

Understanding the diversity of 21 cm cosmology analyses

Miguel F. Morales,^{1,2*} Adam Beardsley,³ Jonathan Pober,⁴ Nichole Barry,^{5,6} Bryna Hazelton,^{1,7} Daniel Jacobs³ and Ian Sullivan⁸

¹Department of Physics, University of Washington, Seattle, WA 98195, USA

²Dark Universe Science Center, University of Washington, Seattle, WA 98195, USA

³School of Earth and Space Exploration, University of Arizona, Tempe, AZ 85287, USA

⁴Department of Physics, Brown University, Providence, RI 02912, USA

⁵School of Physics, University of Melbourne, Parkville, VIC 3010, Australia

⁶ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D)

⁷eScience Institute, University of Washington, Seattle, WA 98195, USA

⁸Department of Astronomy, University of Washington, Seattle, WA 98195, USA

Accepted 2018 October 18. Received 2018 October 10; in original form 2018 July 10

ABSTRACT

21 cm power spectrum observations have the potential to revolutionize our understanding of the epoch of reionization and dark energy, but require extraordinarily precise data analysis methods to separate the cosmological signal from the astrophysical and instrumental contaminants. This analysis challenge has led to a diversity of proposed analyses, including delay spectra, imaging power spectra, m-mode analysis, and numerous others. This diversity of approach is a strength, but has also led to a confusion within the community about whether insights gleaned by one group are applicable to teams working in different analysis frameworks. In this paper, we show that all existing analysis proposals can be classified into two distinct families based on whether they estimate the power spectrum of the *measured* or *reconstructed* sky. This subtle difference in the statistical question posed largely determines the susceptibility of the analyses to foreground emission and calibration errors, and ultimately the science different analyses can pursue. In this paper, we detail the origin of the two analysis families, categorize the analyses being actively developed, and explore their relative sensitivities to foreground contamination and calibration errors.

Key words: methods: data analysis – dark ages, reionization, first stars – cosmology: observations.

1 INTRODUCTION

Radio observations of redshifted 21 cm hydrogen emission have the potential to reveal our cosmic dawn, constrain dark energy, and constrain cosmological parameters over an unprecedented cosmic volume. Major power spectrum (PS) analyses of hundreds of hours of data are underway by PAPER, LOFAR, GMRT, GBT, MWA, and CHIME, and HERA is about to start science data collection (Donald C. Backer Precision Array for Probing the Epoch of Reionization, Parsons et al. 2010; Low-Frequency Array, Haarlem et al. 2013; Giant Metrewave Radio Telescope, Gupta et al. 2017; Green Bank Telescope, Prestage et al. 2009; Canadian Hydrogen Intensity Mapping Experiment, Bandura et al. 2014; Murchison Widefield Array, Tingay et al. 2013; Hydrogen Epoch of Reionization Array, De Boer et al. 2016). Many different PS analyses have been proposed,

and a number have been developed into full-fledged data analysis pipelines.

The community is learning rapidly, and instrument/analysis features such as the ‘foreground wedge,’ the ‘window,’ and the ‘pitchfork’ are major focuses of the literature (see Section 2 and references therein). However, as the effects are always measured or simulated within one of the many analysis frameworks, the details of the analysis become intertwined with the results. Consequently, even domain experts are often uncertain which effects identified by other groups will apply to their own work.

After a careful review of the 21 cm cosmology literature, we find that all of the proposed analyses can be sorted into two ‘families’ based on the mathematics of the final estimator (Section 3). The first family of analyses calculates the PS of the *measured* sky, while the second family calculates the PS of the *reconstructed* sky. This subtle difference in the statistical question being posed has profound effects on the sensitivity of the analyses to foreground emission and calibration errors.

* E-mail: mmorales@phys.washington.edu

The mathematics for how foregrounds contaminate measured (delay-style) and reconstructed (imaging-style) power spectra is fully developed in a paper by Liu, Parsons & Trott (2014). However, the practical ramifications of this distinction have not been fully explored and are underappreciated by the community. The goal of this paper is to clearly illustrate the connections between different analyses. This will enable practitioners both to recognize which effects apply to their efforts, and illuminate the pros and cons of measured and reconstructed power spectra.

After a brief review of interferometric 21 cm measurements with chromatic instruments (Section 2), we identify the two families of analysis (Section 3). We then use this insight to explore how astrophysical foregrounds (Section 4) and calibration errors (Section 5) appear in the two families of analyses. In Section 6, we discuss the pros and cons of the two approaches, and where we see the state-of-the-art 21 cm PS analysis heading in the years to come.

2 INSTRUMENT CHROMATICITY

All interferometric measurements are inherently chromatic. Because the astrophysical foregrounds are very spectrally smooth, in principle the line-of-sight 21 cm fluctuations are free of foreground contamination (see Furlanetto, Oh & Briggs 2006 and Morales & Wyithe 2010 for reviews). However, the natural chromatic response of an instrument leads to a wedge-like region of contamination predicted in Datta, Bowman & Carilli (2010), Morales et al. (2012), Parsons et al. (2012), Trott, Wayth & Tingay (2012), Vedantham, Shankar & Subrahmanyam (2012), Dillon, Liu & Tegmark (2013b), Hazelton, Morales & Sullivan (2013), Thyagarajan et al. (2013), Thyagarajan et al. (2015a) and observed in Pober et al. (2013), Dillon et al. (2013a), Thyagarajan et al. (2015b), Jacobs et al. (2016), and Beardsley et al. (2016). The origin of this contamination can be seen by carefully examining the frequency response of a single visibility. In this section, we briefly review the origin of chromaticity and the associated wedge and window in 21 cm cosmology PS, closely following the developments in Taylor, Carilli & Perley (1999), Thompson, Moran & Swenson (2001), Morales et al. (2012), Liu, Parsons & Trott (2014), and Pober et al. (2016).

The measured visibilities are the cross-correlations of the electric field measured between pairs of antennas. The visibility for one pair of antennas can be written as

$$v_{ij}(f) = \int \left\langle A_i(\theta, f) E(\theta, f) \right\rangle e^{-i2\pi[\Delta r_{ij} f/c]\theta} d^2\theta, \quad (1)$$

where $E(\theta, f)$ is the instantaneous electric field of the sky as a function of direction and frequency, A is the directional electrical sensitivity of antennas i & j , and Δr_{ij} is the physical distance between the antennas. The observed electric field detected by each antenna is given by AE with the exponential describing the direction-dependent time delay of the signal at the second antenna relative to the first (the Green's function propagator of the electric field for a far-field source). The measured visibility is then the angular integral and time average (angle brackets) of the cross-power between the two antennas.

We can move from electric field to brightness description using $\langle EE^* \rangle_i = I$. We can further simplify the resulting equation by assuming the antennas have the same angular response, so we

can use the antenna power sensitivity $B = |A|^2$, and by replacing the physical distance with the antenna separation in wavelengths $\mathbf{u}_{ij} = \Delta \mathbf{r}_{ij} f/c$:

$$v_{ij}(f) = \int I(\theta, f) B(\theta, f) e^{-i2\pi \mathbf{u}_{ij}(f) \cdot \theta} d^2\theta. \quad (2)$$

Note both \mathbf{u} and θ are vectors in this notation. We can immediately see that for a single flat-spectrum source (I is a δ -function at θ_S) the visibility will be proportional to $B(\theta, f) e^{-i2\pi \mathbf{u}_{ij}(f) \cdot \theta_S}$ – the visibility will oscillate with increasing frequency. The longer the baseline length \mathbf{u}_{ij} , the more rapid the frequency oscillations. While there are several sources of instrument chromaticity, fundamentally the array appears larger (in wavelengths) at higher frequencies.

