

Robust Foreground Removal for Hydrogen Intensity Mapping (HIM) Experiments

John Podczerwinski
Department of Physics, UW Madison
Madison, WI, USA

Abstract

This paper describes work done to develop robust methods for foreground removal in hydrogen intensity mapping (HIM) experiments. In HIM, one measures the 21cm line from neutral hydrogen (HI), which in turn provides information about the distribution of matter in the universe. However, foreground emission from the milky way and extra-galactic sources must be removed in order to obtain useful results. Methods developed for removing these foregrounds either have not been thoroughly tested on realistic data, or perform poorly when the instrument is not precisely calibrated. The goal of the work here is to develop a method to remove foregrounds that works on realistic, imperfectly calibrated data.

1. Introduction

This paper describes work intended for hydrogen intensity mapping (HIM) experiments. HIM is an example of a broader technique called line intensity mapping (IM). In intensity mapping, one performs low angular-resolution observations of emission (or absorption) from some spectral line. Moreover, observing the same spectral line at multiple redshifts provides one with 3D maps, as redshift is a function of line of sight distance. In HIM, the redshifted 21cm line from neutral hydrogen (HI) is measured. Measurements of this spectral line in turn provide information about the distribution of matter in the universe, up to a bias term.

There are several types of telescopes one can use for HIM, but the most commonly seen is the radio-interferometer. In radio-interferometry, the telescope consists of a collection of smaller antenna elements. Signals from pairs of these antennas are correlated to form “visibilities”, which in turn can be processed to form maps of the sky. In HIM, these antenna elements are typically dish-based, such as the Hydrogen Epoch of Reionization Array (HERA) experiment [1]. Cylinder-based telescopes also exist, such as in the Canadian Hydrogen Intensity Mapping Experiment (CHIME) experiment [2]. Examples of interferometers can be seen in figures 1 and 2.

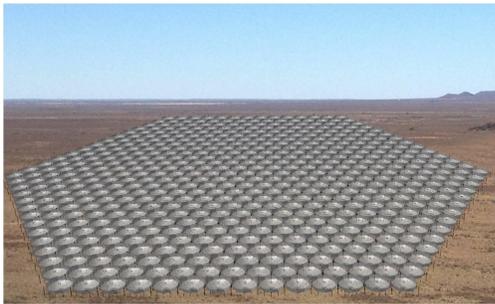


Figure 1: A computer generated image of the HERA experiment, courtesy of UC-Berkeley.



Figure 2: An image of the CHIME experiment, courtesy of The CHIME collaboration.

HIM is of interest to cosmologists because it allows for large volumes of the universe to be measured quickly [3]. Moreover, HIM has the ability to measure epochs of the universe inaccessible to other probes. Measurements of the 21cm line at redshifts $z < 6$ would provide information about the growth of structure, dark energy and inflationary physics. Measurements at redshifts $6 < z < 20$ would provide information about the universe's earliest galaxies and stars. Measurements at $20 < z < 1100$ would probe an era of the universe untainted by astrophysical phenomenon, providing clear information on the physics of inflation.

Although promising, 21cm intensity mapping is a young field, with systematic effects still needing to be understood and controlled. A chief concern of these experiments is foregrounds, which come from both the Milky Way and from extra-galactic sources. In total, these foregrounds are expected to be a factor of 10^5 brighter than the 21cm signal cosmologists wish to measure [3].

A variety of techniques for removing these foregrounds have been proposed. Most of these techniques are “map-based”, removing foregrounds from images produced by the telescope. Examples of such techniques are Generalized Morphological Component Analysis (GMCA) and Independent Component Analysis (ICA), where one assumes that the foregrounds at each frequency channel are a mixture of some set of unphysical sources [4] [5]. Foregrounds can also be removed in a “visibility-based” way, where the map-making step is skipped and foregrounds are removed directly from visibilities. An important visibility-based method is the Karhunen-Loeve transform approach proposed in Shaw et al. (2015) [6].

