The Role of the Instrumental Response in 21 cm Epoch of Reionization Power Spectrum Gridding Analyses

Nichole Barry1,2,3 and Aman Chokshi2,3

1 International Centre for Radio Astronomy Research, Curtin University Perth, WA 6845, Australia
2 The University of Melbourne, School of Physics, Parkville, VIC 3010, Australia
3 ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia

Received 2022 January 24; revised 2022 February 17; accepted 2022 February 26; published 2022 April 13

Abstract

Reconstruction of the sky brightness measured by radio interferometers is typically achieved through gridding techniques, or histograms in spatial Fourier space. For Epoch of Reionization (EoR) 21 cm power spectrum measurements, extreme levels of gridding resolution are required to reduce spectral contamination, as explored in other works. However, the role of the shape of the Fourier space spreading function, or kernel, also has consequences in reconstructed power spectra. We decompose the instrumental Murchison Widefield Array (MWA) beam into a series of Gaussians and simulate the effects of finite kernel extents and differing shapes in gridding/degridding for optimal map making analyses. For the MWA, we find that the kernel must extend out to 0.001–0.0001% of the maximum value in order to measure the EoR using foreground avoidance. This requirement changes depending on beam shape, with compact kernels requiring far smaller extents for similar contamination levels at the cost of less-optimal errors. However, simple calibration using pixelated degridding results, regardless of shape of the kernel, cannot recover the EoR due to catastrophic errors caused by the pixel resolution. Including an opaque horizon with widefield beams also causes significant spectral contamination via a beam–horizon interaction that creates an infinitely extended kernel in Fourier space, which cannot be represented well. Thus, our results indicate that simple calibration via degridded models and optimal map making for extreme widefield instrumentation are not feasible.

Unified Astronomy Thesaurus concepts: Reionization (1383); Astronomy data reduction (1861); Astronomical simulations (1857); Radio interferometers (1345); Radio astronomy (1338)

1. Introduction

Next-generation telescopes like the Square Kilometre Array (SKA; Mellema et al. 2013) are going to drastically increase the size and sensitivity of available radio interferometers, and therefore increase resolution to reach new astronomical capabilities. However, this leap in technology requires a re-evaluation of standard analysis techniques, and whether or not they can achieve the desired scientific goals. In particular, the Epoch of Reionization (EoR) science case requires extreme levels of spectral precision, and thus the EoR has so far yet remained unseen.

By using the 21 cm hyperfine transition of hydrogen, measurements of the EoR during the evolution of the universe from redshifts 6 through 11 can provide answers to questions regarding cosmology (e.g., Furlanetto 2019) and astrophysics (e.g., Mirocha 2019) via the use of the power spectrum. However, the measurement of the EoR is particularly difficult due to its extremely low brightness (~10 mK) when compared to all of the intervening extragalactic and Galactic foregrounds (~300 K and above) (Jelić et al. 2008).

Precision analysis of EoR interferometric data sets is inherently difficult due to the measurement basis, which is neither fully harmonic space nor image space. Thus, spectral and spatial harmonic transformations are required depending on the most sensible basis for each particular analysis technique. These transformations depend on the projection, typically either Cartesian or spherical, resulting in Fourier transforms or spherical harmonic transforms, respectively (e.g., Carozzi 2015).

Most analyses use a Cartesian projection due to the availability of fast and efficient discrete transforms, namely the fast Fourier transform (FFT; Cooley & Tukey 1965). Many standard techniques in radio interferometry rely on FTFs, like power spectra analysis and imaging algorithms based on degridding/gridding methods (e.g., Brouw 1975; Cornwell et al. 2008).

In particular, a technique called gridding can be used to reduce data volume while providing multiple outputs, such as images and power spectra. Measured visibilities are added to a finely resolved grid in spatial Fourier space, essentially creating a Fourier histogram of the observation. The simplest form of gridding treats the contribution of the visibility as a delta function; however, this leads to aliasing, smearing, and decorrelation. Convolutional gridding avoids these issues by adding the visibility with a chosen spreading function, or kernel, at the location of the baseline in question (see Offringa et al. 2019a for more details). If the kernel accurately describes the Fourier response of the all-sky instrumental sensitivity, then this plane is the Fourier transform of the orthographic projected sky domain. Only a double spatial-FFT is required to transform from one space to another in this formulism.