To understand the effect of this visibility oscillation on the power spectra, it is useful to continue to the Fourier representation of the instrument measurement. The integral in equation (2) is equivalent to a 2D angular Fourier transform (FT) from $\theta \rightarrow \mathbf{u}$. Taking the FT we obtain

$$v_{ij}(f) = \int [I(\mathbf{u}, f) * B(\mathbf{u}, f)] \delta(\mathbf{u} - \mathbf{u}'_{ij}) d^2\mathbf{u}, \quad (3)$$

where we have Fourier transformed both the sky brightness I and the antenna brightness sensitivity B ,¹ used the Fourier convolution theorem (* is the convolution operator), and expressed the discrete baselines \mathbf{u}_{ij} as a δ -function. Alternatively, we can use the δ -function to remove the convolution over all \mathbf{u} to obtain

$$v_{ij}(f) = \int I(\mathbf{u}, f) B(\mathbf{u} - \mathbf{u}'_{ij}, f) d^2\mathbf{u}. \quad (4)$$

In other words, the visibility is equal to the FT of the sky brightness integrated by the compact uv-antenna response pattern $B(\mathbf{u})$.

Fig. 1 pictorially shows this process using the angular wavenumber versus frequency space $I(\mathbf{u}, f)$. In the cartoon, the corrugated shading represents the real (or imaginary) component of the emission from a single flat-spectrum source away from the beam centre. An offset δ -function in image space (one point source) becomes an angular corrugation after the angular FT, with constant emission strength as a function of frequency. The visibilities measured the sky emission at the baseline separations indicated by the diagonal black lines. The distance between two antennas measured in wavelengths increases with frequency ($\mathbf{u}_{ij} = \Delta \mathbf{r}_{ij} f/c$), naturally leading to oscillating visibility measurements with frequency (equation 2). As baselines become longer, the speed of oscillation from the foreground source increases. This is the fundamental source of the foreground wedge, as studied in depth in Vedantham et al. (2012), Morales et al. (2012), Trott et al. (2012), Parsons et al. (2012), and Hazelton et al. (2013). However, as we will explore in Section 4, the foreground wedge does appear differently in the two families of PS estimators.

While the foregrounds are spectrally smooth, the cosmological signal has frequency structure due to redshifted line-of-sight fluctuations in the 21 cm brightness temperature. The true PS of the sky is equal to taking the true sky $I(\mathbf{u}, f)$, mapping $\mathbf{u} \rightarrow k_\perp$ with the angular diameter distance and mapping f to line-of-sight distance, and Fourier transforming to obtain $I(k_\perp, k_\parallel)$ and squaring to obtain

¹Often Fourier transformed variables are indicated with a tilde (e.g. $\tilde{I}(\mathbf{u}, f)$), but as we want to use a hat to differentiate between true and estimated quantities later in the paper we omit the tilde and indicate the FT by the variable being a function of the Fourier coordinate \mathbf{u} .

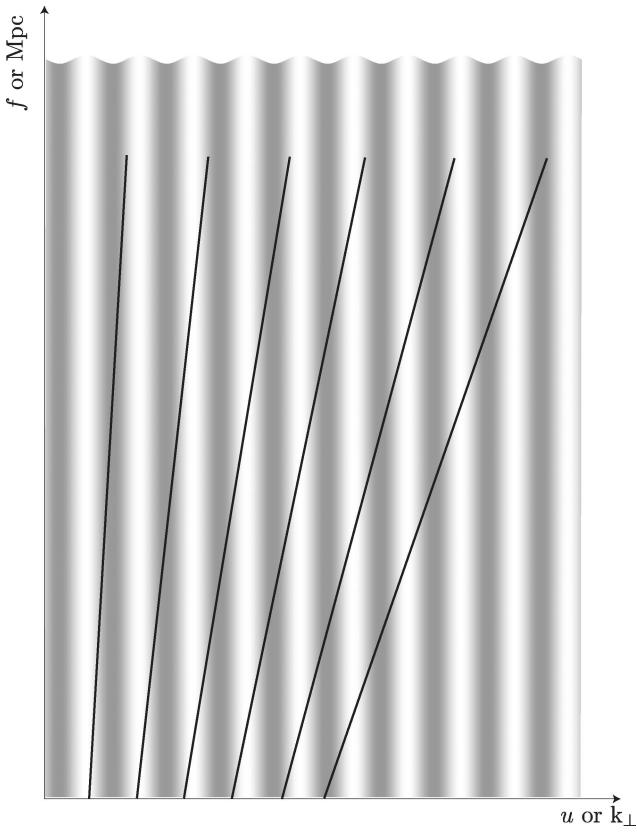


Figure 1. Cartoon adapted from Morales et al. (2012) shows how flat-spectrum foregrounds create oscillatory visibilities, and why the oscillations are faster for longer baselines. Described in detail in Section 2.

$P^T(\mathbf{k})$ (Morales & Hewitt 2004). However, no instrument measures the true sky. The analysis challenge is how to construct an estimate of the *true* PS from the *measured* interferometric visibilities, and it is how this estimate is constructed that determines to which family the PS analysis belongs.

3 POWER SPECTRUM ESTIMATORS

There have been a number of different proposed methods for estimating the cosmological PS \hat{P} from interferometric measurements, including the imaging PS, the variance statistic, the delay spectrum, and m-mode analysis (Morales & Hewitt 2004; Zaldarriaga, Furlanetto & Hernquist 2004; Mellema et al. 2006; Chang et al. 2008; Iliev et al. 2008; Stuart et al. 2008; Pen et al. 2009; Parsons et al. 2012; Vedantham et al. 2012; Masui et al. 2013; Switzer et al. 2013; Dillon et al. 2014; Paul et al. 2014, 2016; Patil et al. 2014; Shaw et al. 2014; Liu, Zhang & Parsons 2016; Trott et al. 2016; Gehlot & Koopmans 2017). While the language used to describe the estimators is highly variable, the mathematics fall into two families depending on whether they measure the PS of the measured visibilities ('delay spectrum' approaches) or the PS of the reconstructed sky ('imaging' PS approaches). In this section we carefully review the proposed measures, and sort the proposals into two families of estimators.

While complete analysis pipelines that transform raw visibilities into PS estimates are often broken into several distinct software packages, here we are considering the complete pipelines and classifying them by the characteristics of the final estimator.

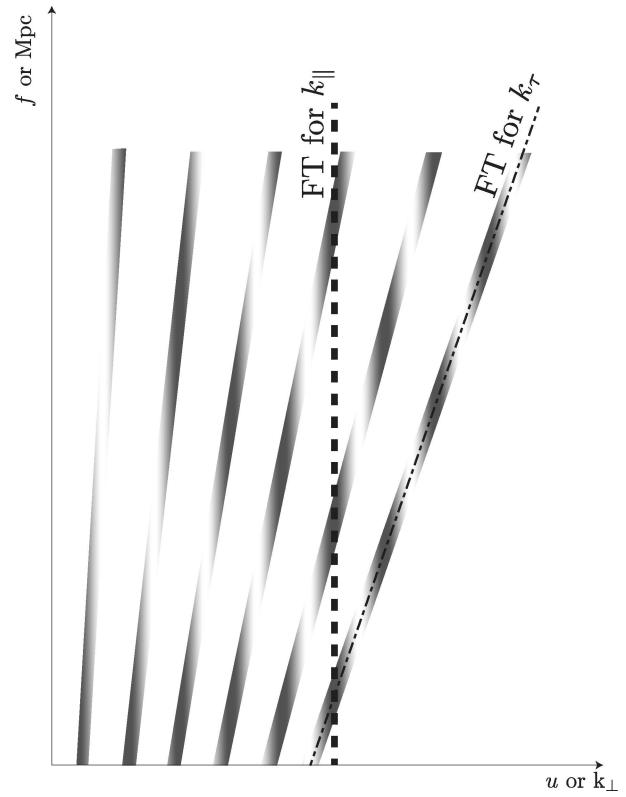


Figure 2. This cartoon shows the oscillations of the different baselines in Fig. 1, showing the faster oscillation for longer baselines. Reconstructed PS estimators $\hat{P}(k_\perp, k_\parallel)$ take the frequency FT along a fixed angular scale, as shown by the vertical thick dashed line, whereas measured sky PS estimators $\hat{P}(k_\perp, k_\tau)$ take the frequency FT along the direction of the baselines as shown by the thin dash-dotted line. This figure is described in detail in Section 3.

