. To date, `pyuvsim` has been compared to `PRISim`, `healvis`, and `FHD` simulations. These comparisons are made using a set of *reference simulations*, which are run at each new version to ensure consistency. One example is a simple mock point source catalog that spells out text on the sky. Data from this simulation can be imaged, to verify that source positions are correct. Fig. 3.1 shows an example of an image generated from one of these reference simulations, which verifies that point sources in a mock catalog are positioned correctly. This tests the data products of `pyuvsim` with the established imaging tools of `casa`.

3. Design for scalability across large computing clusters.

   Evaluating the RIME is a computationally-demanding process, but one that can be parallelized. The parallelization scheme built into `pyuvsim` is designed to scale across distributed systems.

4. Fully-polarized instrument response, floating-point source position accuracy.

   Polarization is an important factor in 21 cm science. Faraday rotation from the ionosphere can impart frequency-dependent polarization on foregrounds, which can be mistaken for an EoR signal if the polarized power leaks into the Stokes-I signal (Martinot et al., 2018; Asad et al., 2015, 2016, 2018). Further, the usefulness of `pyuvsim` beyond 21 cm cosmology requires that it support polarized source models.

5. Full-sky field of view with curvature.

   The simulators `PRISim` and `FHD` (Sullivan et al., 2012) were the first to simulate the sky from horizon to horizon, and found unexpected effects from diffuse sources near the horizon (Thyagarajan et al., 2015a). EoR experiments are generally done with wide fields of view, so this must be accommodated by `pyuvsim`.

6. Integration with established software libraries.

   Whenever possible, calculations are relegated to state-of-the-art external software packages. Coordinate and unit transformations are all handled with `astropy` (Astropy Collaboration et al., 2013; Price-Whelan et al., 2018). Primary beam calculations and data phasing / file management are handled by `pyuvdata`, which is verified to comparable rigor to `pyuvsim` and is under active development within the HERA and MWA collaborations. At the time of writing, source modeling is managed within `pyuvsim`, but these functions are being moved to a separate repository for a more focused effort there. The ultimate goal is for `pyuvsim` to operate *only* as an efficient and verified visibility calculator. As these other tools grow more sophisticated, their new features naturally extend the capabilitites of `pyuvsim`.

7. Support for varied beam models across the array.

   Though `FHD` can, in principle, support different beams for different antennas, this is not a simple feature to use and it hasn't been well-tested. `pyuvsim` offers a simple means to use multiple beam models and assign them to antennas in the array, and can be easily extended to support numerically-defined variations among antennas.

8. Defining a clear, user-friendly standard for simulation design.

As a general-purpose simulator, `pyuvsim` must have a flexible but simple user interface. This has the form of YAML configuration files with clear, well-defined keywords to access all simulator features. This standard is described in detail in publicly-accessible documentation[3].

3.1.2 Structure

`pyuvsim` evaluates eq. (3.2) by directly computing all Jones and coherency matrices and fringe terms, carrying out the matrix products, and then summing over sources. It can also simulate a diffuse model, provided the model is discretized in a compatible way. `pyuvsim` accepts as input:

1. A source catalog, containing source positions in the International Celestial Reference System (ICRS[4]) RA/Dec and Stokes parameters at either one or a set of frequencies. I will refer to this set as $\mathbb{S}$.

2. A list of frequency channels, $\mathbb{F}$.

3. A set of time steps, as Julian dates, $\mathbb{T}$.

4. A list of beam models, $\mathbb{J}$. These can either be analytically-defined (Gaussian, Airy disk, or uniform) or as a path to a *beamfits* file that can be read by `pyuvdata`'s UVBeam class. Each beam is identified with a "beam ID" number.

5. A telescope name and location, the latter provided as a latitude, longitude, and altitude on the Earth's surface in degrees/meters. The altitude is relative the the reference ellipsoid of the Earth as defined in the World Geodetic System 1984 (Kaplan (2005) discusses the relationship between these terrestrial frames and the astronomical coordinate systems).

6. A collection of antennas in an array, $\mathbb{A}$, each specified with a position, a beam ID, a name, and an ID number. The antenna positions are defined in the "East/North/Up" (ENU) Cartesian frame relative to a reference "telescope position", defined in latitude and longitude on the Earth, a number. The beam ID associated with each antenna identifies the primary beam in $\mathbb{J}$ that antenna uses. The *beam dictionary* is the mapping of antenna numbers to beam model.

7. Optional selection parameters for choosing subsets of baselines, frequencies, times, etc. Notably, one can choose to reduce the full set of baselines to only one from each group of redundant baselines, given a redundancy tolerance. This is especially helpful for highly-redundant

[3] `https://pyuvsim.readthedocs.io`

[4] The ICRS frame is defined by the solar system barycenter, with positions fixed in time, and ra/dec are defined to correspond with the previous FK5 system at the J2000 epoch.

arrays like HERA, reducing the number of baselines to be computed from thousands to tens. Specifying this via selection keywords makes for a simpler user interface – one need not make separate antenna layout files or frequency/time specifications in order to simulate a subset of the data.

Baselines are constructed from pairs of antennas (under the convention that the first antenna number is larger than the second):

$$\mathbb{B} = \{(a_i, a_j) \mid a_i, a_j \in \mathbb{J}; i > j\}$$

where subscripts represent the antenna ID numbers. Baseline vectors are defined, as before, $\boldsymbol{b}_{ij} = \boldsymbol{r}_j - \boldsymbol{r}_i$.

Fundamentally, the simulation is carried out in the steps described shown in Pseudocode 1.

Pseudocode 1 `pyuvsim` calculation.

Setup baseline vectors, frequencies, times, beam list, and beam dictionary.
Read catalog.
for all t_i in $\mathbb{T}$ **do**
Calculate local Altitude/Azimuth positions of all sources from ra/dec.
From source altitudes, identify sources that are below the horizon and set their flux to 0.
Rotate the ra/dec coherency matrices of all sources into the topocentric frame.
for all f_i in $\mathbb{F}$ **do**
For each beam in the beam list, evaluate the Jones matrices at all source positions.
Evaluate the flux of the sources at f_i.
for all b_k in $\mathbb{B}$ **do**
Calculate the fringe at each source position.
Multiply Jones matrices with coherencies with fringes.
Sum over sources to obtain $\mathbf{V}_{ij}$.
end for
end for
end for
Collect calculated visibility vectors.
Write the results to a file, using `pyuvdata` file writing methods.

3.1.3 Coordinate Transformations

The visibility calculation step is performed in a local topocentric (Altitude/Azimuth), which means that source coordinates and coherency matrices must be transformed from an equatorial system at each time. The basic coordinate frame transformation is done using `astropy`, which ensures that the transformation takes into account the full effects of the Earth's rotation, precession, and nutation, as well as stellar aberration. Neglecting, for instance, the effects of precession can lead to an error of over 16 arcmin over the 20 years since the J2000 epoch. Importantly, `astropy` automatically tracks updates to the universal time, such as leap seconds.

The transformation of coherency matrices from RA/Dec to topocentric frames must take into account the fact that the electric field components are vector quantities, and so must be properly rotated. The Stokes parameters are initially used to define an "ra/dec coherency" matrix $\boldsymbol{\mathcal{C}}_{\rm eq}$. This is constructed from from electric field components in a $(\hat{e}_\alpha, \hat{e}_\delta)$ basis, where $\hat{e}_\alpha$ points in the direction of increasing right ascension on the celestial sphere and $\hat{e}_\delta$ points toward increasing declination.

$$\boldsymbol{\mathcal{C}}_{\rm eq} = \left\langle \boldsymbol{\mathcal{E}}_{\rm eq} \otimes \boldsymbol{\mathcal{E}}^\dagger_{\rm eq} \right\rangle = \begin{pmatrix} \langle E_\delta E^*_\delta \rangle & \langle E_\delta E^*_\alpha \rangle \\ \langle E_\alpha E^*_\delta \rangle & \langle E_\alpha E^*_\alpha \rangle \end{pmatrix} \tag{3.3}$$

The $\otimes$ symbol denotes a tensor product and the subscripts indicate which component in the (α, δ) basis. This must be transformed to a local coherency matrix, $\boldsymbol{\mathcal{C}}_{\rm top}$ in the topocentric frame.

The transformation may be described by a linear transform at each point on the sky:

$$\boldsymbol{\mathcal{E}}_{\rm top} = \mathbf{R}(\hat{s}) \boldsymbol{\mathcal{E}}_{\rm eq} \tag{3.4}$$

which leads to a transformation of the coherency matrix of the form

$$\boldsymbol{\mathcal{C}}_{\rm top} = \mathbf{R}(\hat{s}) \boldsymbol{\mathcal{C}}_{\rm eq} \mathbf{R}^{\rm T}(\hat{s}) \tag{3.5}$$

where T denotes a transpose. The dependence of $\mathbf{R}$ on $\hat{s}$ is included to indicate that separate rotations are calculated for each source position.

The topocentric frame coherency matrix has the form

$$\boldsymbol{\mathcal{C}}_{\rm top} = \begin{pmatrix} \langle E_{\rm alt} E^*_{\rm alt} \rangle & \langle E_{\rm alt} E^*_{\rm az} \rangle \\ \langle E_{\rm az} E^*_{\rm alt} \rangle & \langle E_{\rm az} E^*_{\rm az} \rangle , \end{pmatrix} \tag{3.6}$$

where subscripts denote components in the direction of increasing altitude and increasing azimuth. Choosing this frame, while convenient for other calculations, does introduce a coordinate singularity at the zenith where the azimuth angle loses definition and the altitude direction flips. The same singularity appears in the Jones matrix, however, and in multiplying them together the effects are cancelled out.

For unpolarized sources, as mentioned in section 2.1.3, the coherency matrix is simply $I_\nu(\hat{s})\mathbf{1}$, so the rotation of the sky has no effect. `pyuvsim` only computes and applies rotation matrices for sources with polarization.

3.1.4 Primary Beams

Antenna beams are defined in two ways in `pyuvsim` – as `UVBeam` or as `AnalyticBeam` objects. The `UVBeam` class, a component of the package `pyuvdata`, defines E-field beams from a set of electric field values at specified angles and frequencies. From these saved values, the E-field is calculated at specified angles using bivariate spline interpolation in various forms. Interpolation in frequency is done using univariate spline interpolation at each angular position. This is discussed in more detail in chapter 5, where a `UVBeam` model of the HERA beam is applied.

The `AnalyticBeam` class is written to be used interchangeably with `UVBeam`, but instead of interpolating fixed values it simply evaluates an analytic function defining the beam. At this time, there are three types of analytic beams defined in `pyuvsim`. Expressed as power beams $A_\nu(\theta)$ (see section 2.1.3), they are:

1. Uniform : A unit response in all directions.

$$A_\nu(\theta) = 1.0 \tag{3.7}$$

2. Airy : An Airy disk pattern, given by the Fourier transform of a circular disk with a given diameter in wavelengths. This is defined from an antenna diameter parameter D.

$$A_\nu(\theta) = \left(\frac{2\ J_1\left[\pi \sin(\theta) D\nu/c\right]}{\pi \sin(\theta) D\nu/c} \right)^2 \tag{3.8}$$

3. Gaussian : A Gaussian function about the pointing center. This is specified with a width parameter σ, representing the standard deviation of the Gaussian, or with an antenna diameter. When given a diameter, the width of the Gaussian is defined to match the main lobe of an Airy disk for an aperutre with that diameter. Gaussian beams may also be defined with a width that varies as a power law.

$$A_\nu(\theta) = \exp(-\theta^2/(2\sigma(\nu)^2)) \qquad \sigma(\nu) = \sigma_0(\nu/\nu_0)^\alpha \tag{3.9}$$

   If a diameter is given:

$$\sigma(\nu) = \arcsin(2.2150894 \times c/\nu/(\pi D))/\sqrt{2\ln 2} \tag{3.10}$$

In the above, θ is the angle from the pointing center, ν_0 is a specified reference frequency, α is a spectral index (0 by default), J_1 is the order 1 Bessel function of the first kind, and c is the speed of light. The constant in eq. (3.10) comes from fitting a Gaussian to an Airy disk function.

These functional forms are motivated by familiar models of power beams, but in `pyuvsim` they must be specified as E-field components. For instance, the Airy beam is understood to be the power pattern of a plane wave diffracted through a circular aperture, but does not strictly correspond with a particular feed design. Therefore, we have some flexibility to choose the E-field form of the analytic beams, provided they behave as expected as power beams.

We define the E-field beams so that off-diagonal terms in the Jones matrix $\mathbf{J}$ are zero, and that the diagonal terms in $\mathbf{J}\mathbf{J}^\dagger$ equal the power beam. In the topocentric frame used in visibility calculations, this has the form

$$\boldsymbol{A}_0 = \begin{pmatrix} \sqrt{A_\nu(\theta)} \\ 0 \end{pmatrix} \qquad \boldsymbol{A}_1 = \begin{pmatrix} 0 \\ \sqrt{A_\nu(\theta)} \end{pmatrix}, \tag{3.11}$$

where $\boldsymbol{A}_i$ is the E-field beam of the i^{th} feed.

3.1.5 Performance and Parallelization

Full evaluation of eq. (3.2) is an enormously demanding task. Letting $(N_b, N_t, N_s, N_f, N_a)$ stand for the number of baselines, times, sources, frequency channels, and antennas, respectively, for a simulated observing session, at least $(2 \times N_b \times N_t \times N_s \times N_f)$ matrix multiplications must be carried out. In addition to these, however, coordinate transformations and beam interpolations must be performed at various steps:

In a worst-case scenario, the following tasks must be performed as many times as indicated by the quantities in parentheses:

- Source position transformation into the coordinate frame of the baseline $(N_s \times N_t)$.
- Jones matrix calculation $(N_s \times N_t \times N_f \times N_a)$
- Coherency matrix transformation into local frame $(N_s \times N_t)$.
- Fringe calculation $(N_b \times N_f \times N_t \times N_s)$.
- The product of Jones matrices, coherency, and exponential fringe must be evaluated and summed over the source axis $(N_b \times N_f \times N_t \times N_s)$.

The calculations that must be done, however, are small and independent of each other, making this a task well-suited to parallel processing. Parallelization is generally difficult in Python, because access to the Python interpreter for running threads is limited by the Global Interpreter Lock (GIL) which only allows one thread at a time to use the interpreter, effectively serializing any threaded application.

Parallelization may be achieved in Python in two ways – by running separate *processes*, which causes the operating system to run several Python interpreters at once, or by running lower-level compiled code internally that is able to circumvent the GIL (Marowka, 2018). The second option is done internally by packages such as `numpy` and `scipy`. The first option may be done through the `multiprocessing` module, or through the Message Passing Interface (MPI) (see, e.g., Gropp et al. (1996)).

`pyuvsim` uses MPI, wrapped for Python compatibility with the `mpi4py` package (Dalcín et al., 2005). In MPI, the same code is run in parallel as separate processes which can transmit data to other ranks via *communicators*. For each communicator, processes are given an identifying number called a *rank*. Processes may be spread across multiple computing *nodes*, and it is possible to make communicators among them that allow for communication across multiple nodes, or within a single node, or through other arbitrary characteristics. It is possible, and often advisable, to allocate several CPUs per process. Each rank has its own address space, which usually means that variables on one process cannot be accessed by others unless those variables are sent. As of MPI version 3, however, there is limited support for shared memory among processes on the same node (Hoefler et al., 2015).

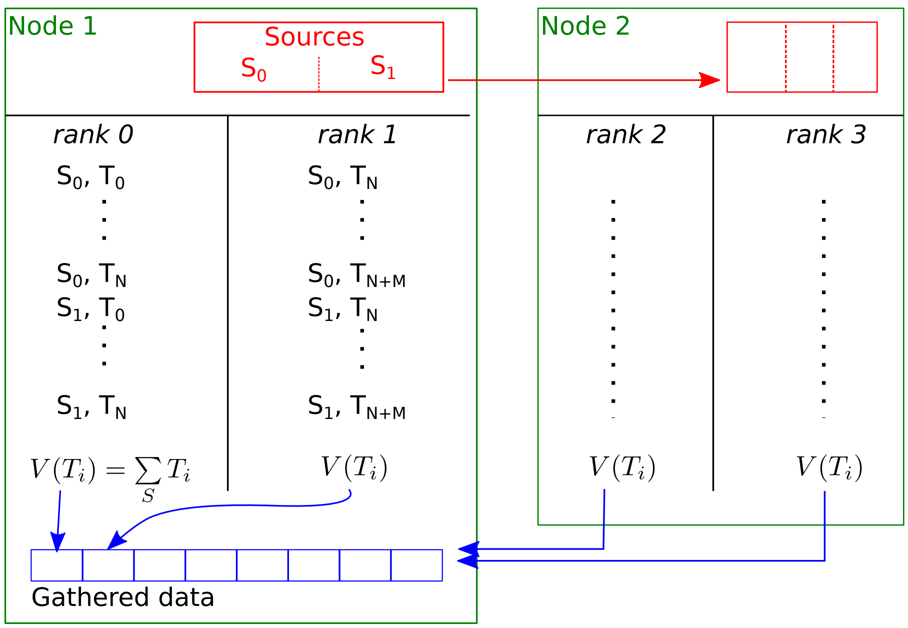


Figure 3.2: A cartoon diagram of the parallelization used in `pyuvsim`. The set of of all combinations of baselines, time, and frequency is divided into subsets of *tasks* T_i such that $T_i = (b, f, t)$ for one particular combination of baseline, time, and frequency. In this example, there are four processes spread across two nodes. On each node, the set of sources is kept in shared memory (indicated by the red box). This set of sources is split into two chunks, S_0 and S_1, on Node 1, but may be split differently on other nodes depending on memory requirements. Arrows denote MPI communications. More details on this may be found in section 3.1.5.

The visibility calculation is broken up into *tasks*, each of which is identified with a single baseline, time, frequency, and a *chunk* of sources. Calculations are parallelized over the source axis internally using `numpy`'s multithreading, so sources are treated in aggregate and separately from the "task" logic. The total number of tasks is, thus, $N_{\text{tasks}} = (N_t \times N_f \times N_b)$.

Initial setup steps, such as reading and parsing the configuration files, initializing time, baseline, and frequency arrays, and reading the source catalog, are all done on the root (rank 0) process and then broadcast to other processes. From the total number of tasks, number of processes, and rank number, each process will identify which of the tasks it is responsible for and execute the loops described in Pseudocode 1 over that subset of times, frequencies, and baselines.

The source array is often the largest object that must be carried in memory. For any reasonably good sky model, there will be orders of magnitude more sources than times, frequencies, or baselines. In the most extreme case, each source carries $4 \times N_f$ values specifying the Stokes parameters at each frequency. Naïvely broadcasting the array of source parameters to all processes means multiplying the memory footprint of one array by the number of processes, creating a strict tradeoff between parallelization and memory consumption. I will refer to this growth of memory usage with parallelization as *memory bloat*.

To alleviate memory bloat, the immutable source data such as source names, IDs, ra, dec, and fluxes are all placed in a *shared memory window* on each node, to be read from by the separate ranks as needed. This ensures that the entire catalog need only be copied once per node. Of course, there are intermediate data products with the same dimensions as the source array, and these need to be generated on each rank separately. The source array is therefore processed in chunks, which are selected to prevent memory overflow through the processing steps.

Fig. 3.2 provides a diagram of the parallization structure. Individual nodes are shown as green boxes (in this case, there are two nodes), and columns mark the separation of tasks (T_i) and source chunks into each rank. Red boxes indicate the shared memory window carrying the source data, with a red arrow marking the "broadcast" operation that copies the source array into a window on each node. Note that this full source list may be split differently depending on the number of processes on a given node and the memory available there. On Node 1, there are N tasks on the zeroth rank and M tasks on the second. Computed visibilities are implicitly summed over sources within each source chunk in each rank, and then summed over source chunks after the loop (denoted by the sum at the bottom of each column). This provides a visibility per task $V(T_i)$ – recall that the "task" refers to a specific baseline/time and frequency. These finished visibilities are then gathered from separate ranks using an MPI gather operation.

Precisely estimating the expected runtime for `pyuvsim` is difficult, because increases along each simulation axis (time, sources, frequencies, baselines) scale differently with available resources. Python "for" loops are inherently slower than `numpy` vectorized operations, and the source axis is treated using numpy vectorization. Some degree of Python parallelization must be maintained in order to keep to the scalability requirement of `pyuvsim`, as a purely `numpy`-accelerated code would be confined to a single node.

As a general rule, more MPI processes are required to accelerate calculations that have more baselines, times, and frequencies. More CPUs may be allocated per process to speed up calculation for more sources. As a benchmark test, a simulation with a 730988 sources,[5] 64 baselines, one time, analytically-defined beams, and a hundred frequencies completed in 30 minutes when given 20 MPI processes and 4 CPUs per process on Brown University's Oscar cluster.

3.2 `healvis`

In the summer of 2018 I started writing another simulator, called `healvis`, out of a need for quick simulations of EoR-like signals. Simulating such signals requires a way of discretizing diffuse maps, and with fluxes specified at all frequencies. At the time I was writing `healvis`, `pyuvsim` was capable of simulating limited numbers of flat-spectrum point sources, and though the restrictive design philosophy of `pyuvsim` ensures that it meets high standards of verification throughout development, the tradeoff is that development is slow.

`healvis` is designed to simulate visibilities from a set of HEALPix[6] maps (Gorski et al., 2005), one for each frequency channel, with pixel values representing Stokes-I surface brightness. HEALPix divides the sky into equal-area pixels with a resolution set by a single parameter, `Nside`, such that the individual pixel area is $\omega = 4\pi/(12{\times}\mathrm{Nside}^2)$ sr. Higher-resolution pixels are nested within lower-resolution pixels in a well-defined hierarchical scheme, which simplifies the process of averaging or dividing pixels when changing map resolution.

Unlike point source models, there is no natural discretization for resolved diffuse emission. In the case of `healvis`, the map is discretized into single pixel values, and it is assumed that the integrand of the scalar RIME eq. (2.13) changes negligibly over each pixel. We can describe the sky pixellization as a set of functions $W_p(\hat{s})$, for pixels p in the map, which is 1 when $\hat{s}$ points into the pixel and 0 otherwise.

$$V(\boldsymbol{b},\nu) = \sum_p \int_{\mathrm{sky}} W_p(\hat{s}) I(\hat{s},\nu) A(\hat{s},\nu) e^{-2\pi i \hat{s}\cdot\boldsymbol{b}\nu/c} d\Omega \tag{3.12}$$

$$\approx \sum_p I(\hat{s}_p,\nu) A(\hat{s}_p,\nu) e^{-2\pi i \hat{s}_p\cdot\boldsymbol{b}\nu/c} \int_{\mathrm{sky}} W_p(\hat{s}) d\Omega \tag{3.13}$$

The integral over each pixel window function then reduces to the solid angle of that pixel, ω_p,[7] and

[5] An Nside 128 HEALPix map, with only sources that will rise above the horizon selected.

[6] Hierarchical Equal-Area Isolatitude Pixellization

[7] It is unfortunately common practice to use capital Ω to refer to several solid angle quantities, including pixel areas and integrals of the beam. To avoid confusion, I will use lowercase ω for pixel areas.

the measurement equation may be written in terms of the pixel values in the maps $I_p(\nu) = I(\hat{s}_p, \nu)$.

$$V_b(\nu) = \sum_p \omega_p I_p(\nu) A(\hat{s}_p, \nu) e^{-2\pi i \hat{s}_p \cdot \boldsymbol{b}\nu/c} \tag{3.14}$$

Equation (3.14) treats the diffuse sky model as a set of point sources at pixel centers with specific flux densities $I_p\omega_p$. We therefore refer to this as the *point source approximation* (PSA). Chapter 4 discusses the limitations of the PSA and explores the sources of errors in simulations of diffuse emission.

3.2.1 Structure

`healvis` accepts as input,

1. A HEALPix *shell*, consisting of a set of HEALPix maps representing the Stokes-I surface brightness of the sky at different frequencies.
2. A set of frequencies, $\mathbb{F}$. These must match a subset of the frequencies in the shell.
3. A set of times to simulate, $\mathbb{T}$.
4. A telescope name and location.
5. A collection of antenna positions, specified with a name, ENU position, and ID number.
6. A beam model, specified as either an analytic power beam or a path to a beamfits file. The same beam is used for all antennas.
7. A Field of View (FoV) parameter. This sets the diameter of a circular selection around zenith that will be included in the integration (e.g., choose 180° to reach the horizon).
8. Options to choose which polarization and frequency interpolation to use with `UVBeam`.
9. Options for smoothing data-defined beams using Gaussian Process Regression (GPR). More details on this are presented in chapter 5.
10. Optional selection parameters for choosing subsets of baselines, frequencies, times, etc. (as in `pyuvsim`).

The configuration is specified using YAML markup files, similar to the inputs of `pyuvsim`, in line with the effort to let `pyuvsim` define the standard. Note that `healvis` has several restrictions in its use, including that the same beam is used by all antennas, the sky model must match the instrument model in frequency, and only one instrument polarization (i.e., among the entries in the matrix $\mathbf{V}$) may be calculated at a time. These limitations reflect the original intent of `healvis` to simulate EoR-like models, which ideally are not interpolated or averaged at any time. Nonetheless, useful results may be obtained within these constraints.

Pseudocode 2 `healvis` calculation.

Setup baseline vectors (as $\mathbb{B}$), frequencies, times, primary beam.
If using a `UVBeam`, convert to a power beam and interpolate to the required frequencies.
Read sky model from file, or generate from given parameters.
for all t in $\mathbb{T}$ **do**
    Calculate the *pointing center* at time t as the ICRS coordinates of the zenith, using `astropy` methods.
    Calculate the position of the Earth's North Pole in the ICRS frame, at time t.
    Select pixels within a radius of FoV/2 of the pointing center.
    From pixel center positions, pointing center, and North pole vector, calculate the topocentric positions (as Alt/Az and as (l, m, n) relative to zenith) of each pixel.
    for all b in $\mathbb{B}$ **do**
        for all f in $\mathbb{F}$ **do**
            Evaluate the fringe at all source positions.
            Evaluate the primary beam at all positions.
            Multiply beam, fringe, and brightness values, summing over the pixel axis.
            Multiply by pixel area ω and convert from Kelvin-sr to Jansky.
        end for
    end for
end for
Construct a UVData object to handle the output data and write to file.

Pseudocode 2 describes the overall structure of `healvis` visibility calculation. Setup is comparable to `pyuvsim`. If an `UVBeam` object is used, it is first interpolated to the frequencies of the simulation before the calculations begin, in order to allow for additional smoothing to be applied if desired. Essentially, a `UVBeam` is just a collection of samples of electric field components or power beam values at various angular positions and frequencies. The frequency interpolation works on a per-pixel basis, essentially making new angular maps at the instrument frequencies. Angular interpolation is done within these maps. All of these interpolation methods are built into `UVBeam`, but unlike `pyuvsim` the frequency interpolation is explicitly performed first.

3.2.2 Coordinate Transformations

A key difference between `healvis` and `pyuvsim` is that `healvis` performs a vectorized calculation of all pixel positions relative to the pointing center, while `pyuvsim` explicitly calculates the RA/dec to topocentric transformation for each source position using `astropy`. In `healvis`, the North pole position and pointing center are both calculated with `astropy`, but then the remaining pixel positions are calculated relative to these points.

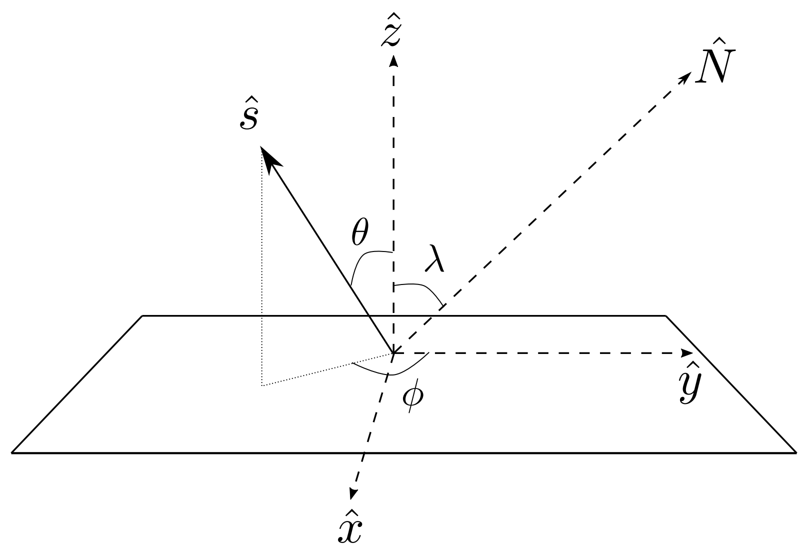


Figure 3.3: Coordinates and unit vectors in the local topocentric frame, used for finding the zenith angle (θ) and azimuth (ϕ) of a source at position $\hat{s}$. $\hat{N}$ points to the Earth's North pole, and $\hat{z}$ to the zenith.

For a given time, the `healvis` calculation requires knowledge of the pointing center and North pole position in the ICRS frame. "North pole" means the North pole of the Earth's rotation axis, which defines the origin of the azimuth angle. This does *not* coincide with the North Celestial Pole (NCP) in the ICRS frame, since the Earth's north pole changes due to precession and nutation. Fig. 3.3 indicates the coordinates and relevant axes. The zenith is in the $\hat{z}$ direction, $\hat{N}$ points due North, and $\hat{x} = \hat{y} \times \hat{z}$. $\hat{N}$ points to the North pole and makes an angle λ with the zenith (i.e., the latitude). A source at position $\hat{s}$ has a zenith angle θ and azimuth ϕ, with respect to the y axis.

To find ϕ and θ, we solve for the (x, y, z) components of $\hat{s}$. We have the following relations:

$$\hat{x} = \frac{\hat{N} \times \hat{z}}{\sin\lambda} \tag{3.15}$$

$$\hat{y} = \hat{z} \times \hat{x} = \frac{\hat{z} \times \hat{N} \times \hat{z}}{\sin\lambda} \tag{3.16}$$

We get the (x, y, z) components by dotting $\hat{s}$ with the respective axes. Solving for the azimuth and zenith angle is then straightforward:

$$\tan\phi = \frac{y}{x} = \frac{\hat{s}\cdot(\hat{N}\times\hat{z})}{\hat{s}\cdot(\hat{z}\times\hat{N}\times\hat{z})} \tag{3.17}$$

$$= \frac{\hat{s}\cdot(\hat{z}\times\hat{N})}{\hat{s}\cdot\hat{z}\cos\lambda} \tag{3.18}$$

$$\cos\theta = \hat{s}\cdot\hat{z} \tag{3.19}$$

In eq. (3.18) we've applied the triple product identity to simplify the expression in the denominator. In **healvis**, the unit vectors are all computed from ra/dec positions directly in the ICRS frame.

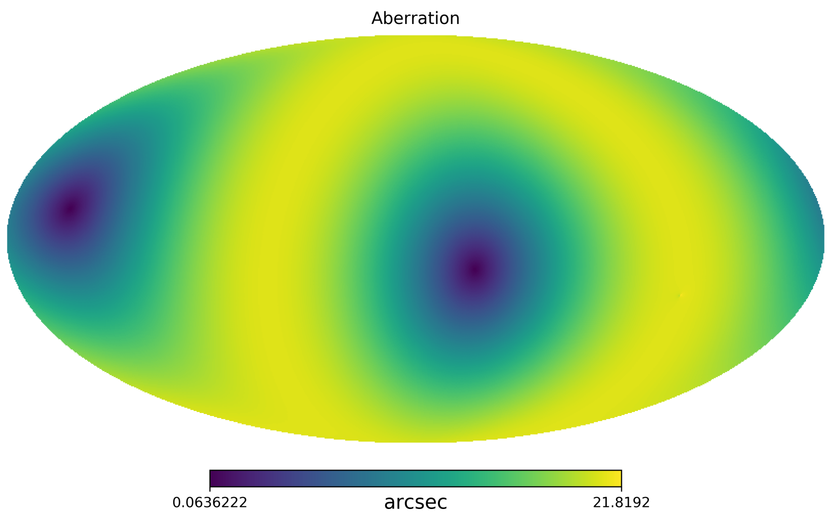


Figure 3.4: Stellar aberration, expressed as the difference between the Geocentric Celestial Reference System (GCRS) and the ICRS frame. The GCRS frame is defined from Cartesian axes parallel to those of the ICRS frame, but with an origin set at the Earth's center of mass instead of the solar system barycenter, such that the GCRS frame includes aberration effects. Aberration introduces a time and angle-dependent deviation on order of tens of arcsec, which is not accounted for in the alt/az calculation used in **healvis**.

Calculating topocentric coordinates from the vectors this way is about twice as fast as applying the **astropy** functions, but implicitly assumes that the change from ICRS to topocentric frame exactly preserves the relative positions of pixel centers. This is not strictly true, as stellar aberration has a varied effect across the sky. Annual and diurnal aberration are relativistic effects due to the rotation and revolution of the Earth, respectively, and can introduce distortions in the apparent positions of objects by up to tens of arcseconds (Fig. 3.4). This can be considered significant for long baseline interferometry. The difference in calculated topocentric positions is noticeable when comparing healvis to pyuvsim calculations, but have no discernable impact on current EoR simulations. Offsets on this scale are well below what most EoR instruments can resolve.