Due to the lighter computational load of gridding, many techniques in image and power spectra analysis rely on it, including w-projection or w-stacking algorithms. However, the gridding kernel needs to be extremely precise for spectral variation studies like the EoR. Offringa et al. (2019a) found that to reach the EoR-level precision, over-resolution factors of...
at least 4000 were required for discrete kernels, in addition to over 500 discrete planes to account for $w$-terms. This is difficult in practice, but obtainable with current computing resources. They found that new algorithms which avoid direct gridding, like image domain gridding (van der Tol et al. 2018), did not require such over-resolution factors.

Kernel resolution is undoubtedly important, but there is a broader question on how well the kernel needs to represent the Fourier dual of the instrumental response. Mathematically, errors are considered optimal, or lowest, when the kernel used to grid describes the instrument (Tegmark 1997a, 1997b). Furthermore, creating model visibilities for standard calibration techniques requires a well-represented instrumental kernel. In practice, however, it is difficult to represent the kernel accurately (e.g., Joseph et al. 2020; Chokshi et al. 2021).

This work explores the limitations with reproducing an accurate instrumental kernel for gridding, both in shape and extent, for EoR power spectrum analyses. We set constraints for kernel extent required in gridding to theoretically measure the EoR, and test if the instrumental response can deviate from the known shape. This is of particular interest to widefield instruments, where extremely low instrumental sensitivity, like secondary sidelobes, can interact with the horizon to produce an infinitely extended kernel in the Cartesian limit. In this particular case, we also explore contamination produced via spatial FFTs and how that propagates to the power spectrum. These results indicate a necessary paradigm shift away from traditional methodologies for the EoR science case, particularly for optimal map-making formulisms.

We outline the role of the kernel in gridding (and degridding) in Section 2. In Section 3, we describe the simulation framework and the control, as well as the quality metric and base level of power spectrum errors within our framework. We show our simulated experiments in Section 4, including the role of the kernel extent, systematic propagation into calibration leakage, and effects of the FFT and beam–horizon interactions.

2. Instrumental Response in Gridding Analyses

A visibility is measured per pair of interferometric elements, where the correlation depends on the separation of the elements in measured wavelengths. If all interferometric elements lie on a 2D plane, then the measurement equation can be expressed as a 2D Fourier transform of the plane-projected sky via the van Cittert–Zernike theorem (Carozzi & Woon 2009 for more details). In this Cartesian formalism, it is easiest to express the sky in directional cosines, $\{l, m\}$, and the visibilities in baseline separation in wavelengths, $\{u, v\}$. The measurement equation is then

$$V(u, v) = \int \int \frac{dl \, dm}{\sqrt{1 - l^2 - m^2}} A(l, m)I(l, m)e^{-2\pi i (ul + vm)},$$

(1)

where $V(u, v)$ are the measured visibilities, $A(l, m)$ is the projected instrumental beam response, and $I(l, m)$ is the specific intensity of the projected sky.

The visibilities are the 2D Fourier transform of the projected sky, convolved with the instrumental beam response. Thus, the complicated response of the instrument is encoded in the visibilities (Hamaker et al. 1996).

The formulation described in Equation (1) is an oversimplification. There is a third dimension, $w$, which breaks the Fourier transform relation and thus further complicates the measurement equation and the correct beam response. In this work, we only focus on a truly coplanar response via a snapshot in time to investigate the base level of error expected. Therefore, a 2D Fourier transform can be used with coordinate distortions (Cornwell & Perley 1992). In addition, we further simplify our simulations by assuming there is no frequency dependence on $A(l, m)$ when in reality there is a nonlinear, complicated dependence as a function of frequency.