3.1 Imaging or reconstructed sky PS

The three dimensional PS of reconstructed 21 cm EoR images was first introduced in Morales & Hewitt (2004). Conceptually this is a very simple estimator—it is just the 3D FT and square of the reconstructed sky estimate $\hat{I}(\theta, f)$ mapped to cosmological coordinates. The resulting PS estimator is $\hat{P}(k_\perp, k_\parallel)$, where k_\perp and k_\parallel are the angular and line-of-sight wavenumbers respectively.

The subtlety is in how the sky estimate is made from interferometric visibility measurements. Because k_\parallel is calculated by taking the FT of $\hat{I}(k_\perp, f)$ perpendicular to k_\perp , the upper and lower frequencies incorporate information from baselines of different physical length as indicated by the vertical thick dashed line in Fig. 2. All of the baselines agree on the amplitude and phase of the foreground sources at that scale (dark colour where the dashed line crosses), so the majority of the power after the line-of-sight FT is at $k_\parallel = 0$. If the estimate of the sky \hat{I} was equal to the true sky, the imaging PS estimator $\hat{P}(k_\perp, k_\parallel)$ would be free of foreground contamination except at the lowest k_\parallel scales. But there are necessarily errors in the estimate of the sky, and these errors leak power from the smooth-spectrum foregrounds into the cosmological volume. This leakage appears in a distinctive 'wedge' pattern with a 'window', which we will explore in Section 4.

A deep literature on how to estimate \hat{I} and $\hat{P}(k_\perp, k_\parallel)$ from interferometric visibilities has been developed, and it has spurred the application of optimal map making and advanced data inversion techniques to interferometric data (e.g. Myers et al. 2003; Bhatnagar et al. 2008; Morales & Matejek 2009; Sullivan et al.

2012; Dillon et al. 2014; Shaw et al. 2014; Zheng et al. 2017). The reconstructed PS has two defining features:

- (i) The line-of-sight coordinate k_{\parallel} is orthogonal to the k_{\perp} coordinate (by construction).
- (ii) Because of this choice of coordinates, data from multiple baselines are linearly combined in forming the values of the $\hat{P}(k_{\perp}, k_{\parallel})$ estimator. The FT of frequency is at a fixed angular scale, and due to the change in array size with frequency (in wavelengths) this involves incorporating information from multiple baselines in a frequency-dependent way. This strongly affects how foregrounds appear in the resulting PS and the sensitivity to calibration errors, as we will discuss in Sections 4 and 5.

It is interesting to note that many of the advanced PS estimation techniques never create an image. But by creating a PS estimate in the k_{\perp}, k_{\parallel} coordinates, they are implicitly working with a reconstructed sky and are combining data from baselines of different length in a frequency-dependent way during the estimation process.

3.2 Delay spectrum or measured sky PS

The delay spectrum was introduced by Parsons et al. (2012) and has been used with PAPER data to provide PS limits. In the delay PS, a visibility – either a single baseline or an average of redundant baselines – is Fourier transformed along frequency and squared to form the PS estimator. Fundamentally, this is a PS of the *measured* sky, and no reconstruction of the true sky is attempted.

From Fig. 2, if k_{\parallel} is the line-of-sight wavenumber we can see that for short baselines the visibility as a function of frequency is sampling sky information *nearly* parallel to the line of sight but not quite. The estimator is really along k_{τ} (diagonal dash-dotted line), where τ is the time delay of the electric field propagation between antennas $\tau_{ij} = \Delta\mathbf{r}_{ij} \cdot \theta/c$, and it is this time delay that gives the delay-spectrum approach its name.

Measurements of multiple baseline lengths can be used to fill out the delay-spectrum estimator $\hat{P}(k_{\perp}, k_{\tau})$, and baselines of different orientation or from different sidereal times can be added incoherently to improve the signal to noise of the estimator. Significant additional work has improved the sensitivity of the technique by optimal time integration fringe-rate filtering (Parsons et al. 2016), leading to a very deep literature and mature estimator. But for our purposes, the measured sky PS estimators have two key features:

- (i) Because the FT from f to k_{τ} follows the baseline migration (diagonal lines in Fig. 2), the same baselines contribute at all frequencies. Redundant baselines can be added together to create a single low-noise measurement before the FT and squaring operations (Parsons et al. 2012), and advanced versions can combine baselines that are not identical (Zhang, Liu & Parsons 2018). But crucially, the baselines contributing to the FT do not change with frequency.
- (ii) The line-of-sight coordinate k_{τ} is nearly, but not exactly, orthogonal to the transverse coordinate k_{\perp} . Because $k_{\tau} \approx k_{\parallel}$, this has a negligible effect on the cosmological PS, and if needed the theory could be rephrased in terms of $\hat{P}(k_{\perp}, k_{\tau})$ (Parsons et al. 2012). However, choosing k_{τ} as the basis of the estimator has a significant impact on how the foreground emission contaminates the PS estimate as explored at length in Section 4.

3.3 Classifying analysis efforts

Table 1 classifies the 21 cm PS analysis efforts that have been developed into full analysis pipelines. The first column lists the measured

sky PS ($\hat{P}(k_{\perp}, k_{\tau})$), and the second column lists the reconstructed sky PS ($\hat{P}(k_{\perp}, k_{\parallel})$). Each of these analysis efforts is backed by many more papers than we have space to list here (see references therein). Other analyses such as Peterson et al. (2009) and Liu et al. (2016) can also be sorted into $\hat{P}(k_{\perp}, k_{\tau})$ or $\hat{P}(k_{\perp}, k_{\parallel})$ families, but have not yet been developed in to full data-ready software pipelines.

The process of classifying all of the analyses highlights two particularly interesting variants: the m-mode analysis by Shaw et al. (2014, 2015) and the physical gridding proposal by Vedantham et al. (2012) and Paul et al. (2014, 2016).

The m-mode analysis was proposed by Shaw et al. (2014), and further developed in Shaw et al. (2015), and is the basis for the CHIME dark energy measurement. On first glance, the m-mode formalism looks very different from most PS analyses, and in many ways it is: no image is ever formed, it is optimized for a drift telescope that sees all right ascensions, and it must work deep in the foreground wedge to achieve CHIME’s dark energy science. In the m-mode analysis, the antenna beams are decomposed into spherical harmonics and are used to map the measured visibilities directly into the l, m spherical harmonic basis. Further, in CHIME the close packing of the antennas along the feed in the N–S (l) direction and the rapid sampling in time (m) leads to almost perfect sampling. Because of the drift scanning strategy that covers all right ascension, these measurements are never mapped to a ‘uv-plane’. But in the chosen harmonics, the measurement coverage is nearly perfect up to the angular resolution of the instrument – perfect ‘uv’ coverage.

While the m-mode analysis never creates an image, it does estimate the \hat{a}_{lm} coefficients of the sky, and is properly classified as a reconstructed sky PS. The estimator is explicitly in the $\hat{P}(k_{\perp}, k_{\parallel})$ frame, and if desired the \hat{a}_{lm} coefficients could be transformed into a reconstructed image of the sky. Further, because the instrument is chromatic there is a frequency-dependent mapping of baselines to $\hat{a}_{lm}(k_{\parallel})$ estimators – baselines of different physical length are linearly combined to form one estimator. These are the two hallmarks of a ‘reconstructed sky’ PS analysis: using the $\hat{P}(k_{\perp}, k_{\parallel})$ basis and combining data from baselines of different physical size in a frequency-dependent manner. m-mode analysis is in the same family as more traditional imaging PS, albeit a very advanced implementation specifically tuned for a unique style of instrument.