In the literature, one will usually see that map-based techniques work quite well [5] [4] [7]. However, such papers work on simple simulated maps that have not gone through any kind of measurement or map-making process. One recent paper found these techniques performing relatively poorly when presented with more realistic data [8]. On the other hand, the visibility-based method proposed in Shaw et al. works extremely well, but only when the beam is precisely calibrated [6].

So, the goal of the work done here was to develop a method for removing foregrounds from 21cm intensity mapping that are more robust to calibration errors than the technique proposed by Shaw et al. (2015), but also providing better results than the techniques analyzed by Hothi et al. [8].

2. Background

The work presented in this paper was conducted assuming that the instrument would be a radio-interferometer. In radio interferometry, the telescope produces a set of “visibilities”, which are defined as

$$V_{ij}(\nu) = \langle v_i(\nu)v_j^*(\nu) \rangle_t . \quad (1)$$

In 1, v_i is the signal from antenna i , v_j is the signal from antenna j and the brackets denote a time average. In interferometry, a pair of antennas is referred to as a “baseline”.

One goal of the work presented here was to develop a method that would work on visibilities rather than maps. Visibilities were preferred as they would allow for the map-making step to be skipped. Another reason was that the results shown in Hothi et al. (2020) seemed to indicate that map-based techniques do not perform well on realistic data [8].

In Shaw et al. (2015), it is shown that visibilities can be described using a linear equation

$$\mathbf{v} = \mathbf{B}\mathbf{a} + \mathbf{n} \quad (2)$$

In equation 2, \mathbf{v} is a vector containing visibilities for all baselines at all frequencies of observation. The variable \mathbf{B} is a matrix representing the measurement process of the antennas. In particular, \mathbf{B} contains coefficients given by

$$B_{lm,ij}(\nu) = \int B_{ij}(\theta, \phi, \nu) Y_{lm}^* d\Omega. \quad (3)$$

In this formula, $B_{ij}(\theta, \phi, \nu)$ is the beam pattern for baseline (i,j). The variable \mathbf{a} is a vector containing spherical harmonic coefficients of the sky, and \mathbf{n} is the contribution of thermal noise to the visibilities.

from this, one can show that the covariance of the visibilities is given by

$$\mathbf{C} = \langle \mathbf{v}\mathbf{v}^\dagger \rangle = \mathbf{F} + \mathbf{S} + \mathbf{N}. \quad (4)$$

In equation 4, the signal matrix \mathbf{S} is given by

$$\mathbf{S} = \mathbf{B}\mathbf{C}_{21}\mathbf{B}^\dagger \quad (5)$$

and the foreground covariance \mathbf{F} is given by

$$\mathbf{F} = \mathbf{B}\mathbf{C}_{FG}\mathbf{B}^\dagger. \quad (6)$$

One should also note that the noise covariance is given by \mathbf{N} .

Contributions of foregrounds to the visibilities can then be removed using a Karhunen-Loeve transform (KLT). In the Karhunen-Loeve transform, the matrices \mathbf{S} and \mathbf{F} are simultaneously diagonalized by solving for a matrix \mathbf{P} . One obtains

$$\mathbf{S} \rightarrow \Lambda = \mathbf{P}\mathbf{S}\mathbf{P}^\dagger \quad (7)$$

and

$$\mathbf{F} \rightarrow I = \mathbf{P}\mathbf{F}\mathbf{P}^\dagger. \quad (8)$$

In these equations, Λ is a diagonal matrix and I is the identity. Then, some eigenvalue λ of Λ is the ratio of HI to foreground for that particular eigenmode. The foregrounds can then be cleaned from the visibilities by projecting out eigenvectors containing more foreground than HI.