3.2.3 Parallelization

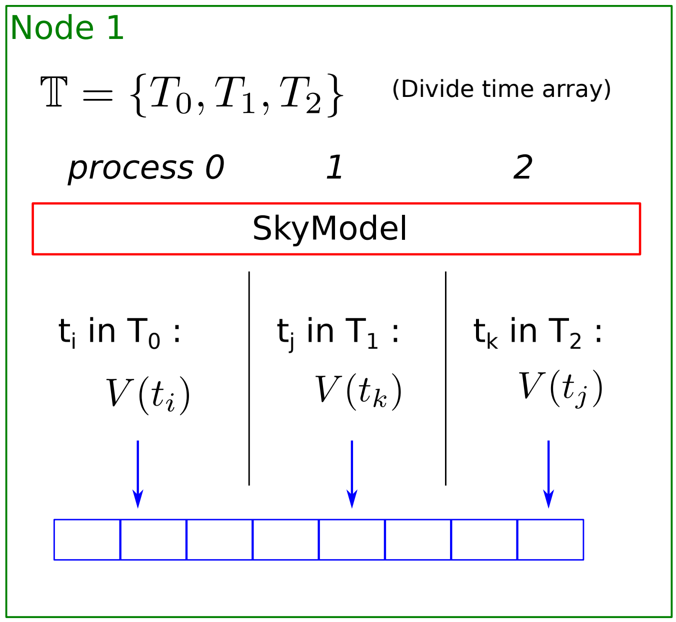


Figure 3.5: A cartoon diagram of the parallelization used in healvis. Separate processes are established to work on subsets of the time array, as denoted by T_0, T_1, and T_2 here. The sky model data is kept in a shared memory array accessible to all processes, while baselines, frequencies, and other data are copied to all processes. The calculated visibilities are gathered with a queue object and inserted into the output data array.

healvis is written with a much simpler and limited parallelization scheme than pyuvsim. It is parallelized along the time axis using the Python multiprocessing module, and uses numpy's implicit multithreading to parallelize along the source, baseline, and frequency axes. The multiprocessing module allows a single calling Python thread to manage multiple concurrent processes, each running a specified function and with its own copy of the data given it. In healvis, the time array is split into N_{proc} parts, where N_{proc} is a specified number of processors.

To reduce memory bloat, the sky model data is kept in a shared memory array, which is provided by the Python multiprocessing module, allowing each process to access the HEALPix maps without causing them to be copied to each process. However, other large arrays do need to be calculated on each process individually, so memory bloat cannot be completely avoided. The FoV selection is key to this tradeoff, as it determines how many pixels will be read from shared memory into the local memory of each process.

This parallelization scheme is shown in Fig. 3.5, in analogy to Fig. 3.2. Note that in this diagram, the capital-T's refer to subsets of the time array only. All other axes are looped over implicitly (ie. using `numpy` vectorization) on each process. The shared memory array is marked again by a red box, but it is not treated in chunks.

There are some downsides to this design. In particular, the `multiprocessing` module is not designed to operate across multiple nodes, and so it is not scalable to large computing clusters. The process-based parallelization is only used to split the job up by times, so running with a large number of processes gains no advantage for simulations that are large on other axes. This was done because originally `healvis` was meant to simulate small numbers of baselines for 24 hour long simulations (see chapter 6). Some advantage can still be gained from allocating more processing cores for the job, as long as `numpy` is compiled for multithreading support.

3.2.4 Testing and Verification

Chapter 4 discusses in detail the comparisons between `healvis` simulations and analytic solutions to the scalar RIME. Most of the main features of `healvis` are covered with unit tests, which verify the following any time a change is made:

1. The visibility of a single nonzero pixel $5°$ off zenith is equal to the analytic solution for an off-zenith point source (mimicking the off-zenith test in `pyuvsim`).

2. The amplitude of the power spectrum of a flat-spectrum EoR-like model is correct, using a simple delay spectrum estimator.

3. For a sequence of times when the galactic center is passing overhead, the visibility amplitude peaks when the galactic center is at zenith. This test uses a low-resolution version of the Galactic Sky Model (GSM) (see section 5.3.2).

In addition to these internal tests, `healvis` has been compared to equivalent simulations in `pyuvsim` and verified with imaging tests. Fig. 3.6 shows a side by side comparison of a test pattern drawn onto a HEALPix map and an image made from simulated observations of this test pattern. The pattern is made as a regularly spaced grid, with pixel brightness set to drop off with latitude. The points on the grid are chosen as the pixel centers of an Nside 12 map, which are still at a pixel center at higher resolution (Nside 128) due to the hierarchical design of HEALPix. The simulation selected a region $50°$ in radius from the zenith, leading to the circular cutoff in the image.

Furthermore, data from simulations of EoR models have been processed with several power spectrum estimation methods, which have all successfully recovered the correct power spectrum

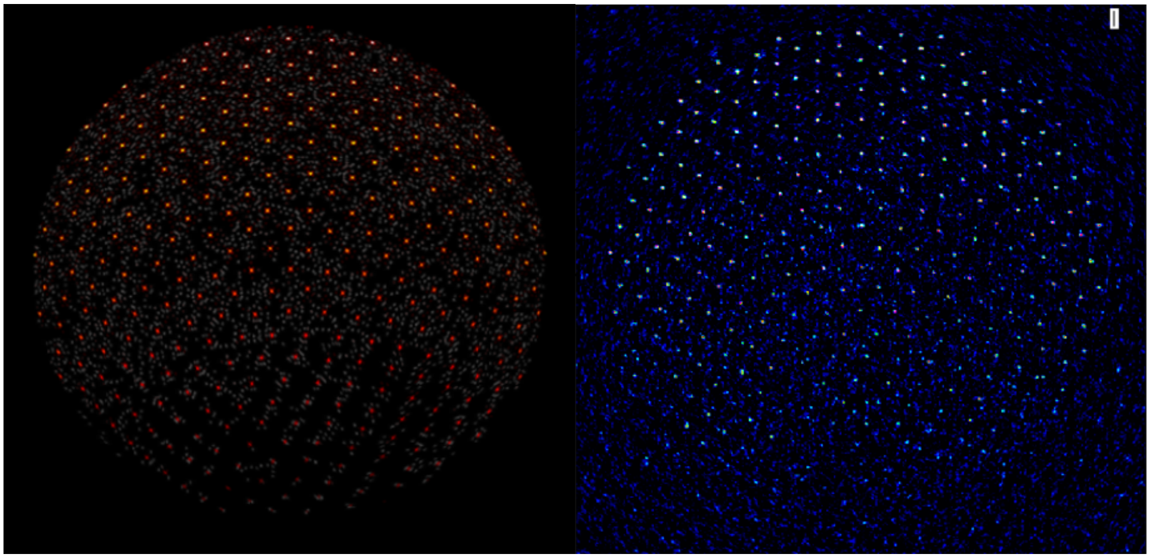


Figure 3.6: On the left, a test pattern of single nonzero pixels with brightness that drops off with latitude on the map. Simulating data for a large array of randomly-distributed antennas in `healvis`, then imaging in `casa`, produces the comparable pattern on the right. The left plot was made by convolving a 6 arcmin fwhm gaussian with the input map and plotted in an orthographic projection, so there is some distortion that doesn't match the right (in a Cartesian projection). This test indicates that `healvis` simulations put sources in the right orientation and brightnesses on the sky (unit tests check this more precisely with the calculated visibilities).

amplitude. Fig. 3.7 shows a comparison of power spectra from simulated data made with `healvis` using three sky models — one is of diffuse foreground power, the second is of an flat-spectrum EoR-like model, and the third combines the two. The power spectra are estimated from delay transforms using a Blackman-Harris window function. The EoR signal is clearly distinguishable above the residual foreground power at high $k_\parallel$, and matches the amplitude of the input model.

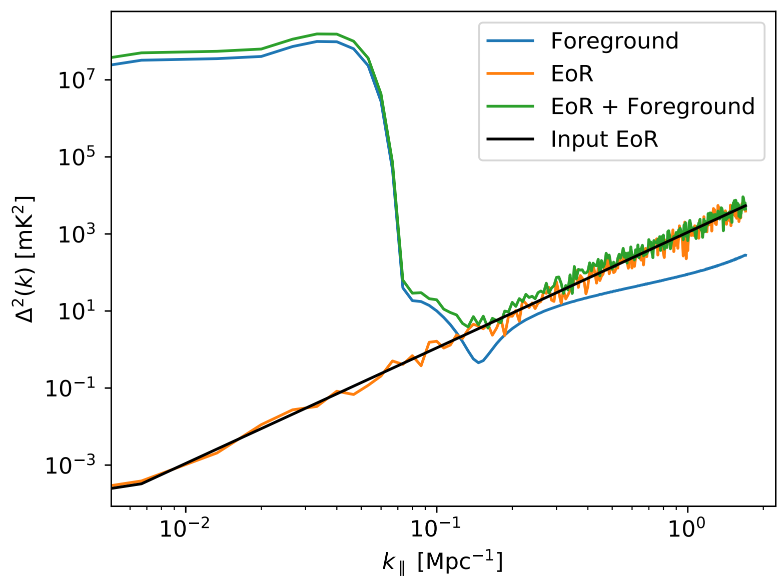


Figure 3.7: Dimensionless power spectra of a single $k_\perp$ of a hexagonal arryay, from three simulations with `healvis`, using an Airy disk beam. One has a flat-spectrum EoR-like model (amplitude 28^2 mK^2 Mpc^3), the second has a diffuse foreground model, and the last combined the two. The black line marks the input power spectrum amplitude. The input power is flat in dimension*al* units, but is presented here as a dimensionless power spectrum to clarify the separation at large $k_\parallel$.

Chapter 4

Analytic Solutions for Diffuse Models[1]

The trustworthiness of `pyuvsim` simulations depends in large part on our ability to verify its calculations against exact analytic solutions for point source models. As discussed in section 3.1, point sources naturally discretize the integral in the measurement equation, which makes numerical evaluation straightforward. Diffuse and extended emission that is resolved by the instrument cannot be treated as a point source. Baselines will record less flux from extended sources when they are sensitive to larger or smaller angular scales than the source itself. This fundamentally different behavior must be tested as well.

Specifically, "extended" sources are contiguous astrophysical objects, such as supernova remnants or radio galaxy lobes, that are resolved. It is generally safe to consider extended sources in the flat-sky approximation of eq. (2.17), which are related to the pattern on the sky by way of a 2D Fourier transform. These structures can be represented in a variety of forms, such as pixel maps, shapelets (Refregier, 2003), or CLEAN components (Cornwell, 2008). Simulations of such sources can be designed to work with whatever representation is optimal. Validating these simulations is more straightforward, as wide-field effects are generally negligible.

In contrast, "diffuse" emission is resolved, large scale emission on the sky that is not associated with a individual astrophysical objects, largely consisting of emission from the Galaxy but also the EoR signal and CMB. This emission cannot be treated in the flat sky approximation, and is often modeled using pixel maps. Instrument simulations, and especially those of wide-field EoR

[1]Much of the content of this chapter was written in collaboration with Steven Murray and Daniel Jacobs at Arizona State University, and is the subject of a forthcoming paper. Text and figures are reproduced with permission of co-authors.

experiments, must be validated for accuracy and precision while taking these effects into account. Chapter 5 looks at some of these wide-field effects in simulation.

Regardless of the level of information in the sky model, simulations of diffuse models are limited by the method with which they integrate the RIME. Analytic solutions for the visibility function can provide an exact point of comparison. In this chapter, I present several such solutions and use them to test the accuracy and precision of `healvis`. I find that `healvis` is capable of modeling diffuse emission accurately, with remaining errors on a negligble scale for the remainder of this thesis.

4.1 Setup

The scalar RIME can be written in direction cosines, as in eq. (2.18):

$$V(u,v,w) = \int \frac{d^2\boldsymbol{l}}{n}\ A(\boldsymbol{l})I(\boldsymbol{l})e^{-2\pi i[\boldsymbol{u}\cdot\boldsymbol{l}+w(1-n)]}, \tag{4.1}$$

where we've written $\boldsymbol{l} = (l,m)$ and $\boldsymbol{u} = (u,v)$, and $n = \sqrt{1-l^2-m^2}$. Within the integral, A and I are entirely interchangeable, and are constrained by the fact that the response falls to zero at the horizon, where $r = |\boldsymbol{l}| = 1$. We thus re-write $AI \implies WT$, with

$$W(r) = \begin{cases} 1 & r \leq 1 \\ 0 & r > 1 \end{cases} \tag{4.2}$$

and T considered to be the "observed" sky intensity (i.e. the beam-attenuated sky intensity).

Noting this, we write

$$V(u,v,w) = \int \frac{d^2\boldsymbol{l}\ W(r)}{n} T(\boldsymbol{l})e^{-2\pi i[\boldsymbol{u}\cdot\boldsymbol{l}+w(1-n)]}. \tag{4.3}$$

4.2 Analytic Derivations

In this section we derive from first-principles a few analytic forms for the visibility, given functional forms for $T(\boldsymbol{l})$. Some of these solutions are defined in terms of certain special functions, such as the hypergeometric function, Bessel function, and Sine integral, which are all defined in Appendix B. There are multiple mathematical routes to these answers, some of which are complementary.

4.2.1 Projected Solutions

The main challenge comes from the factor of $1/\sqrt{1-|\boldsymbol{l}|^2} = 1/n = \cos\phi$ – the Jacobian of the spherical-to-(l,m) coordinate change. This factor projects the brightness $T(\hat{s})$ into the (l,m)-plane, where the fringe term acts approximately as a Fourier transform. Several of our solutions are reached by making a sky model of the form

$$T(l,m) = T'(l,m)\sqrt{1-l^2-m^2}, \tag{4.4}$$

where T' has a simple 2D Fourier transform. We refer to these as "projected" solutions, because they all feature this factor of $\sqrt{1-l^2-m^2}$. These solutions have the disadvantage that, regardless of their form in (l,m) space, their response near the horizon drops off to zero as $\cos(\theta)$ for zenith angle θ, and hence have little response near the horizon.

Assuming I is symmetric about zenith, and thus is purely a function of r, the measurement equation 4.3 can be recast as a function of r, and the Fourier transform can be written as a Hankel transform:

$$V(u) = 2\pi \int_0^1 \frac{rT(r)}{\sqrt{1-r^2}} e^{-2\pi i w(1-\sqrt{1-r^2})} J_0(2\pi ur)\ dr. \tag{4.5}$$

In this subsection, we will denote

$$T'(r) = \frac{T(r)e^{-2\pi i w(1-\sqrt{1-r^2})}}{\sqrt{1-r^2}}, \tag{4.6}$$

as the "de-projected" sky intensity, with $n = \sqrt{1-r^2}$. This allows us to write

$$V(u) = 2\pi \int_0^1 rT'(r)J_0(2\pi ur)\ dr. \tag{4.7}$$

Note that $r = \sin\theta$. Thus, neglecting the w-term, we may write the form of the sky in angular coordinates (i.e. that required to simulate in `healvis` to achieve the results presented here) as

$$T(\phi) = \cos\phi\, T'(\sin\phi). \tag{4.8}$$

It is often convenient to substitute $x = 2\pi ur$:

$$V(u) = \frac{1}{2\pi u^2}\int_0^{2\pi u} xT'(x/2\pi u)J_0(x)\ dx. \tag{4.9}$$

Relatively few solutions to this equation exist in analytic form, but we present a few.

Cosine sky

The first solution is uniformly bright in the projected sense, so that

$$T'(r) = T_0 \implies T(r) = T_0\cos\phi \tag{4.10}$$

for constant T_0. Inserting this into eq. (4.9) gives the visibility

$$V_{\cos}(u) = \frac{1}{u}J_1(2\pi u). \tag{4.11}$$

Polynomial Dome

Another general solution is for the form in which

$$T'(r) = 1 - r^{2n} = 1 - (x/2\pi u)^{2n}, \tag{4.12}$$

which describes an inverted dome centered at zenith and falling to zero at $r = 1$. Inserting this into eq. (4.9),

$$V_{\text{poly,n}}(u) = \frac{1}{2\pi u^2} \int_0^{2\pi u} x \left(1 - \left(\frac{x}{2\pi u}\right)^{2n}\right) J_0(x)dx \tag{4.13}$$

$$= V_{\cos}(u) - \frac{2\pi}{(2\pi u)^{2(n+1)}} \int_0^{2\pi u} x^{2n+1} J_0(x)dx \tag{4.14}$$

The remaining integral evaluates to an expression in terms of the generalized hypergeometric function ${}_pF_q$, (see Appendix B),

$$\int_0^{2\pi u} x^{2n+1} J_0(x)dx = \frac{(2\pi u)^{2(n+1)}}{2(n+1)} {}_1F_2(n+1; 1, n+2; -\pi^2 u^2) \tag{4.15}$$

which leaves us with

$$V_{\text{poly,n}}(u) = V_{\cos}(u) - \frac{\pi}{(n+1)} {}_1F_2(n+1; 1, n+2; -\pi^2 u^2) \tag{4.16}$$

The ${}_1F_2$ hypergeometric functions are closely related to Bessel functions. In particular, for $n = 0$ we get

$$\pi \ {}_1F_2(1; 1, 2; -\pi^2 u^2) = J_1(2\pi u)/u, \tag{4.17}$$

which means $V_{\text{poly,0}}(u) = 0$, because this cancels with the first term $V_{\cos}$. Recall that $T'(r) = 1 - r^{2n} = 0$ for $n = 0$.

For $n = 1$ we find

$$V_{\text{poly,1}} = V_{\text{unif}}(u) - \frac{J_2(2\pi u) - \pi u J_3(2\pi u)}{\pi u^2}. \tag{4.18}$$

Higher-order polynomial solutions may also be expressed in combinations of Bessel functions, using eq. (4.16).

Projected Gaussian

A particularly desirable form for which to have a solution is the *projected* Gaussian,

$$T'(r) = e^{-r^2/\sigma^2}. \tag{4.19}$$

Using our basic integral, eq. (4.9), we have

$$V_{\text{proj.gauss}}(u) = \frac{1}{2\pi u^2} \int_0^{2\pi u} x \exp\left(\frac{-x^2}{4\pi^2 u^2 \sigma^2}\right) J_0(x)dx. \tag{4.20}$$

We can use a converging Taylor expansion on either the exponential or the Bessel function (or both). Each of them yields different forms for the solution, which admit different insights into the structure.

For *large* σ (i.e. $\sigma \gtrsim 1$), expanding the exponential makes sense, as the integration then ranges over values for the exponential which have magnitude smaller than unity. This yields

$$\begin{aligned} V_{\text{proj.gauss}}(u) &= \frac{1}{2\pi u^2} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(2\pi u\sigma)^{2k}} \int_0^{2\pi u} dx\ x^{2k+1} J_0(x) \\ &= \pi \sum_{k=0}^{\infty} \frac{(-1)^k}{\sigma^{2k}(1+k)!}\, {}_1F_2(1+k; 1, 2+k; -\pi^2 u^2), \end{aligned} \tag{4.21}$$

where the second equality follows from eq. (4.15). A useful outcome of this equation is that if $\sigma \to \infty$, only $k = 0$ contributes at all, and we obtain $V_{\text{proj.gauss},\infty}(u) = J_1(2\pi u)/u$, as we expect from the cosine sky solution (see eq. (4.11)). This expansion converges quickly for $\sigma > 1$, yielding solutions that resemble the cosine model. Furthermore, it is numerically stable even for large u. In practice, this expansion works well down to $\sigma \sim 0.25$, and is preferred over the small-σ expansion for these values because it is stable for large u.

For small $\sigma \lesssim 0.25$, this expansion is not efficient, and it is better to expand the Bessel function in its Taylor series:

$$\begin{aligned} V_{\text{proj.gauss}}(u) &= \sum_{k=0}^{\infty} \frac{(-1)^k}{2^{2k}(k!)^2} \frac{1}{2\pi u^2} \int_0^{2\pi u} x^{2k+1} \exp\left(-\frac{x^2}{(2\pi\sigma u)^2}\right) \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(k!)^2} \pi\sigma^2 (\pi u\sigma)^{2k} \left[\Gamma(k+1) - \Gamma(k+1,1)\right] \\ &= \pi\sigma^2 \left[e^{-\pi^2\sigma^2u^2} - \sum_{k=0}^{\infty} \frac{(-1)^k}{(k!)^2} (\pi u\sigma)^{2k} \Gamma(k+1, \sigma^{-2}) \right], \end{aligned} \tag{4.22}$$

where $\Gamma(a, b)$ is the upper incomplete Gamma function,

$$\Gamma(a, b) = \int_b^{\infty} t^{a-1} e^{-t} dt. \tag{4.23}$$

The second sum comes up in similar form from other solutions in this chapter, so we'll take a moment now to look at its convergence properties. Note that the incomplete Gamma function can be written as

$$\Gamma(k+1, b) = k! - \gamma(k+1, b)$$

where $\gamma(k+1, b)$ is the *lower* incomplete gamma function,

$$\gamma(a, b) = \int_0^b t^{a-1} e^{-t} dt.$$

We've used the fact that k is an integer to express the complete Gamma function as a factorial instead. The terms in the expansion are then

$$a_k = (-1)^k \frac{(\pi u \sigma)^{2k}}{k!} - (-1)^k \frac{(\pi u \sigma)^{2k}}{(k!)^2} \gamma(k+1, \sigma^{-2}) \tag{4.24}$$

For small σ, $\gamma(k+1, \sigma^{-2}) \to \Gamma(k+1) = k!$, and $a_k \to 0$, leaving just the exponential term. For larger σ, $\gamma(k+1, \sigma^{-2}) \to \sigma^{-2(k+1)}/(k+1)$, so the second term drops off rapidly. The first term then dominates,

$$a_k \approx (-1)^k \frac{(\pi u \sigma)^{2k}}{k!} \tag{4.25}$$

The factorial dominates over the power law as $k \to \infty$, but the series coefficients are not necessarily monotonically decreasing for all k. The ratio of subsequent coefficients

$$\frac{|a_{k+1}|}{|a_k|} = \frac{(\pi \sigma u)^2}{k+1} \tag{4.26}$$

is greater than one for $(\pi \sigma u)^2 > k + 1$, so at least $N \sim (\pi \sigma u)^2$ are needed to ensure that adding additional terms improves the estimate. This means that long baselines u, relative to the parameter σ, will require a large number of correction terms for comparison with simulated visibilities.

In the small sigma case, $a_k \to 0$ and we're left with

$$V_{\text{proj.gauss}}(u|\sigma \ll 1) = \pi \sigma^2 e^{-\pi^2 \sigma^2 u^2}, \tag{4.27}$$

which is exactly the Fourier transform of I', as expected. This solution thus characterizes deviations from this limit by the term containing the sum.

4.2.2 Convolution Solutions

Starting again with eq. (4.3), we note that the integral bounds in principle extend to infinity.

We may treat the combined W, w term and n as a modified circular window function. In this case, it is clear that the integral has the form of a 2D Fourier transform in (l, m) to Fourier duals (u, v). It is the transform of a product of two functions: the window $W \exp(-2\pi i w n)/n$, and the sky surface brightness in response to the instrument, I. By the convolution theorem, the Fourier transform of this product is the convolution of the respective Fourier transforms of each term.

The window function is relatively easy to transform since it has circular symmetry:

$$\begin{aligned}
\widetilde{W}(u,v) &= \int \frac{W(l,m)}{\sqrt{1-l^2-m^2}} e^{2\pi i w\sqrt{1-l^2-m^2}} e^{-2\pi i(ul+vm)} dldm \\
&= 2\pi \int_0^\infty \frac{rW(r)e^{2\pi i w\sqrt{1-r^2}}}{\sqrt{1-r^2}} J_0(2\pi ur)dr \\
&= 2\pi \int_0^1 \frac{re^{2\pi i w\sqrt{1-r^2}}}{\sqrt{1-r^2}} J_0(2\pi ur)dr \\
&= 2\pi \sum_{n=0}^{\infty} \frac{(2\pi i w)^n}{n!} \int_0^1 r(1-r^2)^{\frac{n-1}{2}} J_0(2\pi ur)dr \\
&= 2\pi \sum_{n=0}^{\infty} \frac{(2\pi i w)^n}{n!(1+n)} {}_0F_1\left(\frac{3+n}{2}, -\pi^2u^2\right) \qquad (4.28)
\end{aligned}$$

$$(4.29)$$

The expansion coefficients also go as $(2\pi x)^n/(n!)$, as in eq. (4.25), and requires $\gtrsim (2\pi w)^2$ terms are needed for convergence. When $w = 0$ (the fiducial case), all of the summands are zero except for $n = 0$, giving

$$\begin{aligned}
\widetilde{W}(u) &= 2\pi {}_0F_1(3/2, -\pi^2u^2) \\
&= 2\pi \,\mathrm{sinc}(2\pi|u|). \qquad (4.30)
\end{aligned}$$

Taking $w = 0$, this leaves us with the following expression for the visibility:

$$V(\boldsymbol{u}) = 2\pi \mathrm{sinc}(2\pi|u|) * \widetilde{I}(\boldsymbol{u}), \qquad (4.31)$$

where I and $\widetilde{I}$ are defined over the entire plane. This has the advantage that in principle, I is not constrained to be symmetric around zenith. We also no longer have the $\cos\theta$ tapering toward toward the horizon, for zenith angle θ, so models of strong emission near the horizon can be found. This form reveals some natural closed-form solutions that the direct integration method did not. However, it is not any easier to find functions whose Fourier-transforms that naturally convolve with the sinc function in 2D, so we are limited in the solutions that can be found.

If we assume that I is symmetric about zenith, we can re-write the equation as

$$V(u) = \int_0^\infty du' \int_0^{2\pi} d\theta \;\, \mathrm{sinc}(2\pi u')\widetilde{I}(\sqrt{u^2+u'^2-2uu'\cos\theta}), \qquad (4.32)$$

taking (without loss of generality) θ to be the angle between $\boldsymbol{u}$ and $\boldsymbol{u}'$.

Though we are still limited in the analytic solutions that can be found with this approach, this does provide a quick route to accounting for the horizon with numerical models. Convolutions are relatively easy to carry out numerically, as are 2D Fourier transforms, so the effects of the horizon cutoff on a given diffuse model can be studied by convolving a sinc function the Fourier transform of the model.

Monopole

The simplest possible diffuse sky model is a monopole:

$$T(r) = T_0, \tag{4.33}$$

which has the FT $\widetilde{I}(u) = \delta(u)$. Using eq. (4.32) directly results in

$$V_{\mathrm{monopole}}(u) = 2\pi \,\mathrm{sinc}(2\pi u). \tag{4.34}$$

Gaussian

Taking

$$T(r) = e^{-l^2/2\sigma^2}, \tag{4.35}$$

so that

$$\widetilde{I}(u) = a^2 e^{-a^2u^2}, \quad a^2 = 2\pi^2\sigma^2 \tag{4.36}$$

we find

$$V(u) = 2\pi a^2 e^{-a^2u^2} \int_0^\infty du' \;\sin(2\pi u') e^{-a^2u'^2} I_0(2uu'a^2), \tag{4.37}$$

where I_0 is the modified Bessel function of the first kind.

The remaining integral does not have a known closed-form solution, but we can consider some limiting cases to ensure it behaves appropriately. First, we note that the integrand converges for large u', because

$$I_0(x) \sim \frac{e^x}{\sqrt{2\pi x}} \qquad \text{for } x \gg 1 \tag{4.38}$$

which grows more slowly than the $\exp(-a^2u'^2)$ term drops off. We also have $I_0(x) \sim 1$ for $x \ll 1$.

The solution for $a \ll 1$ should converge to the Gaussian $\widetilde{I}$, as the bulk of its mass is then in the flat-sky approximation. Conversely, $a \to \infty$ corresponds to a uniform sky – a solution which we have already obtained, i.e. the solution is

$$V(u|a \to \infty) = 2\pi^2 \,\mathrm{sinc}(2\pi u), \tag{4.39}$$

where the extra factor of π here comes from the fact that $\lim_{a\to\infty} \widetilde{I}(u)$ is not a Dirac-delta distribution, but is inflated by a factor of π.

Expanding I_0 in its power series we have

$$\begin{aligned}
V(u) &= 2\pi a^2 e^{-a^2u^2} \sum_{k=0}^{\infty} \frac{(\pi u a)^{2k}}{\Gamma^2(k+1)} \int_0^{\infty} du' \;\sin(2\pi u') e^{-a^2 u'^2} u'^{2k} \\
&= 2\pi^2 a^2 e^{-a^2u^2} \sum_{k=0}^{\infty} \frac{(\pi u a^2)^{2k}}{\Gamma^2(k+1)} \frac{\Gamma(1+k)}{a^{2(k+1)}} {}_1F_1\left(1+k; 3/2; -\frac{\pi^2}{a^2}\right) \\
&= 2\pi^2 e^{-a^2u^2} \sum_{k=0}^{\infty} \frac{(\pi u a)^{2k}}{k!} {}_1F_1\left(1+k; 3/2; -\frac{\pi^2}{a^2}\right) \qquad (4.40)
\end{aligned}$$

Note that for $u = 0$ we have simply $V(0) = 2\pi^2$, as required by our boundary solutions.

The next question is whether this series converges. Recall that we expect this to converge to a Gaussian for small a and to a sinc for large a. The expansion above has the form of a multiplicative correction to a Gaussian, so it is unlikely that it will neatly converge to a form resembling a sinc function. The hypergeometric function converges to 1 for large a, but becomes small at small a. With increasing k, it grows smaller at small a. The convergence of the expansion is therefore dominated by the coefficients $a_k = (\pi u a)^{2k}/(k!)$, which have already been discussed with eq. (4.25).

Unfortunately there doesn't seem to be an expansion for $a >> 0$ that integrates neatly, as in the case of the projected Gaussian. We leave this to future work.

Off-zenith Gaussian

The convolutional formalism allows us to find a solution for an off-zenith Gaussian. If we imagine a Gaussian centered off-zenith, we get by the Fourier shift theorem that

$$\widetilde{I}_{SG}(\boldsymbol{u}) = e^{-2\pi i \boldsymbol{l}_0 \cdot \boldsymbol{u}} \widetilde{I}_G(|u|), \qquad (4.41)$$

with $\boldsymbol{l}_0$ the centre of the blob.

Then we have

$$V(u) = e^{-2\pi i \boldsymbol{l}_0 \cdot \boldsymbol{u}} \int_0^{2\pi} d\theta \int du' \;\sin(2\pi u') e^{2\pi u' i (l_0^x \cos\theta + l_0^y \sin\theta)} \widetilde{I}(|u|). \qquad (4.42)$$

Using the Gaussian form, we arrive at the same formula as the Gaussian solution above, but with a^4u^2 in the hypergeometric function replaced with

$$4u^2a^4 + iua^2\pi l_0^x - 4\pi^2 |l|^2, \qquad (4.43)$$

as well as the extra leading phase term $\exp(-2\pi i \boldsymbol{l} \cdot \boldsymbol{u})$.

4.2.3 Separable Solution

The solutions in the previous sections are all axisymmetric in UV space, and hence not affected by baseline orientation. In this section, we derive a solution that is separable in (l,m) coordinates and not axisymmetric. The main obstacle to finding separable solutions like this is the horizon boundary, which is circular in (l,m) space and hence not explicitly separable. We overcome this by making an artificial "square" horizon, defining the brightness to be zero outside of a square region on the sky. The model defined within that square has a closed form solution.

Let x and y be orthogonal axes rotated some angle ξ from the l and m axes, such that $\boldsymbol{x} = (x, y)$ is related to $\boldsymbol{l}$ via

$$\boldsymbol{x} = R_\xi \boldsymbol{l} \tag{4.44}$$

for rotation matrix R_ξ. Define,

$$T(x,y) = \begin{cases} \mathrm{sinc}(ax)\,\mathrm{sinc}(ay)\sqrt{1-x^2-y^2} & |x| \text{ and } |y| < \frac{1}{\sqrt{2}} \\ 0 & \text{otherwise} \end{cases} \tag{4.45}$$

where a is a constant. The limits on $|x|$ and $|y|$ describe a square region inscribed within the horizon, which lets us carry the integral to the horizon while maintaining separability of the function. Inserting this into eq. (4.1), we obtain

$$\begin{aligned} V(\boldsymbol{u}) &= \int \mathrm{sinc}(ax)\,\mathrm{sinc}(ay)\sqrt{1-x^2-y^2}\,e^{-2\pi i \boldsymbol{l}\cdot\boldsymbol{u}} \frac{d^2 l}{\sqrt{1-l^2-m^2}} \\ &= \int \mathrm{sinc}(ax)\,\mathrm{sinc}(ay) e^{-2\pi i \boldsymbol{l}\cdot(R_\xi^{-1}\boldsymbol{u})} dx dy \\ &= \int_{-1/\sqrt{2}}^{1/\sqrt{2}} \mathrm{sinc}(ax) e^{-2\pi i x u'_x} dx \int_{-1/\sqrt{2}}^{1/\sqrt{2}} \mathrm{sinc}(ay) e^{-2\pi i y u'_y} dy \end{aligned} \tag{4.46}$$

In the last step, we've defined $\boldsymbol{u'} = (u'_x, u'_y) \equiv R_\xi^{-1}\boldsymbol{u}$, and switched the integration variables. Each integral has a closed form solution in terms of sine integrals (see Appendix B):

$$\int_{-1/\sqrt{2}}^{1/\sqrt{2}} \mathrm{sinc}(ax) e^{-2\pi i u'_x x} dx = \frac{\mathrm{Si}[(a+2\pi u'_x)/\sqrt{2}] + \mathrm{Si}[(a-2\pi u'_x)/\sqrt{2}]}{a} \tag{4.47}$$

This solution has roughly the form of a square rotated by an angle ξ in the UV plane. This is not azimuthally symmetric, and can thus be used to test how well the simulator handles baseline orientation.

4.3 Comparison of Simulation to Theory

We construct the sky models for some of the solutions derived in section 4.2 as HEALPix maps. For each model, we construct maps with Nside parameters of 32, 64, 128, 256, 512, and 1024, to demonstrate the effect of changing resolution on the accuracy of the simulation.