Another variant first proposed by Vedantham et al. (2012) and further developed by Paul et al. (2014, 2016) and Gehlot & Koopmans (2017), is to grid visibilities by their physical separation (\mathbf{r} instead of \mathbf{u}). This can be visualized as a skewed version of Fig. 1, where the black baselines now run vertically but the fringes from a flat-spectrum source run diagonally (same crossing angle, fringes run from top left to bottom right). One can efficiently image this kind of analysis using the Chirp-Z Transform in place of the FT, but the proposed PS estimator is to take the FT along the black lines of set physical separation. While gridding is normally associated with imaging and reconstructing a sky, in this case the PS estimator is implicitly in $\hat{P}(k_{\perp}, k_{\tau})$, and because the FT aligns with the physical baseline separation, data from baselines of different physical length are never combined in a frequency-dependent way. So despite the gridding, this matches the characteristics of a measured sky delay-spectrum analysis.

Which family an estimator belongs to is of more than academic interest, as it directly determines how foregrounds and calibration errors affect the final PS measurement. In general, any lesson learned can be directly applied to all of the analyses in the same family. There may need to be some translation, but even technical effects such as *uv* beam clipping (how the contribution of a baseline goes

Table 1. Major 21 cm PS analysis efforts.

Measured PS	Reconstructed PS $\hat{P}(k_{\perp}, k_{\parallel})$
PAPER & HERA delay PS (Parsons et al. 2012; Ali et al. 2015; Jacobs et al. 2015; Kohn et al., in preparation)	GMRT (Paciga et al. 2011)
MWA gridded delay PS (Vedantham et al. 2012; Paul et al. 2016)	GBT auto and cross-correlation (Masui et al. 2013; Switzer et al. 2013)
LOFAR gridded delay PS (Gehlot & Koopmans 2017)	LOFAR variance (Patil et al. 2014)
	LOFAR PS (Patil et al. 2017)
	MWA FHD + EmpCov (Dillon et al. 2015; Ewall-Wice et al. 2016)
	MWA FHD + eppsiilon (Jacobs et al. 2016; Beardsley et al. 2017)
	MWA RTS + CHIPS (Trott et al. 2016; Beardsley et al. 2017)
	CHIME m-mode (Shaw et al. 2014, 2015)

to zero as it passes away from the angular mode of interest, e.g. Figs 2 and 4) translate generally from FHD to CHIPS and m-mode analysis and all of the other $\hat{P}(k_{\perp}, k_{\parallel})$ -style analyses. Similarly the ‘pitchfork’ effect identified by Thyagarajan et al. (2015a) can be directly translated to all of the measured sky $\hat{P}(k_{\perp}, k_{\tau})$ analyses.

In the next two sections we will explore how residual foregrounds and calibration errors generically affect analyses within the two families.

4 FOREGROUND FEATURES

One of the significant developments in 21 cm cosmology has been the recognition and understanding of the ‘foreground wedge’ and the associated ‘EoR window’. The appearance of the foreground wedge in both reconstructed and measured PS has led to the erroneous conclusion that the approaches are effectively identical. While the foreground wedge appears in the same *location* in both formalisms, the *amplitude and characteristics* of the wedge differ between them.

To illustrate these effects, we have analytically calculated the visibilities (equation 2) for a single flat-spectrum source towards the edge of the field of view for an idealized array with uniform *uv*-density.² We then use these analytic visibilities with full PS estimation pipelines to produce both measured and reconstructed sky PS shown in Fig. 3.³

Referring back at equation (2), we can see that a single source will cause frequency ripple in a visibility proportional to $e^{-i2\pi[\Delta r_{ij}f/c]\theta} = e^{-i2\pi u_{ij}(f)\theta_s}$, where u is proportional to the angular cosmological distance k_{\perp} . For the measured PS, we simply FT each visibility along frequency to get a δ -function in delay

²Uniform *uv*-density is not realizable with a physical antenna layout, but is useful for illustration purposes as the baseline density does not depend on k_{\perp} .

³At the dynamic range shown, all measured PS estimators will give the left-hand plot, and all reconstructed PS estimators will give the right-hand plot. At higher dynamic ranges, implementation details will become important. For these figures, the MWA beam was used to simulate the visibilities, a custom python implementation of the delay PS was used for the left-hand plot, and the FHD-eppsiilon reconstruction package was used to create the right-hand plot (detailed in Barry et al., in prep.). Both incorporate a Blackman-Harris spectral window function.

(convolved with a spectral window function associated with the bandpass and frequency-dependent beam, Parsons et al. 2012). For our array with many baselines and one contaminating source, the resulting measured sky PS $\hat{P}(k_{\perp}, k_{\tau})$ is shown in the left-hand panel of Fig. 3. The power from the contaminating source appears at a constant diagonal line at $\tau \propto k_{\perp} \cdot \theta_s$.

The primary feature of the foreground contamination in the measured sky PS is that **all** of the power from a contaminating source appears at the associated k_{τ} . The delay transform is a form of 1D ‘imaging’, with every source having a specific delay. With many sources at different locations the wedge will fill in via superposition, and as it is the apparent brightness of the sources that matters the amplitude of the foreground wedge will fall in k_{τ} proportional to the antenna beam with no meaningful contamination beyond the Earth’s horizon (Thyagarajan et al. 2015a) carefully explores how diffuse structure appears near the horizon in measured sky PS). As the foreground sources are approximately four orders of magnitude brighter than the EoR signal, the contamination of the foreground in the wedge is much brighter than the expected signal. In data, this bright foreground emission with a clear drop in emission at the horizon is most clearly seen in Pober et al. (2013) and confirmed in simulations and observations by Thyagarajan et al. (2015a,b).

For reconstructed sky PS, the frequency FT is taken at a fixed angular scale (u) instead of along individual visibilities. Often this is accomplished by gridding to a *uvf*-cube prior to the FT, but the same effects can be incorporated with frequency-dependent visibility mappings such as in the m-mode and CHIPS analyses. This transform at a fixed angular scale is shown visually by the thick dashed vertical line in Fig. 2. Along that line, the different baselines agree on the amplitude and phase of a contaminating source, so the dominant term in the FT is the flat spectrum $k_{\parallel} = 0$ DC mode. The right-hand panel of Fig. 3 shows the reconstructed sky PS for the same analytic visibilities. There is contamination along the same diagonal band as seen in the left-hand measured PS but at a significantly reduced amplitude.

The amplitude of the foreground wedge in a reconstructed PS is related to the accuracy of the sky reconstruction – with perfect reconstruction all of the foreground emission is correctly mapped to $k_{\parallel} \approx 0$ and the foreground wedge disappears. In practice, the accuracy of the sky reconstruction depends on both the analysis