In order to illustrate the need for precise beam calibration, we present a quick test case. This test case is a simplified version of the CHIME telescope with only 100 baselines operating from 400MHz up to 500MHz. This simplified CHIME is looking up at a sky that contains only HI and unpolarized galactic synchrotron radiation. This telescope also does not have any thermal noise. In figure 3, we show the results of KLT cleaning when the antenna beams are perfectly well understood. This figure shows results in the delay space, which is the Fourier dual to spectral frequency. Figure 4 shows results when the beam-widths of the antennas have been perturbed randomly with a 1 percent standard deviation. These plots illustrate the need for very precise calibration when using the KLT method. the commonly quoted figure is that beam-widths need to be understood to the 0.1 percent level in order for this technique to work, which is quite difficult to achieve in practice when dealing with large numbers of antennas.

So, foreground removal is a solved problem in the case where the beams are precisely calibrated. However, achieving calibration this precise is quite difficult and perhaps not even possible in light of the mutual coupling that occurs in tightly packed arrays [9]. Thus, it would be useful to find an approach to foreground removal that does not require precise beam calibration.

3. Attempting to Model Covariances

As seen in the previous section, foregrounds may easily be removed as long as one has accurate models for the foreground and signal covariance matrices. One observation found early on is that the signal covariance \mathbf{S} is not too sensitive to beam calibration issues. An example of this can be seen in figure 5. In this figure, the green curve uses incorrect models for \mathbf{F} and \mathbf{S} . On the other hand, the red curve uses a correct model for \mathbf{F} and incorrect model for \mathbf{S} .

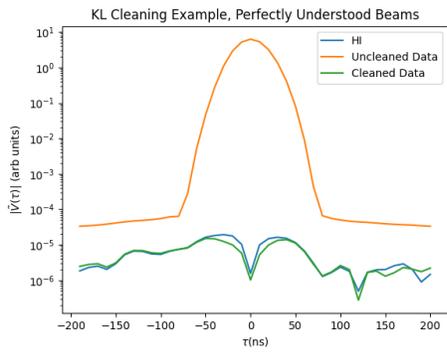


Figure 3: An example of the KLT cleaning method when the beams are perfectly calibrated. Note that the cleaned data and HI signal match up quite well.

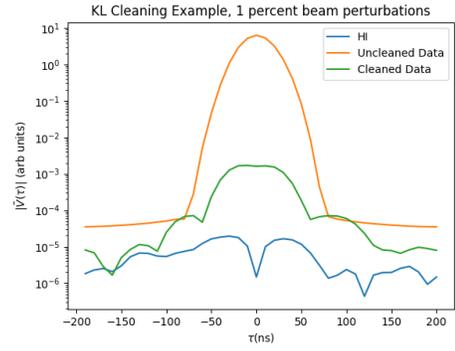


Figure 4: An example of the KLT cleaning method when the beam widths have been perturbed by 1 percent. Note that the cleaned data no longer matches up with the HI signal.

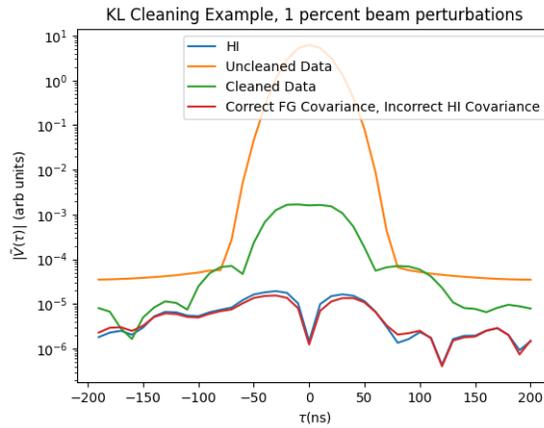


Figure 5: An updated version of figure 4. The red curve uses an incorrect model for \mathbf{S} but a correct model for \mathbf{F} .

This example illustrates the fact that the foreground cleaning process is much more sensitive to errors in \mathbf{F} than in \mathbf{S} . This makes sense, as even proportionally small errors in foregrounds will dwarf the HI signal due to the difference in their magnitudes.