The process of gridding involves placing all visibilities from a frequency/time measurement onto a single, pixelated $uv$-plane to create a $uv$-histogram. However, visibilities will contribute to multiple $uv$-pixels given their baseline locations and the resolutions required to access modes of interest. Thus, visibilities are added to the histogram with a kernel given by $A(u, v)$, or the instrumental response in $uv$-space. This process can be reversed in an action called degridding. Given a $uv$-plane filled with the contribution of model sources, model visibilities can be generated via integration—again, where the integration kernel used is $A(u, v)$. These model visibilities encode the instrumental response, and thus can be used for calibration or subtraction.

The process of including a kernel is sometimes referred to as convolutional gridding/degridding. Conversely, the contribution of a visibility can be treated as a delta function by disregarding the complicated instrumental sensitivity or, in the specific case of degridding, incorporating the beam via the apparent intensity $A(l, m)I(l, m) \rightarrow I(l, m)$. However, the discrete nature of the $uv$-plane can result in associated errors, like aliasing, without a kernel.

2.1. Optimal Map Making

Optimal map making is a process of lossless data compression which does not destroy cosmological information and produces the smallest (optimal) error bars, initially developed for cosmic microwave background data analysis (Tegmark 1997a). The mathematical formulism of optimal maps has since been adapted for radio interferometry (Morales & Matejek 2009), and is of particular importance to the precise cosmological measurement of the EoR (Dillon et al. 2015b). Optimal map making has been used to analyze data from the Murchison Widefield Array (MWA) (Dillon et al. 2015a; Beardsley et al. 2016), the Hydrogen Epoch of Reionization Array (i.e., The HERA Collaboration et al. 2022), and the Donald C. Backer Precision Array for Probing the Epoch of Reionization (i.e., Ali et al. 2015; Kolopanis et al. 2019).

An important outcome of optimal map making is the ability to propagate errors from the initial measurements through the entirety of the data analysis. Naturally, this means that all operations on the data must be trackable via the variances. This restricts the type of operations that can be used in the optimal map making formulism, and thus has consequences on how we treat the $uv$-plane.

First, the $uv$-plane is only truly optimal when the kernel $A(u, v)$ used in gridding is the actual instrumental response. One can choose to apply an anti-aliasing window function, either directly or as a modification to the instrumental response (i.e., Tapered Gridded Estimator; Choudhuri et al. 2016), to preferentially weight the visibility to the center of the beam at the consequence of higher error bars. While no longer truly
optimal, instrumentation with complicated beams can benefit
from the trade-off (Barry et al. 2019b). However, in order to
track this weighting in the propagation of uncertainties, it
must be applied in the $A(u, v)$ kernel rather than directly in image
space.

This requirement can be visualized simply by the convolu-
tion theorem. If some weighting, $W$, is convolved with the data,
$V$, as is the case with gridding, then the image space operation
is the multiplication of their Fourier transform equivalents,
$\tilde{V}(W)$ and $\tilde{V}(V)$. Thus, this weighting on the data can be
applied in either $uv$-space via a convolution or image space via
a multiplication. However, for the propagated variances, $\sigma^2$,
the weighting must be applied per $\sigma$, and thus there are two
convolved $W$ factors. The same window weighting on the
variances in image space is then

$$
(W * \sigma) \cdot (W * \sigma) \rightarrow (\tilde{V}(W) \cdot \tilde{V}(\sigma)) \star (\tilde{V}(W) \cdot \tilde{V}(\sigma)),
$$

which is notably not $\tilde{V}(W^2) \cdot \tilde{V}(\sigma^2)$.

Image space corrections that are typical of radio astronomy,
like using the conjunction of image space padding with
window functions to remove unstable edges (see van der Tol
et al. 2018 for more details), cannot be used when propagating
variances in an optimal map-making formulism. In this work,
we only apply methodologies that allow for the ability to
propagate variances from the measurement, and thus we are
restricted in the operations that can be used.

3. Simulation Framework

To investigate the effects of the instrumental response in
gridding for an optimal map-making framework, we employ an
in situ simulation technique using a data analysis pipeline. This
allows us to simulate a real analysis response with all the
potential contamination effects while maintaining control over
inputs and experimental variations. This framework was first
introduced in this pipeline by Barry et al. (2016), which goes
into greater detail. In this section, we summarize the key points
necessary to understand the experimental procedure used in
Section 4.