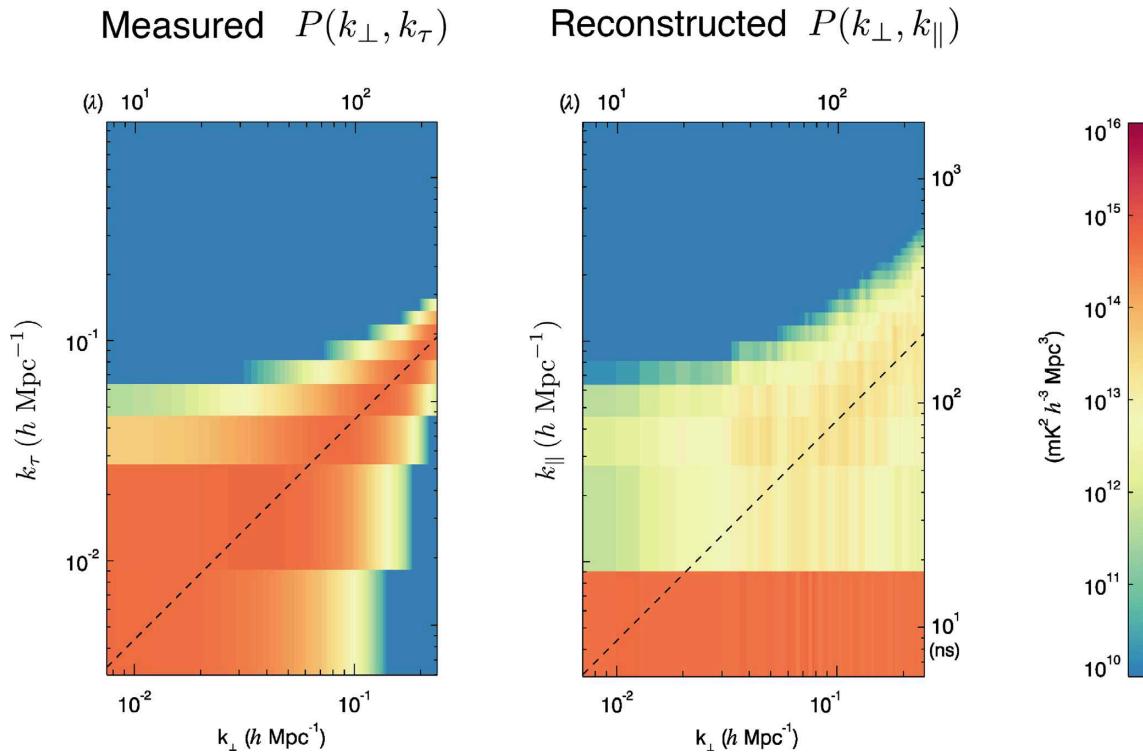


Figure 3. This figure shows the measured (left) and reconstructed (right) PS for the same simulated data. Visibilities for a single flat-spectrum foreground source near the edge of the field of view were analytically generated for an instrument with uniform baseline density. In both cases of PS, we have taken a 2D slice through the 3D PS volume with k_{\perp} aligned with the source. In the left-hand measured PS, all of the power from the contaminating source appears at the associated band in the PS (the dashed line shows the theoretical contamination location of the source). In the right-hand reconstructed PS, most of the power is reconstructed at $k_{\parallel} = 0$, with a smaller amount of power appearing in the same diagonal band due to imperfect sky reconstruction. While both PS have foreground contributions at the same location in the wedge, the amplitude of that contamination and the power at $k_{\parallel} \approx 0$ is quite different as described in the text. If we were to add many more sources at different locations, the wedge would fill in as seen in data.

and the instrument characteristics. For smooth spectrum sources, the instrumental effects indicated in Fig. 1 can be inverted to create a catalogue or model, either through traditional deconvolution (e.g. Carroll et al. 2016; Hurley-Walker et al. 2017) or more formal matrix inversion techniques (e.g. Shaw et al. 2014). Foreground emission captured in these models can be fully removed (modulo calibration and instrument model errors), with no associated emission in the foreground wedge. Residual foreground emission not captured in these models has its power split between the $k_{\parallel} \approx 0$ modes and the wedge as shown in the right-hand panel of Fig. 3. The fraction of the residual emission in the diagonal band is determined by the point-spread function (PSF) of the instrument – the better the PSF the more accurate the (dirty) sky reconstruction and correspondingly less emission in the foreground wedge. As the uv coverage and calibration precision of an array increases, both the PSF and the depth of foreground source catalogues improve and the magnitude of the emission in the foreground wedge decreases.

While both measured and reconstructed sky PS see a foreground wedge, the wedge has qualitatively different properties:

(i) For the measured sky PS, the foreground naturally reconstructs to $k_{\perp} \cdot \theta_S$. $k_{\tau} = 0$ simply refers to sources at phase centre (zero delay) and is not preferred over other locations in the wedge. The amplitude of the foreground emission in the wedge is much brighter in a measured PS because it contains the full amplitude

of the foregrounds. This is clearly seen in Pober et al. (2013) and Thyagarajan et al. (2015b).

(ii) For the reconstructed sky PS, most of the foreground power appears at $k_{\parallel} \approx 0$, with leakage due to imperfect foreground isolation forming a line at the associated $k_{\perp} \cdot \theta_S$. The wedge will fill in with the superposition of many sources, but the amplitude of the wedge is strongly associated with the instrument PSF and the quality of the foreground removal. This is seen in PS measurements in Jacobs et al. (2016) and Beardsley et al. (2016).

Measured and reconstructed PS estimators are asking different statistical questions. The measured PS estimator $\hat{P}(k_{\perp}, k_{\tau})$ is the PS of the calibrated visibilities and does not explicitly project sources to a specific location on the sky. A particular k_{τ} in the wedge can be generated by either a flat-spectrum source at a particular position or by sources with spectral variations at other locations – the estimator does not distinguish. In contrast, $\hat{P}(k_{\perp}, k_{\parallel})$ is the PS of the reconstructed sky. The reconstruction actively tries to remove the effects of the chromatic instrument and place the emission at its true location. The errors in this reconstruction due to unsubtracted sources coupling with the chromatic PSF lead to the emission in the wedge. In both estimators, the foreground wedge is generated by the baseline movement of the instrument (Fig. 1), so the wedge appears in the same location. But because the estimators are asking different questions, the foreground characteristics and the amplitude of the wedge are different.

The differences in how the foregrounds appear in the two classes of estimators impact the kinds of science that can be pursued. But to fully understand these effects, we must also understand how the analyses respond to calibration errors.

5 CALIBRATION

The importance of calibration has long been recognized in 21 cm cosmology and is a very active area of current research (e.g. Barry et al. 2016; Patil et al. 2016; Trott et al. 2016; Ewall-Wice et al. 2017). The purpose of calibration is to measure the as-built instrument performance – $A_i(\theta, f)$ in equation (1) – and include it in the analysis. A full understanding of the causes and effects of calibration errors is well beyond the scope of this paper, however, it is instructive to qualitatively understand how calibration errors affect our two families of PS estimators.

Because calibration is tightly associated with array layout, we find the clearest general argument is analytic. Assuming the broad calibration worked well, we will focus only on small differences between the true antenna performance and the antenna calibration used in the analyses – the calibration errors. In the limit of small calibration errors, it is conceptually helpful to factor the error in the antenna model into three terms:

$$\Delta A_i(\theta, f) = \overbrace{\Delta a_i e^{i \Delta \phi_i}}^{\text{f indep.}} \times \overbrace{\Delta a_i(f) e^{i \Delta \phi_i(f)}}^{\text{f dep.}} \times \overbrace{\Delta G_i(\theta, f)}^{\text{beam shape}}. \quad (5)$$

These terms represent the overall frequency-independent error in the sensitivity (Δa) and timing ($\Delta \phi$) of an antenna, the frequency-dependent sensitivity and timing errors, and the beam shape error. In the following subsections, we will explore how these different kinds of calibration errors affect the reconstructed and measured sky PS. The major result of this analysis is that reconstructed PS are much more sensitive to frequency-independent calibration errors than measured PS.

5.1 Frequency-independent calibration errors

Our first case is a small frequency-independent error in either the sensitivity or phase of one antenna (Δa or $\Delta \phi$). For the delay spectrum of a single baseline, frequency-independent calibration errors have almost no effect. The amplitude makes a minute error in the strength of the PS. The phase error corresponds to a small shift in τ . If many baselines are added together incoherently, the phase error will shift and blur the delay PS very slightly, but will have no material effect on the location and amplitude of the foreground emission and the EoR window should remain clean.