So, the main concern should be finding a way to model \mathbf{F} or obtain it from the data.

3.1. Projection Onto DPSS Functions One useful fact is that the unpolarized part of the foregrounds are spectrally smooth, while the HI signal has quite a bit of structure [3]. This can be seen in figure 3, where the uncleaned data drops off quickly as a function of delay. On the other hand, the polarized components of the foregrounds are less smooth. Thus, if one designs an instrument with minimal polarization leakage, then one can produce a rough estimate for the foregrounds by projecting the data onto spectrally smooth functions.

The spectrally smooth functions used here are discrete prolate spheroidal sequences (DPSS) [10]. Given a data vector with N frequency samples and some window in delay space of width T , one can generate a set of DPSS functions

$$\{U_{i,N,T}(\nu) \text{ for } i = 1, 2, \dots, N\}. \quad (9)$$

These functions will form an orthonormal basis and will be ordered in terms of how much energy they concentrate in the delay window $(-T/2, T/2)$. Thus, if one has a rough estimate of the width of the foregrounds in delay space, then they can generate a DPSS basis in which the foregrounds will be sparsely represented.

Since the DPSS functions form an orthonormal basis, functions can be expressed as linear combinations of them in the usual way

$$C_i = U_i \cdot f \quad (10)$$

$$f = \sum_{i=1}^N C_i U_i. \quad (11)$$

Figure 6 shows the DPSS coefficients of the foregrounds and HI for a simple test case. In this case, the telescope is a pair of dishes pointed to zenith. These dishes have a 6 m diameter and are spaced apart by 8 m along the east-west axis. Observations are conducted between 700MHz and 800MHz , with samples every 1MHz . The foregrounds contain only the unpolarized component. One can see in the figure that foregrounds are constrained to the first 50 coefficients, while the HI is spread out fairly evenly across all coefficients. Thus, we can approximate the foregrounds from our data by keeping only the first 50 coefficients.

Let v denote the visibility, and \bar{v} be the visibility with HI dominated modes removed. Next, assume that the foregrounds are statistically isotropic. In reality they are not isotropic, but this serves as an adequate approximation. So, assuming the foregrounds are isotropic, the foreground covariance can be estimated by taking an average over many LST times.

$$\hat{\mathbf{F}} = \frac{1}{N} \sum_{i=1}^N \bar{v}_i \bar{v}_i^\dagger. \quad (12)$$

In equation 12, the variable i represents an azimuthal pointing direction, and N is the total number of pointings observed as the earth goes through one full rotation.

Then, one can perform the KL cleaning using this approximate foreground covariance ($\hat{\mathbf{F}}$) and the usual signal covariance \mathbf{S} . An example of this cleaning is shown in 7. This test was performed on the simple two dish telescope described earlier in this section.

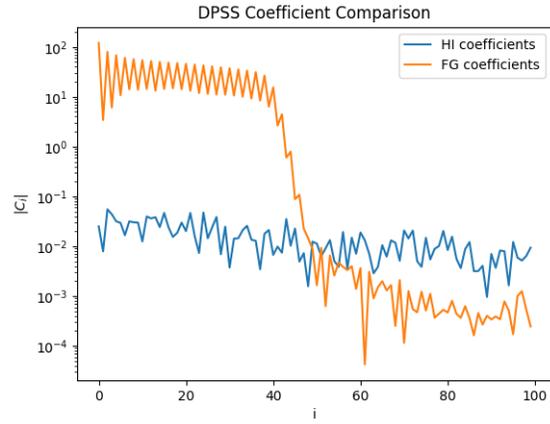


Figure 6: DPSS coefficients for a simple test case.