3.1. Analysis Pipeline Source

The analysis framework for our experiment is an open-
source, EoR-specific pipeline composed of (1) Fast Holo-
graphic Deconvolution$^4$ (FHD; Sullivan et al. 2012; Barry
et al. 2019a) and (2) Error Propagated Power Spectra with
Interleaved Observed Noise$^5$ (epsilon; Barry et al. 2019a),
hereafter called FHD/epsilon.

FHD/epsilon has been used to generate competitive EoR
upper limits (Beardsley et al. 2016; Barry et al. 2019b; Li et al.
2019), investigate calibration systems (Barry et al. 2016; Li
et al. 2018; Byrne et al. 2019; Zhang et al. 2020), and probe
subtleties in foreground and radio frequency interference
subtraction (Carroll et al. 2016; Pober et al. 2016; Kerrigan
et al. 2018; Wilensky et al. 2019, 2020; Byrne et al. 2022).
This analysis pipeline has been proven to be able to quantitatively
characterize instrumental effects and spectral precision require-
ments for EoR science.

Of course, how these effects propagate to the final
measurement space depends on the specific analysis techniques
used. For the purposes of our simulations on single observa-
tions, these can be summarized chronologically as the
following steps.

Beam kernel. We generate a hyper-resolved table of the spatial
Fourier transform $(u, v)$ from the instrumental beam
response on the sky in directional-cosine space $(l, m)$. This
flattens the 3D sky response of the instrument into a
2D Fourier representation via orthogonal projection, and is
the estimation for the instrumental contribution to visibility
measurements. For EoR analysis, this is generally
calculated at a high hyper-resolution (Offringa et al.
2019b). We reduce resolution errors further by performing
binlinear interpolation of the hyper-resolved table during
analysis.

Point-source model. Each source’s celestial position and flux
density is analytically Fourier transformed to a uv-grid. All
of the sources’ uv-grids are added together to generate a
single, discrete model uv-plane in the absence of
instrumental effects (Kerrigan et al. 2018). Since there
are commonly over 50,000 sources per field, we hold the
sky fixed in order to only generate one model $uv$-plane per
observation, and correct for a mean spectral index after
model visibility estimation.

Model visibility estimation. We perform degridding of the
singular model $uv$-plane to create model visibilities. In
brief, we multiply the hyper-resolved beam kernel with the
model $uv$-plane at each baseline location and integrate over
all pixels per baseline. Since we held the sky fixed, we
must then move each baseline for each time and frequency
bin and re-integrate. The response of the instrument is
encoded via the beam kernel, and represents an estimation
for how the instrument measures the point-source sky.

Calibration. The gains of each individual station/tile are found
via a linear least-squares reduction of a system of
equations between the data visibilities and the newly
made model visibilities (Mitchell et al. 2008) for baselines
longer than 50 wavelengths (Patil et al. 2016). We do not
constrain spectral smoothness in this work in order to show
the base level of error that must be mitigated.

Discrete $uv$-plane estimation. We then grid each data/model
visibility to a $uv$-plane per frequency. This operation is
performed by multiplying the hyper-resolved beam kernel
with the visibility value and adding the pixelated result to
the $uv$-plane. This is analogous to a histogram of the data,
where the spread of the visibility in the $uv$-plane is
encoded by the beam kernel.

Time-interleaved sets. The $uv$-plane is split into two sets which
are interleaved in time. This choice allows us to estimate
the cross-power spectrum via sums and differences in
addition to other diagnostics (Barry et al. 2019a).

Residual $uv$-plane. Optionally, we subtract the model $uv$-plane
from the data $uv$-plane for each frequency to generate a
residual $uv$-plane.

Frequency Lomb–Scargle transform. We apply a low-sidelobe
window function along each spatial $uv$-pixel and perform a
spectral Lomb–Scargle transform to go to full harmonic
space. The Lomb–Scargle periodogram ensures that we
capture any power that is hidden in the covariances due to
uneven sampling (Lomb 1976; Scargle 1982).