In contrast, for a reconstructed sky PS the same calibration error can move power from the $k_{\parallel} = 0$ line into the foreground wedge and potentially beyond. We can show this pictorially in Fig. 4 for simple interferometric imaging, though the same mathematical effect occurs in all reconstructed PS, even those in which no image is explicitly formed. In simple interferometric imaging, a visibility is first gridded to the uvf -cube ($\hat{I}(\mathbf{u}, f)$) using a gridding kernel with a small width in $\Delta u \Delta v$, as shown in the left-hand side of Fig. 4. Mathematically, the estimated apparent sky in uvf coordinates is

$$\hat{I}(\mathbf{u}, f) = \frac{\sum_{ij} v_{ij}(f) B^*(\mathbf{u} - \Delta \mathbf{r}_{ij} f/c, f)}{\sum_{ij} B^*(\mathbf{u} - \Delta \mathbf{r}_{ij} f/c, f)}, \quad (6)$$

where B^* is the gridding kernel. The term $\Delta r_{ij} f/c$ captures the frequency-dependent movement of the baseline. Because of the

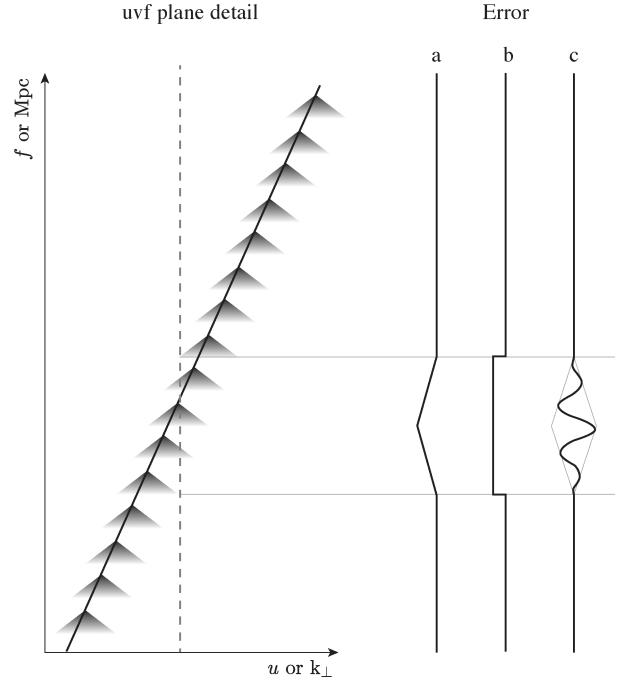


Figure 4. The contribution of a single miscalibrated baseline to the reconstructed uvf -cube. In the left-hand panel, a single miscalibrated baseline (diagonal line) contributes to the estimate at a particular angular scale (vertical dashed line) over a range of frequencies. At every frequency, the baseline contributes to a range of angular scales as indicated by the small triangles (the gridding kernel in an imaging analysis or the covariance scale in other reconstructed sky estimators). The error contribution of the miscalibrated visibility to the angular scale of interest is shown to the right in three scenarios: (a) frequency-independent calibration error in the high baseline number limit, (b) frequency-independent calibration error in the single baseline limit, and (c) frequency-dependent calibration error in the high baseline number limit.

finite width of the gridding kernel, each baseline contributes to a range of frequencies with a weight associated with the uv -space beam shape. The range of frequencies depends on the size of the gridding kernel (antenna size, inverse of FoV) and the baseline length since longer baselines cross more quickly. We see that along any given angular scale \mathbf{u} the uvf -cube $\hat{I}(\mathbf{u}, f)$ is a weighted average of the visibilities.

The question is how are $\hat{I}(\mathbf{u}, f)$ and $\hat{P}(k_{\perp}, k_{\parallel})$ affected by a single miscalibrated visibility measurement. Because $\hat{I}(\mathbf{u}, f)$ is a weighted average, the effect depends on how many visibilities are contributing.

In the limit of many visibilities, a frequency-independent error takes on the shape of the uv -beam as it crosses the angular scale \mathbf{u} of interest, as shown in (a) on the right-hand side of Fig. 4. This produces an enveloped error. The shape of the envelope depends on the gridding kernel, and how close the baseline comes to directly crossing the angular scale \mathbf{u} (impact parameter), but for a single miscalibrated baseline that does directly cross and an angularly symmetric frequency-independent beam, the error in the reconstruction at the angular scale \mathbf{u} is

$$\Delta \hat{I}(\mathbf{u}, f) = \Delta v_{ij} B^*(\Delta f \Delta \mathbf{r}_{ij}/c), \quad (7)$$

where Δf is the difference from the frequency at the centre of the baseline crossing. The frequency shape of the error along the angular scale of interest is just the shape of the gridding kernel times

a scaling factor r_{ij}/c related to how quickly the baseline moves with frequency. In most analyses, the gridding kernel is the angular FT of the antenna beam shape, so the contamination is a scaled version of the antenna beam in uv coordinates. In a nice bit of symmetry, when we perform the frequency FT $f \rightarrow k_{\parallel}$ to calculate $\hat{P}(k_{\perp}, k_{\parallel})$, this translates the uv beam back to scaled angular coordinates. The calibration contamination in the PS is enveloped by a scaled version of the angular antenna beam, centred at $k_{\parallel} = 0$. So in the limit of many baselines, frequency-independent calibration errors move power from $k_{\parallel} \approx 0$ into the wedge for a reconstructed PS. (A full mathematical description can be obtained by substituting in the visibility error Δv into the mathematics of Liu et al. 2014, and can be qualitatively understood with a similar substitution into Hazelton et al. 2013 and the animated version their fig. 2 available online).

There is, however, the other limit of very few baselines. In this limit, the weighting provided by the beam in the numerator and denominator of equation (6) divide out, and we get a sharp change in the reconstructed $\hat{I}(\mathbf{u}, f)$ as the miscalibrated visibility crosses, as shown in case (b) of Fig. 4. This sharp change leads to calibration errors contributing to all k_{\parallel} , including those in the window.

These effects can be seen in data. For the MWA, the baseline density is very high at small k_{\perp} and becomes sparse at large k_{\perp} . In the PS published in Jacobs et al. (2016) and Beardsley et al. (2016), the contamination at low k_{\perp} is mostly confined to the wedge because the baseline density is high, whereas it leaks into the window at high k_{\perp} where the baseline density becomes low and the opposite limit is reached.

Sparse baselines usually correspond to low sensitivity, so only areas of high baseline density materially contribute to most PS measurements. So, the frequency-independent calibration errors are confined to the wedge for most reconstructed PS. However, the baseline density is not infinite, so some of the contamination can bleed into the window.

The frequency-dependent contribution of a measurement to a single angular scale – the hallmark of a reconstructed sky PS estimator – is what causes frequency-independent calibration errors to move power from $k_{\parallel} \approx 0$ into the wedge and potentially into the window. This is fundamental to the approach, as it was exactly this frequency-dependent baseline contribution that enabled the power to be reconstructed to $k_{\parallel} \approx 0$ in the first place. This kind of calibration error, where baselines of different physical length must agree very precisely on the amplitude and phase, is a very active area of current research.

In conclusion, measured sky PS estimators $\hat{P}(k_{\perp}, k_{\parallel})$ are much less sensitive to frequency-independent calibration errors than reconstructed sky PS estimators.

5.2 Frequency-dependent calibration errors

There are two other classes of calibration errors from equation (5). The first of these is frequency-dependent calibration errors. Effects such as a cable reflection can source sinusoidal ripples in the antenna sensitivity (and phase) that are nearly impossible to calibrate at the necessary precision (Barry et al. 2016; Beardsley et al. 2016). These frequency-dependent calibration errors can be described with a set of δ -functions in τ at the corresponding time delays. For the delay spectrum, these errors act like a large additive delay $\Delta\tau$ and can directly move power from the foreground wedge into the window.

For a reconstructed sky analysis in the large baseline limit, the impact of a sinusoidal calibration error is shown in (c) on the right-hand

side of Fig. 4. The sinusoidal calibration error appears, enveloped by the gridding kernel. This will throw power into the EoR window, as in the measured PS case, but it will not be as sharp a contribution due to the modulation of the sinusoid by the gridding kernel. The effect of frequency-dependent calibration errors on reconstructed PS was simulated and explored in depth in Barry et al. (2016).

As both classes of PS estimators are extremely sensitive to frequency-dependent calibration errors, eliminating spectral features is driving the instrumental requirements of HERA, MWA III, and SKA-Low (DeBoer et al. 2016).