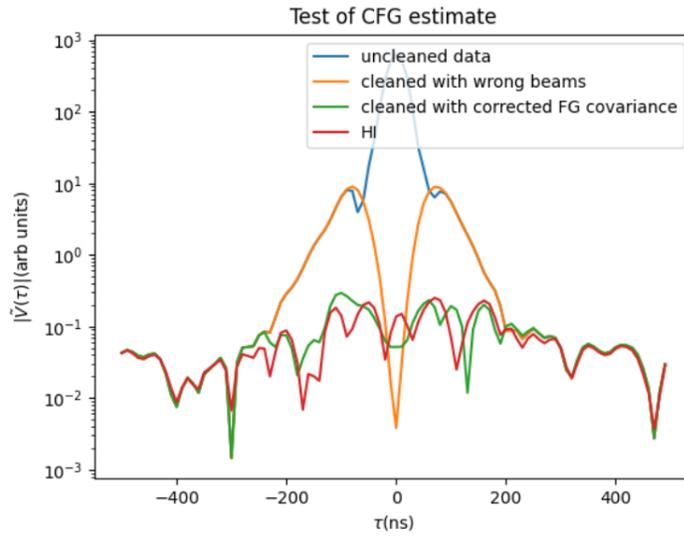


Figure 7: A test of the DPSS foreground estimation described. One can see that the foreground covariance provided by DPSS cleaning (result shown in green) performs much better than the standard one (shown in orange).

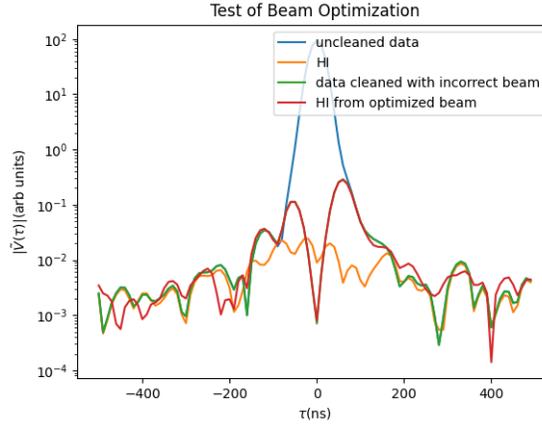


Figure 8: An example of cleaning using a maximum likelihood beam.

3.2. Forming an Estimate for \mathbf{B} Another approach considered was obtaining \mathbf{F} and \mathbf{S} by correcting the estimate for the beam matrix \mathbf{B} . This is found by maximizing the probability of \mathbf{B} given \mathbf{d} , with a prior requiring that \mathbf{B} be spectrally smooth and close to the original estimate for \mathbf{B} . So, the goal is to find a \mathbf{B} maximizing

$$p(\mathbf{B}|\mathbf{d}) = p(\mathbf{d}|\mathbf{B})p(\mathbf{B}). \quad (13)$$

The prior for \mathbf{B} is assumed to be a Gaussian. It is assumed that beam spherical harmonic coefficients have a covariance

$$\langle B_{lm}(\nu)B_{l'm'}(\nu') \rangle = \delta_{ll'}\delta_{mm'}C_{lm}(\nu\nu'). \quad (14)$$

For some spherical harmonic mode (l, m) , it was assumed that the frequency-frequency covariance of the antenna spherical harmonic coefficients would be modeled by a matern kernel

$$C_{lm}(\nu, \nu') = \sigma^2 \frac{2^{1-\eta}}{\Gamma(\eta)} (\sqrt{2\eta} \frac{|\nu - \nu'|}{\rho})^\eta K_\eta(\sqrt{2\eta} \frac{|\nu - \nu'|}{\rho}). \quad (15)$$

In this formula, η is a free parameter which was chosen to be $3/2$. The variable ρ is meant to set the scale of correlation and was chosen to be $100MHz$ to encourage smooth variation with frequency. The variable σ was chosen as an upper bound for how much that particular B_{lm} might vary when the beam is perturbed. The prior on the beam can then be expressed as a product over all l, m .

$$p(\mathbf{B}) \propto \prod_{l,m} e^{-\frac{1}{2}(\mathbf{B}_{lm} - \bar{\mathbf{B}}_{lm})^\dagger \mathbf{C}_{lm}^{-1}(\mathbf{B}_{lm} - \bar{\mathbf{B}}_{lm})}. \quad (16)$$