The maps are defined with a center at ra/dec of $(0, 0)$, and the pointing center in all simulations is assigned exactly there. In this way, the azimuthal symmetry of the models used in simulation is guaranteed. The sampling however, as determined by the distribution of HEALPix pixel centers about this point, may not be perfectly symmetric. These simulations therefore examined the errors due to the pixelization and point source approximation of the sky model and didn't look at the effects of pointing errors or errors with sky rotation.

4.3.1 Flat Array

The simulator is given a fixed layout of 128 antennas distributed according to a radially-Gaussian profile with FWHM of 228 m, with 10 frequencies in steps of 80 kHz starting at 100 MHz. This provides a good UV sampling that is consistent across simulations, and especially on scales near $|u| \ll 1$ that are most sensitive to large angular scales. The models are all set with an amplitude of $T_0 = 2.0$ K.

Our models are labelled as follows:

- COSZA : Cosine of the zenith angle.
- QUADDOME : The polynomial sky model with $n = 1$ (hence, quadratic).
- MONOPOLE : Uniform brightness in all pixels.
- PROJGAUSS : A projected Gaussian blob in (l,m) coordinates, with width parameter a. We simulate $a = 0.05$, 0.5, and 2.5, which show a clear transition in behavior.
- XYSINC : Sinc functions in X/Y directions on the sky, defined at an angle $\xi = 45°$ with respect to the l axis. The parameter $a = 64$.

For the azimuthally-symmetric models (i.e., all except XYSINC), we compare simulated visibilities to the analytic solutions as a function of baseline length. Fig. 4.1 shows the comparison of the four azimuthally-symmetric models to the analytic solution. The imaginary and real components of the visibility are plotted separately. Each of these plots shows data from simulations run at Nside 1024. For these plots, the visibilities are converted to units K sr. The analytic solution for the projected Gaussian is evaluated to 15 terms, whic is sufficient for convergence. The three width parameters shown ($a = 0.05$, 0.5, and 2.5), show a clear transition from a Gaussian shape to a jinc pattern with increasing a. The solutions are fully real due to their symmetry on the sky, so the imaginary components are small, tending to zero, and can be used to check the integrator at small values.

Fig. 4.2 shows the results of the xysincs model simulation at Nside 1024, side by side with the analytic model. Fig. 4.3 shows the comparison to the analytic model as in Fig. 4.1 for a slice along $y \sim 0$.

4.3.2 Non-coplanar Array

The full expansion in eq. (4.29) must be convolved with the 2D Fourier-transform of the sky brightness to obtain the visibility for baselines with nonzero w terms. Though difficult in general, this can be done trivially for a monopole sky model, whose Fourier transform is a 2D Delta function. When we assumed previously that $w = 0$, this gave us a sinc function. Now, we relax that assumption and see how well eq. (4.29) describes the visibilities of a non-coplanar array viewing a monopole sky.

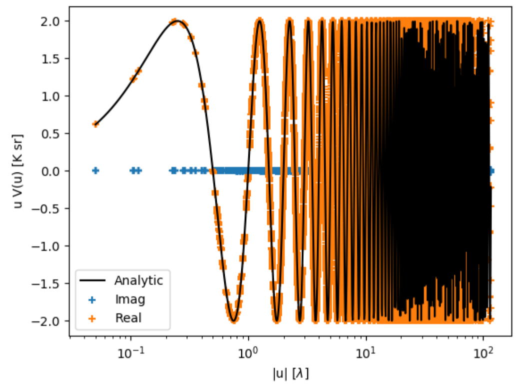
The terms in the expansion eq. (4.29) are fairly well-behaved for most cases. As a result of the factor of i^n, odd-numbered terms are imaginary, while even numbered terms are real, and every other odd term and every other even term has opposite sign. The expansion was taken about $w = 0$, so we expect low-order truncations to be accurate when w is small compared with the baseline length u. To determine the number of required terms for a good estimate, we'll consider the factors in the summand separately. The hypergeometric function ${}_0F_1((3+n)/2, -\pi^2u^2)$ is observed to go to 1 for large n, with a rate that increases with decreasing u. In light of this, the convergence is dominated by the behavior of the factor $a_n = \frac{w^n}{(n+1)!}$ as in eq. (4.25), so for large w we argue a decent estimate may be found by choosing $N_{\max} \sim 2|w|$ terms in the sum.

We define a test layout consisting of 30 antennas arranged in a line from southwest to northeast, duplicated twenty times at heights ranging from zero to 10 meters. Baselines are defined from each antenna to the southwest-most antenna at height 0. This is unphysical, since antennas are stacked on top of each other, but it provides a set of baselines with equivalent lengths in the UV plane but different w terms.

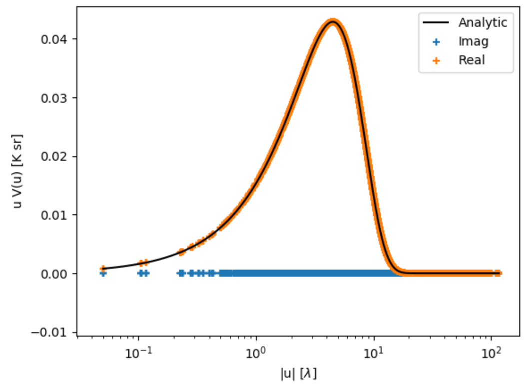
With the test layout, we simulate the monopole sky model at Nside 256. The real and imaginary components of this simulation are plotted in Fig. 4.4, along with the analytic result expanded to twenty terms. Each color indicates a different w-term.

4.4 Error Analysis

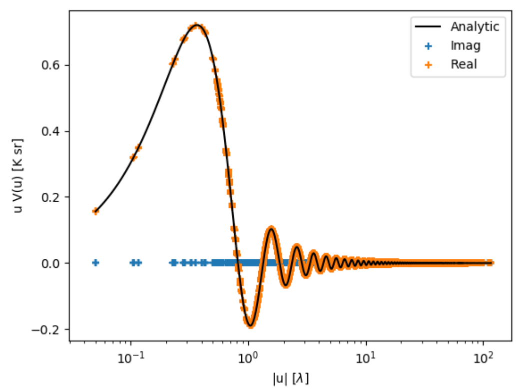
We'll look now at the error $\varepsilon(\boldsymbol{u}) = V(\boldsymbol{u}) - V_a(\boldsymbol{u})$, where V_a is the analytic solution and V is the simulated visibility. This error reflect the effects of the point source approximation made in



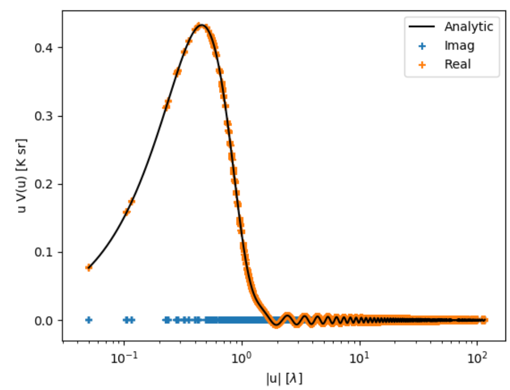
(a) MONOPOLE



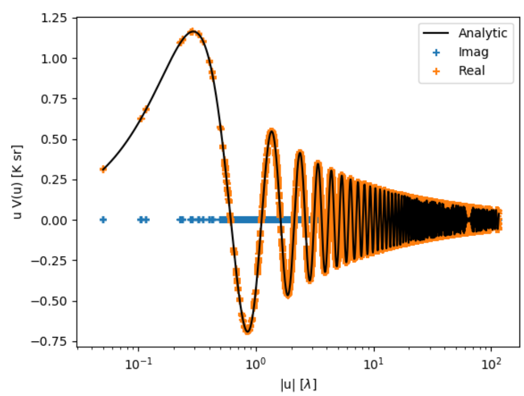
(b) PROJGAUSS, a=0.05



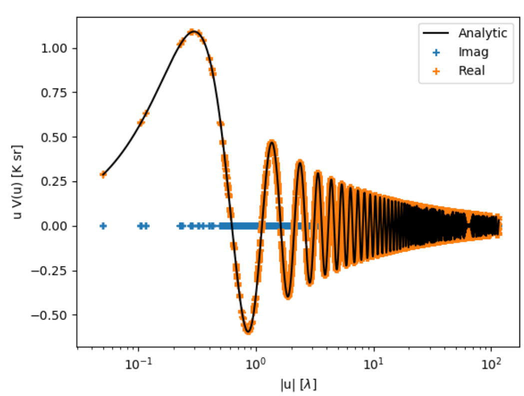
(c) QUADDOME



(d) PROJGAUSS, a=0.5

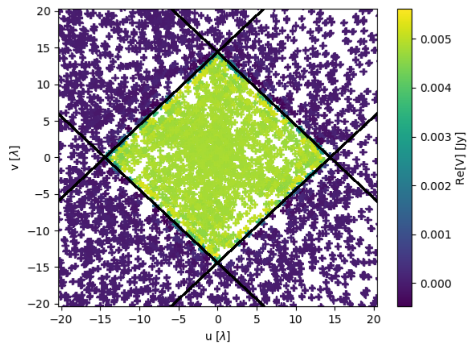


(e) COSZA

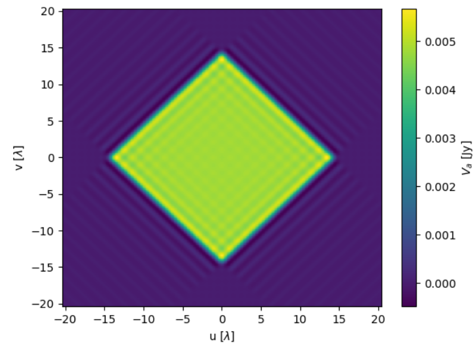


(f) PROJGAUSS, a=2.5

Figure 4.1: Simulated visibilities four axisymmetric models plotted against baseline length $|u|$. Note that the visibilities are scaled by a factor of u, to better show the structure at large u. The visibilities are all real due to the symmetry of the sky – the imaginary component of the fringe is antisymmetric, and so integrates to zero for a symmetric sky model. The analytic solution is plotted in black. All of these plots are made from simulations with Nside=1024, the highest resolution tested. Visibilities are converted to K sr from Jy.



(a)



(b)

Figure 4.2: The xysincs model, plotting visibility amplitude in UV space for the simulated baselines (a) and the analytic model (b). The black lines mark the boundaries of the rectangle with side length $a/2\pi$, rotated by $\xi = \pi/4$.

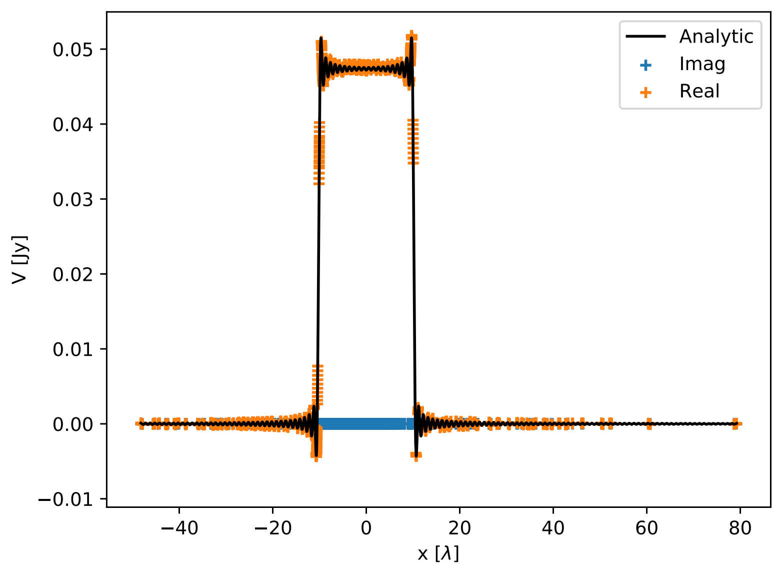


Figure 4.3: A slice through the xysincs results for $y \sim 0$, for comparison with Fig. 4.1.

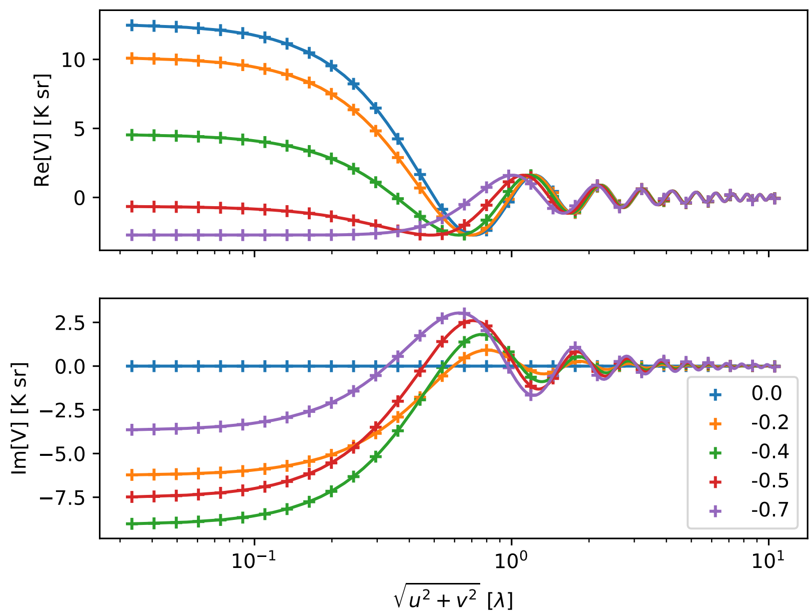


Figure 4.4: Non-coplanar array simulation results, with analytic calculations. Real components are shown on top, imaginary on the bottom. The color of the line indicates the w term. Points are simulated values, curves are the analytic calculations from eq. (4.29) to 15 terms.

`healvis` eq. (3.14), which assumed that the change in fringe, brightness, and beam are negligible across a given pixel. Most of the analytic models derived in this chapter are very smooth across the sky, and so the most of the error should be due to the fringe. One exception is the XYSINCS model, which oscillates with a period of $2\pi/a$. Additional error should emerge when the sinc period $2\pi/a$ is smaller than twice the pixel resolution $2 \times \sqrt{4\pi/12 \times \mathrm{Nside}^2}$, sinced the pixels would be undersampling this fringe.

The most intuitive error source is due to the assumption that the interferometric fringe is constant across the pixel scale. The angular scale to which a baseline of length b is sensitive is roughly λ/b for observing wavelength λ. We therefore expect error to grow when the baseline length approaches $\sim 1/\sqrt{\omega}$, where ω is the pixel size in steradians.

An interesting feature of many of the analytic models is that they are zero for some baseline lengths – The monopole model has zeros whenever u is a half integer; the cosine model has zeros when $u = x_n/2\pi$, where x_n is the n^{th} root of the 1st order Bessel function; etc. – In contrast, visibilities from point sources are nonzero for all baseline lengths. The oscillating fringe in the RIME makes the numerical integration a *poorly-conditioned* problem, highly sensitive to the sampling points, such that integrating down to a very small value requires that the sampling of this oscillating term be very nearly symmetric. For example, consider a 1D real analogy of the visibility equation:

$$V(u) = \int_{-\infty}^{\infty} T(x)\cos(2\pi ux)dx \approx \sum_{x_i \in \mathbb{X}} T(x_i)\cos(2\pi ux_i) \tag{4.48}$$

Assuming $T(x)$ is very slowly varying and $V(u) \approx 0$, then we need to choose the set of sampling points $\mathbb{X}$ such that for each $x_i \in \mathbb{X}$ then $(x_i - 1/2u)) \in \mathbb{X}$ for the cosine to cancel out. In the monopole model, $V(u) = 0$ when $2u$ is an integer. If we take the sampling scale as $\Delta x = \sqrt{\omega}$, and $u = n/2$ for integer n, then the residual error will be

$$\varepsilon = T_0 \sum_i \cos(\pi n \Delta_x \times i), \tag{4.49}$$

for monopole amplitude T_0 and $x_i = i\Delta_x$. If $\Delta x = 1/2u$, then the terms all cancel out, but if Δx deviates from $1/2u$ then the sum does not go to zero, and the remaining error scales as T_0.

As a first comparison, we'll look at the errors on the lower-resolution simulations with the monopole and cosine models. Fig. 4.5 shows the error vs baseline length u for the Nside 32, 64, and 128 simulations with the monopole and cosine models. At these low resolutions, there is a clear spike in the errors at the limit $1/\sqrt{\omega}$, indicated by the vertical black lines, which strongly suggests that these errors are due to the long baselines resolving scales smaller than a pixel.

Simulations at higher resolution avoid this error, allowing us to explore the behavior of other errors. Fig. 4.6 shows histograms of the errors for the higher resolution simulations with the MONOPOLE, COSZA, and QUADDOME models. For each model, the errors show a similar distribution with a width that scales as ω.

The histograms of the monopole model show two distinct peaks (at $\pm 10\mu$K sr for Nside 256) in the distribution, indicating a systematic offset for both positive and negative visibilities. Furthermore, the monopole case shows very different behavior in its imaginary visibilities. For the other models, the imaginary visibilities are all effectively zero (to machine precision), but for the monopole model they are substantially larger. We will discuss this more in section 4.5.

4.4.1 xysincs

For the xysincs model, we plot the error as a function of baseline length and baseline angle in Fig. 4.7. The vertical black line marks the width of the box, $l = a/2\pi$. The overall scale of the errors decreases rapidly with resolution. For the lowest-resolution simulations, large errors appear at longer baselines, as was the case for the other models.

For the high resolution cases, the errors do still seem to grow with baseline length, but are largest in general in the space outside the box. The change in error distribution with angle shows

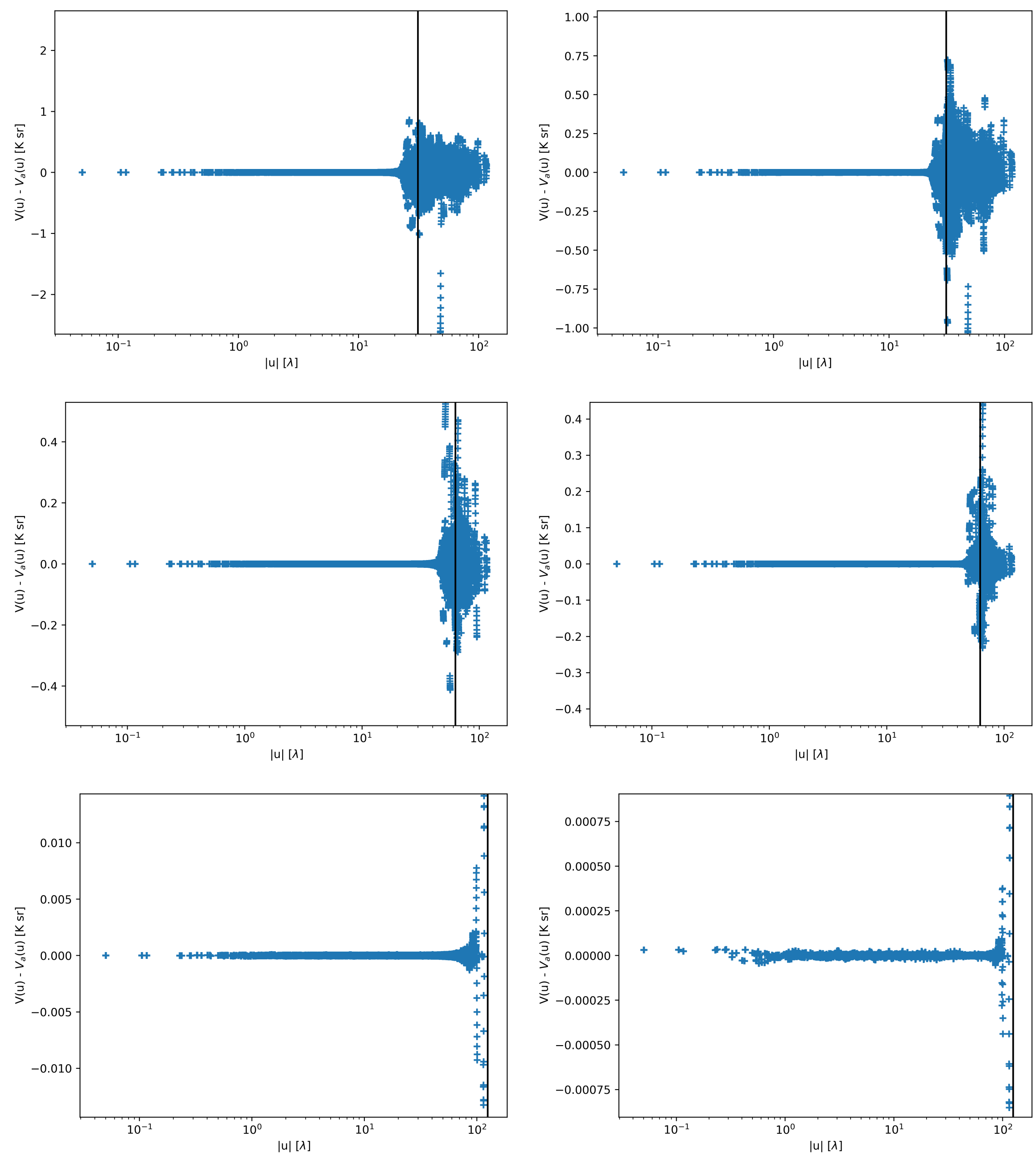
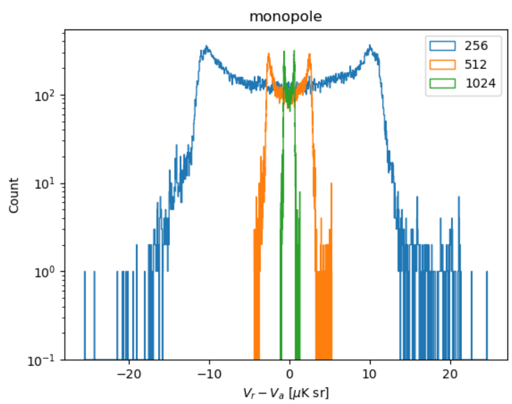
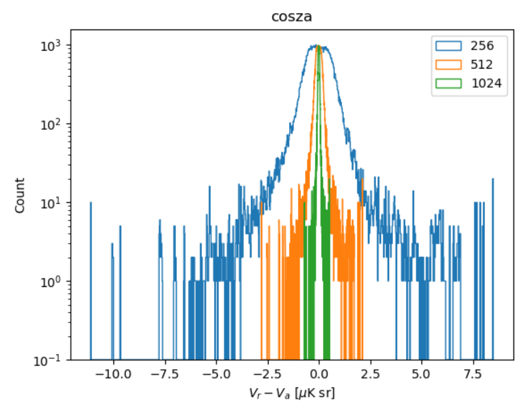


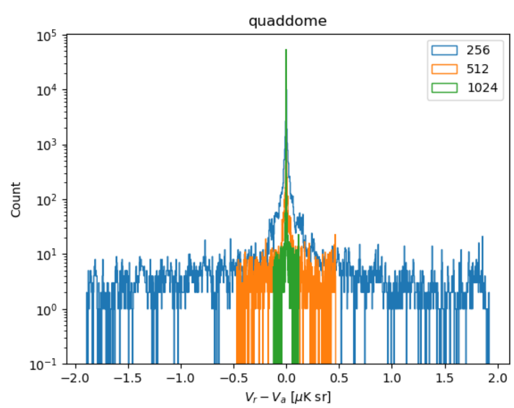
Figure 4.5: The error vs. baseline length for the monopole (left) and cosine (right) models. From top to bottom, Nside 32, 64, and 128. The vertical line marks the point $1/\sqrt{\omega}$. Note that the y axis range is changing. The cosine and monopole models are set to the scale of 2 K.



(a) MONOPOLE



(b) COSZA



(c) QUADDOME

Figure 4.6: Histograms of the errors on the real parts for the three highest-resolution simulations. The errors are expressed in terms of μ K sr, while the actual visibilities are on the scale of K sr.

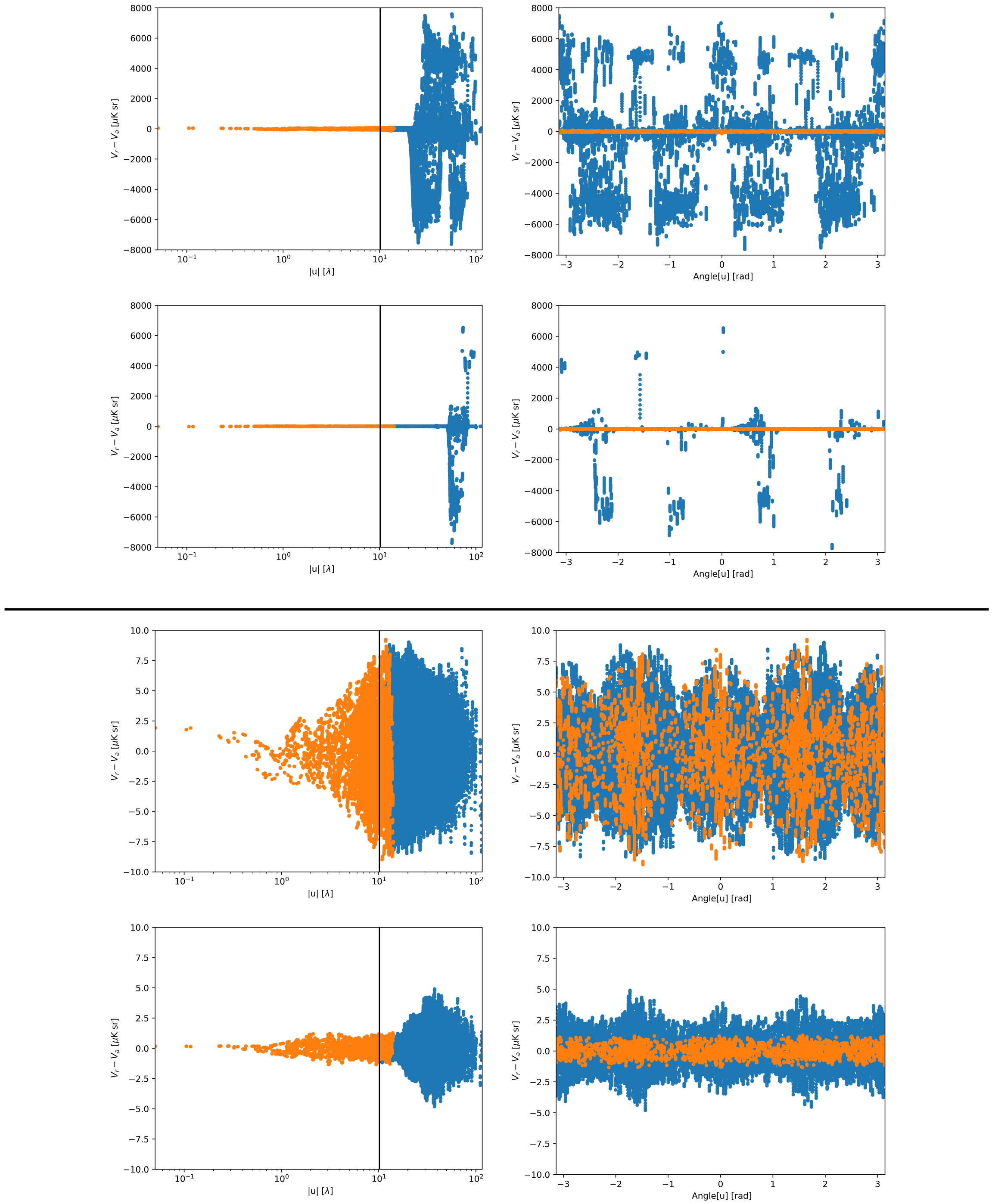


Figure 4.7: Absolute error of the xysincs model vs. baseline length (left) and angle (right). Points are marked orange if they are inside the box with side length $2l = a/\pi$ (see text), and blue if they are outside. From top to bottom, each row is for a simulation with Nside = 32, 64, 128, and 256. The vertical black lines mark the box side length l. The horizontal black line marks a transition in y axis range.

that the error is largest in the "triangular gaps" – If we think of the box as the overlap between two broad stripes forming an X, the "gaps" refers to the space outside of the stripes. Within these gaps, the interfererence between the two sine integral functions must exactly cancel to zero. Generally, these results suggest that the error is larger when the visibility is closer to zero.

4.5 Summary and Discussion

In this chapter, we have presented several analytic solutions to the full, curved-sky radio interefer-ometric measurement equation, and compared the results of `healvis` simulations of these models with the solutions. These models test several important aspects of simulator, including the accuracy of baseline orientations (XYSINCS), the response near the horizon (MONOPOLE), or varying size of a feature on the sky (PROJGAUSS, COSZA, QUADDOME).

With poor map resolution, the expected jump in error on long baselines does occur as expected, as the fringe oversamples the poorly resolved map. The remaining error scales as the pixel area, going as $1/\mathrm{Nside}^2$. The monopole model shows a somewhat even spread of errors, resulting in two peaks in the error histograms of Fig. 4.6. This is also noticeable in the distribution of the imaginary component of the simulated visibility, which should be zero for a symmetric sky model.

The monopole model is interesting for a couple of reasons. It has no structure on the sky for the simulator to under-resolve, so errors are entirely due to integrating the fringe term. In addition, of the models used it is the only one with a strong response near the horizon. We suspect that this pixelization is the cause of this remaining error – `healvis` simply selects pixels out to a radius of 90° from the zenith. At some azimuths, the farthest pixel selected may not reach the horizon exactly, since the HEALPix pixelization is not perfectly symmetric about zenith. Any asymmetry in the sampling will disrupt the cancellation of symmetric terms when the visibility is small.

Another clue to this is in the XYSINCS results, which also has a nonzero response near the horizon in some directions. Recall that the crossed-sinc functions on the sky are confined to a square region inscribed in the circular horizon. The directions that reach the horizon are those 45° off from the defined x and y axes, and are where the oscillating sine integral functions need to cancel out perfectly. These are also the directions in which we see the largest spread of errors (in the "triangular gaps" of uv space).

Disentangling the various sources of the remaining errors remains a topic of ongoing exploration. These tests show that we can avoid large errors as long as the map resolution is finer than $1/b_{\mathrm{max}}$, for the longest baseline b_{max}. Remaining errors are typically six orders of magnitude smaller than the visibility amplitude, and decrease proportionally to the pixel area. These results sufficiently validate the accuracy of `healvis` for the science cases described in chapters 5 and 6.

Chapter 5

Diffuse Foreground Contamination[1]

5.1 Introduction

Though in theory foreground power should be confined to the foreground wedge (see section 2.2.3), other sources of spectral structure, including the intrinsic spectra of foreground sources, the finite bandpass of the instrument, and the chromaticity of the antenna gains can spread foreground power beyond the wedge into the window. This suprahorizon power is usually accounted for in foreground avoidance approaches by avoiding $k_{\|}$ modes within some buffer distance of the horizon line.

Estimates of an appropriate buffer distance have been largely motivated by simple, conservative models and observations in data. Parsons et al. (2012b) argued that foreground power spillover in delay spectra is caused by the convolution of the spectral responses of the beam and sky brightness, and so is limited by the inverse bandwidth of the instrument. Through simplified foreground simulations they demonstrate that this spillover does not exceed $k_{\|} \sim 0.13$ Mpc^{-1}. Thyagarajan et al. (2013) examined the effect of spectral windowing on this spillover using numerical formulations of point source and noise power spectra, finding that increased foreground suppression came at the expense of narrowing the "effective" bandwidth, and hence increasing the extent to which foreground power spills into the EoR window. Pober et al. (2013b) observed suprahorizon power in PAPER data, finding that the spillover is not constant with $k_{\perp}$ (baseline length), and is larger on short baselines. Imaging the data after applying a high-pass filter to remove the wedge, they found that

[1] A version of this chapter has been submitted for publication in MNRAS, and is currently shared on arXiv (Lanman et al., 2019a). Content is reprinted here with permission of the co-authors.

the remaining contamination is dominated by unresolved diffuse power.

More recently, Thyagarajan et al. (2016) used wide-field visibility simulations with delay spectrum analysis to measure the levels of foreground power in the EoR window for baselines in the 19-element deployment of HERA. These simulations used a HERA beam model made through electromagnetic simulation of a HERA dish and receiver element, which incorporates spectral structure due to reflections among the various components of the antenna. They found that, except when the galaxy is directly overhead, foreground power can be suppressed below the EoR amplitude for $k_{\|} \geq 0.2$ h Mpc^{-1}.

Though these earlier works predicted that foreground power should extend as far as $\Delta k_{\|} \sim 0.2$ h Mpc^{-1}, published power spectrum limits usually choose narrower buffers when binning measurements in k. Dillon et al. (2015) used a moderate buffer of $\Delta k_{\|} \sim 0.02$ h Mpc^{-1} when binning MWA power spectra in 1D, which noticeably left in some suprahorizon emission. Ali et al. (2015) applied a buffer of 15 ns when binning power spectra for PAPER-64 data, corresponding with $\Delta k_{\|} \sim 0.008$ h Mpc^{-1} at their redshift of $z \sim 8.74$. This buffer was set by the inverse of the bandwidth used for the analysis, and did not take into account any expectations of spectral structure in the instrument or sky. Both Beardsley et al. (2016) and the more recent Barry et al. (2019) account for suprahorizon emission by increasing the slope of the horizon line by 14%. This choice makes for a wider buffer at large $k_{\perp}$, which runs contrary to the expectation that foreground spillover is larger for small $k_{\perp}$.

Avoiding foreground-contaminated modes is essential to making a robust EoR power spectrum measurement, so it is advisable to set a wide buffer beyond the horizon limit. However, setting too wide a buffer means potentially throwing out the most sensitive measurements. It is therefore essential to have precise predictions of how foreground power spreads into the EoR window for different instruments in order to use foreground avoidance techniques.