5.3 Beam shape errors

The last calibration error from equation (5) is the antenna beam shape. In particular, due to manufacturing and electronic differences one antenna may have a different angular response. Measuring the antenna beam using celestial sources (e.g. Newburgh et al. 2014; Berger et al. 2016), satellites (Neben et al. 2015, 2016a,b, Line et al. in preparation), and drones (Jacobs et al. 2017) is an active area of research.

Using modern techniques (e.g. Myers et al. 2003; Bhatnagar et al. 2008; Morales & Matejek 2009), sky reconstruction analyses can correct for the direction-dependent polarized response of individual antennas. Effectively, the analysis can project visibilities to the sky and apply direction-dependent corrections during image reconstruction. The PS of the reconstructed image is then free of these errors to the precision of the direction-dependent calibration.

In contrast, because measured sky PS have no sky reconstruction step, it is more difficult to apply a direction-dependent gain correction. Parsons et al. (2016) show how limited direction-dependent information can be propagated to the PS estimator, but there is no way to completely correct for angular differences in the polarized antenna beams. The measured sky PS inherently assumes that the angular gain of all antennas are similar and that the polarization at the relevant spectral scales is small (e.g. Moore et al. 2013; Kohn et al. 2016; Moore et al. 2017).

In both classes of analyses, purely angular differences in the antenna beam shape will be confined to the wedge. This can be seen by substituting perturbed beams into Liu et al. (2014), or by noting that beam errors will make certain areas of the sky appear artificially bright or dim but not affect the spectral structure. In reconstructed analyses, the contamination of the wedge can be reduced using knowledge of the individual antenna beams to improve the reconstruction. There is no fully analogous technique for measured PS, but the foreground wedge is already extremely bright for these measurements.

5.4 Calibration effects reviewed

The two classes of PS estimators have different responses to the three kinds of calibration errors described in equation (5).

(i) Both measured and reconstructed PS are extremely sensitive to frequency-dependent calibration errors (Section 5.2). These errors can move foreground power into the window. This sensitivity to spectral ripple is driving the intrinsically smooth bandpass requirements for antennas in next generation 21 cm instruments.

(ii) Frequency-independent calibration errors (Section 5.1) have negligible impact on measured PS, but in reconstructed PS can move power from $k_{\parallel} \approx 0$ into the wedge and window. Amplitude and phase calibration between baselines of different physical length must be very precise to enable working within the wedge.

(iii) Direction-dependent calibration can be incorporated in reconstructed sky PS, but there is no natural way of incorporating direction-dependent effects in measured sky PS.

6 DISCUSSION

By realizing that all of the proposed 21 cm analyses can be categorized into two broad classes of PS estimator (Section 3, Table 1), we can start to ask more generic questions about the different approaches. In this paper, we have identified two major features: the brightness of the wedge (Section 4) and the sensitivity to frequency-independent calibration errors (Section 5.1).

The location of the foreground wedge is determined by the chromatic nature of an interferometer (Fig. 1), and is the same for both measured and reconstructed PS estimators. However, the wedge is much brighter for measured PS estimators. Because measured PS analyses make no attempt to reconstruct the true location of the emission, all of the foreground power in a visibility is mapped to the corresponding location in the wedge (Fig. 3a). In contrast, for reconstructed PS estimators most of the smooth spectrum emission is mapped to $k_{\parallel} \approx 0$ (Figs 2 and 3b). The more precise the reconstruction, the less power appears within the wedge. Reconstruction errors due to residual sources, instrumental PSF, and calibration errors determine the brightness of the foreground wedge in a reconstructed PS analysis.

A particular advantage of measured PS estimators is their immunity to frequency-independent calibration errors (Section 5.1). Because the frequency FT is taken along k_{τ} (Fig. 2), frequency-independent calibration errors have almost no effect on the resulting PS. In contrast, because different baselines contribute at different frequencies for the FT to k_{\parallel} , reconstructed PS are quite sensitive to frequency-independent calibration errors (Figs 2 and 4, Section 5.1).

For analyses that limit themselves to the window, there is no clear advantage to either the measured or reconstructed PS approaches – contamination in the window is primarily associated with frequency-dependent calibration errors. Instrument specific differences such as the smoothness of the bandpass, antenna layout, and calibration approach quickly determine the precision that can be reached in the window.

Things become more interesting if one wishes to work within the wedge. Both LOFAR and CHIME are predicated on being able to work deep within the wedge, and there is a large boost in both sensitivity and science reach that can be obtained if the cosmological modes within the wedge can be used (e.g. Pober et al. 2013).

In practice, working within the wedge requires a reconstructed PS analysis. PS of a reconstructed sky have the advantage that the foreground contamination within the wedge is attenuated (Fig. 3b). In contrast, the full foreground emission appears in the measured PS wedge (Fig. 3a). Working within the wedge with a measured PS would require much more precise foreground models because there is no natural suppression from the instrumental PSF. Worse, determining the foreground model to subtract would require an imaging analysis. Within the delay-spectrum approach, there is no way to identify whether the flux in a k -space bin is foreground or cosmology – both smooth-spectrum foregrounds and the spectrally varying cosmological signal map to the same bins. Measuring the foregrounds to subtract in order to work within the wedge would require an imaging analysis stage, and it would require imaging to a significantly higher precision than is needed for a reconstructed PS analysis. To work within the wedge with the measured PS, one would have already been able to perform the analysis on lower precision sky estimates with a reconstructed PS estimator.

The measured PS estimators do enjoy a smaller software development burden. While all of the PS efforts involve building large custom software analysis pipelines, the complexity of the codes needed to accurately reconstruct the sky is significantly higher (e.g. Sullivan et al. 2012; Jacobs et al. 2016; Trott et al. 2016; Hazelton et al. in prep; Pindor et al. in preparation). Because of the unprecedented precision required for 21 cm imaging, extraordinarily small errors can lead to unexpected deleterious results, such as those described in Trott et al. (2016), Beardsley et al. (2016), Patil et al. (2017), and Barry et al. (2016). Reconstructing the sky at the precision needed for 21 cm cosmology is incredibly difficult.

There are nascent efforts to develop hybrid analyses such as Kerrigan et al. (2018). It may be possible to filter foregrounds in one style of analysis and construct a PS in the other, or transfer calibrations. These methods are in their infancy but crossing the divide between the measured and reconstructed PS analyses may capitalize on the unique advantages of both.

7 CONCLUSIONS

The GBT, PAPER, MWA, and LOFAR teams are all actively reducing hundreds of hours of EoR data, and the progress of all the associated analyses is progressing rapidly. CHIME is just starting to collect its first science data, major hardware upgrades to the MWA and LOFAR are underway, HERA construction has begun, and SKA-low is on the horizon.

Looking to the future of 21 cm cosmology analysis, it is the belief of many in the community that we will need to incorporate lessons from all of the different approaches, and that both measured and reconstructed PS approaches have key roles to play. HERA, MWA-II, and CHIME have all been designed to have both redundant baselines (best for measured sky PS) and excellent uv coverage (best for reconstructed PS), and HERA and MWA-II are explicitly planning to perform both styles of analysis.

This is an exciting time for 21 cm cosmology as we develop the foundational analysis tools and techniques. In this paper, we have endeavoured to conceptually organize the different proposed analyses (Section 3), characterize how foregrounds and calibration errors appear (Sections 4 and 5), and highlight the relative advantages of the different approaches (Section 6).

In future work, we intend to further explore the subtle differences in the shape and extent of the window between the two classes of analysis, and we hope our categorization will help best practices be more quickly shared across the many groups working on 21 cm cosmology.

ACKNOWLEDGEMENTS

This work was directly supported by National Science Foundation (NSF) grants #1613855, #1613040, #1506024, and #1636646 and National Aeronautivcal and Space Administration grant 80NSSC18K0389. AB is supported by an NSF Astronomy and Astrophysics Postdoctoral Fellowship under #1701440. We'd also like to thank the anonymous referee, Adrian Liu, Cathryn Trott, and Michael Eastwood for their helpful comments in improving this paper.