In this formula, $\bar{\mathbf{B}}_{lm}$ is the unperturbed value of \mathbf{B}_{lm} . The most likely value for \mathbf{B} can then be estimated by minimizing

$$-\text{Log}(p(\mathbf{B}|\mathbf{d})) \propto \frac{1}{2}\mathbf{d}^\dagger \mathbf{C}^{-1}\mathbf{d} + \frac{1}{2}\text{Log}(|\mathbf{C}|) + \frac{1}{2}\sum_{lm}(\mathbf{B}_{lm} - \bar{\mathbf{B}}_{lm})^\dagger \mathbf{C}_{lm}^{-1}(\mathbf{B}_{lm} - \bar{\mathbf{B}}_{lm}). \quad (17)$$

In this case,

$$\mathbf{C} = \mathbf{S} + \mathbf{F} + \mathbf{N}. \quad (18)$$

Readers should recall from equations 5 and 6 that \mathbf{S} and \mathbf{F} are both functions of the beam matrix \mathbf{B} . This method was tried out on the same test case as used in subsection 3.1. The results are shown in figure 8. One can see in the figure that the beam generated by this optimization procedure yielded poor results, performing the same as the incorrect beam at low delays and worse at high delays.

4. Discussion and Conclusion

In order to clean foregrounds from visibilities, one needs to obtain adequately accurate estimates for \mathbf{S} and \mathbf{F} . It was found that the standard approach to estimating \mathbf{S} will still produce reasonably good results in the KLT method. The foreground covariance \mathbf{F} , on the other hand, is much more of a challenge.

One approach is to estimate the foreground contribution to the data by projecting on the DPSS basis and average over many different azimuthal pointings. This approach however, relies on several assumptions. One is that the polarized foregrounds have already been suppressed to a level below the HI. The other is that the instrument has adequately smooth spectral properties to allow for a separation between the HI and foregrounds in this basis. Such conditions are unlikely to be true in practice, and thus this method will likely not be useful.

Another option explored was to correct the model for \mathbf{B} by maximizing the likelihood subject to a prior on \mathbf{B} . This approach yielded poor results and was computationally expensive, meaning it would be difficult to scale to a full-sized experiment. This approach likely did not work for a few reasons, for one the optimization problem presented is non-convex, meaning that one can easily fall into local minima. In addition, the beam may improve the total likelihood while still producing estimates for \mathbf{S} and \mathbf{F} that do not describe the foregrounds or HI accurately. There may still be hope for such a method, however. In Nasirudan et al. (2022), a PCA basis was used to characterize beam perturbations [11]. Such an approach would alleviate the computational burden taken on by the optimization process. Moreover, a more careful model for the beam covariance could improve performance further.

To summarize, it was found that obtaining a good estimate for \mathbf{F} is required to clean foregrounds from data when working in visibility space. Efforts were made this summer to estimate this covariance with minimal knowledge of the antenna beam patterns. Such efforts, however, have produced poor results.

Due to this difficulty, it may be beneficial in the future to try working in map-space rather than visibility space. The reason for this is that many map-space approaches do not require a model of the foregrounds to be provided. Although Hothi et al. (2020) seems to indicate relatively poor performance of map-space methods, there are still options to try, such as scale dependent cleaning [8].

Acknowledgments

I would like to thank WSGC for funding this work, and Peter Timbie for providing feedback.