Cross-power estimation and binning. Finally, we combine the
sums and differences of the interleaved time sets in $k$-space

$^4$ https://github.com/EoRImaging/FHD
$^5$ https://github.com/EoRImaging/epsilon
to form the cross-power spectrum and bin to either cylindrical or spherical power spectra.

This is a simplified set of modules; we did not perform image processing with our single, short-time interval observation simulations in order to reduce the complexity of the analysis. Our simulations use a MWA Phase I layout (Tingay et al. 2013), which has a well-covered central core in $uv$-space with slowly decreasing coverage after 50 wavelengths.

3.2. Quality Metric

Our metric for statistical 21 cm EoR detection in this work is the power spectrum. By harnessing the ability to prescribe an almost exact redshift to the narrow line emission of the 21 cm hyperfine transition of hydrogen, we can form measurements of the power for each mode on the sky as a function of $z$.

In this work, we use a reconstructed power spectrum as described by Morales et al. (2019). This entails binning the sky image into a fully perpendicular cube in comoving megaparsecs along the spatial $and$ frequency dimensions. Other forms of power spectra estimation (i.e., delay spectra) and other statistical measurements (i.e., the bispectrum) experience contamination differently, and are not explored in this work.

We bypass creating images and instead directly form a $(u, v, f)$-plane of our simulated data, which only requires a simple coordinate transform to create a $(k_u, k_v, f)$-plane. A further frequency transform (specifically in the case of FHD/epssilon, a Lomb–Scargle periodogram) along with another simple coordinate transform creates a fully perpendicular $k$-cube (Morales & Hewitt 2004). An anti-aliasing window is used during the frequency transform, which reduces unwanted systematics at the cost of a smaller effective bandwidth. Using the assumption of isotropy and homogeneity of the EoR, we average the $k$-cube in spherical shells to create 1D power spectra as a function of $|k_i|$.

Low $k$-modes will be contaminated by the chromatic response of the instrument (Datta et al. 2010; Morales et al. 2012; Parsons et al. 2012; Trott et al. 2012; Vedantham et al. 2012; Hazelton et al. 2013; Pober et al. 2013; Thyagarajan et al. 2013; Liu et al. 2014). Higher $k$-modes are theoretically free from this contamination, and thus this region, known as the EoR window, holds the modes of interest for foreground avoidance techniques. Since our simulations are noiseless and without an underlying EoR signal, any remaining power in high $k$-modes indicates analysis-based systematics. Figure 1 shows the 1D power spectrum for the simulations described in Section 3.3 along with a theoretical realisation of the EoR from Barry et al. (2019b) in units of mK$^2$. These simulations have varying degrees of error, but all could theoretically measure the 21 cm signal in the EoR window since their level of error is below the expected power.

3.3. In Situ Simulation

The FHD/epssilon pipeline was built to analyze observations, but can also be harnessed as an in situ instrumental simulation to test for various spectral precision effects.

By default, FHD/epssilon runs an instrumental simulation with point-source foregrounds for every analysis run—this constitutes the model visibilities. These model visibilities encode the instrument via the beam kernel, and are made with typically 50,000 known sources or more. We can use these model visibilities as an input into the pipeline to perform an ad hoc, in situ simulation.

The now-input visibilities are treated the same as data; all of the steps described in Section 3.1 are performed, and new model visibilities are made alongside our input visibilities. By varying the amount of known point-sources in the new model, we can simulate our lack of knowledge of the flux density of the sky.

More importantly, we can simulate our lack of knowledge about the instrumental beam response. This will form the basis of our experiments in Section 4. We can create the input simulated visibilities using a specific beam kernel, and then during the in situ simulation, use a different, modified beam kernel in analysis. Errors that arise from using a gridding kernel that is different than the encoded instrumental response will then become apparent in the power spectra. When compared to a control run, we can then quantify the excess error.

3.3.1. Gaussian Decomposition and Semi-analytic Input

The input model visibilities are treated as our simulated data in the FHD/epssilon in situ simulation. These were created from a model $uv$-plane of an EoR0 target field (centered around R.A. = 0°, decl. = −27°) zenith observation with 49,572 point sources from the GLEAM catalog (Hurley-Walker et al. 2017). There is no thermal noise, no ionospheric effects, and no diffuse emission.