In this chapter, we run `healvis` simulations of a zenith-pointing array with realistic, frequency-dependent antenna beams, observing a variety of diffuse foreground models. These models reflect realistic angular and spectral structures of known diffuse foregrounds at low frequencies. From these simulated observations we make power spectra using delay-transform methods and measure how far beyond the horizon foreground power spills over as a result of the effects included in simulations. Such effects include the spectral structure of the primary beam or foreground model, baseline length, primary beam shape, and proximity of the brightest emission to the beam center. We focus on diffuse foreground models because EoR power spectrum estimations typically only use the shortest baselines in the array, which are both most sensitive to the expected EoR signal and to diffuse foregrounds.

The following section shows where suprahorizon power arises, within the context of delay spectrum estimation, and discusses the origin of the foreground pitchfork. Section 5.3 discusses the numerical simulator and the instrument and sky model configurations tested. Sections 5.4.1 to 5.4.4 presents measurements of the effects of baseline foreshortening near the horizon, spectral windowing, source and beam spectral structure, and source position and baseline length. Lastly, we show the same results with a more realistic primary beam model in section 5.4.5.

5.2 Diffuse Foreground Power

5.2.1 Foreground Power Spillover

The scalar RIME eq. (2.13) may be expressed as

$$V(\boldsymbol{b}, \nu) = \int I_\nu(\hat{s}) A_\nu(\hat{s}) e^{-2\pi i \tau_g} d\Omega, \tag{5.1}$$

where

$$\tau_g = \frac{(\hat{s} \cdot \boldsymbol{b})\nu}{c}$$

is the *geometric delay* of the point at $\hat{s}$ for a baseline vector $\boldsymbol{b}$. This delay corresponds with the time it takes a plane wave front from the direction $\hat{s}$, having reached one antenna in the baseline, to reach the other.

In the case of a single point source at position $\hat{s}_p$, with corresponding geometric delay τ_p, and spectrum $I_p(\nu)$, eq. (5.1) reduces to

$$\widetilde{V}(\boldsymbol{b}, \tau) = (\widetilde{\phi} * \widetilde{A} * \widetilde{I}_p)(\tau - \tau_p) \tag{5.2}$$

where $*$ denotes convolution in delay and tildes indicate a delay-transformed function. If the beam, source, and window function all have flat spectra, and the bandpass is infinite, then this becomes a delta function centered at the geometric delay of the point source. The maximum geometric delay is limited to the baseline length $b = |\boldsymbol{b}|$ over the speed of light.

$$|\tau_g| \leq \frac{b}{c} \tag{5.3}$$

Under these conditions, the power from a point source on the horizon has the form of a delta function centered at this maximum delay. Hence, this maximum delay is known as the "horizon line", and separates flat-spectrum power from spectrally-structured emission. In cosmological terms, this horizon limit is given by

$$k_\parallel \leq \frac{H(z)\chi(z)}{c(1+z)} k_\perp \tag{5.4}$$

where $k_\perp = |\boldsymbol{k}_\perp|$.

This limit applies in the ideal point source case as discussed — infinite bandwidth and a flat spectrum. This power will spread out, possibly beyond the horizon limit, once these assumptions are relaxed. For example, assume that the primary beam A and point source flux I_p have no spectral structure, but the window function has a limited domain. In this case, $\widetilde{A}(\tau) = A_p\delta_D(\tau)$ and $\widetilde{I}_p(\tau) = I_p\delta_D(\tau)$, where δ_D is the Dirac delta function, and eq. (5.2) reduces to

$$\widetilde{V}(\boldsymbol{b}, \tau) = I_p A_p \widetilde{\phi}(\tau - \tau_p) \tag{5.5}$$

If ϕ is a rectangular window function, then $\widetilde{\phi}$ is a sinc function centered on $\tau = \tau_p$, which has strong ringing sidelobes that extend beyond the horizon. The sinc function is known to lack the dynamic range needed to access the expected EoR amplitude, so most analyses use other common window functions such as Hamming, Blackman-Harris, or Blackman-Nuttall. These windows suppress their sidelobe amplitude at the expense of widening the main lobe. In any case, the finiteness of the bandpass necessarily causes power to spread.

5.2.2 Foreground Pitchfork

Resolved diffuse emission behaves in a fundamentally different way from unresolved point-like sources in interferometric measurements. The difference can be illustrated in angular Fourier space by applying the convolution theorem to eq. (2.17), we can see that the visibility measured by a baseline is given by the convolution of the primary beam and sky brightness in Fourier space:

$$V(u, v) \simeq \int A(l, m)T(l, m)e^{-2\pi i(ul+vm)}dldm \tag{5.6}$$

$$= (\widetilde{A} * \widetilde{T})(u, v) \tag{5.7}$$

Here, (l, m) are direction cosines, tildes denote the 2D Fourier transform and (u, v) are orthogonal components of the baseline vector measured in wavelengths. These are also the Fourier-duals to l and m. A point source has a surface brightness I in the form of a delta function, which is constant in Fourier space. A diffuse source, however, will be more compact in Fourier space, centered on $u, v \approx 0$. Short baselines will therefore measure more diffuse power than long baselines, while the measured power from point sources will be the same for both.

The increased brightness of diffuse power on short baselines causes a stronger response to diffuse emission near the horizon. This is because the baseline length enters the visibility equation as $\hat{s} \cdot \boldsymbol{u} \approx 0$ when $\hat{s}$ is parallel to $\boldsymbol{u}$. In delay space this causes brightening along the horizon lines, making a pitchfork pattern (a third "tine" along the $k_\parallel = 0$ line, comes up when there is a monopole term in the model). This was observed in simulation data (Thyagarajan et al., 2015a), once horizon to horizon sensitivity was achieved, and was later found in MWA data (Thyagarajan et al., 2015b). More recently, a pitchfork has been observed in HERA data (Kern et al., 2019a).

The pitchfork effect can potentially change our intuition for which factors dominate foreground spillover. For a zenith-tracking array, measured power is highest when bright sources are near zenith, because the primary antenna beam is strongest there. In delay space, however, this power is clustered around $\tau = 0$, so how much of that power spreads beyond the horizon depends on the spectral structure of the emission and the instrument. Power in the pitchfork is very close to the horizon, but is attenuated by the weaker primary beam there. This raises a question – Does a bright patch of the sky contaminate the window more when it's at zenith or near the horizon?

To answer this, we will look at how much power spreads beyond the horizon as a function of the hour angle of the galactic center, for a variety of primary beam models, in simulated data. Some of these beam models have effectively zero response near the horizon, which prevents any pitchfork from appearing. Comparing these, we can demonstrate that the pitchfork is the dominant source of power near the horizon for most observing time, while spectral structure in the beam or source affects how that power is spread. By taking both effects into consideration, it is possible to set bounds on the extent of foreground contamination.

5.3 Simulations

Simulations were carried out using `healvis`. Table 5.1 summarizes the various simulation configurations used, including antenna layouts, primary beam models, sky brightness distribution models, observing time ranges and time step sizes. The same bandpass and channelization, covering 100 to 150 MHz at 97 kHz channel resolution, is used for all simulations. Generally, the "Off-Zenith" time configuration is used whenever LST is not considered important to the result. This time set consists of only five 11 s integrations, from a time when the Galactic center is in the sky but not near the beam center (off by 4 hours). This is considered as a "typical" sky brightness, an assumption supported by the tests with time dependence (using the "transit" time set). More details on the sky and beam models are given in the subsequent subsections.

5.3.1 Primary Beam Models

We use three analytically-defined beam models and one calculated through electromagnetic simulations of a HERA dish. The analytic models are described in eqs. (3.8) and (3.9). The first analytic beam is an Airy disk beam defined with a diameter of 14 m. The other two are Gaussian beams defined with the same FWHM as the Airy disk at 100 MHz. Of these, one is an "achromatic Gaussian", with a fixed width across the band, and the other is designated the "chromatic Gaussian" and has a spectral index $\alpha = -1$, to resemble the spectral evolution of the Airy beam. When other spectral indices are used, these are indicated. We note that the Gaussian beams strongly resemble the Airy beam, but lack any response near the horizon.

Beam Models	
Airy Disk	Radiation pattern of a circular aperture with a 14 m diameter.
Achromatic Gaussian	$\exp(-\theta^2/(2\sigma_0^2))$, with $\sigma_0 = 7.37°$. This is set to match the FWHM of the main lobe of the Airy beam.
Chromatic Gaussian	The same as achromatic, but with $\sigma = \sigma_0(\nu/\nu_0)^{-1}$
HERA CST	Power beams made from far-field electric fields measured in CST simulations of a HERA dish with crossed-dipole feeds.
Sky Models	
GSM	The Global Sky Model in HEALPix, generated from three principal components per pixel (Zheng et al., 2017).
sym	A model with the same angular power spectrum as the GSM, defined to be azimuthally symmetric about the galactic center. Each pixel is given a power law frequency spectrum.
Layout	
37 hex	A thirty-seven element hexagon, with 14.6 m spacing.
Imaging	80 antennas distributed randomly, with a Gaussian radial density profile with FWHM of 230 m. The longest baseline is 608 m.
Line	A line of 300 antennas, evenly spaced, east to west. Baselines are formed by pairing each of these antennas with a 301st antenna, such that the whole array provides East/West baselines with lengths ranging from 10 to 150 m.
Time	
Off-Zenith	Five 11s timesteps when the Galactic center is 4 hours off of zenith.
Transit	24 hours at 5 minute timestep, such that the galactic center transits 6 hours from the start.
Frequency	
50 MHz	100 to 150 MHz with channels of width 97.65 kHz. The channel width is chosen to match that of the first generations of HERA.

Table 5.1: A summary of parameters used for foreground simulations in this chapter.

The last beam model is defined numerically by a set of CST Suite[2] simulations of a realistic HERA dish and receiver done by Fagnoni & Acedo (2016). The simulated antenna included a parabolic reflecting dish with a 14 m diameter, two 1.3 m copper dipole receivers with a pair of aluminum discs acting as sleeves, and a cylindrical mesh "skirt" 36 cm high around the feed to protect from reflections of other antennas in the array (see Fagnoni & Acedo (2016) for more details). Reflections among the various elements of the antenna introduce spectral structure into the beam, which is largely captured by the CST simulations.

These simulated beam models are saved as a set of electric field components at altitude and azimuth positions on a spherical grid with steps of 1° in both latitude and longitude, at frequencies between 50 and 250 MHz in steps of 1 MHz. For use in simulations of unpolarized (i.e., purely Stokes I) sky models, we convert this electric field beam into a *power beam* using the definition in section 2.1.3:

$$A^{pp}(\hat{s},\nu) = |A^p_\theta(\hat{s},\nu)|^2 + |A^p_\varphi(\hat{s},\nu)|^2 \tag{5.8}$$

where the beam for feed p is written as components in the zenith angle and azimuthal directions for each position.

The CST simulated beams have two crossed-dipole feeds, one oriented East/West (designated X) and the other oriented North/South (Y). For comparison, we may think of the analytic beams defined in this section as representing the response of an idealized single feed to an unpolarized sky model.

CST Beam Interpolation

Our simulations are done at much higher spectral and angular resolution than the CST simulations, so we need to interpolate the CST simulated beam models values to the required angles and frequencies. The beam data is handled using the new `pyuvdata` class UVBeam (Hazelton et al., 2017).[3] This class supports frequency interpolation at each pixel using the methods accessible to the `scipy.interpolate.interp1d` method, which includes linear, polynomial splines, and nearest-neighbor interpolation among other methods. UVBeam handles angular interpolation using bivariate spline interpolation.

Interpolating to a higher frequency resolution means adding information in the beam model on higher delay scales. Since the main results of this paper are to measure the behavior of foreground power in delay space, we need to be sure that the frequency interpolation does not have a substantial effect on simulated data. In this section, we interpolate the UVBeam to a finer frequency resolution

[2]https://www.cst.com

[3]https://github.com/RadioAstronomySoftwareGroup/pyuvdata

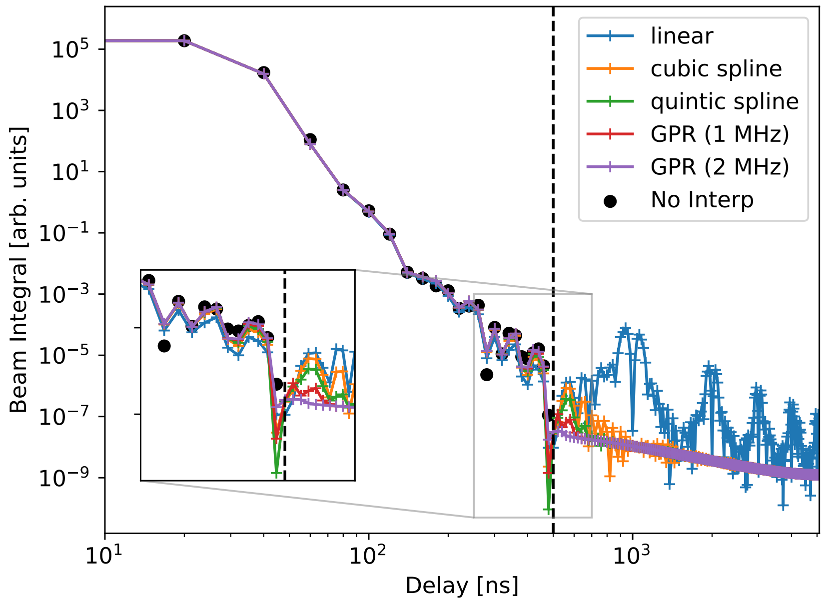


Figure 5.1: Comparison of the delay-transformed beam integral of the (black dots) non-interpolated beam to the interpolated beam with five methods of interpolation. The vertical dashed line marks 500 ns, the Nyquist sampling rate of the 1 MHz resolution beam. The inset zooms in on the region around the Nyquist limit. There is some disagreement among the curves at and just before this limit, so it seems the effects of interpolation are not strictly limited to interpolated delay modes.

using a variety of interpolation methods and compare these interpolated beams to the original beam data under a delay transform.

The power beam is obtained from the E-field using eq. (5.8), then the values at each frequency are normalized to their peaks. The peak-normalized power beam is then, at each pixel, interpolated to the 50 MHz bandpass (see table 5.1) used for simulations. There are two general types of interpolation used: polynomial spline interpolation, which we do to linear, third, and fifth order, and Gaussian process regression (GPR) using the `scikit-learn` package. The GPR method uses a radial basis function (RBF) kernel, also known as a squared exponential, as the covariance function of the Gaussian process. For a fixed point on the sky, we can describe the covariance of the beam at two frequencies ν_1 and ν_2 as

$$k(\nu_1, \nu_2|\ell) = \exp\left[-\tfrac{1}{2}(\nu_1 - \nu_2)\ell^{-2}(\nu_1 - \nu_2)\right], \tag{5.9}$$

where ℓ is a hyperparameter that sets the characteristic scale of the covariance function. Given a covariance function, the GPR maximum a posteriori estimate of the underlying signal constrained by the data can be calculated at any new frequency ν' (Eq. 2.23 of Rasmussen & Williams, 2006).

The RBF kernel has the property of producing smoothly varying estimates with frequency structure set by the length-scale hyperparameter. By keeping this length-scale fixed (rather than solving for its optimal value given the data), the RBF kernel can act as a low-pass filter, filtering out structure at scales above the length-scale (e.g. Kern et al., 2019b). Given the inherent Nyquist sampling rate of the CST output, we set $\ell = 1$ MHz for our GPR interpolation method. Altogether, this leaves us with five methods of interpolation to test.

To test the effect of interpolation on the beam, we construct a single metric to compare in delay space. At each frequency, the interpolated beam values are summed over pixels. This list of 512 sums is then is multiplied by a Blackman-Harris window function and then Fourier transformed from frequency to delay using an inverse DFT. The result of these steps is a quantity analogous to the power spectrum measured at $k_\perp = 0$. Fig. 5.1 shows this quantity for each of the interpolation types. The GPR results in the plot are shown with the chosen smoothing scales in parentheses. Overall, all of the interpolation schemes preserve the power on the known scales, marked with black dots, up to the Nyquist limit (shown by the vertical dashed line). Beyond that, the spline interpolations introduce noticeable structure. Further testing has shown that this is likely due to some sharp features in the spectra of pixels far off from the beam center, which can cause aliasing under Fourier transform. The GPR-interpolation seems to smooth out these effects well, and so the curves for the GPR methods are relatively smooth beyond the Nyquist limit.

To avoid any potential issues that may depend on the beam interpolation scheme, we avoid using delay modes beyond the Nyquist limit of 500 ns for any results. Simulations are still carried out to the full frequency resolution of HERA, with beams interpolated using the GPR method with a 1 MHz smoothing scale. This is a conservative choice, because the smoothing scale avoids accidentally smoothing over sub-Nyquist scales in the beam while suppressing the interpolation artifacts. It is of course possible that the HERA beam has structure on scales smaller than 1 MHz, but that cannot be determined from the existing CST simulation data.

5.3.2 Sky Models

With the `healvis` simulator, there are trade-offs between resolution, run-time, and memory requirements, which are exacerbated by simulating wide fields of view and bandpasses. For this work we use maps at an `Nside` of 256, which corresponds with a resolution of 13.7 arcmin. The highest resolution probed by any of our simulation configurations is 14.13 arcmin, corresponding with the longest baseline and highest frequency, and shorter baselines are only sensitive to larger scales.

We use two diffuse foreground models. The first is the 2016 Global Sky Model (GSM), which is compiled from 29 different sky maps covering a total bandpass of 10 MHz to 5 TH (Zheng

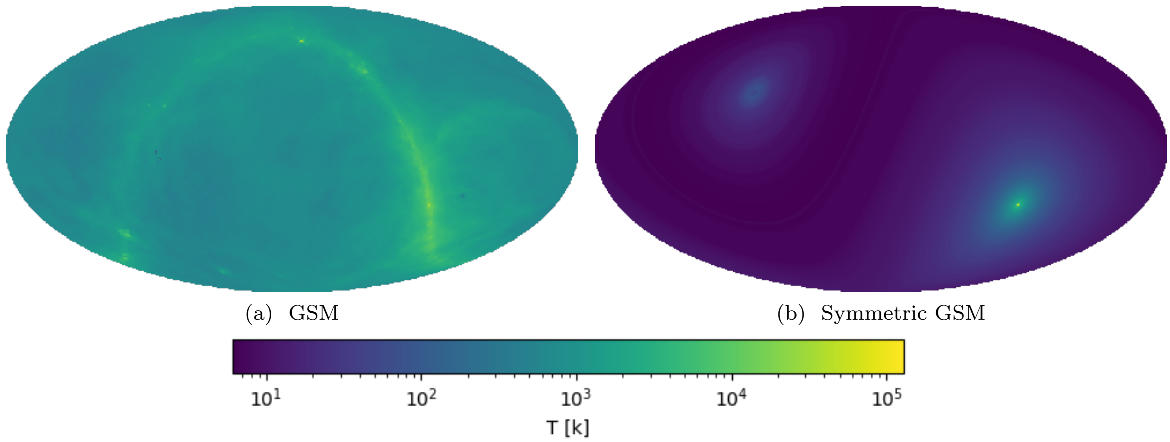


Figure 5.2: The GSM (a) and symmetric GSM (b) models in a Mollweide projection of equatorial coordinates.

et al., 2017). The GSM is implemented in the Python package `PyGSM`[4], which generates HEALPix GSM maps at an `Nside` of 1024 using a principal component decomposition to model the spectral structure at each pixel. We degrade this to an `Nside` of 256 by averaging together high-resolution pixels nested within low-resolution pixels, which can be done without interpolating or splitting pixel values due to the hierarchical structure of HEALPix maps. In our chosen bandpass, diffuse foreground power is dominated by Galactic synchrotron emission, which generally follows a power law with a spectral index of -2.5. Most pixels in the GSM model used here are observed to have power law spectra with index ~ -2.5.

The second model is the "symmetric GSM" or "sym", which is made to have the same angular power spectrum as the GSM but given azimuthal symmetry about the Galactic center position. The sym model thus provides a well-defined locus of bright emission on the sky, which will make for a more meaningful test of how spillover relates to the "position of bright emission". The sym model is constructed in the following way:

1. Calculate the angular power spectrum $\mathcal{C}_l$ of the GSM at 100 MHz from the Nside 256 GSM.

2. Generate a set of spherical harmonic expansion coefficients a_{lm}, with $a_{l0} = (2l+1)\sqrt{\mathcal{C}_l}$ and $a_{lm} = 0$ for $m \neq 0$.

3. Apply a rotation to the a_{lm} using Wigner D-matrices so that the $(\theta, \phi) = (0, 0)$ point is at the Galactic center.

[4]https://github.com/telegraphic/PyGSM

4. Convert the a_{lm} to an Nside 256 map.

5. Scale each pixel in this map by a power law in frequency: $(\nu/\nu_0)^\alpha$ for ν_0 =100 MHz and spectral index α.

We set $\alpha = 0$ for the symmetric model unless explicitly testing the effect of model spectral structure. This allows us to test other effects in isolation. The angular power spectrum calculation and spherical harmonic transforms are done using `healpy`'s `anafast` and `alm2map` functions, respectively. The rotation is done using `rotate_map_alms`. The GSM and symmetric GSM models at 100 MHz are shown in ICRS equatorial coordinates in Fig. 5.2, with a log scale on each plot.

5.3.3 Power Spectrum Estimation and Measurements

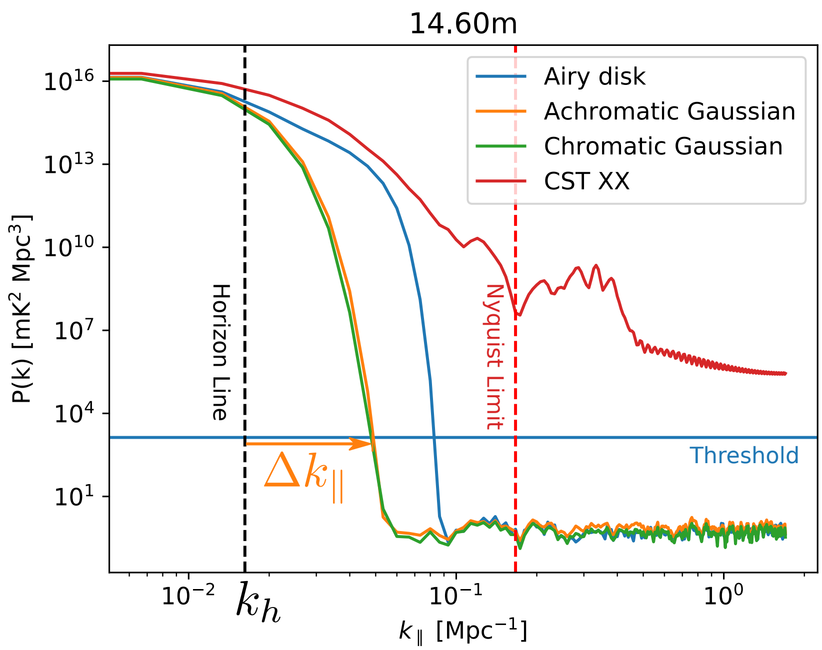


Figure 5.3: The delay spectra of the GSM for a single 14m baseline with a variety of beam types. The fiducial EoR amplitude is shown by the horizontal line, the black vertical dashed line shows the horizon limit, and the foreground power spillover $\Delta k_{\|}$ for the Gaussian beams is indicated by the horizontal green arrow. A vertical dashed red line marks the Nyquist limit for the 1 MHz resolution of the CST beam. For the results with the CST beam, we will use a higher threshold to avoid measuring spillovers past the Nyquist limit. Such measurements would be highly-dependent on the frequency interpolation scheme used.

The form of the delay spectrum estimator was discussed in section 2.2.2, in particular with eq. (2.29). In this section, we will briefly discuss the process of estimating power spectra from simulated visibilities, and define the power spillover and fiducial EoR amplitude that will be used to quantify the spread of power beyond the horizon.

The simulator provides visibilities $V[b,t,\nu]$ for each baseline (b), time step (t), and frequency (ν), in units of Jy, as well as the beam integral Ω_{pp} at the zeroth frequency in units of steradians. The visibilities are converted from Jy to mK sr by weighting each channel with a frequency-dependent conversion factor:

$$V_{\mathrm{mK\ sr}}[b,t,\nu] = 10^{-20}\left(\frac{c^2}{2k_B\nu^2}\right)V_{\mathrm{Jy}}[b,t,\nu] \tag{5.10}$$

For each baseline and time, the delay transform is done by weighting each channel with the spectral window function and then doing an inverse discrete Fourier transform (IDFT), and taking the absolute square:

$$\widetilde{V}[b,t,\tau_q] = \frac{B}{N_f}\sum_p \phi(\nu_p)V_{\mathrm{mK}}e^{2\pi i\nu_p\tau_q} \tag{5.11}$$

$$P[b,t,k_\parallel] = \frac{X^2Y}{B_{pp}\Omega_{pp}}|\widetilde{V}[b,t,\pm\tau_q]|^2 \tag{5.12}$$

In the above, ν_p is the p^{th} frequency, and the delay modes are $\tau_q = q/(2B)$ for bandwidth B. Note that there are two delay modes for every $k_\parallel$, since the inverse DFT goes to both positive and negative delays. Since the power spectrum is a real quantity, the positive and negative delay modes are equivalent, so we average them together.

This leaves us with an estimate of the power spectrum for each baseline, time, and $k_\parallel$. From these, we average together power spectra from baselines with similar lengths (to within a half meter). Additional tests show no significant variation in the power spectrum with baseline orientation. For simulations using the Off-Zenith time configuration, we also average power spectra incoherently over the simulated time span (55 seconds). The spectrum is not observed to change significantly over such a short period.

For comparison with the foregrounds, we consider a fiducial EoR power spectrum amplitude of $P_{21} = 29^2$ mK^2 Mpc^3, which is approximately the amplitude at $k \sim 1$ Mpc^{-1} at the band-center redshift $z \sim 10.4$. This amplitude is derived from a `21cmFAST` EoR simulation with default settings (Mesinger et al., 2011). The true EoR power spectrum is stronger at smaller k at this redshift, so this is a conservatively low choice.

Rather than look at the power at specific $k_\parallel$ modes, we are interested in the range of $k_\parallel$ that are contaminated by foreground power. Using the fiducial EoR amplitude, we can define the power spectrum *spillover* $\Delta k_\parallel$ as the distance in $k_\parallel$ that the foreground power extends beyond the horizon. More precisely, the spillover is defined as the largest $k_\parallel$ where the power spectrum crosses below a set threshold of P_{21}, such that $P(k_\parallel) < P_{21}$ for all $k_\parallel > \Delta k$, less the horizon k_h:

$$\Delta k_\parallel(P_{21}) \equiv \max\left(k_\parallel | P(k_\parallel) = P_{21}\right) - k_h \tag{5.13}$$

Defining the spillover relative to the horizon makes it easier to compare the results among different length baselines, which have different horizon limits.

Fig. 5.3 shows power spectra for a single baseline with several different beam models, with the threshold/horizon/spillover labelled. These spectra are from a simulation with the Off-Zenith time configuration, and with the 14.6 m baselines binned from the hex layout. The delay transform was done with a Dolph-Chebyshev spectral window applied (see 5.4.2).

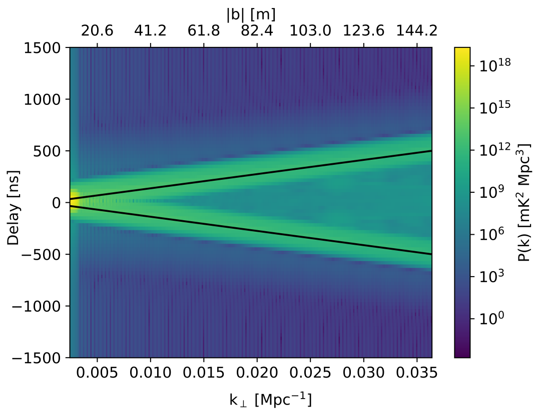
5.4 Results

5.4.1 Pitchfork Confirmation

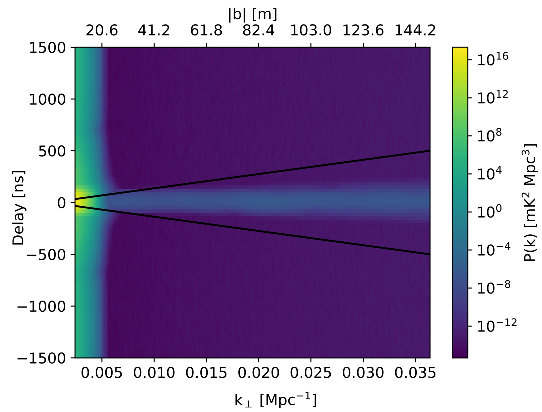
The pitchfork foreground signature in Fourier space was first found by simulations in `PRISim` (Thyagarajan et al., 2015a), and later confirmed in data form the Murchison Widefield Array (MWA) experiment (Thyagarajan et al., 2015b). `PRISim` was the first interferometer simulator that could handle diffuse sky models down to the horizon, and was thus sensitive to the effects of baseline foreshortening there. `healvis` is also sensitive to sources near the horizon, and so we expect to see pitchfork power in our simulations. The first set of simulations used a monopole sky signal to confirm the existence of the pitchfork, since a monopole should have minimal power away from the horizon for most baseline lengths.

The simulated array consists of 300 East-West oriented baselines with lengths linearly spaced from 10m to 150m. This provides a range of $k_\perp$ modes to study. We do see the pitchfork for randomly distributed antennas of all orientations as well. We ran two simulations – first using the Airy beam, and then with the achromatic Gaussian beam. Fig. 5.4 shows the resulting delay spectra. The pitchfork appears for the Airy disk beam, since that has a nonzero response near the horizon. The Gaussian beam is effectively zero at the horizon, and so no pitchfork can be seen. We also observe a strong signal on the shortest baselines. This is evidence that short baselines of an interferometer may be made sensitive to a global signal, as discussed in Presley et al. (2015).

The observed pitchfork appears to extend beyond the horizon by a consistent amount for all baseline lengths, suggesting that the pitchfork represents a constant minimum level of foreground spillover for antennas with a nonzero horizon response. Most realistic antennas used in EoR experiments have some response near the horizon, since the width of the beam varies inversely with the effective diameter of the antenna, and most of these experiments use relatively small antenna elements. As we will see, the pitchfork is an important factor of foreground spillover for most of the cases studied here.



(a)



(b)

Figure 5.4: Delay spectra of a monopole signal for a variety of baseline lengths (top axis), and their corresponding $k_\perp$ mode (bottom axis). (a) shows the results with an Airy beam, which has strong sidelobes and nonzero response down to the horizon. (b) is the same for a Gaussian beam, with width set to the width of the main lobe of the Airy beam. Black lines show the delay corresponding with sources on the horizon (b/c). Note the difference in color scales.

5.4.2 Window Function Comparison

As discussed in section 5.2.1, the shape and width of the bandpass has an effect on the spread of power beyond the horizon, as the intrinsic power spectrum of the source is convolved in delay with the spectral window function and the beam. By choosing an appropriate window function to weight the frequency channels, one can reduce the amplitude of power beyond the horizon at the expense of widening the main lobe of the foreground signal. In this section, we explore the effects of several window functions on power spectra derived from a single symmetric-GSM simulation. The simulations used the Airy beam, Off-Zenith time configuration, and 37 hex layout. The window function is applied to the delay transform for each baseline separately, and the resulting power spectra were averaged over time and binned in $k_\perp$ such that spectra from baselines of the same length to within 10 cm are binned together. The results presented here are for 14.6 m baselines.

The most basic window function is the rectangular window, sometimes called a Dirichlet window, which applies even weighting to all samples in the bandpass. As mentioned, in Fourier space this has the form of a sinc function, which has a limited dynamic range in power. We include it here for comparison, though it is not used in most 21 cm power spectrum analysis.

Note that applying a window function, other than the rectangular, involves down-weighting channels toward the end of the bandpass. This reduces the effective bandwidth B_{pp}, and hence increases the width of the main lobe of the Fourier transform, which increases the basic spillover beyond the horizon. There is thus an implicit trade-off between the level of suppression in the foreground sidelobes and the loss of low-k modes to foreground spillover.

We test two window functions, in addition to the unweighted rectangular spectral window. The first is the Blackman-Harris window, which is defined by a sum of three cosine functions optimized to reduce the level of sidelobes in Fourier space (see e.g. Harris (1978)). The Blackman-Harris window reduce has a dynamic range of about 120 dB in power, which is generally considered sufficient for 21 cm EoR experiments. Thyagarajan et al. (2016) applied a squared Blackman-Harris window to reach an additional 10 orders of magnitude in sidelobe suppression.

The second window is the Dolph-Chebyshev, which is designed to minimize the sidelobe amplitude for a choice of main lobe width and thus has a tunable dynamic range (Smith, 2011). Functionally, it is constructed from Chebyshev polynomials in Fourier space and then inverse-transformed to direct space. Unlike the Blackman-Harris and rectangular windows, the sidelobes do not drop off with k, and so usually appear as flat lines in our power spectra. We use Dolph-Chebyshev windows set to 120 dB, 150 dB, and 180 dB ranges, which we refer to as DC120, DC150, and DC180 respectively.

Fig. 5.5 shows the comparison of power spectra with the different window functions. The rectangular window function, which has the form of a sinc function in delay space, does not have sufficient dynamic range to keep spilled foreground power below the expected EoR amplitude. The three Dolph-Chebyshev windows demonstrate the expected tradeoff between main lobe width and sidelobe suppression, as can be seen in the slight spread among the curves around $P(k) \sim 10^9$ mK^2 Mpc^3). The Blackman-Harris window displays similar behavior to the DC120 window, and continues to drop off at higher $k_{\|}$ while the Dolph-Chebyshev window function stays flat.