REFERENCES

Ali Z. S. et al., 2015, *ApJ*, 809, 61
 Bandura K. et al., 2014, in Stepp L. M., Gilmozzi R., Hall H. J., eds, Proc. SPIE Conf. Ser. Vol. 9145. SPIE, Bellingham, p. 914522

Barry N., Hazelton B., Sullivan I., Morales M. F., Pober J. C., 2016, *MNRAS*, 461, 3135

Beardsley A. P. et al., 2016, *ApJ*, 833, 102

Beardsley A. P., Thyagarajan N., Bowman J. D., Morales M. F., 2017, *MNRAS*, 470, 4720

Berger P. et al., 2016, in Hall H. J., Gilmozzi R., Marshall H. K., eds, Proc. SPIE Conf. Ser. Vol. 9906, Ground-based Airborne Telescope. SPIE, Bellingham, p. 99060D

Bhatnagar S., Cornwell T. J., Golap K., Uson J. M., 2008, *A&A*, 487, 419

Carroll P. A. et al., 2016, *MNRAS*, 461, 4151

Chang T. C., Pen U. L., Peterson J. B., McDonald P., 2008, *Phys. Rev. Lett.*, 100, 091303

Datta A., Bowman J. D., Carilli C. L., 2010, *ApJ*, 724, 526

DeBoer D. R. et al., 2016, *PASP*, 129, 045001

Dillon J. S. et al., 2013a, *Phys. Rev. D*, 89, 023002

Dillon J. S., Liu A., Tegmark M., 2013b, *Phys. Rev. D*, 87, 043005

Dillon J. S. et al., 2014, *Phys. Rev. D*, 91, 1

Dillon J. S. et al., 2015, *Phys. Rev. D*, 91, 123011

Ewall-Wice A. et al., 2016, *MNRAS*, 460, 4320

Ewall-Wice A., Dillon J. S., Liu A., Hewitt J., 2017, *MNRAS*, 470, 1849

Furlanetto S. R., Oh S. P., Briggs F. H., 2006, *Phys. Rep.*, 433, 181

Gehlot B. K., Koopmans L. V. E., 2017, *Proc. Int. Astron. Union*, 12, 106

Gupta Y. et al., 2017, *Curr. Sci.*, 113, 707

van Haarlem M. P., Wise M. W., Gunst A. W., Heald G., McLean J. P., Hessels J. W. T., Bruyn A. G. D., 2013, *A&A*, 556, A2

Hazelton B. J., Morales M. F., Sullivan I. S., 2013, *ApJ*, 770, 156

Hurley-Walker N. et al., 2017, *MNRAS*, 464, 1146

Iliev I. T., Mellema G., Pen U. -L., Bond J. R., Shapiro P. R., 2008, *MNRAS*, 384, 863

Jacobs D. C. et al., 2015, *ApJ*, 801, 51

Jacobs D. C. et al., 2016, *ApJ*, 825, 114

Jacobs D. C. et al., 2017, *PASP*, 129, 35002

Kerrigan J. et al., 2018, *ApJ*, 864, 131

Kohn S. A. et al., 2016, *ApJ*, 823, 88

Liu A., Parsons A. R., Trott C. M., 2014, *Phys. Rev. D - Part. Fields, Gravit. Cosmol.*, 90, 1

Liu A., Zhang Y., Parsons A. R., 2016, *ApJ*, 833, 242

Masui K. W. et al., 2013, *ApJ*, 763, L20

Mellema G., Iliev I. T., Pen U. -L., Shapiro P. R., 2006, *MNRAS*, 372, 679

Moore D. F., Aguirre J. E., Parsons A. R., Jacobs D. C., Pober J. C., 2013, *ApJ*, 769, 154

Moore D. F. et al., 2017, *ApJ*, 836, 154

Morales M. F., Hewitt J., 2004, *ApJ*, 615, 7

Morales M. F., Matejek M., 2009, *MNRAS*, 400, 1814

Morales M. F., Wyithe J. S. B., 2010, *ARA&A*, 48, 127

Morales M. F., Hazelton B., Sullivan I., Beardsley A., 2012, *ApJ*, 752, 137

Myers S. T. et al., 2003, *ApJ*, 591, 575

Neben A. R. et al., 2015, *Radio Sci.*, 50, 614

Neben A. R. et al., 2016a, *ApJ*, 820, 44

Neben A. R. et al., 2016b, *ApJ*, 826, 199

Newburgh L. B. et al., 2014, in Stepp L. M., Gilmozzi R., Hall H. J., eds, Proc. SPIE Conf. Ser. Vol. 9145, Ground-based Airborne Telescope. SPIE, Bellingham, p. 91454V

Paciga G. et al., 2011, *MNRAS*, 413, 1174

Parsons A. R. et al., 2010, *AJ*, 139, 1468

Parsons A. R., Pober J. C., Aguirre J. E., Carilli C. L., Jacobs D. C., Moore D. F., 2012, *ApJ*, 756, 165

Parsons A. R., Liu A., Ali Z. S., Cheng C., 2016, *ApJ*, 820, 51

Patil A. H. et al., 2014, *MNRAS*, 443, 1113

Patil A. H. et al., 2016, *MNRAS*, 463, 4317

Patil A. H. et al., 2017, *ApJ*, 838, 65

Paul S. et al., 2014, *ApJ*, 793, 28

Paul S. et al., 2016, *ApJ*, 793, 17

Pen U. -L., Staveley-Smith L., Peterson J. B., Chang T. -C., 2009, *MNRAS*, 394, L6

Peterson J. B. et al., 2009, preprint ([arXiv:0902.3091](https://arxiv.org/abs/0902.3091))

Pober J. C. et al., 2013, *ApJ*, 768, L36

Pober J. C. et al., 2016, *ApJ*, 819, 13

Prestage R. M., Constantinescu K. T., Hunter T. R., King L. J., Lacasse R. J., Lockman F. J., Norrod R. D., 2009, *Proc. IEEE*, 97, 1382

Shaw J. R., Sigurdson K., Pen U. L., Stebbins A., Sitwell M., 2014, *ApJ*, 781, 57

Shaw J. R., Sigurdson K., Sitwell M., Stebbins A., Pen U. L., 2015, *Phys. Rev. D*, 91, 083514

Stuart J., Wyithe B., Loeb A., Geil P. M., 2008, *MNRAS*, 383, 1195

Sullivan I. S. et al., 2012, *ApJ*, 759, 17

Switzer E. R. et al., 2013, *MNRAS*, 434, L46

Taylor G. B., Carilli C. L., Perley R. A., eds, 1999, ASP Conf. Ser. Vol. 180, *Synthesis Imaging in Radio Astronomy II*. Astron. Soc. Pac., San Francisco

Thompson A. R., Moran J. M., Swenson G. W., Jr., 2001, *Interferometry and Synthesis in Radio Astronomy*, 2nd edn. Wiley, New York

Thyagarajan N. et al., 2013, *ApJ*, 776, 6

Thyagarajan N. et al., 2015a, *ApJ*, 804, 14

Thyagarajan N. et al., 2015b, *ApJ*, 807, L28

Tingay S. J. et al., 2013, *AJ*, 146, 103

Trott C. M., Wayth R. B., Tingay S. J., 2012, *ApJ*, 757, 101

Trott C. M. et al., 2016, *ApJ*, 818, 139

Vedantham H., Shankar N. U., Subrahmanyam R., 2012, *ApJ*, 745, 176

Zaldarriaga M., Furlanetto S. R., Hernquist L., 2004, *ApJ*, 608, 622

Zhang Y. G., Liu A., Parsons A. R., 2018, *ApJ*, 852, 110

Zheng H. et al., 2017, *MNRAS*, 465, 2901

This paper has been typeset from a *TeX/LaTeX* file prepared by the author.