The choice of instrumental beam kernel when degridding this model $uv$-plane will determine the encoded instrumental response in the input visibilities. Thus, we want to ensure that we are including as few known systematics as possible. We create a new representation of the beam for this purpose—the Gaussian decomposition of the instrumental sky response and its analytic Fourier transformation.

We create a hyper-resolved directional cosine image from outputs of a FEKO simulation of an averaged embedded element beam with mutual coupling (Sutinjo et al. 2015). This is a 2D representation of the sky response, and is the Fourier
dual of our desired $uv$-space instrumental kernel. However, instead of performing FFTs to transform the instrumental model, we fit a series of Gaussians to the image to create an analytic decomposition of the instrumental sky response. Thus, the Fourier transform is fully analytic and calculable.

The sum of analytic Fourier transforms of each Gaussian produces an alias-free $uv$-kernel. In addition, we can calculate the exact beam values at every pixel for each unique baseline location using the analytic decomposition function, rather than relying on interpolation or nearest-neighbor methods to calculate the beam values per pixel from a discrete kernel. We use this to reduce gridding-resolution effects, in addition to calculating beam values for all included pixels above machine-level precision. This is a semi-analytic formulism; the visibilities are analytic in every way except that the $uv$-plane to be degridded is inherently pixelated. Unfortunately, this semi-analytic formulism is not a realistic option for large-scale data analyses for extreme widefield instruments, and thus we only use this for our input.

Other functions can be chosen in place of Gaussians as long as they are analytic, such as prolate spheriodals, Bessel functions, or spherical harmonics. The main reason for choosing Gaussians with our simulation is that they are extremely simple while still achieving the desired end result. In the future, finding an analytic form of the projection of spherical harmonics to 2D Cartesian $uv$-space would avoid fitting procedures since FEKO simulations can output spherical harmonics. However, this is non-trivial, potentially very time-consuming, and adds unnecessary complexity to our basic simulations.

The Gaussian decomposition of the instrumental sky response is composed of 18 individual Gaussians, and fits the FEKO simulation to within 0.3%. Nevertheless, one important feature is excluded: the horizon. Gaussians are infinite and thus the Gaussian decomposition of the beam has no horizon. The effect of including a horizon is explored in Section 4.3.

### 3.3.2. Discrete Control and Associated Errors

The control run determines the level of base error that is expected when the 2D Gaussian decomposition of the beam is used throughout the analysis. Whatever systematics are present in the control will also be present in the experiments. To first order, we can assume these systematics are additive in the experiment. Given our simplified approach (i.e., no imaging with a smooth $uv$-coverage and a perfectly 2D array), this basic assumption should hold.

For our control, we use a representative gridding resolution, which will result in a base level of error. For these experiments, we use a beam with bilinear interpolation to pixel centers, hyper-resolved by a factor of 250, resulting in the control simulation (blue) in Figure 1. If we were to decrease the resolution by five times (yellow, Figure 1), however, we would still be below the level of the EoR.

Throughout the analysis, we perform calculations at double precision. If the gridding operation is performed at floating-point precision (green, Figure 1), this would have resulted in a small increase in error. We also use a seven-term Blackman–Harris (BH) spectral window function in transforming from $\{k_x, k_y, f\}$ to $\{k_x, k_y, k_z\}$ due to the modest increase in the base level of errors with a four-term BH spectral window function (red, Figure 1).

We can optionally allow the in situ simulation to subtract 11,188 of the brightest sources in the sky as seen by the instrument, or roughly 70% of the apparent flux density (purple, Figure 1). This represents a realistic subtraction of flux density on the sky that we can expect to achieve in real data analysis. In addition, by only subtracting a subset of sources, we enable a realistic dynamic range problem. Errors that arise in the analysis that couple the foregrounds to the measurement modes are only observable when foregrounds or foreground residuals are present. As seen in Figure 1, we finally reach an unknown, varying systematic with the foreground subtraction that is six orders of magnitude below the expected EoR power.