For the rest of this paper, we use the DC180 window for all delay spectrum estimation. The Dolph-Chebyshev is the only window tested that has sufficient dynamic range to reach the EoR amplitude without any assumed foreground subtraction,[5] and does so with comparable foreground spillover to the Blackman-Harris used in other work (see, e.g., Thyagarajan et al. (2013); Parsons et al. (2012a); Vedantham et al. (2012)).

[5]The Blackman-Harris window does drop below this flat EoR threshold at the high $k_{\|}$ for some limited cases, but not enough to be useful for our results

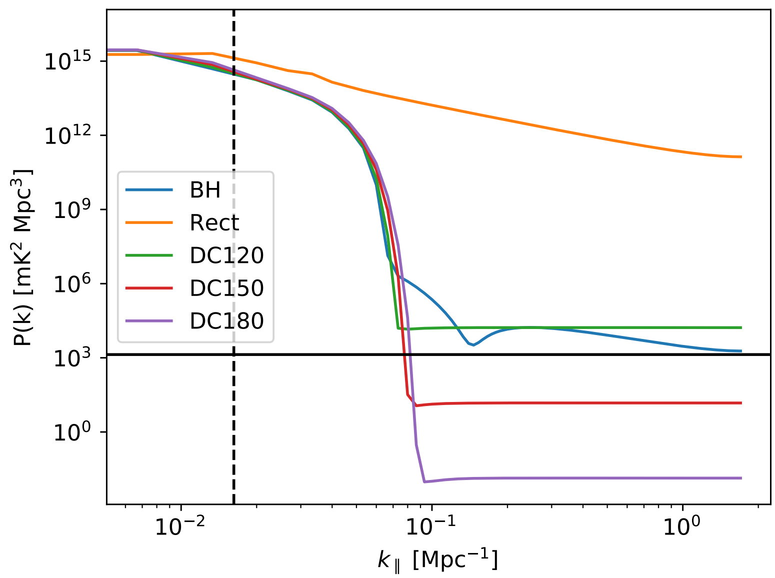


Figure 5.5: Power spectra with rectangular (rect), Blackman-Harris (bh) window functions, as well as Dolph-Chebyshev window functions with dynamic range set to 120 dB, 150 dB, and 180 dB (DC120, DC150, and DC180, respectively). The horizontal black line marks a fiducial EoR amplitude (see section 5.4.2). The vertical dashed line marks the horizon for a 14.6 m baseline. The Blackman-Harris function alone does not provide sufficient dynamic range to push the leaked foreground power below the EoR level. By contrast, applying the right Dolph-Chebyshev window can suppress the leaked foreground power below the EoR threshold.

5.4.3 Spectral Index Comparison

The next tests explored the effects of power law spectral structure in the primary beam to structure in the sky model itself. As mentioned in section 5.3.2, diffuse power typically has a power law spectrum. The angular width of a beam typically evolves linearly with frequency, as for the Airy disk. Here we look to slightly steeper evolution of the beam width and sky brightness. Note that the structure is applied to the model and to the beam in slightly different ways – the beam is normalized to its peak value at all frequencies, but its width changes with frequency. The sky model's total brightness is evolving with frequency instead.

We ran simulations with the following configurations:

- Chromatic Gaussian beam with power law index $\alpha = \{-1, -2, -3\}$, Flat-spectrum sym model. (Solid lines in Fig. 5.6).

- Achromatic Gaussian beam, sym model with spectral index $\alpha = \{-1, -2, -3\}$. (Dashed lines in Fig. 5.6)

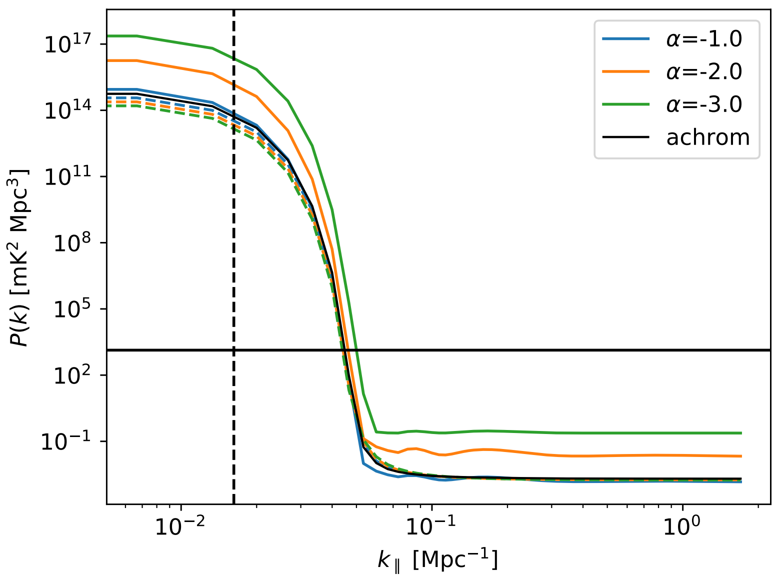


Figure 5.6: Power spectra of a single 14.6 m baseline with a Gaussian beam. The color indicates the spectral index of either the beam or the sym sky model. For solid lines, the spectral index is applied to the beam and the sky model is flat in frequency. For dashed lines, the spectral index is applied to the sky model and the beam is flat-spectrum. The black line shows the result of an achromatic beam and sky model.

These simulations were run with the 37 hex antenna layout with the Off-Zenith times (see table 5.1). Power spectra were averaged in time and binned in baseline length. Fig. 5.6 shows the power spectrum of the 14.6 m baselines for these different sky and beam models. Solid lines are for simulations with a chromatic beam, while dashed lines are for simulations with the spectral sky model. The amplitude at low-k shifts with spectral index for a couple of reasons. When the spectral index is applied to the sky model, the model has less power at higher frequencies as compared with the flat-spectrum model, so the total power drops. When the spectral index is applied to the beam width, the beam is narrower at higher frequencies. A narrower beam in image space corresponds with a wider footprint in Fourier space, and hence integrates more power. This causes the low-k modes of the solid curves to shift higher with decreasing (increasingly negative) spectral index.

The effects of the beam chromaticity are much more noticeable, shifting the overall amplitude higher, than the skymodel chromaticity. Since the diffuse foreground power is known to largely follow a clean power law within this range, we take this to mean that the spectrum of the sky has a negligible effect on the foreground spillover. Increasing beam chromaticity does also shift the spillover out to slightly higher $k_\parallel$, although not to the same extent that can be seen in a more complicated beam response as with the CST beam as seen in Fig. 5.3. We conclude that typical foreground spectral structure has negligible effect on spillover when compared with beam effects.

5.4.4 Foreground Spillover

Here we quantify the spillover and test how it changes with time and with baseline length. The foreground spillover is estimated by fitting a polynomial spline to the power spectrum measurements and finding the points where the spline crosses the threshold. The spline fit is done in log-log space where it is most stable. The largest $k_\parallel$ of intersection, less the $k_\parallel$ of the horizon, is taken as the measured foreground spillover $\Delta k_\parallel$, as defined in eq. (5.13).

Time

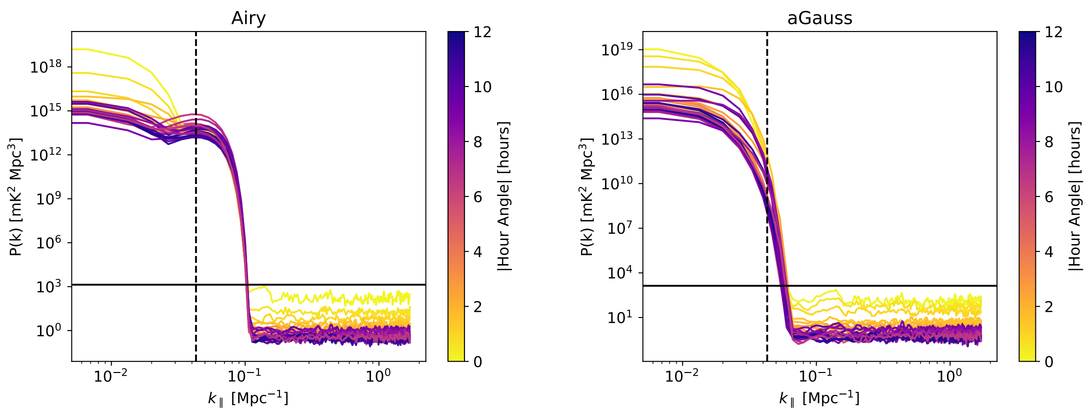


Figure 5.7: 1D power spectra for a 38.6 m baseline vs. the absolute value of the galactic center hour angle. Each curve corresponds with a different hour angle of the galactic center, for the symmetric GSM model. The vertical dashed line marks the horizon limit, and the horizontal line marks the fiducial EoR amplitude. The left is for the Airy beam; the right is for the achromatic Gaussian beam. There Airy beam shows a persistent "shoulder" near the horizon that is due to the foreground pitchfork, while power spectra with the gaussian beam do not. The pitchfork keeps the spillover at the EoR threshold line very nearly constant for all hour angles, despite the changing amplitude within the horizon.

We ran a set of simulations that covered 24 hours of LST in steps of 5 minutes (the "transit" time set), in order to measure how foreground power spillover on single baselines changes with the hour angle of the galactic center (recall that the "galactic center" also refers to the center of the sym model). As the galactic center moves away from the zenith, the brightest part of the sky is shifted toward a weaker part of the primary beam, but the delay spectrum peak moves closer to the horizon. Brightness near the horizon is further enhanced on all baselines by the pitchfork effect, so it is important to test how the power spillover vs. time changes with different baseline lengths.

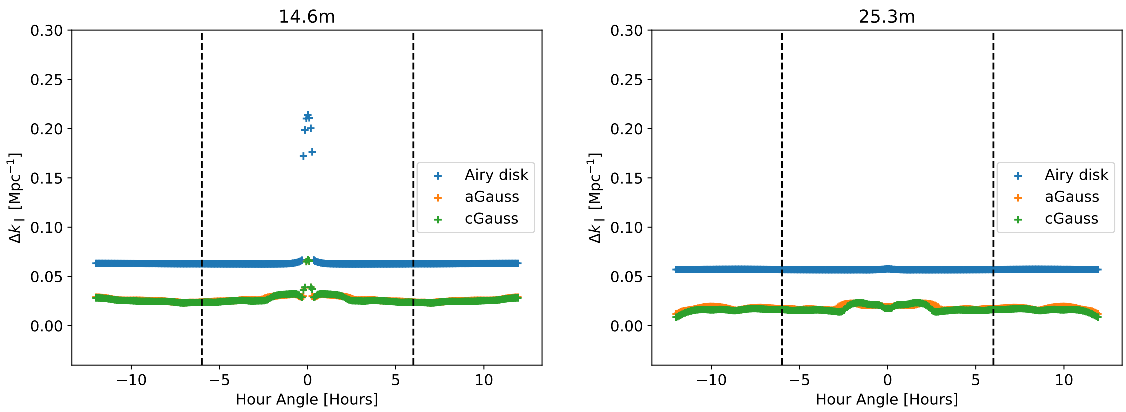
Fig. 5.7 shows the power spectra vs. galactic center hour angle for a 14.6 m baseline for the symmetric GSM model with Airy (left) and achromatic Gaussian (right) beams. The vertical line marks the horizon and the horizontal line marks the chosen threshold $P_{\mathrm{thresh}} = 29^2$ mK^2 Mpc^3. There is a clear increase in total power as the galactic center (hence, brightest part of the symmetric sky model) transits overhead. With the aGauss beam, the spillover shifts slightly as the total power changes. For the Airy beam, the pitchfork effect is apparent as a persistent bump near the horizon line, and the spillover stays relatively constant except for when the galactic center is overhead.

Fig. 5.8 shows the measured foreground spillover vs. galactic hour angle for the symmetric GSM (5.8a) and the actual GSM (5.8b) sky models for a 14.6 m baseline. As will be shown in the next section, the spillover is highest on short baselines because the foreground power is strongest there. Vertical dashed lines mark the rise and set times for the galactic center, which in both cases is the brightest part of the sky. The symmetric GSM model shows roughly the same behavior as the GSM, with a few minor differences. The symmetric GSM is flatter and more symmetric with hour angle, as expected. The GSM is less peaked at hour angles near zero, probably because the full power is more spread out across the sky than in the sym model. For the Airy beam, which has a non-negligible response near the horizon and thus has a consistent pitchfork, the spillover is higher. It is clear that the pitchfork dominates suprahorizon emission during times when the galaxy is out of the primary beam. The two Gaussian beams show very similar behavior, with minimal spillover at times when the galactic center is away from the zenith.

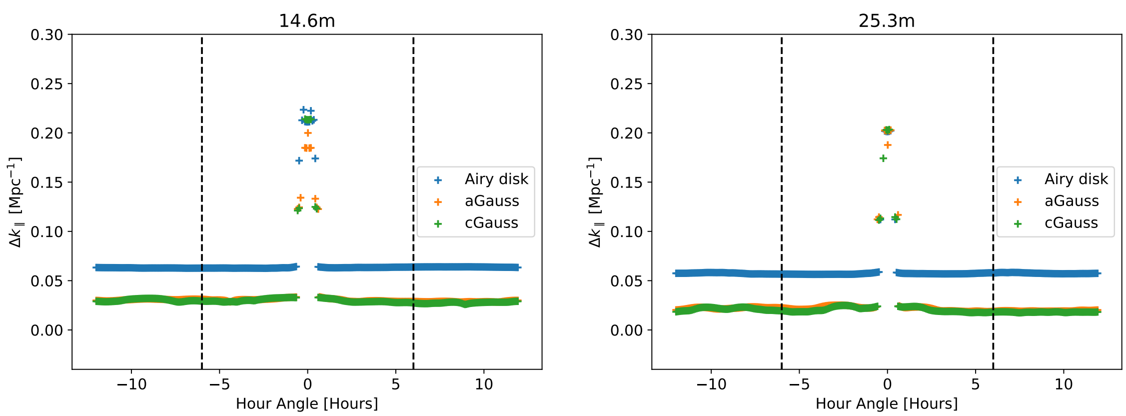
Baseline Length

The last set of simulations looked at effect of baseline length on foreground power spillover. For this set we use the "Line" antenna layout, which provides 300 East-West baselines with lengths spanning 10 to 150 m. The 150 m baseline here is the longest baseline used here for foreground spillover measurements. These simulations used the Off-Zenith time set, and power spectra were averaged over that time span.

Fig. 5.9 shows the results with the symmetric GSM (5.9a) and full GSM (5.9b). In both cases, the foreground spillover is dominated by the pitchfork for the Airy beam above about 50 m, and is pushed up by the first sidelobe of the BH window on shorter baselines. For the Gaussian beams, which have no pitchfork, the spillover simply drops off with increasing baseline length. In all cases the spillover decreases with increasing baseline length, albeit more slowly for the Airy beam, due



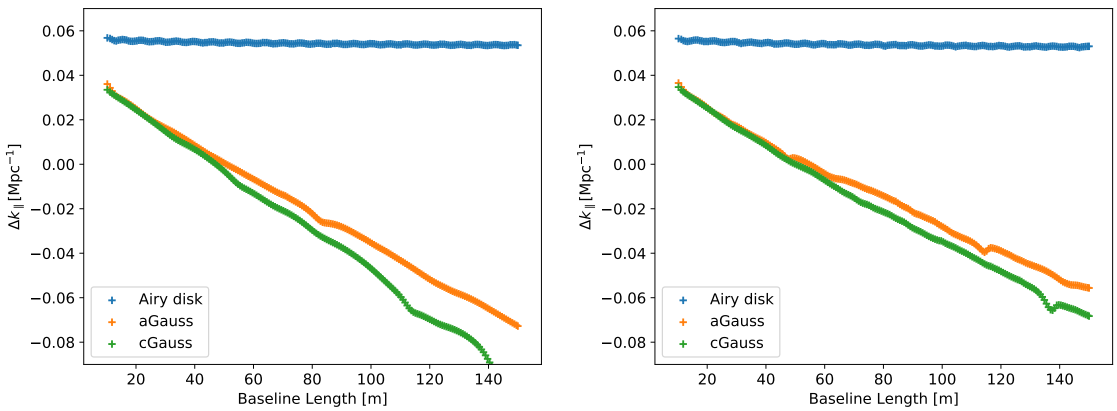
(a) sym



(b) GSM

Figure 5.8: The foreground spillover at the threshold shown in Fig. 5.7, for a 14.6 m baseline (left) and a 25.3 m baseline (right), plotted against the hour angle, for the three analytic beams. 5.8a shows the results with the sym model, while 5.8b shows the GSM results. Vertical dashed lines mark ±6 hours, when the galactic center (or, symmtric GSM center) is rising or setting. The spikes near $\simeq 0$ hrs arises from the Dolph-Chebyshev sidelobe ripples rising above the threshold. Overall, the spillover remains fairly constant with time, with the Airy disk beam showing consistently higher spillover due to the constant presence of the pitchfork. On a longer baseline, the overall amplitude is lower since longer baselines collect less power from diffuse emission.

to the decrease in power measured with longer baselines, as is expected for diffuse models.



(a) sym (b) GSM

Figure 5.9: The foreground spillover at the threshold shown in Fig. 5.7, plotted against East-West baseline length, for the three analytic beams. The black line marks the horizon. These simulations were done with the Off-Zenith time configuration. Spillover increases steadily with decreasing baseline length, but remains larger due to the pitchfork. Note that a negative spillover means that foreground power is contained within the horizon limit at the EoR amplitude.

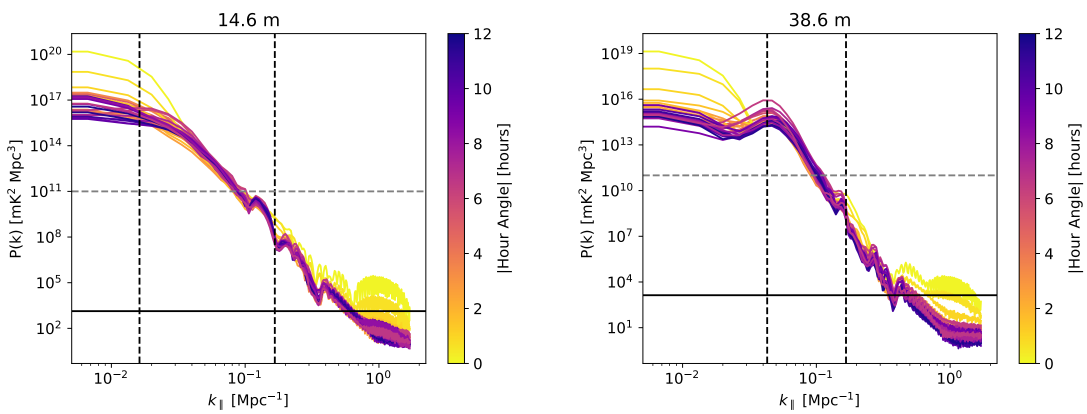


Figure 5.10: Power spectra vs. galactic center hour angle for the achromatic sym model and the CST XX beam. The vertical dashed lines in each plot are the horizon and Nyquist limit, while the horizontal lines show the EoR amplitude threshold used before (solid, black) and the new threshold for section 5.4.5.

5.4.5 CST Beam Results

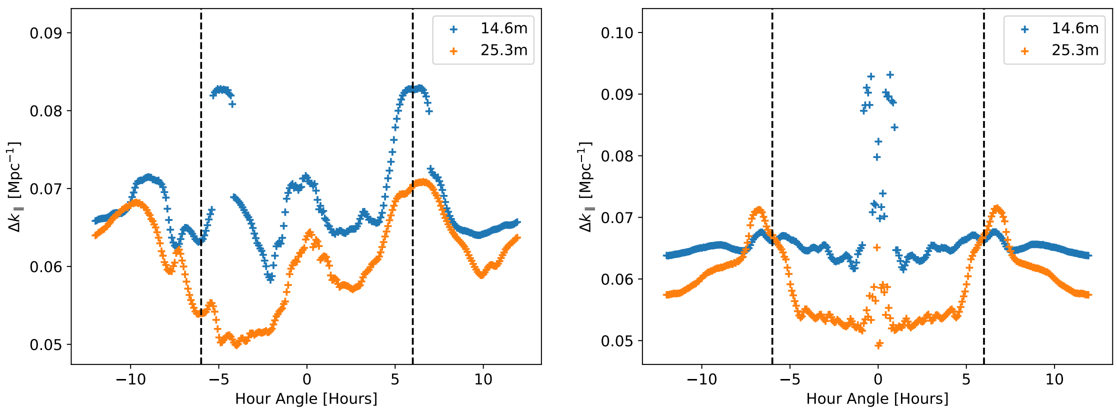
In this section we share the results of simulations with the CST beam. Fig. 5.10 shows the power spectra vs galactic center hour angle for the achromatic sym model with the CST XX beam at two baseline lengths. The pitchfork is present for both baseline lengths, but is clear persistent on the longer baseline at a higher amplitude than the power at smaller $k_\parallel$. As we have seen with the analytic beams, the pitchfork is largely independent of baseline length. Note also that for almost all times, the power spectrum drops below the EoR amplitude threshold (solid black horizontal line) at $k_\parallel$ well past the Nyquist limit. We therefore choose a higher threshold of 10^{11} mK^2 Mpc^3 (the horizontal, dashed, gray line in Fig. 5.10) for measuring foreground spillover with the CST beam. This avoids measuring modes that were only reached by interpolation. Though these results are not at the EoR amplitude, they demonstrate the general behavior of the foreground spillover with the CST beam.

Fig. 5.11 shows the spillover vs. hour angle for the CST beam with the GSM Fig. 5.11a and sym Fig. 5.11b models. The symmetric model more clearly demonstrates a pattern – the spillover is generally higher for the shorter baseline, because total power is higher on the shorter baseline. When the galactic center is near the horizon, the baseline length becomes less important to the spillover and the curves come together. This is because the strongest foreground component, near the galactic center, is near to the horizon in delay space, so its power extends further beyond the horizon than when it is at a smaller delay.

Fig. 5.12 shows the spillover vs. baseline length for the GSM and sym models. The overall decrease with baseline length is apparent, as in Fig. 5.9, but is much lower overall than for the Airy beam. The black line marks the Nyquist limit, beyond which the spillover is measured at Fourier modes that were interpolated. The continued trend there suggests that the interpolation / GPR smoothing is producing appropriate results. The sym model, because of the enforced azimuthal symmetry of the power, has a more concentrated bright point, which may explain why longer baselines measure more power from it as long as the “galactic center” is still within the primary beam. This would explain the slight divergence between the two curves in Fig. 5.12.

5.5 Summary

Smooth-spectrum foreground sources are expected to be well-behaved in Fourier space, which leaves a clear and well-defined window in which to seek an EoR detection. Spectral structure of the foregrounds, chromatic effects of the baseline and primary beam, and chosen bandpass and window function all contribute to the breakdown of this simple picture and the spread of power into the EoR window.



(a) GSM

(b) Sym

Figure 5.11: The spillover vs. hour angle for the (a) GSM and (b) sym models, for 14.6 m (blue) and 25.3 m (orange) baselines. Horizon lines are marked in the same colors. The strong response of the CST beam near the horizon causes the spillover to increase when the galactic center is near the horizon, at ± 6 hours (marked with dashed vertical lines). The longer baseline has less power spillover than the shorter for most hour angles, likely because the total power it measures is less, but at the horizon the curves approach each other. This pattern is especially clear for the sym model.

We carried out a series of simulations to characterize and measure this power spread for diffuse models in a variety of controlled cases. We used the Global Sky Model, and a symmetrized version of it with varied spectral steepness, in comparison to a fiducial EoR amplitude threshold. The analytic beams, antenna layouts, and frequency channelization and time step sizes are all chosen to resemble the design of HERA.

The choice of a spectral window function has the strongest effect on suppression of foreground power beyond the horizon; windowing along the frequency axis is required to reach the EoR level. Applying a window function effectively narrows the bandpass, widening the main lobe of foreground power in delay space. The lowest $k_{||}$ modes are lost to foreground spillover, but with a suitable window function the traeoff between foreground supression and spillover can be optimized to provide access to lower $k_{||}$ modes. The Dolph-Chebyshev window offers an optimized tradeoff between width and dynamic range, and may be the best choice to achieving this.

For most observing time, the most significant source of suprahorizon power is the foreground pitchfork, which appears when the primary beam has a response near the horizon, and is caused by the shortening of projected baselines near the horizon. As short baselines are more sensitive to diffuse structure, even faint monopole power near the horizon produces a strong response which

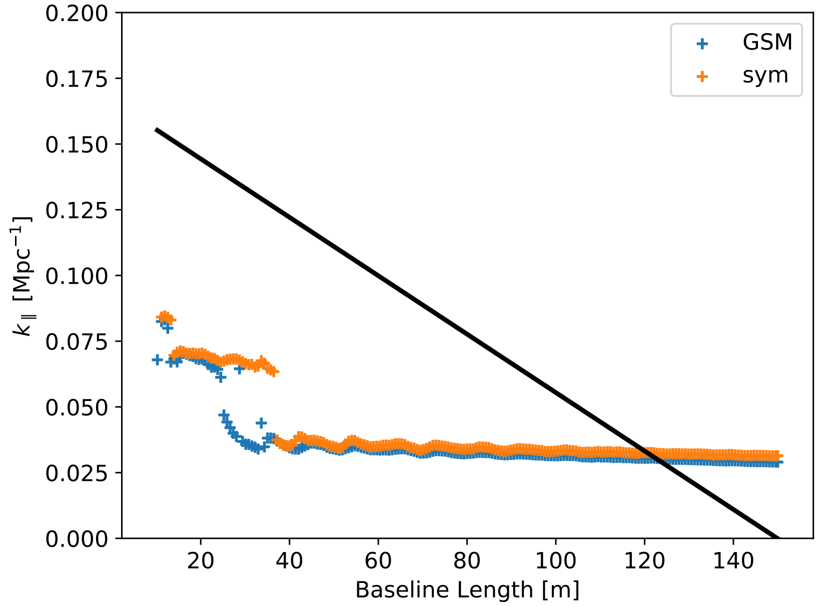


Figure 5.12: The spillover vs baseline length for the CST beam, for GSM (blue) and sym (orange) models. The black line indicates the Nyquist frequency cutoff of the CST beam model. On especially long baselines, the spillover exceeds this limit, but continues in the general trend seen before. The jumps near small baseline length are due to the increase spectrum amplitude shifting some of the features at higher $k_{\parallel}$ above the threshold.

can be observed for long baselines which are normally insensitive to diffuse power.

It is worth noting that the simulations used here do not take into account mutual coupling effects among antennas, as discussed in Fagnoni et al. (2019). In a real array, the response of an antenna to sources near the horizon will be affected by the presence of the other antennas that interfere with the propagation of light. To take an extreme case, a bright source exactly in line with both antennas in a baseline will only be visible to the antenna nearer to it, since the farther antenna is in the near one's shadow. In this case, the correlation between the two antennas is zero. However, it may be that the effect of baseline projection is still significant enough to produce a pitchfork when emission is near the horizon but not in the shadow of other antennas. the observation of a pitchfork in HERA data by Kern et al. (2019a) would seem to confirm this expectation. The full effects of mutual coupling are too sophisticated an effect to model with `healvis` at this time, and so we leave this to future work.

The Airy disk beam and CST simulated power beam both have a strong response near the horizon, and so have a constant foreground spillover due to this pitchfork for most times and baseline lengths. Avoiding LSTs when the galactic center is near the beam center, and choosing an appropriate constant buffer beyond the foreground wedge, appears to be sufficient to avoid diffuse foreground power, absent other systematic effects. It is difficult to draw a conclusion about foreground suppression at the level of the EoR for the CST beam, since those power spectra do not reach the EoR amplitude within the Nyquist limit of the CST simulations. Finer resolution in frequency and an improved frequency interpolation will be necessary to obtain useful simulation results at high delays.

If foreground power can be successfully avoided, the next obvious question is, "What EoR power spectra can be measured, and to what precision?". The next chapter looks at `healvis` simulations of EoR models, and studies how the baseline and primary beam affect the variance of power spectrum measurements.

Chapter 6

Sample Variance of EoR Power Spectrum Measurements[1]

6.1 Introduction

Choosing the optimal interferometer design to maximize power spectrum sensitivity requires balancing the effects of antenna positions, receiver element, and computational requirements of the design. The most significant factors that need to be reduced are thermal noise and foreground power (Franzen et al., 2016; Yatawatta et al., 2013; Bernardi et al., 2010; Jelić et al., 2008). Short baselines are more sensitive to diffuse structures like the EoR signal, and a random antenna distribution typically improves imaging capability, which helps foreground modeling and subtraction (Lidz et al., 2008). This motivated the Phase I layout of the Murchison Widefield Array (MWA), in Murchison Radio Observatory in Western Australia, which consisted of a random distribution of 128 elements, featuring a dense core but with several baselines reaching as far as a few kilometers. On the other hand, a highly-redundant array can be made more sensitive to line of sight Fourier modes with relatively few antennas, because redundant baselines directly sample the same angular Fourier modes ($k_\perp$) (Dillon & Parsons, 2016; Parsons et al., 2012a). This has the additional benefit of improving calibration through redundant techniques (Liu et al., 2010).

With these considerations, many current EoR experiments feature highly redundant layouts, fine spectral resolution, and drift-scanning, zenith-pointing antennas. The Hydrogen Epoch of Reionization Array (HERA), under construction in the Karoo desert of South Africa, comprises a

[1]A version of this chapter has been published in MNRAS (Lanman & Pober, 2019). Material has been reprinted here with the permission of co-authors.

densely packed hexagonal grid of 14.6 m parabolic dishes with dual-polarization feeds at the prime focus (DeBoer et al., 2017). It is sensitive to a bandpass of 100 − 200 MHz with 1024 frequency channels (97 kHz channel width). When finished, it will have 350 dishes, 30 of which are outriggers placed farther away from the compact core to improve imaging capabilities.

The Phase II deployment of the MWA is a hybrid design (Wayth et al., 2018) with 54 of the original core tiles of Phase I left in place, and the remaining 74 tiles re-purposed into two 37 element hexagonal grids with a 14 m spacing. Each "tile" is a phased array of 16 dual-polarization dipole antennas acting as a single element in the correlator. This hybrid design allows for a combination of calibration and foreground removal techniques to be applied, and provides data with well-tested hardware for comparison with HERA.

The trade-off to these redundant designs is that for more antennas there are fewer independent samples of each Fourier mode, due to the redundancy of many measurements. Statistical quantities such as the power spectrum have an intrinsic uncertainty, called *sample or cosmic variance*, due to the limited size of the volume surveyed by an instrument. For a Gaussian signal, this is proportional to the power spectrum amplitude over the number of independent samples N (see McQuinn et al. 2006). This holds true for non-Gaussian signals when N is sufficiently large, due to the central limit theorem. The root-mean-square (RMS) error is the square root of the sample variance, and drops off as $1/\sqrt{N}$. Assuming no correlation between neighboring times, the number of independent samples of each $k_{\|}$ mode is just the number of integration times, but real observations are correlated in time. Each radio interferometric measurement (visibility) is an integral over the primary beam across the sky, and visibilities at neighboring times will integrate over overlapping regions. The effective number of independent samples in a given survey volume is decreased by this correlation, so the sample variance decreases more slowly.

Previous sensitivity estimates (Zhang et al., 2018; Pober et al., 2014; Thyagarajan et al., 2013; Beardsley et al., 2013; Parsons et al., 2012a) have argued that sample variance is negligible except in a handful of k-modes, and so subsequent analysis has focused on the more significant issues of thermal noise and foreground contamination. Since the wedge is narrower at short baselines, many published power spectrum limits use data from short baselines only. For example, the power spectrum limits from PAPER (Kolopanis et al., 2019; Cheng et al., 2018; Ali et al., 2015) were made using only the three shortest baselines in the array.

Ultimately, measurements of 21 cm power spectra from the cosmic dawn will be used to constrain cosmological and astrophysical model parameters using such methods as Markov Chain Monte Carlo (MCMC) (Greig & Mesinger, 2015) and machine learning (Schmit & Pritchard, 2017; Gillet et al., 2019). Greig & Mesinger (2015) found that a 25% uncertainty in the measured power spectrum can increase uncertainty in model parameter estimates by a factor of a few, significantly degrading the constraints on the properties of galaxies during the EoR (although sampling the 21 cm power spectrum over a wide range of redshifts can reduce the effect). Such uncertainty must be understood and accounted for in the era of precision cosmology.

We argue that the sample variance of single-baseline power spectrum estimation does not fall below this 25% threshold for a typical data volume. In other words: even if all foreground and noise power can be properly mitigated, the fundamental uncertainty of the estimated 21 cm power spectrum will be too high for it to be used for precision cosmology. Through instrument simulations of a mock EoR signal, we explore how this sample variance is affected by baseline length and beam width, and how much it can be reduced by combining data from multiple non-redundant baselines.

6.2 Power Spectrum Sensitivity

Many previous measures of instrument sensitivity estimated the power in Fourier space of thermal noise and residual foreground contamination, and made simplified approximations of sample independence to calculate the variance in each Fourier mode (e.g. Beardsley et al., 2013; Pober et al., 2014). The approximate integration time per Fourier mode is calculated by dividing the UV plane into bins of size given by the effective antenna aperture in wavelengths and calculating the motion of baselines with time and frequency. The variance is estimated by modeling thermal noise and residual foreground power and dividing by this integration time. Sample variance is estimated as a component proportional to the expected power spectrum amplitude.