References

- [1] David R. DeBoer and HERA. Hydrogen Epoch of Reionization Array (HERA). In *American Astronomical Society Meeting Abstracts #225*, volume 225 of *American Astronomical Society Meeting Abstracts*, page 328.03, January 2015.
- [2] Kevin Bandura, Graeme E. Addison, Mandana Amiri, J. Richard Bond, Duncan Campbell-Wilson, Liam Connor, Jean-François Cliche, Greg Davis, Meiling Deng, Nolan Denman, and et al. Canadian hydrogen intensity mapping experiment (chime) pathfinder. *Ground-based and Airborne Telescopes V*, Jul 2014.
- [3] Cosmic Visions 21 cm Collaboration, Réza Ansari, Evan J. Arena, Kevin Bandura, Philip Bull, Emanuele Castorina, Tzu-Ching Chang, Shi-Fan Chen, Liam Connor, Simon Foreman, Josef Frisch, Daniel Green, Matthew C. Johnson, Dionysios Karagiannis, Adrian Liu, Kiyoshi W. Masui, P. Daniel Meerburg, Moritz Münchmeyer, Laura B. Newburgh, Andrej Obuljen, Paul O’Connor, Hamsa Padmanabhan, J. Richard Shaw, Christopher Sheehy, Anže Slosar, Kendrick Smith, Paul Stankus, Albert Stebbins, Peter Timbie, Francisco Villaescusa-Navarro, Benjamin Wallisch, and Martin White. Inflation and early dark energy with a stage ii hydrogen intensity mapping experiment, 2019.

- [4] Emma Chapman, Filipe B. Abdalla, J. Bobin, J.-L. Starck, Geraint Harker, Vibor Jelić, Panagiotis Labropoulos, Saleem Zaroubi, Michiel A. Brentjens, A. G. de Bruyn, and L. V. E. Koopmans. The scale of the problem: recovering images of reionization with generalized morphological component analysis. *Monthly Notices of the Royal Astronomical Society*, 429(1):165–176, nov 2012.
- [5] Emma Chapman, Filipe B. Abdalla, Geraint Harker, Vibor Jelić, Panagiotis Labropoulos, Saleem Zaroubi, Michiel A. Brentjens, A. G. de Bruyn, and L. V. E. Koopmans. Foreground removal using fastica : a show-case of LOFAR-EoR . *Monthly Notices of the Royal Astronomical Society*, 423(3):2518–2532, 06 2012.
- [6] J. Richard Shaw, Kris Sigurdson, Michael Sitwell, Albert Stebbins, and Ue-Li Pen. Coaxing cosmic 21 cm fluctuations from the polarized sky using m -mode analysis. *Phys. Rev. D*, 91:083514, Apr 2015.
- [7] F G Mertens, A Ghosh, and L V E Koopmans. Statistical 21-cm signal separation via gaussian process regression analysis. *Monthly Notices of the Royal Astronomical Society*, may 2018.
- [8] Ian Hothi, Emma Chapman, Jonathan R Pritchard, F G Mertens, L V E Koopmans, B Ciardi, B K Gehlot, R Ghara, A Ghosh, S K Giri, I T Iliev, V Jelić, and S Zaroubi. Comparing foreground removal techniques for recovery of the LOFAR-EoR 21 cm power spectrum. *Monthly Notices of the Royal Astronomical Society*, 500(2):2264–2277, nov 2020.
- [9] Nicolas Fagnoni, Eloy de Lera Acedo, David R DeBoer, Zara Abdurashidova, James E Aguirre, Paul Alexander, Zaki S Ali, Yanga Balfour, Adam P Beardsley, Gianni Bernardi, and et al. Understanding the hera phase i receiver system with simulations and its impact on the detectability of the eor delay power spectrum. *Monthly Notices of the Royal Astronomical Society*, 500(1):1232–1242, Oct 2020.
- [10] D. Slepian. Prolate spheroidal wave functions, Fourier analysis, and uncertainty. V - The discrete case. *AT T Technical Journal*, 57:1371–1430, June 1978.
- [11] Ainunbilal Nasirudin, David Prelogovic, Steven G Murray, Andrei Mesinger, and Gianni Bernardi. Characterizing beam errors for radio interferometric observations of reionization. *Monthly Notices of the Royal Astronomical Society*, 514(3):4655–4668, jun 2022.