The resulting error in the measurement modes of interest is the base level expected in Sections 4.1 and 4.2 (blue for unsubtracted power, purple for subtracted power, Figure 1). For Section 4.3, we relax the hyper-resolution parameters for computational reasons (yellow, Figure 1). Any additional errors can be quantified as resulting from the specific variable changes in our analysis.

### 4. Simulated Experiments

We present three simulated experiments to investigate the effects of the gridding/degridding kernel on the EoR power spectrum, namely the choice of gridding kernel in optimal map making (Section 4.1), calibration bias (Section 4.2), and inclusion of the horizon (Section 4.3).

Each experiment uses three types of $uv$-kernels: the Gaussian decomposition of the beam (top row, Figure 2), the Gaussian decomposition of the beam with all components at half power except the primary lobe component (second row, Figure 2), and a single Gaussian matched to the FWHM of the primary lobe (bottom row, Figure 2). These kernels represent a perfect instrumental beam, an imperfect instrumental beam, and a non-instrumental beam, respectively. As seen in Figure 2, the presence of sidelobes causes an interference pattern in the analytic Fourier transform.

We only show the E–W instrumental amplitude response for brevity.

#### 4.1. Gridding Kernels in Optimal Map Making

The $uv$-kernel is infinite in extent yet highly peaked, and therefore a choice must be made in producing a finite representation. This choice is influenced by efficiency requirements; the larger the kernel extent, the more pixels involved in operations that are repeated millions of times, and therefore a compact yet accurate representation is preferred.

However, choosing a finite extent can create a sharp discontinuity in the $uv$-kernel at the boundary. As baselines move as a function of frequency, the pixels in the $uv$-plane that contribute will change. Thus, if there exists a sharp discontinuity in the $uv$-kernel, there will be abrupt transitions as a function of frequency, which will propagate to the power spectrum as a systematic.

A beam kernel cut is equivalent to convolving a sinc-like function in image space. This will undoubtedly generate contamination in the power spectrum. However, our options are limited. We cannot apply image-space corrections easily, a common mitigation technique, due to our choice of optimal map making (see Section 2.1). In addition, we cannot realistically grid all pixels for every single baseline given the extreme widefield nature of the MWA sidelobes. Therefore, the
The best option for optimal map making is to find a level of contamination caused by the extent of the beam kernel that is sufficiently below the power level of the EoR. This can be partially mitigated if the extent is chosen to be a contour. The beam value of the contour can be subtracted off the entire kernel to reduce the step to zero, and then the $uv$-kernel can be renormalized. This deforms the $uv$-kernel slightly, making it less instrumental at the cost of reducing systematics. Nevertheless, this technique will not change the fact that the discontinuity remains. We apply this mitigation method to show the minimum level of error that is expected. The effects of the small deformation via the renormalization appear minimal, as will become evident.

We choose three different extents to highlight the dependence on the edge discontinuity, as shown in third column of Figure 2. For the Gaussian decomposition (top right), these occur at contours of 0.01% (red), 0.0001% (green), and 0.000001% (purple) of the maximum $uv$-kernel value, corresponding to approximately $10\lambda$, $20\lambda$, and $30\lambda$ extents in the $u$ direction, respectively. In order to reproduce similarly behaved responses in the other beam kernel types, we calculate the mean derivative along the subtracted contour edge in the