Beardsley et al. (2013) used a method like this to demonstrate that the MWA Phase I would be capable of detecting the EoR power spectrum, but relied on optimistic assumptions of the degree to which residual foreground power contaminated the signal (Dillon et al., 2015). Pober et al. (2014) estimated the sensitivity of HERA in this way by combining the thermal noise uncertainty, from a formula presented in Pober et al. (2013a), and residual foreground power from three foreground models.

Thyagarajan et al. (2013) considered the contributions of thermal noise, residual foreground power, and sample variance to the uncertainty of MWA Phase I observations in comparing two observing modes: a long integration on a single field, or (for an equal total observing time) fewer observations of multiple fields. As a phased array, each MWA tile is capable of pointing in discrete directions. This allows for a “drift and shift” observing strategy, where the array can approximately track a field by phasing to a position just west of it, letting the field drift through, then “shifting” the pointing west again. Thyagarajan et al. (2013) also took into account a more realistic model of the antenna power pattern (beam) in UV space, providing a more realistic look at how primary beam effects (especially sidelobe sources) contribute to foreground power leakage. Sample variance was assumed to drop off as one over the number of fields, i.e., for perfectly independent fields.

Trott (2014) performed a similar analysis for MWA Phase I, but instead of looking at discrete fields also considered the effects of covariance between visibilities in a constantly-evolving *drift scan*. In this case, the independence of observing fields is harder to determine. Trott (2014) found that the choice of ideal observing strategy depended on the available observing time and survey volume. The drift-scanning mode reduced sample variance the most for a given observing time, as one might expect since drift-scanning covers a larger area, but did not reduce thermal noise as significantly. For a drift scanning approach, they estimate a best signal to noise ratio of 9.6 (or an RMS error of about 10%) for the power spectrum amplitude.

We take an exact approach to study the effects of sample variance on realistic 21 cm observations for redundant arrays. By simulating the observations of an idealized interferometer over 24 hours of observing time, we can directly measure the variance of the measured power spectra and the covariance of delay-transformed visibilities. We simulate observations of a full sky model with a known power spectrum, and measure the sample variance of power spectra estimated from the simulated datasets.

We ignore the effects of foregrounds and thermal noise in this work, assuming that foreground power has been sufficiently removed through sky modeling techniques. Thermal noise can in principle be reduced as much as desired by including data from multiple days. Ionospheric effects should also average down by combining data from multiple days, or at least can be avoided by selecting data from days of low solar activity (see Datta et al. (2016)). The sample variance ultimately is the most fundamental uncertainty, because we are limited to one sky.

We introduce our simplified delay spectrum estimator, similar to that used by PAPER and HERA, and discuss its statistics in section 6.3. Section 6.4 further explores the statistics of the estimator through controlled numerical tests, examining how visibility correlation relates to the distribution of estimator values. We describe our EoR-like sky model in section 6.5. Section 6.6 presents our main questions and the simulations used to address them. In section 6.7 we present the results of several sets of simulations, exploring the correlations of visibilities with time and the sample variance of power spectrum estimates.

6.3 Delay Spectrum Statistics

Starting with the delay spectrum estimators of section 2.2.2, we can combine estimates from multiple times and baselines to improve the signal to noise ratio. These simulations do not include thermal noise, so a simple delay-transform estimator may be obtained from eq. (2.29) for a single baseline,

$$\widehat{P}(\boldsymbol{k}_{b,\tau}, t) = \left(\frac{\lambda^2}{2k_B}\right)^2 \frac{X^2 Y}{B_{pp}\Omega_{\mathrm{pp}}} |\widetilde{V}(b,\tau,t)|^2 \tag{6.1}$$

where we include the explicit time and baseline dependence (t and b) into the visibility. Cosmological isotropy requires that the power spectrum be independent of direction, which for a drift-scanning instrument is equivalent to independence in time. We can therefore improve our estimator by averaging samples over observing time:

$$\widehat{P}(\boldsymbol{k}_{b,\tau}; N_t) = \frac{1}{N_t} \sum_{i=0}^{N_t} \widehat{P}(\boldsymbol{k}_\tau, t_i) \tag{6.2}$$

where t_i are the separate integration times of the instrument, of which we are averaging together delay spectra from N_t time samples, and $\boldsymbol{k}_{b,\tau}$ is the k-mode corresponding with baseline b and delay mode τ. Note that the basis for $\boldsymbol{k}$ is changing over time, as the telescope observes different parts of the sky, but corresponding k modes may be combined due to the isotropy of the EoR signal.

For a single baseline ($\alpha = \beta$), we can combine eqs. (6.1) and (6.2) into a single expression.

$$\widehat{P}(k_{b,\tau}, N_t; b) = \frac{\Phi}{N_t} \sum_{n=1}^{N_t} |\widetilde{V}(b, \tau, t_n)|^2 \tag{6.3}$$

where $\Phi = (\lambda^2/2k_B)^2 X^2 Y / B_{pp} \Omega_{pp}$ is a scalar combining the cosmological scaling factors of eq. (2.29). As discussed in section 2.2.2, squaring $\widetilde{V}$ for a single baseline will cause a bias due to thermal noise, so data from redundant baselines or from different days (but the same LST) are usually cross-multiplied, but the analysis presented in this work doesn't include thermal noise, so we can use eq. (6.3) as an unbiased estimator.

The delay-transformed visibilities may be written as a measurement vector $\boldsymbol{v}_\tau = (v_0, \ldots, v_N)$, with $v_n = \widetilde{V}(\tau, t_n)$. For a Gaussian sky signal, $\boldsymbol{v}_\tau$ follows a complex multivariate Gaussian distribution, with mean vector $\boldsymbol{\mu} = 0$ and covariance matrix $\mathbf{C} = \left\langle \boldsymbol{v}_\tau \boldsymbol{v}_\tau^\dagger \right\rangle$. $\mathbf{C}$ is an $N_t \times N_t$ symmetric, real, positive semi-definite matrix.

We may write our estimator (6.3) as

$$\widehat{P}_N \equiv \widehat{P}(k_{b,\tau}, N_t; b) = \boldsymbol{v}_\tau^\dagger \mathbf{A} \boldsymbol{v}_\tau \tag{6.4}$$

where $\mathbf{A} = (\Phi/N_t)\mathbf{1}_N$ for $(N_t \times N_t)$ identity matrix $\mathbf{1}_N$, and we let $\widehat{P}_N$ stand for eq. (6.3) to simplify notation with the understanding that it refers to a single baseline and delay mode. Similarly, we drop the τ subscript on $\boldsymbol{v}$. In this form, $\widehat{P}_N$ is recognizable as a real quadratic form. In the case that visibilities are uncorrelated and have similar noise, eq. (6.3) follows a χ^2 distribution with N_t degrees of freedom. More generally, this quadratic form may be shown to be χ^2 distributed with degrees of freedom $d = \text{rank}(\mathbf{AC})$ under the condition that the matrix $\mathbf{AC}$ is idempotent (Mathai & Provost, 1992; Searle, 1971). We find experimentally that eq. (6.3) is well-described by a χ^2 distribution (see section 6.4), despite a lack of idempotency of $\mathbf{AC}$. The underlying distribution of eq. (6.3) is not important to our results, but is of interest for interpreting the measured uncertainty. We will fit χ^2 distributions to the measured power spectra.

The variance of the estimator may be found directly from the covariance matrix using Theorem 1 of Chapter 5 of Searle (1971). The proof presented there readily extends to the complex case with the result:

$$\mathrm{Var}[\boldsymbol{v}^\dagger \mathbf{A} \boldsymbol{v}] = \mathrm{Tr}[(\mathbf{AC})^\dagger (\mathbf{AC})] + 2\boldsymbol{\mu}^\dagger \mathbf{ACA}\boldsymbol{\mu} \tag{6.5}$$

For the delay-transformed visibilities, the mean $\boldsymbol{\mu} = 0$, so this expression reduces to the trace of the squared covariance matrix with the scalar factors. Recalling that C_{ij} is Hermitian, it follows that the variance becomes the sum of the squares of all entries:

$$\mathrm{Var}[\widehat{P}_N] = \left(\frac{\Phi}{N}\right)^2 \sum_{i,j}^{N_t} |\mathbf{C}_{ij}|^2 \tag{6.6}$$

For the case of uncorrelated visibilities with variances σ^2, this reduces to $\mathrm{Var}[\widehat{P}_N] = \Phi^2\sigma^2/N$, as expected. To understand the variance of our estimator, therefore, we need to look at the covariance between visibility measurements at different times.

6.3.1 Visibility covariance

A full expression for the covariance between two delay-transformed visibilities has been worked out in Zhang et al. (2018) in detail. For a single baseline and drift-scanning visibilities, their expression reduces to the form

$$\begin{aligned} \mathbf{C}_{ij} = \frac{P(k_{b,\tau})}{X^2 Y} \left(\frac{2k_B}{\lambda^2}\right)^2 \int\int A^*(\hat{s},\nu) A(\mathbf{R}_{ij}\hat{s},\nu) \\ \times \exp\left[2\pi i \frac{\nu}{c}(\hat{s} - \mathbf{R}_{ij}\hat{s}) \cdot \boldsymbol{b}\right] d\Omega\, d\nu \end{aligned} \tag{6.7}$$

Equation (6.7) assumes the power spectrum is approximately constant across the bandpass (i.e. ignoring light-cone effects). $\mathbf{R}_{ij}$ is a three-dimensional rotation matrix describing the motion of source from time t_i to t_j in the observer's topocentric frame (e.g., altitude/azimuth).[2] Note that in the case of $t_i = t_j$, $\mathbf{R}_{ij} = \mathbf{1}$ and we recover the estimator (2.29) from eq. (6.7).

We can manipulate eq. (6.7) further to gain some better intuition. Let Θ_ν represent the sky integral, such that,

$$C_{ij} = \int d\nu \Theta_\nu$$

[2] A notational note – The subscripts on $\mathbf{R}_{ij}$ are not indices of the matrix itself.

Then

$$\begin{aligned}\Theta_\nu &= \int d\Omega A^*(\hat{s}) A(R_{ij}\hat{s}) \exp\left[-2\pi i \frac{\nu}{c}(\mathbf{1} - \mathbf{R}_{ij})\hat{s}\cdot\boldsymbol{b}\right] \\ &= \int d\Omega A_\nu^*(\hat{s}) A(\mathbf{R}_{ij}\hat{s}) \exp\left[-2\pi i \frac{\nu}{c}\hat{s}\cdot\boldsymbol{b}'\right] \\ &\quad \text{where } \boldsymbol{b}' = (\mathbf{1} - \mathbf{R}_{ij}^T)\boldsymbol{b}\end{aligned} \tag{6.8}$$

In this form, we can see that Θ_ν is given by the overlap of the beam between the pointings, weighted by a fringe term. The fringe width is given by length of the baseline projected onto the displacement vector of the sky, $|\boldsymbol{b}'|$.

If we choose a fixed phase center $\hat{s}_0$, we can rewrite eq. (6.8) under a flat-sky approximation. We define $\vec{b}'$ as the projection of $\boldsymbol{b}'$ onto the plane orthogonal to $\hat{s}_0$, and $\vec{l} = \hat{s} - (\hat{s}\cdot\hat{s}_0)\hat{s}_0$ is the projection of the sky position vector into the sky plane.[3] Finally, we denote by $\mathbf{\Gamma}_{ij}$ the transformation matrix of $\vec{l}$ corresponding with the 3D transformation R_{ij}.

In these terms, eq. (6.8) may be written

$$\Theta_\nu = \int A_\nu^*(\vec{l}) A_\nu(\mathbf{\Gamma}_{ij}\vec{l}) \exp\left[-2\pi i \frac{\nu}{c}\vec{l}\cdot\vec{b}'\right] d^2 l \tag{6.9}$$

In this form, the covariance is a 2D Fourier transform of a product of the primary beams, which may be expressed as the convolution of the Fourier-transformed primary beams, which we call *beam kernels*:

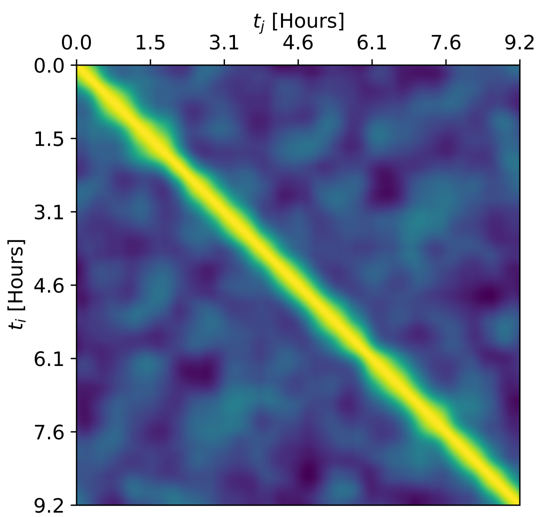
$$\Theta_\nu \simeq \int \widetilde{A}^*\left(\frac{\nu}{c}\vec{b}' - \vec{u}\right) \widetilde{A}(\mathbf{\Gamma}_{ij}^T \vec{u})\, d^2 u \tag{6.10}$$

This is a specific case of the more general beam kernel used in Patwa & Sethi (2019), for a single baseline and delay mode. This quantity depends only on the baseline vector, beam kernel width, and time between visibilities (*lag*). Since beam kernels are compact in Fourier space, the correlation is likewise limited in range, and can be characterized by a *correlation time* t_c. For our purposes, we define t_c to be the full-width at half maximum of the correlation vs. lag. Note that in the case of nearly-Gaussian primary beam functions, the convolution in eq. (6.10) will give a Gaussian function, which can be characterized entirely by its FWHM.

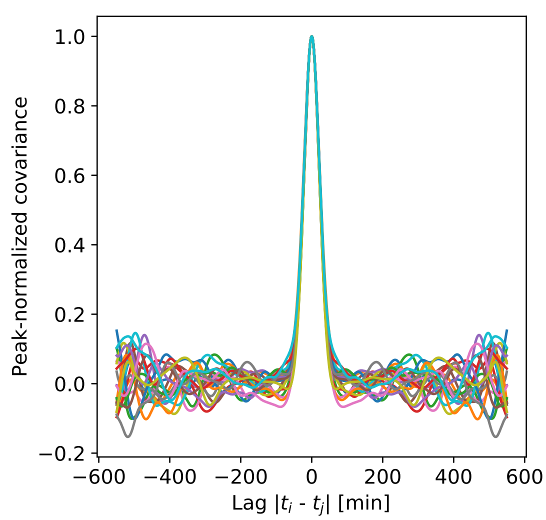
The width of beam kernel, typically equal to the antenna size in wavelengths, scales inversely with the primary beam width on the sky. From eq. (6.10) we therefore expect that measurements from shorter baselines should stay correlated longer than those from long baselines. In addition, observations with wider beams (corresponding to narrower beam kernels) should stay correlated for shorter times than measurements from narrower beams.

[3]We denote spatial 2D vectors with overhead arrows to distinguish from 3D spatial vectors.

This may seem counter-intuitive, because one may think of the primary (antenna) beam as selecting a "patch" of sky observed at a time sample. In this picture, nearby times will be correlated due to the overlap of their primary beams on the sky, and so one would expect the correlation time to be the time it takes the sky to pass through the primary beam. The fringe term in eq. (6.8) complicates this interpretation. For a given point on the sky, the fringe moves across the sky as the Earth rotates. The correlation at a point between times t_i and t_j goes to zero when the fringe at time t_i destructively interferes with the fringe at t_j, weighted by the primary beam at both times. With this consideration, it is simpler to think about the Fourier-transformed formula eq. (6.10), which is more in line with what previous sensitivity analyses have done. We will show in section 6.7.1, however, that the Fourier-space interpretation of the correlation is not always correct.



(a)



(b)

Figure 6.1: (a) The covariance matrix at a single delay, for a single baseline, from simulated data. Both axes are time in hours. (b) The measured correlation function, made from binning the covariance matrix in lag for each delay mode, scaled to 1 at a lag of 0. The bins of (b) correspond with lines parallel to the diagonal of (a). There are fewer entries in the covariance matrix that correspond with the large lags, which leads to the increased scatter away from the central peak. The central peak is well-described by a Gaussian function.

Fig. 6.1a is an example covariance matrix C_{ij} obtained from simulated data (see section 6.5 for simulation details). This is the correlation in time for a single baseline at a single delay and shows the expected symmetry due to the isotropy of the sky signal. Binning the covariance matrices for each delay mode in time separation, or *lag*, we get the curves Fig. 6.1b, which we call the *correlation* *function*. Each curve in Fig. 6.1b is essentially Fig. 6.1a for a given delay, averaged in bins parallel to the diagonal. These curves show that the central stripe in the covariance matrix has a nearly Gaussian shape.

6.3.2 Theoretical sample variance

We can construct an analytic function for the variance, eq. (6.6), under some assumptions for the visibility correlation. We first assume that the covariance matrix C_{ij} is symmetric and constant parallel to the diagonal (also known as a symmetric Toeplitz matrix). This matrix symmetry is equivalent to the assumption that the EoR signal is isotropic since the correlation matrix depends only on the separation in time index,

$$\gamma_n = \Phi^2 |C_{ij}|^2 \qquad \text{such that } n = |i-j|$$

The factor of Φ, from eq. (6.3), is included to absorb the cosmological and unit scalars into the definition of γ_n. We can express the sum in eq. (6.6) by counting the number of matrix entries at each n.

$$\mathrm{Var}[P_k] = \frac{1}{N}\left(2\sum_{n=0}^{N-1}\gamma_n - \gamma_0\right) - \frac{2}{N^2}\sum_{n=0}^{N-1} n\gamma_n \tag{6.11}$$

To continue, we will treat γ_n as a discretization of a continuous correlation function $\gamma(t)$. Assuming the time is discretized into time steps of length Δt, such that $t_n = n\Delta t$, with a total time $T = N\Delta t$, the sums in eq. (6.11) become:

$$\frac{1}{N}\sum_{n=0}^{N-1}\gamma_n \approx \frac{1}{N\Delta t}\int_0 \gamma(t)dt = \frac{1}{T}\int_0^T \gamma(t)dt \tag{6.12}$$

$$\frac{1}{N^2}\sum_{n=0}^{N-1} n\gamma_n \approx \frac{1}{N^2\Delta t^2}\int_0 t\gamma(t)dt = \frac{1}{T^2}\int_0^T t\gamma(t)dt \tag{6.13}$$

Approximating the correlation function as a Gaussian with FWHM t_c, we can evaluate these integrals. Letting $w_c = t_c/2.355$, the correlation function is

$$\gamma(t) = \gamma_0 \exp\left(\frac{-t^2}{2w_c^2}\right) \tag{6.14}$$

Carrying out the integrals in eq. (6.11) with this Gaussian, we're left with the following for the variance of the power spectrum estimator at averaging time T:

$$\mathrm{Var}[\widehat{P}_k(T)] = \gamma_0^2\left(\frac{w_c}{T}\sqrt{2\pi}\,\mathrm{Erf}\left[\frac{T}{\sqrt{2}w_c}\right] - 2\left(\frac{w_c}{T}\right)^2\left[1 - e^{-T^2/2w_c^2}\right]\right) \tag{6.15}$$

Erf[·] is the error function. This expression holds only under the assumption that $\Delta t \ll T$. In section 6.9 we show that this approximate expression fits with the sample variances measured from simulated data.

6.4 Toy Model

To better understand the relationship between the distribution of $\widehat{P}_N$ and the correlation of visibility data, we ran a series of simple tests using mock data with a known correlation. We generated an ensemble of M arrays of complex Gaussian random variables $\boldsymbol{x} = (x_0, \ldots, x_{N_t})$, each with a length N_t, and convolve each with a normalized Gaussian function to introduce a correlation. We then take progressively longer averages of the absolute square of these arrays to mimic the form of the power spectrum estimator (6.3).

$$q(l) = \frac{1}{l}\sum_{j=1}^{l} |x_j|^2 \qquad \text{for } l \leq N_t \tag{6.16}$$

For each averaging length l, this gives us an ensemble of M averages $\{q(l)\}$. We can fit a scaled χ^2 distribution to this set using maximum likelihood estimation. The distribution function has the form

$$f(q; k, \alpha) = \frac{(\alpha q)^{k/2-1}\exp(-(\alpha q)/2)}{2^{k/2}\Gamma(k/2)} \tag{6.17}$$

where α is a scaling parameter and k is the degree of freedom parameter, and $\Gamma(k)$ is the gamma function. When defined from a sum of squares of unit variance variables, the χ^2 distribution has a mean equal to its degrees of freedom (dof). In the form (6.17), the scaling parameter reflects that x is not unit variance and that q is an average, not a sum.

We expect that the dof in the fits should be related to the correlation time, and we do indeed find that

$$\frac{l}{k} \to l_c \text{ as } l \to N_t$$

where l_c is the full-width at half maximum of the convolution kernel. This is shown in Fig. 6.2. The scaling parameter α converges to $\alpha \to \sigma^2/(k-1)$, where σ^2 is the variance of x.

To make for a better comparison with our simulations, we define a time step $\Delta t = 11$ sec, and use $N_t = 7854$ samples as described in Table 6.1. The convolution kernel width is set in terms of time as $l_c = t_c/\Delta_t$ for $t_c \in \{20 \text{ min}, 30 \text{ min}, 50 \text{ min}, 100 \text{ min}\}$. For these cases, we try to recover the input t_c through the estimator

$$\widehat{t_c} = k\,\Delta t \tag{6.18}$$

where k is the fitted dof. Fig. 6.2 shows the result.

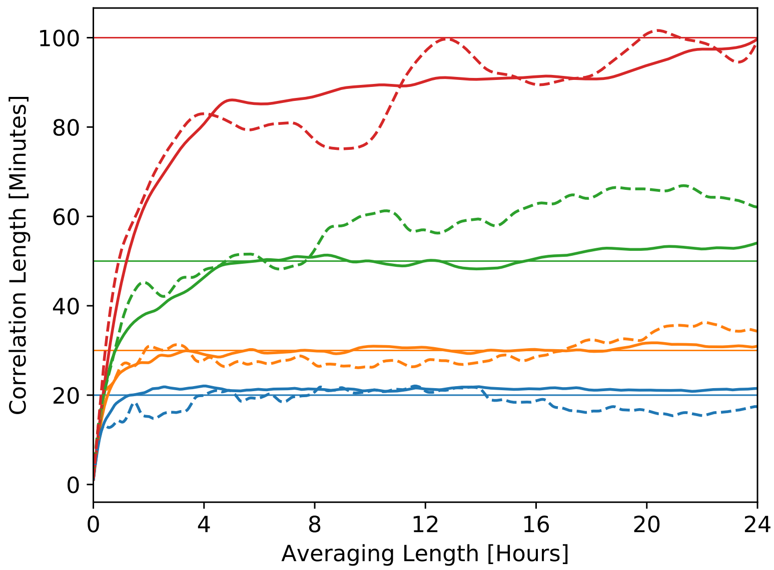


Figure 6.2: Estimated correlation time vs. averaging length for the toy model of section 6.4. The correlation time is estimated as in each averaging length by fitting the scaled χ^2 distribution of eq. (6.17) and using eq. (6.18). The horizontal lines are the true correlation times. The solid lines were made with an ensemble size $M = 1000$, the dashed lines are for $M = 100$. The correlation times shown are typical of the baseline cases described in section 6.5. The poor convergence for longer correlation times and smaller ensembles indicates that we need to be careful interpreting the results of fitting a χ^2 distribution to power spectrum estimates.

For the simple, short-range correlations of the delay-transformed visibilities we can still expect the $\widehat{P}_N$ to be χ^2 distributed, and that we will be able to relate the correlation time. However, the convergence of $\hat{t}_c$ to the true value is slow, and depends on the number of samples going into each fit and how the correlation time compares with the total time available (24 hours). For longer correlation times and fewer samples, the convergence is worse.

Overall, we interpret the results of this toy model to mean that for the baseline types, beam widths, and averaging lengths under consideration, the distribution of $\widehat{P}_N$ will be best described by a χ^2 distribution with relatively few degrees of freedom. The dof is connected to the correlation time, but it may not be possible to *reliably* estimate t_c from fitting the dof. For the rest of this work, we use more robust methods to estimate the correlation time and compare with the results of a χ^2 fit for consistency.

6.5 Simulation

We carry out simulations using `healvis`, which supports simulating multiple realizations of a statistical sky model within a single job and is optimized for a long time axis. These simulations cover 24 hours of LST at 11 s time steps, and 30 MHz of bandpass at 78 kHz channel width (equivalent to the MWA channelization). For EoR simulations, we do not require complete horizon to horizon coverage for realistic results – only suitably far for potentially long correlation times. For foreground simulations, the FoV should be set to extend to the horizon, (see Thyagarajan et al. 2015a), but since the choice of FoV increases the number of points to be integrated (especially near the horizon) increased FoV comes at the expense of increased runtime. We choose the FoV diameter to be 110°, at which distance the widest beam used in this analysis drops to 0.01% of its maximum.

For all simulations in this chapter we used an achromatic Gaussian beam. From here on, "beam width" refers to the FWHM of this Gaussian function ($\approx 2.355\sigma$). This is chosen for computational simplicity and to allow us to explore the effects of varied beam width directly. More realistic beam models, especially those obtained from calibrator sources or EM modeling, are important when modeling foregrounds, because bright sources moving through the beam and its sidelobes can introduce a spectral structure (see (Pober et al., 2016) and chapter 5). Since the signal we analyze here is isotropic, there won't be any drastic changes in integrated flux due to, say, a bright source near the edge of beam dropping into a null at lower frequency. We therefore don't expect these subtle beam effects to be important here.

We note that, though well-defined in sky coordinates, this beam function does not directly correspond with a physical antenna on the ground, such that we are ignoring the potential impacts of beam chromaticity. Within the chosen bandpass, the effects of baseline chromaticity and the signal structure are much more important.

6.5.1 Mock signal model

Instrument simulation tools are generally designed to read sky models in instrument coordinates (angle and frequency). Since frequency maps to distance, the survey volume forms a spherical shell with thickness corresponding with the bandwidth. This shell is divided into voxels with perpendicular area $\chi_\nu^2\omega$ and thickness $\delta\chi_\nu = |\chi(\nu + \delta\nu) - \chi(\nu)|$, where χ_ν is the comoving distance corresponding with frequency ν, ω is the pixel solid angle, and $\delta\nu$ is the frequency channel width. Within the HEALPix pixellization scheme, the Nside parameter sets the pixel solid angle to $\omega = 4\pi/(12{\times}\mathrm{Nside}^2)$ sr, such that all pixels have the same area and pixel centers are evenly spaced (Gorski et al., 2005).

For our sky model, we select a brightness for each voxel sampled from a Gaussian distribution with frequency-dependent variance $\sigma^2(\nu)$. A space with equal voxel volumes has a flat power spectrum with amplitude $\sigma^2 dV$, since there are no spatial correlations among the values. The comoving voxel volume changes with distance for two reasons. The radial thickness of the voxel changes because the frequency channels are evenly spaced and are nonlinearly related to comoving distance. In addition, the transverse comoving size of the voxel grows quadratically with distance because the solid angle remains the same. To account for this, we scale the variance at each frequency to ensure that the power spectrum amplitude is the same across the band.

$$\sigma^2(r) = \frac{\sigma_0^2 dV_0}{dV(r)} \tag{6.19}$$

Here, $dV(\nu) = \chi_\nu^2 \omega \delta\chi_\nu$ is the comoving voxel volume as a function of frequency. Quantities subscripted 0 refer to a reference frequency, chosen to be the center of the bandpass.

This model has a flat power spectrum, $P(k) = \sigma_0^2 dV_0$ (the proof of this is in Appendix C), which has some advantages for further analysis and is simple to generate in large quantities. The problem of generating a more realistic EoR model involves establishing spatial correlations that are isotropic and homogeneous within the spherical shell geometry, and is not trivial. It will be of interest to simulate data from EoR models that takes into account both the spherical nature of the sky and the expected non-Gaussian features of the late-stage EoR signal, especially since these nonlinearities can increase uncertainty on power spectrum measurements (Mondal et al., 2015). We leave full-sky modeling of EoR signals to future work.

6.6 Key Questions

We ran four sets of simulations to explore the following questions:

1. What is the relationship between correlation time and baseline length?

   We explore the validity of eq. (6.7) by running a suite of short simulations with a variety of baseline lengths and beam widths, and calculate the correlation times from them.

2. What is the relationship between correlation time and baseline orientation?

   Baseline orientation affects how the fringe is oriented with respect to sky motion, and can therefore affect visibility correlation. We run simulations similar to those of question (i), but fixing the baseline length and varying its angle in the topocentric frame.

3. How does the root-mean-square (RMS) error of single baseline delay spectra drop off with increasing averaging time for HERA-like and MWA-like beams?

   Here we wish to know how well a single baseline can measure the EoR power spectrum through a delay transform method.

4. What is the lowest RMS error that can be achieved using all the elements of a 37 element hexagon, like MWA Phase II?

   Combining data from multiple independent baselines can improve the sensitivity. Within the basic delay spectrum formalism, we are limited, however, to combining baselines of similar lengths, which correspond to the same $k_\perp$ mode. With all antennas in the 37-element array, how much does this binning improve the RMS error?

Table 6.1 summarizes the common parameters used by the simulations. Each simulation used the same bandpass, channel width, and array location. The latitude of the array does affect how baselines move through Fourier space over time, which may impact results. We use the HERA array position for all simulations, which is within 5° of latitude from the MWA site, so we expect our results apply to both cases. We use the MWA channel width, which is the finer of the two instruments, and use a 30 MHz bandpass. Most delay analyses focus on a fiducial 10 MHz band, to avoid redshift evolution effects, but for our purposes choosing a wider bandwidth increases the number of $k_\parallel$ probed.

For our sky models, we choose an Nside of 128 and a band-center pixel brightness variance $\sigma_0^2 = (31\text{ mK})^2$, based on a fiducial EoR signal amplitude at $z \sim 13$, the bottom of our bandpass (Pritchard & Loeb, 2008). Since we're looking at the relative error, the actual signal amplitude is inconsequential. The choice of Nside is a compromise with available computational resources. The simulator had 120 GB of memory available, which was sufficient for a shell with an Nside of 256 (resolution of about 13 ′) and a 30 MHz bandpass, or an Nside of 512 (resolution 6.8 ′) with a 10 MHz bandpass. Both higher resolutions come at the expense of significantly longer run times, and since many tasks involved running hundreds of simulations we choose a modest resolution of Nside=128, which corresponds with an angular resolution of about 27.8 ′. This is still small relative to the smallest interferometric fringes in the simulations, and comparisons with higher resolution simulations showed no noticeable impact on results.

6.7 Results

6.7.1 Correlation time vs. Baseline length and Beam Width

The first set of simulations was done to explore the effect of baseline length and beam width on visibility covariance. For 86 beam widths ranging from 10° to 50° we simulated an array of 50 coplanar baselines pointing East, with lengths evenly spaced from 20 to 100 m. These simulations were repeated for 50 separate sky realizations, and run for 2 hours worth of simulated time.

For each simulated dataset, we delay-transform the visibility array and compute the covariance

σ_0	31 mK
Bandpass	100 – 130 MHz
Channel Width	78.125 kHz
Duration	24 Hours
Integration Time	11 seconds
Latitude	-30.72°
Longitude	21.428°
MWA beam FWHM	31.5°
HERA beam FWHM	12.13°
Field of View	110°

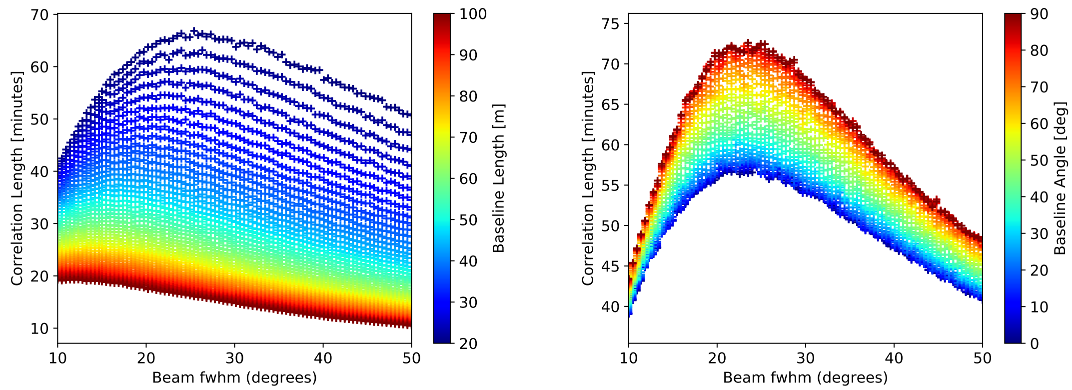
Table 6.1: Summary of EoR simulation parameters.

by cross-multiplying and averaging over delay. By treating the delays as an ensemble axis this way, we implicitly assume that all visibilities follow the same distribution at all delays, which we will show later to be a fair assumption due to the flatness of the power spectrum. This cross-multiplication gives us an $N_t \times N_t$ covariance matrix which is symmetric across both diagonals. We then bin the covariance matrix in lag, producing a correlation function that has a nearly Gaussian shape.

Fig. 6.3a shows the results of this set of simulations. As expected from eq. (6.7), correlation time decreases rapidly with increasing baseline length. For most baseline lengths, correlation time generally decreases with beam width. However, for beam widths less than 25° for baselines shorter than 50 m, correlation time *increases* with beam width. This is unexpected, and counter to our interpretation in section 6.3.1. This is also the regime in which we find the shortest HERA and MWA baselines.

There are several factors at play that may be causing this change in behavior at short baselines and narrow beams. As discussed in section 6.3.1, if we ignore the fringe term in eq. (6.8) then we would expect the correlation time to increase with beam width. The fringe width is determined by the baseline length, and for short baselines will grow more slowly with time than for a long baseline. It may be that, for short baselines and narrow beams, the decoherence due to narrow beams has a greater effect than the fringe term.

Patwa & Sethi (2019) have recently published a paper deriving an analytic expression for the time-correlation of drift-scan visibilities. They find a similar behavior in the correlation timescale vs. baseline length and beam width, showing a turnover at narrow beams. They attribute the turnover in behavior to a change in dominant factors – For large beams, the correlation time goes as $\simeq 1/(u\sigma \sin\lambda)$, for latitude λ and Gaussian width σ, because the correlation is dominated by the interactions of the Fourier-transformed beams. For small beams, the correlation time goes as $\simeq \sigma/\cos\lambda$, independent of baseline length, and the correlation is dominated by the overlap of



(a) Varying baseline length.

(b) Varying baseline orientation.

Figure 6.3: Correlation times vs primary beam width for 2 hour simulations. (a) shows the effect of varying baseline length for a purely East-West baseline. (b) shows the effect of varying the orientation of a 25m baseline, where the orientation is the angle north of east.

the beams on the sky. From these considerations they see a similar turnover in behavior when $\sigma^2 |\tan \lambda| u \simeq 1$.

The main takeaway here is that our intuition for correlation, as the overlap of beam kernels in Fourier space, typically leads to the behavior we expect: longer baselines and wider beams give us a shorter correlation time, and so visibilities will decorrelate faster with time and sample variance will drop off faster. However, the shortest baselines in HERA and the MWA do fall into the counter-intuitive regime where increasing beam width increases correlation time.

6.7.2 Correlation time vs. Baseline orientation and Beam Width

The second set of simulations explored the effect of baseline orientation. The parameters and methods used here were identical to those used in section 6.7.1, except that we fix the baseline length to 25m and vary its angle North of East in 30 steps from 0° to 90°. The results are shown in Fig. 6.3b. Increasing the baseline angle has the effect of increasing the correlation length. For the 25m baseline, this increase never exceeds about 20 minutes.

This behavior can be explained by considering the correlation length in the sky frame, eq. (6.8). The fringe frequency in the integral is set by the baseline displacement vector $\vec{b}'$, which points East to West due to the Earth's rotation. At the equator, a perfectly North-South baseline would experience no rotation at all, so $b' = 0$. In this extreme case, the correlation length would be set only by the beam width because the exponential term in eq. (6.8) would go to one. Similarly, at nonzero latitudes the correlation length is maximized for a perfectly North-South baseline because the displacement b' is minimized in this case.

6.7.3 Sample Variance

The second set of simulations was done to measure the RMS error of the power spectrum estimator (6.3) for a range of observation times for individual HERA- and MWA-like baselines. By "observation time," we mean the total contiguous simulated time that is averaged together in the estimator eq. (6.3). The RMS error $s_k(t)$, with different observation times t, as a fraction of power spectrum amplitude, for each $k = |\boldsymbol{k}_{b\tau}|$ mode, is defined as

$$s_k(t) = \frac{\sqrt{\left\langle |\widehat{P}(k; N_t) - P_{\text{theory}}|^2 \right\rangle}}{P_{\text{theory}}} \tag{6.20}$$

where $P_{\text{theory}} = \sigma_0^2 dV_0$ is the input power spectrum amplitude.

The MWA/HERA beam widths used here are set based on the true HERA/MWA beam widths at the lowest frequency of the selected bandpass, where they are widest. This is a conservative approximation, since the covariance is smaller for wider beams, so any error from the lack of beam chromaticity is likely to cause us to underestimate sample variance. The HERA beam width is set to 12.13°, which is the FWHM of an Airy disk pattern for a circular aperture of diameter 14.6 m (the HERA dish diameter) at 100 MHz. Estimating a good effective beam width for the MWA is less straightforward, since each MWA tile is a phased array of 16 dipole antennas and so has a more complicated response than a parabolic dish. We choose a FWHM of 31.4°, extrapolating from the 23° at 137MHz estimated in Neben et al. (2016).

We simulate 100 sky realizations with each beam width and with two baseline configurations. The first is an equilateral triangle of 14.6 m baselines, one oriented East, the others pointing NW and NE. The second configuration is three long baselines: (136,143) points East and is 102.2 m, (2,136) points NW and is 139.2 m, and (2,143) points NE and is 115.9 m. These baselines were selected from the current 65 element layout of HERA, and are plotted in Fig. 6.4.

For each baseline, we obtain one delay spectrum per time step by doing an inverse FFT along the frequency axis, taking the absolute square, and applying the cosmological scaling factors as in eq. (2.29). We want to measure the variance of progressively longer averages of these power spectra (i.e., vary N_t in the sum of eq. (6.3)). We consider a set of observation times $\{T\}$, which range from one time sample (11s) to the full day (24 hours).

To measure the statistics of the estimator, we need an ensemble of averaged power spectra. We build this ensemble by taking separate advantage of all parts of the dataset that are independent of each other. For each observation time T, we partition the time array into non-overlapping segments

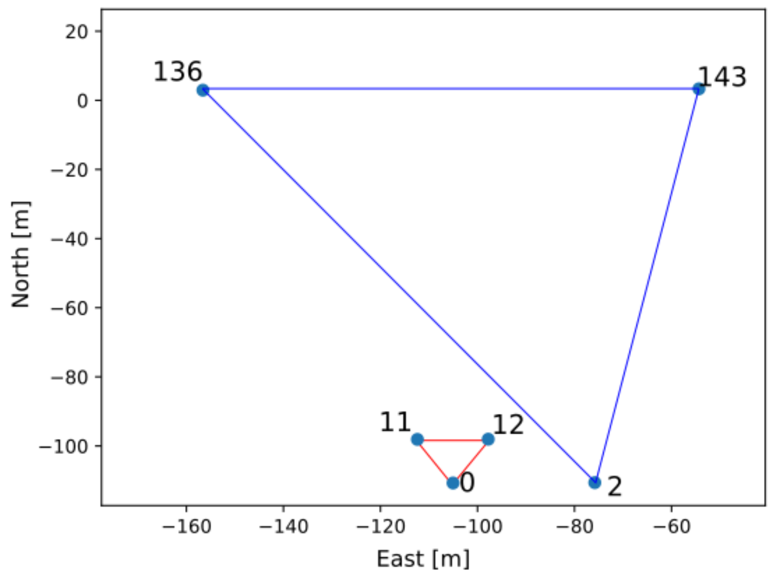


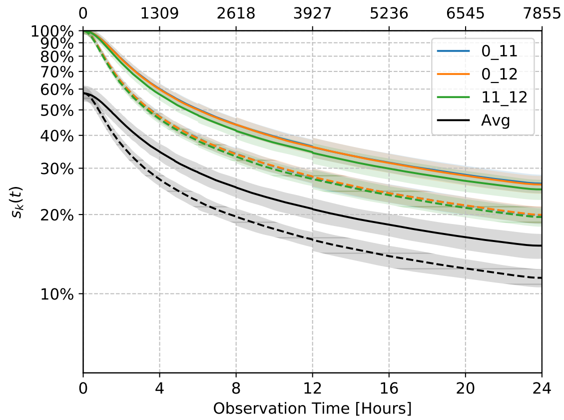
Figure 6.4: The two baseline configurations, taken from the current layout of HERA.

of length T, such that $N_{\text{parts}} = \text{floor}(24\text{hrs}/T)$ where the "floor" function rounds its argument down to the nearest integer. The single time-step delay spectra are averaged within each segment, giving us $N_{\text{parts}} \times N_{\text{skies}}$ independent samples of power spectra averaged at a length T, where $N_{\text{skies}} = 100$ is the number of sky realizations. We then measure the RMS error of each $P(k)$ across this ensemble, giving us the $s_k(t)$ of (6.20). This partitioning method naturally gives more samples at shorter averaging time, since there are more short non-overlapping segments within 24 hours than long. Using multiple sky realizations ensures that there are always at least N_{skies} samples in the ensemble.

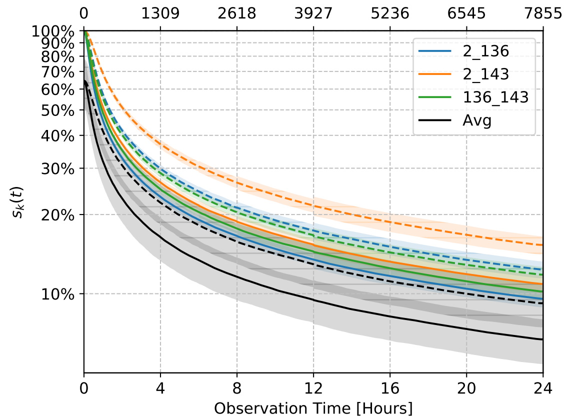
Fig. 6.5 shows the RMS of power spectrum estimates at different time observation times. The lines show the mean of $s_k(t)$ across all k, while the error shading show the standard error calculated across k.[4] The discontinuities at 12 hours, 8 hours, and elsewhere are due to the sudden changes in the number of partitions N_{part} available.

For the short baselines, the RMS drops off more quickly for the HERA beam (dashed) than for the MWA beam (solid). This is consistent with the conclusions of Fig. 6.3, since the 14.6 m baselines are in the regime where increasing beam width increases correlation time, which means a slower drop-off. This behavior is reversed in the long baseline case. The black lines and grey shading are the corresponding results after binning power spectra in $k_\perp$. Averaging together the three baselines

[4]Note that this is effectively the variance of the variance of the power spectrum, which is itself a variance.



(a) Three short baselines.



(b) Three long baselines.

Figure 6.5: RMS error vs. averaging time. (a) shows the three 14.6m baselines individually in color, and the result from averaging across the three baselines in black. Dashed lines are for the HERA-like beams and solid lines are for MWA-like beams. (b) shows the same results for the three long baselines. Shaded regions are 1σ errors, estimated across the delay axis. The labels on top indicate the number of integration times.

in each case drops the error by a factor of $\sim \sqrt{3}$, as expected for independent baselines.

The baseline length appears to have the most significant effect on RMS error, yet even with a 139 m baseline the RMS error doesn't get below 10%. The situation is much worse for the short baselines, which individually don't reach below 19% for HERA beams or below 25% for MWA beams. Single-baseline estimators, particularly for short baselines, cannot be used reliably for precision cosmology.

6.7.4 Full array simulations

Results for individual baselines of specific lengths and orientations motivated us to determine the sample variance limits of a 37 element hexagonal array, i.e., the layout of the MWA Phase II East hexagon. Antenna elements are spaced by 14 m in this configuration. Baselines are grouped by redundancy and one baseline is selected from each redundant group. We ran a series of simulations of 100 sky realizations, calculating the visibilities only for these selected baselines. For the 37 element hexagonal array, there are 66 unique baseline types.

Each baseline samples the same range of $k_{\|}$ for a given $k_{\perp}$. The top half of Fig. 6.6 shows the RMS errors of single-baseline power spectrum estimates for a selection of unique baseline lengths in the array vs. observation time.

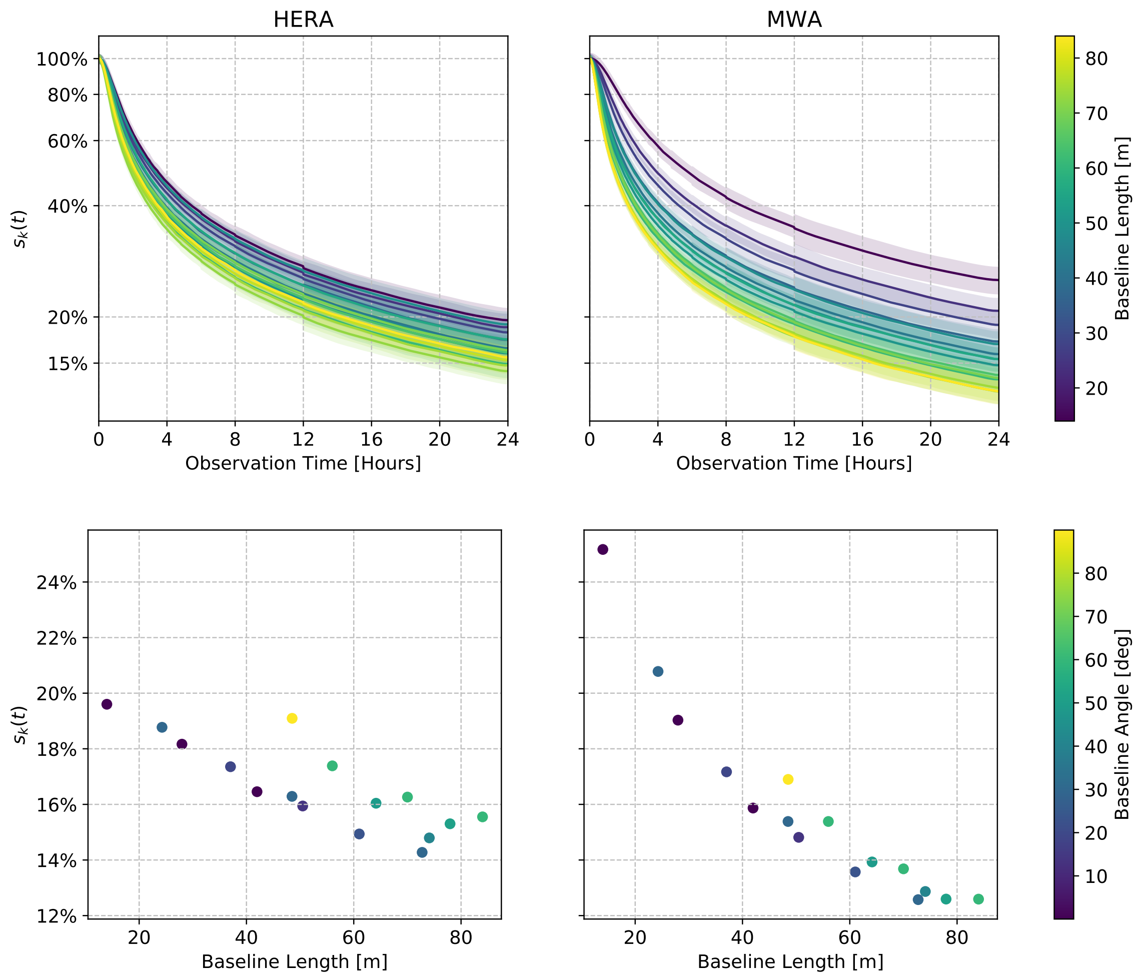


Figure 6.6: (Top) RMS Error vs averaging time for the unique baseline lengths in the 37-element hexagon simulations. (Bottom) The RMS vs. baseline length and angle at 24 hours of observation time.

The spread of curves for the MWA beam, compared with the HERA beam curves, is consistent with the change in behavior observed in Fig. 6.3: for short baselines the narrower HERA beam has a shorter correlation time than the MWA, and vice versa for long baselines. The general trend of longer baselines having faster variance dropoff is clear for both instruments. The bottom of Fig. 6.6 shows the RMS error vs. baseline length at 24 hours of observation time, with points colored according to the absolute value of the angle of the baseline (ie., deviation from the East/West axis). The variety of baseline angles in the array complicates the trend, as baselines with larger North-South components will have longer correlation lengths, and hence drop off more slowly. This trend is more obvious for the HERA beam width, where in the most extreme cases it almost undoes the trend of effect of increasing baseline length.

We can see from Fig. 6.6 that, for individual baselines in the 37-element hexagonal layout, s_k stays above about 13% for MWA beams and above 15% for HERA beams. This is in the optimistic case that all 24 hours of LST can be used. With only 8 hours of data, which is how much time was used by Ali et al. (2015) for PAPER-64, the error for a single baseline of HERA is around 35%, and is upwards of 45% for the MWA.

We can extend eq. (6.3) by binning separate $k_{b,\tau}$ modes, combining estimates from multiple baseline lengths to reduce the variance dramatically. Each $k_{b,\tau}$ is a single delay mode for a single baseline, and coresponds with a point in the cylindrical power spectrum space $(k_\perp, k_\parallel)$. We can define bins in the spherical-averaged $k = \sqrt{k_\parallel^2 + k_\perp^2}$, where $k_\parallel = \tau/Y$ and $k_\perp = b/(\lambda X)$, and average together $\widehat{P}(k_{b,\tau}, N_t; b)$ in each bin.

Choosing a set of k bins is a subtle problem that depends on the expected shape of the power spectrum. For a highly-structured power spectrum, one needs to choose fine k bins to resolve the structure. For a smoother spectrum, one can combine measurements into wider k bins with little loss of information. Choosing wider bins means more measurements are combined, which can reduce sample variance, so there is a trade-off between the precision with which we can constrain the power spectrum shape and the per-bin sample variance. For a flat power spectrum model, like the ones we're using, one could in principle average all k modes with impunity, since there is no shape to constrain.

We choose to bin in logarithmically spaced k bins, the default configuration in `21cmFAST`, which provide a representative range of scales and k space resolution for describing the information in power spectra typical of the EoR (Mesinger et al., 2011). The bin edges are shown as contours in Fig. 6.7, over the power spectrum $P(k_\perp, k_\parallel)$ averaged 24 hours for a single sky realization.

Fig. 6.8 shows the RMS error in each k bin, at different observation times and for the different instruments, using all baselines in the 37 element array for both HERA and MWA-like beams and both 8 and 24 long observations. The variance drops off with increasing k due to the logarithmic

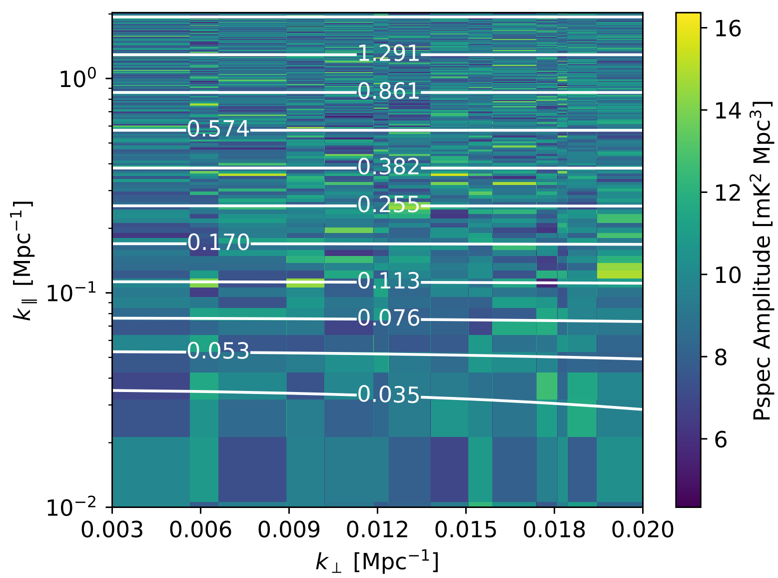


Figure 6.7: The cylindrical power spectrum from a HERA beam simulation with a 37 element hexagonal array, with contours indicating the boundaries of $|k|$ bins used in section 6.7.4. The full range of $k_\parallel$ goes up to about 2 Mpc^{-1}, but the contours get crowded at high $k_\parallel$ on this log-scale and so we only plot up to 0.6 Mpc^{-1}. For most $k_\parallel$, the spherical-binning effectively combines all $k_\perp$ in the array.

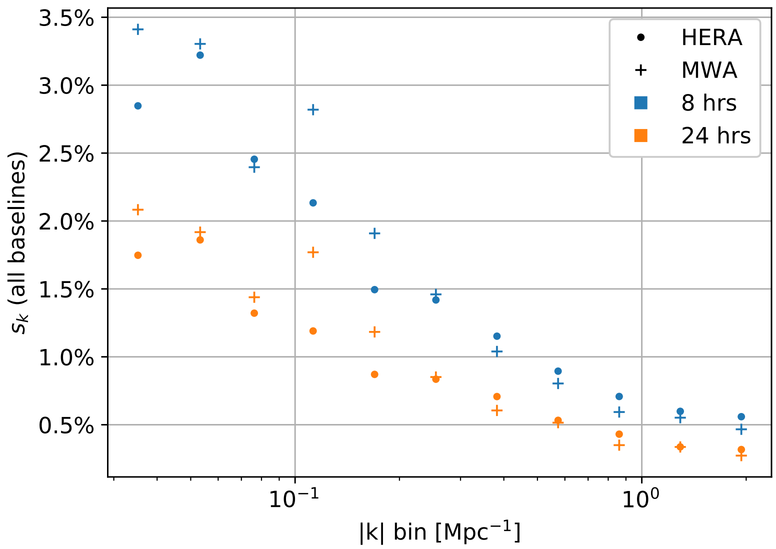


Figure 6.8: The RMS error of the spherically-binned power spectrum in each logarthmically-spaced $|k|$ bin, including baselines of all lengths in the 37 element hexagonal array. Blue points are the result of averaging 8 hours of LST, and orange are for 24 hours of averaging. Crosses are for the MWA-beam simulations, and circles are for the HERA beam simulation. More modes are included in bins with higher $|k|$ since the bin width increases logarithmically, which results in a reduced sample variance compared with smaller bins at lower $|k|$. For a k binning scheme typical of an EoR experiment, the 37 element hexagonal array brings the sample variance error below 3.5% for all bins, and can bring it under a percent for some.

bin sizes, which become wider at high k and therefore include more samples. We note that for all bins considered, the error is under 3.5%, and can be brought under a percent for the highest bins. Therefore, we expect that even with a relatively small number of antennas, a maximally redundant array can place high precision constraints on the power spectrum, as long as all baseline types are used in the measurement.

6.8 Comparing Correlation Time Measures

The correlation time t_c specifies the approximate separation in time of independent samples of the power spectrum, and can in principle be calculated from the baseline length and primary beam model. This allows one to estimate the sample variance error for a given baseline and beam model directly from a formula like eq. (6.15), without running expensive simulations. In this section, we

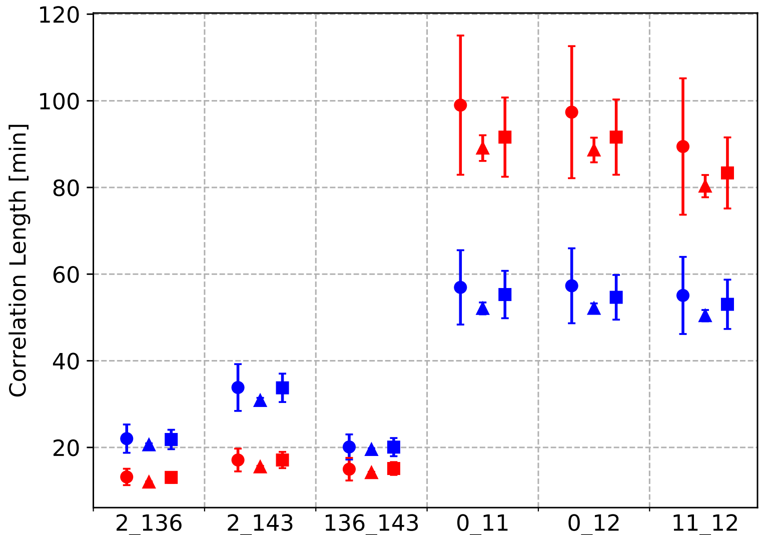


Figure 6.9: Correlation times measured with three different methods, for the six baselines in Fig. 6.4. Red and blue are the results for MWA and HERA beams, respectively. Circles, triangles, and squares are results from the 24 hour RMS measurement, the binned covariance matrices, and from fitting χ^2 distributions to data at 12 hours of averaging, respectively (see section 6.8). All three estimates give consistent results.

show that the correlation time as defined in section 6.7.1 gives the RMS error for longer averaging times, and is consistent with the number of degrees of freedom in a χ^2 distribution, using the six baselines simulated in section 6.7.3. We also show that using the correlation time calculated here fits the RMS error vs. time plots, using eq. (6.15). The results are summarized in Fig. 6.9.

6.8.1 Binned covariance

The first t_c measurement method is to calculate the covariance matrix, bin it in lag, and find the width of the central peak. Since we have N_{skies} simulations in each configuration, we can calculate the covariance matrix separately for each delay mode. This is in contrast to the method in section 6.7.1, where we used the delay axis as an ensemble when computing covariance matrices. We switch to this method now because it is more consistent with the per-k RMS error measurement of section 6.7.3. The results of this method are plotted as triangles in Fig. 6.9. The points and error bars are the means and standard deviation, across k, of the measured correlation times.

We note that there is no technical reason why the FWHM of the covariance should set the correlation time. The correlation beyond the half-maximum distance is not negligible. For this reason, the FWHM is likely underestimating t_c. This explains why the triangles in Fig. 6.9 are a bit lower, though still within the uncertainty, of the results from the other methods.

6.8.2 RMS in the central limit

For a large number of independent samples, the estimator $\widehat{P}_t$ is approximately Gaussian-distributed under the central limit theorem. We therefore expect RMS of the power spectrum estimator should go roughly as $s_k(t) = 1/\sqrt{M(t)}$, where M is the number of independent time samples within the time t, for $t >> t_c$. From M we can estimate t_c as

$$t_c \approx t/M = t\,[s_k(t)]^2 \tag{6.21a}$$

$$\Delta t_c \approx 2ts_k(t)\Delta s_k(t) \tag{6.21b}$$

Equation (6.21b) is the error on the estimate, propagating from the measured error $\Delta s_k(t)$. The squares in Fig. 6.9 show this estimate from the 24 hour averaging time, using the results in Fig. 6.5 with eqs. (6.21a) and (6.21b).

6.8.3 χ^2 distribution fit

The third method, discussed in section 6.4, is to fit a scaled χ^2 distribution to the measured power spectra and treat the fitted degrees of freedom parameter κ as the number of independent samples in the measurement, such that the correlation time is $t_c = t_{\text{avg}}/\kappa$[5] We take the ensemble of $\widehat{P}_N$ measurements for a chosen observation time, and estimate the parameters of the scaled χ^2 distribution separately for each k-mode using a maximum likelihood estimator. The square points on Fig. 6.9 show the mean of these per-k estimates, with error bars given by the standard error across k.

Two factors influence how well this method may be used. For shorter averaging times, the ensemble of $\widehat{P}_N$ measurements is much larger, making fewer samples to fit with the χ^2 distribution. However, as Fig. 6.2 shows, for longer correlation times (e.g. the red curve), there is a risk of underestimating the sample variance in this case. With much longer averaging times, there is the opposite problem. Since the other methods suggest the correlation times to be between 20 and 100 minutes, we look at estimates from 12 hours of averaging. Fig. 6.10 shows histograms of the averaged power spectra for three observation times, along with the fitted χ^2 for that time. With

[5] To address a possible confusion in terminology: In most cases, the phrase "χ^2 fitting" refers to fitting model parameters to data using a chi-squared goodness-of-fit test. Here, we mean we are fitting for the parameters of a χ^2 distribution that describes the statistics of our data.

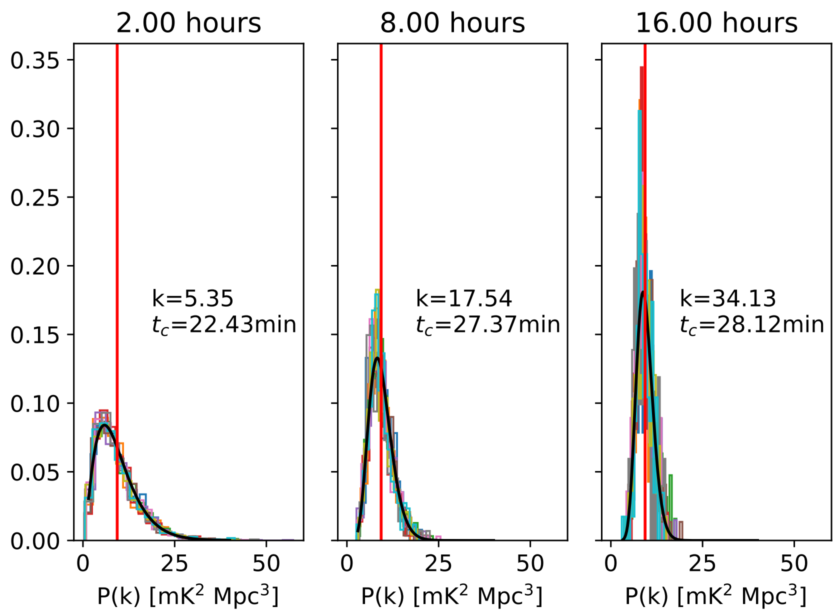


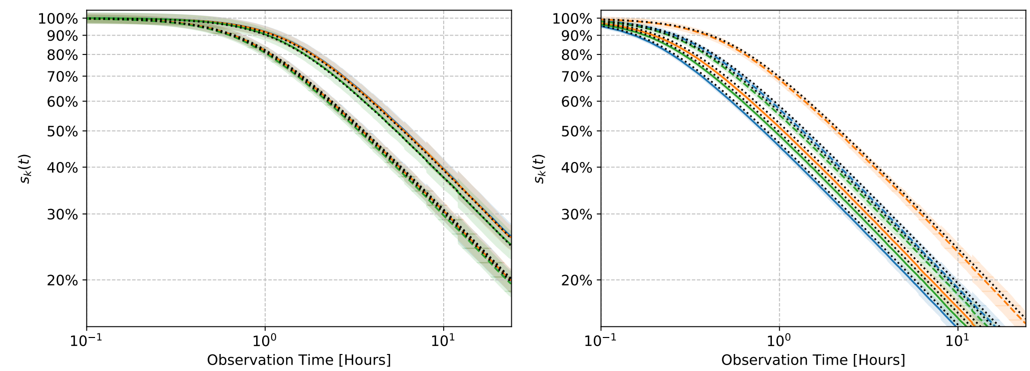
Figure 6.10: Histograms of the power spectra at each k for the three short HERA baselines configuration, at different observation times. The vertical red line is the theoretical power spectrum amplitude. Each k-mode has approximately the same χ^2 distribution, with effective degrees of freedom increasing with increasing observation time. Note that for longer averaging lengths the ensemble is smaller, which causes the P_k measurements to be noisier.

longer averaging times, the degrees of freedom increases and the distributions become less skewed, as expected. We note that there is still a noticeable skew around 8 hours.

6.8.4 Theoretical RMS error

Fig. 6.11 shows the RMS vs. averaging time curves of Fig. 6.5, but with the theoretical sample variances for those baselines indicated with dashed black lines. This theoretical curve is obtained from eq. (6.15) by dividing out by $\gamma_0 = P_{\text{theory}}$ and taking the square root. The $t_c = 2.355 w_c$ used here are the correlation times measured by χ^2 distribution fitting (i.e. the squares in Fig. 6.9).

The good agreement shown in Fig. 6.11 indicates that the form of eq. (6.15) can be used to estimate the single-baseline sample variance at different averaging times, knowing only t_c, in the case that the correlation function is approximately Gaussian in time. The correlation time may be calculated directly from eq. (6.8) or eq. (6.7). It is thus possible to estimate the sample variance of an experimental setup without carrying out simulations. This may prove invaluable for future instrument design.



(a) (b)

Figure 6.11: The same rms curves as Fig. 6.5, shown with the theoretical RMS curves of eq. (6.15) using the correlation times measured from the χ^2 distribution fitting (dashed black lines). The axes are shown in log-log scale to emphasize the flattening at short averaging time.

6.9 Discussion and Conclusions

We have used wide-field instrument simulations with Gaussian beams to estimate the statistical independence of delay-transformed visibilities over time, and calculate the uncertainty per-baseline for the simple delay-spectrum estimator of eq. (6.3). We have further looked further at the case of a 37 element hexagonal array, like the MWA Phase II, and measured the uncertainty of spherically-binned power spectrum estimates. This was done under the assumption that thermal noise and foreground power have been fully mitigated, and so represents a fundamental limit on the uncertainty that this estimator can reach.

Table 6.2 summarizes sample variance estimates in several of the simulated scenarios. The shortest baselines in HERA and the MWA cannot bring the sample variance below 20%, even in the ideal case that all 24 hours of LST are available. Spherically averaging power spectrum estimates across all baselines in the 37 element hexagonal array brings the RMS error under 3.5% for all cases.

As mentioned in section 6.7.4, choosing spherical-k bins for the power spectrum estimator does come with its own difficulties. In the case of a flat power spectrum, the only parameter we have to constrain is the amplitude, which does not depend on k. For a non-flat spectrum, however, binning in k can potentially smooth over small scale features, which will propagate through to constraints on the model parameters. We chose the logarithmic bins of `21cmFAST` as a typical binning used in 21 cm analysis, but the actual choice of k bins will affect the available improvement in sample variance.

	MWA		HERA	
	8 hrs	**24 hrs**	**8 hrs**	**24 hrs**
14.6m	43.86% ± 2.94%	26.22% ± 2.13%	34.01% ± 2.15%	19.89% ± 1.50%
140m	16.55% ± 0.92%	9.57% ± 0.69%	21.23% ± 1.12%	12.37% ± 0.92%
37hex	3.41%	2.04%	3.22%	1.86%

Table 6.2: Results at 8 hours and 24 hours for MWA and HERA beams, for single baselines, and for the spherically-binned power spectra of the 37 element hex. The binning used is discussed in section 6.7.4. The maximum uncertainty across all k bins is shown for the 37hex binned case, so no uncertainty is quoted.

We note also that these results largely relied on the Gaussianity of the tested signal. The EoR signal, especially at late times, is expected to be highly non-Gaussian due to the evolution of structure and appearance of ionized regions in the field. Mondal et al. (2015), and more recently Shaw et al. (2019), have shown that higher-order statistics such as the trispectrum enhance the errors of power spectrum estimates by $50\% - 100\%$ at redshifts up to $z \lesssim 10$. Future simulations with a more realistic EoR model will be able to examine the effect of these non-Gaussianities on realistic power spectrum estimates.

Nonetheless, our results bode well for HERA and the MWA's usefulness for precision cosmology. As suggested by Greig & Mesinger (2015), getting the sample variance under 25% can be enough to constrain EoR parameters within a factor of a few, though the tolerable level of power spectrum uncertainty remains to be determined. As HERA construction continues, improvements in foreground modeling and removal will clean up more of Fourier space, allowing longer baselines to be used. Further, the split-hexagon final layout of HERA, discussed in Dillon & Parsons (2016), will add more non-redundant short baselines to the analysis, providing more independent samples of low $k_\perp$ modes. The MWA phase II also consists of another hexagonal array and a number of randomly-placed tiles, which greatly increases the numbers of available baselines. Most importantly, none of these sample variance uncertainties prevent the possibility of an EoR power spectrum detection, which is the most immediate goal of these instruments.

To summarize our main conclusions:

- Sample variance is an important factor on single baselines, and severely limits the precision of cosmological model constraints that can be made from power spectrum measurements.
- Including measurements from longer baselines can help to mitigate the impact of sample variance, which drops rapidly with increasing baseline length. Combining measurements from multiple baselines by binning across $k_\perp$ will further reduce the sample variance down to tolerable levels, while still avoiding the use of very long baselines.

- For a Gaussian signal, the time-averaged delay spectrum estimator follows a scaled chi-squared distribution. For small effective degrees of freedom this distribution is skewed, which may further bias parameter constraints by underestimating the error toward larger P_k values. Analyses attempting to make cosmological parameter constraints should account for this potential skew.

- The sample variance of a Gaussian or near-Gaussian signal may be estimated from the time-correlation of visibilities. This correlation does not scale simply with beam width and baseline length, but can be evaluated analytically or numerically from knowledge of the primary beam shape and baseline vector.

- The time-correlation of delay-transformed visibilities depends on the primary beam width and the baseline length, but can behave in unexpected ways. The correlation time of short baselines will increase with increasing beam width for narrow beams, and decrease for wider beams.

Chapter 7

Closing Remarks

When working at an untested level of sensitivity with such complicated instruments, simulations provide a means of testing different effects in isolation to root out the causes of unexplained systematic errors. Simulations of signals provide a means of validating analysis methods. With a sufficiently detailed source model, simulations can be also aid calibration, be used for analysis pipeline validation, and aid in foreground subtraction.

The simulators described in chapter 3 have served as validation tools for the HERA collaboration. Flat-spectrum EoR simulations from `healvis` have been used to test the power spectrum estimator pipeline of HERA for signal loss, confirming that the expected power spectrum is returned and behaves appropriately. Foreground simulations (of the GSM) with `healvis` have also been used as a starting point for modeling mutual antenna coupling and reflections along the analog signal chain of HERA, for a study on how these effects could be removed (Kern et al., 2019c,d).

At the center of all this validation work is `pyuvsim`, which is itself used to validate the other simulators in the collaboration. The slow, meticulous development method in `pyuvsim` ensures that it covers the broadest possible range of cases while maintaining its trustworthiness. Comparisons between `healvis` and `pyuvsim` revealed slight errors in `healvis`'s azimuth calculation, which, while not significant to the results presented here, could be important in the future. The important takeaway is that the collaboration now has a simulation tool that is *known* to be correct, built from the ground up with each component carefully tested, such that errors in other tools may be understood and mitigated.

Going forward, `pyuvsim` continues to see improvements both in performance and usefulness. The diffuse simulation methods developed for `healvis` have been written into `pyuvsim` now, so now it can simulate diffuse models by treating them as collections of point sources. The analytic solutions presented in chapter 4 are being implemented into the unit testing framework, ensuring that any modifications to the diffuse simulation methods maintain comparable or better precision and accuracy in their results. With this diffuse simulation capability, we have recently been able to simulate an EoR model with `pyuvsim` and recover the correct power spectrum.

The biggest challenges now lie in the primary beam and source modeling. Most `pyuvsim` results have been produced with analytic beam models, but it has the capability of simulating data-defined models like the CST-simulated beams used in `healvis` in chapter 5. There are two major problems with these models, however – the beams are relatively large in memory, which creates memory bloat as they're copied around, and the beam data must be interpolated in both frequency and angle. The interpolation has been shown to be inconsistent and introduce spectral structure, which undermines the usefulness of these models. Further improvements in beam interpolation, both for computational performance and accuracy.

Beyond beam modeling, there is a lack of effective software for handling diverse source models. Point source catalogs have a long history, and pixellized diffuse maps such as HEALPix have become important for CMB work, but there is no generalized software container for handling point, extended, and diffuse sources altogether. As `pyuvsim` aims to be a general-purpose simulator, it needs an abstract interface for getting, simply, the Stokes parameters from an arbitrary point on the sky. To this end, a new software package called `pyradiosky` is being developed out of source modeling methods previously written in `pyuvsim`. This new package aims to generalize the idea of a "sky model" to accommodate all the ways in which source data is represented. As `pyuvsim` is revised to work with `pyradiosky`, we can ensure that source models handled by `pyradiosky` can be immediately simulated with `pyuvsim`.

Simulated data has been used here to study two effects of the instrument on measurements – the spillover of diffuse foreground power into the EoR window, and the correlation in time of visibility measurements (as it affects the statistics of the EoR signal). More sophisticated foreground and EoR models can reveal much more. As discussed in chapter 6, simulating flat-spectrum EoR leaves open the question of how correlations in visibilities affects the recovered power spectrum shape – Spherically-binning power spectrum estimates can reduce sample variance, but at the cost of losing precision in k. How this translates into uncertainties on EoR or cosmological model constraints is a potentially interesting question. Future work will require developing more realistic EoR models in the large volume and spherical geometry seen by instrument simulators. `pyuvsim`'s support for polarized sources allows us to study another important factor – Faraday rotation imparted on incoming radiation by its passage through an ionized medium can cause frequency-dependent polarization, potentially introducing spectral structure that can be confused with an EoR signal.

In the not-too-distant future, Cosmic Dawn experiments will need to go above the Earth's ionosphere to view the earlier stages of the Dark Ages, since the ionosphere is effectively opaque below 30 MHz, ($z \gtrsim 46$). Experiments placed on the far side of the moon, like the proposed Cosmic

Dawn Mapper (CDM), or in orbit around the moon, like the Dark Ages Radio Explorer (DARE), cannot be upgraded and commissioned *in situ*, and will need to be extensively tested well before any launch is done. Simulations will be needed to optimize design choices to provide the most scientifically useful instrument with the least launch cost.

Appendix A

Simulator Comparison[1]

The design approach used for `pyuvsim` prioritizes precision, design clarity, and flexibility over speed and efficiency. As such, it has relatively slow speed and large memory usage compared with several other simulators. Deciding which tool to use will ultimately depend on what factors are important to the user, and the computing resources available. `pyuvsim` is most useful at present as a point of reference for validating methods of other simulators We recommend pyuvsim as a reference tool to verify the results of other simulation tools.

In this appendix, we compare `pyuvsim`'s features and performance to some other visibility simulators to highlight key differences.

A.1 Precision Radio Interferometry Simulator (PRISim)

A Python package for generating full-sky simulations.
Advantages:

1. Faster due to vectorized calculations.
2. Well-tested through years of use.
3. Support for a range of sky models including point sources, diffuse emission, and spectral cubes.
4. Support for tracking as well as transit telescopes.
5. Parallelized over baselines, frequencies, or sky model/catalog.

[1] Much of the content of this appendix has been published in the Journal of Open Source Software (JOSS) (Lanman et al., 2019b) and on the `pyuvsim` ReadTheDocs page, `https://pyuvsim.readthedocs.io`.

6. Has been tested using imaging with CASA.

Disadvantages:

1. MPI overhead is large if time axis in simulation is long.
2. Limited to accurately simulating the total intensity (Stokes I) radio interferometer measurement equation, which is accurate if the polarized sky emission is negligible and the off-diagonal terms in the beam pattern (e.g. XY or YX for linear polarization) are negligible.
3. Does not support non-identical antenna patterns.

A.2 `healvis`

For simulating visibilities from unpolarized HEALPix maps. This is a new simulator discussed in detail in section 3.2.
Advantages:

1. More efficient, as it only handles unpolarized models with a single pair of feeds.
2. Verified against analytic visibility calculations, as pyuvsim does, for both point[2] and diffuse models (chapter 4).
3. Interfaces with `pyuvdata` module and `astropy`.
4. Capable of using Gaussian Process Regression for smoothing interpolated primary beam models in frequency (see section 5.3.1)

Disadvantages:

1. Parallelized in Python's `multiprocessing` module, which does not support jobs spread across multiple nodes (SMD architecture).
2. Does not support point source catalogs.
3. Does not support polarized source models.

[2] `healvis` technically does not support point sources, but for the purposes of the tests a point source was made as a single nonzero pixel on a HEALPix map. The result is the same under the pixellization scheme of `healvis`

A.3 Fast Holographic Deconvolution (FHD)

FHD is a data calibration and modeling framework written in IDL by researchers by researchers at the University of Washington, originally developed for analyzing EoR data from the Murchison Widefield Array (MWA) in Western Australia. It calculates sky models by convolving a high-resolution psf with the exact DFT of point sources.
Advantages:

1. Optimized for smooth-spectrum sources.
2. Well-tested through years of use.
3. Support for diffuse foreground models.
4. Support for discrete pointings.
5. Supports non-identical antenna patterns (can be memory intensive if there are many different patterns).
6. Efficiently uses IDL's native parallelization.

Disadvantages:

1. Written in the proprietary IDL language.
2. The discrete convolution in Fourier Space can introduce aliasing and ringing artifacts at a level relevant to 21 cm cosmology (see Fig. A.1).
3. Memory usage scales with the size of the uv-plane. It is very fast for compact 21 cm cosmology arrays, but for higher resolution arrays with sparse uv coverage the memory scaling can be poor.
4. Primary beam motion on the sky is limited to the "snapshot" size, not the time step size.
5. Much slower for simulating complex spectral structure.

The figure above shows a comparison of a `pyuvsim` simulation of several point sources with a corresponding FHD simulation. The visibilities have been delay-transformed with a Blackman-Harris window function, squared, and then averaged in time. For this smooth-spectrum foreground, the power at high delays should be very low. The FHD simulation (shown in orange) introduces excess power at high delays, which can be a significant problem for 21 cm cosmological research.

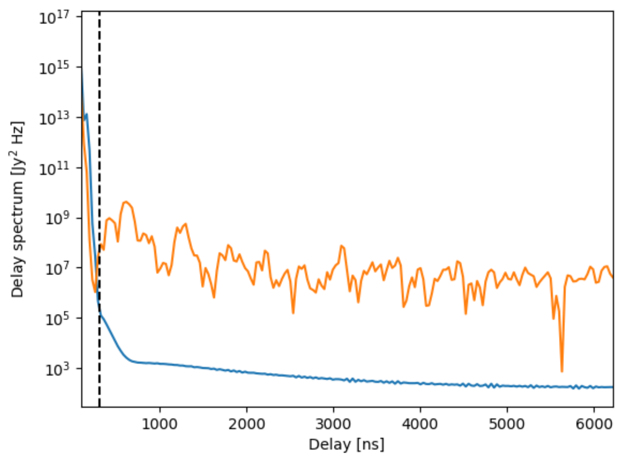


Figure A.1: Comparison of delay transformed visibilities from `pyuvsim` and `FHD`, showing excess structure at high delays in the FHD simulation. The high delay structure should not be there.

A.4 Common Astronomy Software Applications (CASA)

CASA is a set of tools published by the National Radio Astronomical Observatory (NRAO), and has tools available for visibility simulation. Fully parallelized by partitioning Measurement Set (MS) files into smaller tasks and distributed with MPI across any number of available processing units. Non-trivial parallelization also available via OpenMP for shared memory computations on a single node.

Advantages:

1. Uses compiled C code internally, which is faster.
2. Established in the field already and very well-documented.
3. OpenMP utilizes shared memory on a single node if the calculation an be decomposed efficiently. MPI for all other parallel processing needs.
4. Support for a source component lists and FITS image source models.

Disadvantages:

1. Limited support for user-defined primary beam models.
2. Internal UVW rotation is known to be incorrect, affecting coherence far from the phase center (CASA helpdesk ticket 2291, listed as closed but apparently not fixed).

3. In its default (and fastest) mode of operation, point sources are gridded to pixel locations so an FFT can be performed. This pixel-scale imprecision can introduce point source subtraction errors that are significant to 21 cm cosmology experiments (Trott et al., 2012).

4. Full direction-dependent Jones matrices can only be simulated if the beam times sky model calculation is carried out in separate software (Jagannathan et al., 2017).

5. Does not support non-identical antenna beam patterns.

Appendix B

Special Functions

Bessel Function

The Bessel function of the first kind $J_\alpha(z)$ is a solution to Bessel's differential equation.

$$z^2 \frac{d^2 f}{dz^2} + z\frac{df}{dz} + (z^2 - \alpha^2)f = 0$$

has solution

$$f(z) = J_\alpha(z)$$

J_α may be expressed as an integral:

$$J_\alpha(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ix(\sin t - nt)} dt$$

The modified Bessel functions $I_\alpha(z)$ are related to $J_\alpha(z)$ by

$$I_\alpha(z) = i^{-\alpha} J_\alpha(iz).$$

Hypergeometric Function

The generalized hypergeometric function is defined as [1]

$$ {}_pF_q(a_1, \cdots, a_p; b_1, \cdots, b_q; z) = \sum_{n=0}^{\infty} \frac{(a_1)_n (a_2)_n \cdots (a_p)_n}{(b_1)_n (b_2)_n \cdots (b_q)_n} \frac{z^n}{n!}, \tag{B.1}$$

[1] http://dlmf.nist.gov/15

with $(a)_n$ the Pochhammer symbol or "rising factorial":

$$(a)_n = a(a+1)(a+2)\cdots(a+n-1) = \frac{\Gamma(a+n)}{\Gamma(a)} \tag{B.2}$$

The expansion converges under certain conditions:

1. If $p \leq q$, the series converges for all z.

2. If a_i is zero or negative for one or more of the a's, then the $(a_i)_k$ is zero for all k's large than some k, so the series terminates and the function is a polynomial.

3. For $p \geq q+1$, the series may be redefined in terms of contour integration or analytic continuation. None of the examples considered here fall into this regime.

In this text, three forms of hypergeometric function have come up – ${}_0F_1$, ${}_1F_1$, and ${}_1F_2$. All three of these fall into the first category, and so are convergent.

The case of $p = 0$ and $q = 1$ are called the *confluent hypergeometric limit functions*, and are related to Bessel functions of the first kind.

$${}_0F_1\left(;\alpha+1;-\frac{x^2}{4}\right) = \Gamma(\alpha+1)\left(\frac{x}{2}\right)^{-\alpha} J_\alpha(x) \tag{B.3}$$

With $p = 1$ and $q = 1$ we get the *confluent hypergeometric functions of the first kind*, and is a solution to Kummer's differential equation:

$$z\frac{d^2f}{dz^2} + (b-z)\frac{df}{dz} - af = 0$$

such that

$$f(z) = {}_1F_1(a;b;z)$$

Sine Integral

The sine integral function is defined as

$$\mathrm{Si}(x) \equiv \int_0^x \frac{\sin t}{t}dt \tag{B.4}$$

Appendix C

Flat-Spectrum EoR Model

Here I present a quick proof that the model described in section 6.5.1 results in a flat power spectrum. We have a set of spherical maps (a "shell") consisting of pixels at angular positions (θ_i, ϕ_i) with uniform size ω, where $i = 0, \ldots, N_{\mathrm{pix}}$. If this is a HEALPix shell with resolution set by the Nside parameter, then $N_{\mathrm{pix}} = 12 \times \mathrm{Nside}^2$. Each map j corresponds with a distance r_j separated by $\delta r_j = r_j + 1 - r_j$. Each pixel/distance combination gives a single *voxel* $V_{ij} = (\theta_i, \phi_i, r_j)$ with volume $dV_{ij} = r_j^2 \omega \delta r_j$.

Let $p = (i, j)$ denote a single voxel index. We insert a brightness $T(V_p)$ into each voxel according to a Gaussian distribution with variance

$$\sigma_p^2 = \sigma_0^2 dV_0 / dV_p \tag{C.1}$$

We're scaling the variance this way in anticipation of the result. We can turn this set of discrete voxel values into a continuous quantity by writing a sum of delta functions

$$T(\boldsymbol{x}) = \sum_p \delta_D^3(\boldsymbol{x} - \boldsymbol{v}_p) G(0, \sigma_p^2) dV_p \tag{C.2}$$

$G(0, \sigma_p^2)$ is a Gaussian random variable of, mean 0 and variance σ_p^2, which has units of K. The vector $\boldsymbol{v}_p$ specifies the center of voxel p. The voxel volume dV_p is included because the delta function has units of inverse volume. By writing the voxel brightnesses in this form, it is enforced that there are no spatial correlations among voxel brightnesses, since G has no dependence on position.

The Fourier transform of T is

$$\widetilde{T}(\boldsymbol{k}) = \int d^3x \, T(\boldsymbol{x}) e^{-2\pi i \boldsymbol{k} \cdot \boldsymbol{x}}$$

From this, we get a power spectrum estimator.

$$P(\boldsymbol{k}) = \frac{1}{V} \int d^3x \, d^3x' \, \langle T^*(\boldsymbol{x}') T(\boldsymbol{x}) \rangle \, e^{-2\pi i \boldsymbol{k} \cdot (\boldsymbol{x} - \boldsymbol{x}')} \tag{C.3}$$

where $\langle\rangle$ is an ensemble average and V is the total volume $V = \sum_p dV_p$.

Inserting (C.2) into (C.3),

$$P(\boldsymbol{k}) = \frac{1}{V}\int d^3x\, d^3x' \sum_{p,p'} \delta_D^3(\boldsymbol{x} - \boldsymbol{v}_p)\delta_D^3(\boldsymbol{x'} - \boldsymbol{v}_{p'}) \left\langle G(0,\sigma_p^2)G(0,\sigma_{p'}^2)\right\rangle dV_p dV_{p'} e^{-2\pi i \boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}')}$$

$$= \frac{1}{V}\sum_{p,p'}\int \delta_D^3(\boldsymbol{x} - \boldsymbol{v}_p)\delta_D^3(\boldsymbol{x'} - \boldsymbol{v}_{p'}) \left\langle G(0,\sigma_p^2)G(0,\sigma_{p'}^2)\right\rangle dV_p dV_{p'} e^{-2\pi i \boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}')} d^3x\, d^3x'$$

$$= \frac{1}{V}\sum_{p,p'} \left\langle G(0,\sigma_p^2)G(0,\sigma_{p'}^2)\right\rangle dV_p dV_{p'} e^{-2\pi i \boldsymbol{k}\cdot(\boldsymbol{v}_p-\boldsymbol{v}_{p'})}\delta_{p'}^{p} \tag{C.4c}$$

$$= \frac{1}{V}\sum_{p} \sigma_0^2 \frac{dV_0}{dV_p}(dV_p)^2 \tag{C.4d}$$

$$= \sigma_0^2 dV_0 \sum_p \frac{dV_p}{V} = \sigma_0^2 dV_0 \tag{C.4e}$$

In (C.4c) I've taken the integral over the delta functions, which sets $\boldsymbol{x}' = \boldsymbol{v}_{p'}$ and $\boldsymbol{x} = \boldsymbol{v}_p$ and goes to zero unless $\boldsymbol{v}_p = \boldsymbol{v}_{p'}$. This is conveyed in the Kronecker delta $\delta_{p'}^{p}$. With (C.4d) I've applied the Kronecker delta in the sum, which leaves us with a single sum over p. Since $p \to p'$, the average of the product of the two random variables goes to σ_p^2, which I've substituted for the scaled quantity defined in (C.1).

In summary, we end up with a power spectrum that is simply $P(\boldsymbol{k}) = \sigma_0^2 dV_0$, where dV_0 is a voxel volume.

Bibliography

Ali Z. S., et al., 2015, The Astrophysical Journal, 809, 61

Asad K. M. B., et al., 2015, Monthly Notices of the Royal Astronomical Society, 451, 3709

Asad K. M. B., et al., 2016, Monthly Notices of the Royal Astronomical Society, 462, 4482

Asad K. M. B., Koopmans L. V. E., Jeli V., de Bruyn A. G., Pandey V. N., Gehlot B. K., 2018, Monthly Notices of the Royal Astronomical Society, 476, 3051

Astropy Collaboration et al., 2013, A&A, 558, A33

Barkana R., Loeb A., 2005, The Astrophysical Journal Letters, 624, L65

Barry N., Hazelton B., Sullivan I., Morales M. F., Pober J. C., 2016, Monthly Notices of the Royal Astronomical Society, 461, 3135

Barry N., et al., 2019, arXiv:1909.00561 [astro-ph]

Beardsley A. P., et al., 2013, Monthly Notices of the Royal Astronomical Society: Letters, 429, L5

Beardsley A. P., et al., 2016, arXiv:1608.06281 [astro-ph]

Becker R. H., et al., 2001, The Astronomical Journal, 122, 2850

Becker G. D., Bolton J. S., Lidz A., 2015, arXiv:1510.03368 [astro-ph]

Bernardi G., et al., 2010, Astronomy & Astrophysics, 522, A67

Born M., et al., 1999, Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light, 7 edn. Cambridge University Press, doi:10.1017/CBO9781139644181

Bowman J. D., et al., 2013, Publications of the Astronomical Society of Australia, 30

Bradley R. F., Tauscher K., Rapetti D., Burns J. O., 2019, The Astrophysical Journal, 874, 153

Carroll P. A., et al., 2016, Monthly Notices of the Royal Astronomical Society, 461, 4151

Chapman E., et al., 2015, arXiv:1501.04429 [astro-ph]

Chapman E., Zaroubi S., Abdalla F. B., Dulwich F., Jeli V., Mort B., 2016, Monthly Notices of the Royal Astronomical Society, 458, 2928

Cheng C., et al., 2018, The Astrophysical Journal, 868, 26

Clark B. G., 1999, in Taylor G. B., Carilli C. L., Perley R. A., eds, , Vol. 180, Synthesis Imaging in Radio Astronomy II. `https://ui.adsabs.harvard.edu/abs/1999ASPC..180.....T/abstract`

Cornwell T. J., 2008, IEEE Journal of Selected Topics in Signal Processing, 2, 793

Dalcín L., Paz R., Storti M., 2005, Journal of Parallel and Distributed Computing, 65, 1108

Datta A., Bowman J. D., Carilli C. L., 2010, The Astrophysical Journal, 724, 526

Datta A., Bradley R., Burns J. O., Harker G., Komjathy A., Lazio T. J. W., 2016, The Astrophysical Journal, 831, 6

DeBoer D. R., et al., 2017, Publications of the Astronomical Society of the Pacific, 129, 045001

Dillon J. S., Parsons A. R., 2016, arXiv:1602.06259 [astro-ph]

Dillon J. S., et al., 2015, Physical Review D, 91

Dodelson S., 2003, Modern Cosmology. Academic Press

Fagnoni N., Acedo E. D. L., 2016, in 2016 International Conference on Electromagnetics in Advanced Applications (ICEAA). pp 629–632, doi:10.1109/ICEAA.2016.7731474

Fagnoni N., et al., 2019, arXiv:1908.02383 [astro-ph]

Fan X., et al., 2006, The Astronomical Journal, 132, 117

Field G. B., 1959, The Astrophysical Journal, 129, 536

Franzen T. M. O., et al., 2016, Monthly Notices of the Royal Astronomical Society, 459, 3314

Furlanetto S. R., McQuinn M., Hernquist L., 2006a, Monthly Notices of the Royal Astronomical Society, 365, 115

Furlanetto S., Oh S. P., Briggs F., 2006b, Physics Reports, 433, 181

George E. M., et al., 2015, The Astrophysical Journal, 799, 177

Gillet N., Mesinger A., Greig B., Liu A., Ucci G., 2019, Monthly Notices of the Royal Astronomical Society

Gorski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, The Astrophysical Journal, 622, 759

Greig B., Mesinger A., 2015, Monthly Notices of the Royal Astronomical Society, 449, 4246

Greig B., Mesinger A., Haiman Z., Simcoe R. A., 2016, Monthly Notices of the Royal Astronomical Society, p. stw3351

Gropp W., Lusk E., Doss N., Skjellum A., 1996, Parallel Computing, 22, 789

Gunn J. E., Peterson B. A., 1965, The Astrophysical Journal, 142, 1633

Hamaker J. P., Bregman J. D., Sault R. J., 1996, Astronomy and Astrophysics Supplement Series, 117, 137

Harris F. J., 1978, Proceedings of the IEEE, 66, 51

Hazelton B. J., Morales M. F., Sullivan I. S., 2013, The Astrophysical Journal, 770, 156

Hazelton B. J., Jacobs D. C., Pober J. C., Beardsley A. P., 2017, The Journal of Open Source Software, 2, 140

Hilborn R. C., 1982, American Journal of Physics, 50, 982

Hills R., Kulkarni G., Meerburg P. D., Puchwein E., 2018, Nature, 564, E32

Hoefler T., Dinan J., Thakur R., Barrett B., Balaji P., Gropp W., Underwood K., 2015, ACM Transactions on Parallel Computing, 2, 1

Howlett C., Lewis A., Hall A., Challinor A., 2012, Journal of Cosmology and Astroparticle Physics, 2012, 027

Hu W., Sugiyama N., 1995, The Astrophysical Journal, 444, 489

Jagannathan P., Bhatnagar S., Rau U., Taylor A. R., 2017, The Astronomical Journal, 154, 56

Jelić V., et al., 2008, Monthly Notices of the Royal Astronomical Society, 389, 1319

Kaplan G., 2005, USNO Circular No. 179, The IAU Resolutions on Astronomical Reference Systems, Time Scales, and Earth Rotation Models, Explanation and Implementation. US Naval Observatory, Washington

Kern N. S., et al., 2019a, arXiv:1910.12943 [astro-ph]

Kern N. S., Parsons A. R., Dillon J. S., Lanman A. E., 2019b, arXiv e-prints,

Kern N. S., Parsons A. R., Dillon J. S., Lanman A. E., 2019c, arXiv:1909.11732 [astro-ph]

Kern N. S., et al., 2019d, arXiv:1909.11733 [astro-ph]

Kolopanis M., et al., 2019, arXiv:1909.02085 [astro-ph]

Lanman A. E., Pober J. C., 2019, Monthly Notices of the Royal Astronomical Society, 487, 5840

Lanman A. E., Pober J. C., Kern N. S., Acedo E. d. L., DeBoer D. R., Fagnoni N., 2019a, arXiv:1910.10573 [astro-ph]

Lanman A., Hazelton B., Jacobs D., Kolopanis M., Pober J., Aguirre J., Thyagarajan N., 2019b, Journal of Open Source Software, 4, 1234

Lewis A., Challinor A., Lasenby A., 2000, The Astrophysical Journal, 538, 473

Lidz A., Zahn O., McQuinn M., Zaldarriaga M., Hernquist L., 2008, The Astrophysical Journal, 680, 962

Liu A., Shaw J. R., 2019, arXiv:1907.08211 [astro-ph]

Liu A., Tegmark M., 2012, Monthly Notices of the Royal Astronomical Society, 419, 3491

Liu A., Tegmark M., Morrison S., Lutomirski A., Zaldarriaga M., 2010, Monthly Notices of the Royal Astronomical Society, 408, 1029

Liu A., Parsons A. R., Trott C. M., 2014,] 10.1103/PhysRevD.90.023018

Marowka A., 2018, in Wyrzykowski R., Dongarra J., Deelman E., Karczewski K., eds, Parallel Processing and Applied Mathematics. Springer International Publishing, Cham, pp 203–212

Martinot Z. E., Aguirre J. E., Kohn S. A., Washington I. Q., 2018, arXiv:1808.05968 [astro-ph]

Mathai A. M., Provost S. B., 1992, Quadratic forms in random variables: theory and applications. Dekker

Matteo T. D., Perna R., Abel T., Rees M. J., 2002, The Astrophysical Journal, 564, 576

McQuinn M., Zahn O., Zaldarriaga M., Hernquist L., Furlanetto S. R., 2006, The Astrophysical Journal, 653, 815

Mellema G., et al., 2013, Experimental Astronomy, 36, 235

Mesinger A., Furlanetto S., Cen R., 2011, Monthly Notices of the Royal Astronomical Society, 411, 955

Mondal R., Bharadwaj S., Majumdar S., Bera A., Acharyya A., 2015, Monthly Notices of the Royal Astronomical Society: Letters, 449, L41

Morales M. F., Hewitt J., 2004, The Astrophysical Journal, 615, 7

Morales M. F., Matejek M., 2009, Monthly Notices of the Royal Astronomical Society, 400, 1814

Morales M. F., Hazelton B., Sullivan I., Beardsley A., 2012, The Astrophysical Journal, 752, 137

Morales M. F., Beardsley A., Pober J., Barry N., Hazelton B., Jacobs D., Sullivan I., 2019, Monthly Notices of the Royal Astronomical Society, 483, 2207

Mortlock D. J., et al., 2011, Nature, 474, 616

Neben A. R., et al., 2016, The Astrophysical Journal, 826, 199

Oh S. P., Mack K. J., 2003, Monthly Notices of the Royal Astronomical Society, 346, 871

Paciga G., et al., 2011, Monthly Notices of the Royal Astronomical Society, 413, 1174

Paciga G., et al., 2013, Monthly Notices of the Royal Astronomical Society, 433, 639

Parsons A. R., Backer D. C., 2009, The Astronomical Journal, 138, 219

Parsons A. R., et al., 2010, The Astronomical Journal, 139, 1468

Parsons A., Pober J., McQuinn M., Jacobs D., Aguirre J., 2012a, The Astrophysical Journal, 753, 81

Parsons A. R., Pober J. C., Aguirre J. E., Carilli C. L., Jacobs D. C., Moore D. F., 2012b, The Astrophysical Journal, 756, 165

Parsons A. R., et al., 2014, The Astrophysical Journal, 788, 106

Patil A. H., et al., 2017, arXiv:1702.08679 [astro-ph]

Patwa A. K., Sethi S., 2019, arXiv e-prints, p. arXiv:1905.05772

Planck Collaboration et al., 2016, A&A, 594, A13

Pober J. C., et al., 2013a, The Astronomical Journal, 145, 65

Pober J. C., et al., 2013b, The Astrophysical Journal Letters, 768, L36

Pober J. C., et al., 2014, The Astrophysical Journal, 782, 66

Pober J. C., et al., 2016, The Astrophysical Journal, 819, 8

Presley M., Liu A., Parsons A., 2015, The Astrophysical Journal, 809, 18

Price-Whelan A. M., et al., 2018, AJ, 156, 123

Pritchard J. R., Loeb A., 2008, Physical Review D, 78

Pritchard J. R., Loeb A., 2012, Reports on Progress in Physics, 75, 086901

Rasmussen C. E., Williams C. K. I., 2006, Gaussian processes for machine learning. MIT Press

Refregier A., 2003, Monthly Notices of the Royal Astronomical Society, 338, 35

Reichardt C. L., 2016, Observing the Epoch of Reionization with the Cosmic Microwave Background. Springer International Publishing, Cham, pp 227–245, doi:10.1007/978-3-319-21957-8˙8, https://doi.org/10.1007/978-3-319-21957-8_8

Schmit C. J., Pritchard J. R., 2017, arXiv:1708.00011 [astro-ph]

Searle S. R., 1971, Linear Models (Wiley Series in Probability and Statistics - Applied Probability and Statistics Section). Wiley

Shaw A. K., Bharadwaj S., Mondal R., 2019, arXiv:1902.08706 [astro-ph]

Sims P. H., Lentati L., Alexander P., Carilli C. L., 2016, Monthly Notices of the Royal Astronomical Society, 462, 3069

Smirnov O. M., 2011, Astronomy & Astrophysics, 527, A106

Smith J. O., 2011, Spectral audio signal processing. W3K publishing

Sullivan I., et al., 2012, The Astrophysical Journal, 759, 17

Thompson A. R., Moran J., Jr G. W. S., 2017, Interferometry and Synthesis in Radio Astronomy, 3 edn. Astronomy and Astrophysics Library, Springer International Publishing, https://www.springer.com/us/book/9783319444291

Thyagarajan N., et al., 2013, The Astrophysical Journal, 776, 6

Thyagarajan N., et al., 2015a, The Astrophysical Journal, 804, 14

Thyagarajan N., et al., 2015b, The Astrophysical Journal, 807, L28

Thyagarajan N., Parsons A. R., DeBoer D. R., Bowman J. D., Ewall-Wice A. M., Neben A. R., Patra N., 2016, The Astrophysical Journal, 825, 9

Trott C. M., 2014, Publications of the Astronomical Society of Australia, 31

Trott C. M., Wayth R. B., Tingay S. J., 2012, The Astrophysical Journal, 757, 101

Vedantham H., Udaya Shankar N., Subrahmanyan R., 2012, The Astrophysical Journal, 745, 176

Wayth R. B., et al., 2018, Publications of the Astronomical Society of Australia, 35

Wouthuysen S., 1952, The Astronomical Journal, 57, 31

Yatawatta S., et al., 2013, Astronomy & Astrophysics, 550, A136

Zaroubi S., 2013, The Epoch of Reionization. Springer Berlin Heidelberg, `http://link.springer.com/chapter/10.1007/978-3-642-32362-1_2`

Zhang Y. G., Liu A., Parsons A. R., 2018, The Astrophysical Journal, 852, 110

Zheng H., et al., 2017, Monthly Notices of the Royal Astronomical Society, 464, 3486