**Figure 2.** Three types of projected image-space beam shapes (left column) used in this work and their respective analytic 2D Fourier transforms (amplitude only, middle column) along with derivative-matched extents (right column; red: 0.01%, green: 0.0001%, purple: 0.000001% of the Gaussian decomposition maximum). The top row is the Gaussian decomposition of the instrumental Murchison Widefield Array beam, accurate to within 0.3%, which is used as our control in our experiments. The middle row is a Gaussian decomposition where all components except the primary lobe component have been reduced by half. The bottom row is a single Gaussian with a FWHM matched to the instrumental beam. These kernels represent a perfect instrumental beam, an imperfect instrumental beam, and a non-instrumental beam, respectively. Fewer pixels are required for comparable levels of power spectrum contamination for the non-instrumental beam.
analyses could expect Gaussian decomposition as shown in Figure 2. Power must be within the contours correspond approximately the same mean edge-derivative as the Gaussian decomposition density subtracted to be representative of measurements. The extents of Figure 3. 1D power for the various extents and shapes with 70% of the flux density subtracted to be representative of measurements. The extents of the Gaussian decomposition (solid lines) are 0.01% (red), 0.0001% (green), and 0.000001% (purple) of the maximum beam value. For the half-power sidelobe beam (dotted lines) and matched-FWHM Gaussian beam (dashed lines), the contours correspond approximately the same mean edge-derivative as the Gaussian decomposition as shown in Figure 2. Power must be within the shaded region in order to theoretically detect the EoR. Thermal noise (1σ, yellow) for 333 hr of data for each beam shape shows there is very little degradation of optimal uncertainties.

Gaussian decomposition kernel and then find the corresponding contour value for the kernel in question. The edge derivatives are therefore roughly matched between the different kernel shapes.

Matched-derivative contours result in vastly different kernel extents. The single, matched-FWHM Gaussian (bottom right, Figure 2) has a very compact kernel where the maximum extent is less than 15 λ. This is extremely useful in gridding/degridding analyses, as the number of pixels required per visibility is four times fewer. In contrast, the kernel with half-power sidelobes (middle right, Figure 2) has not changed dramatically in extent, only decreasing by a few wavelengths. This indicates that it is difficult to reduce the contribution of sidelobes to the uv-kernel extent, and that even a 50% decrease in sidelobe contribution has very little effect.

The kernel extent due to widefield sidelobes has consequences in instrumental design. Significantly spaced sidelobes create an interference pattern in Fourier space, which extends the contribution of the kernel out much further than if those sidelobes were not present. Decreasing the sidelobes by 50%, a technical feat in-of-itself when designing an instrument, would not significantly change this extent. However, if sidelobes were able to be sufficiently suppressed, $O(N^2)$ analyses could expect a speed-up factor of 16 given the reduction of pixels per visibility required in analysis.

The resulting power spectra from each uv-kernel shape and extent are shown in Figure 3. In each case, the input simulation was calculated with a Gaussian decomposition beam in our semi-analytic formulism described in Section 3.3.1. We have removed 70% of the flux density through a degridded model. These power spectra are without calibration errors, and therefore only probe the effect of the uv-kernel specified in gridding/degridding.

The high $k$-modes, or EoR window, decrease in power as the extent is increased for each kernel shape. The Gaussian decomposition kernel and the kernel with half-power sidelobes decrease at roughly the same rate given a derivative-matched contour, whereas the matched-FWHM Gaussian kernel decreases quicker. This can be attributed to, in part, the differences in the derivative around the contour in the Gaussian decomposition kernel, which vary and are not fully captured by the mean. All kernel shapes could theoretically detect the EoR with a derivative-matched contour of $\sim 0.0001\%$ for a majority of EoR window modes.

The low $k$-modes, however, indicate the importance of faithfully reconstructing the beam in foreground removal. The Gaussian decomposition kernel reduced the foregrounds more effectively than the other kernel shapes because it better represented the input, and thus has lower power on low $k$-modes. Nevertheless, there is still a significant amount of power left in this region because the Gaussian decomposition kernel used did not perfectly match the input, which was semi-analytic in nature. This may indicate that optimal map-making analyses that use pixelated gridding/degridding procedures will be unable to remove enough power to measure the EoR on low $k$-modes without significantly finer-resolution $uv$-planes. Additional procedures, like optimal map making with covariance weighting (e.g., Ali et al. 2015) could potentially remove enough power, but may suffer from signal loss (Cheng et al. 2018).

Misrepresenting the beam to reduce systematics at high $k$-modes (for example, via the single matched-FWHM Gaussian beam) essentially results in a convolutional image-space weighting (Choudhuri et al. 2016). This may result in non-optimal errors (Tegmark 1997a, 1997b), and no longer give the orthographic projection of the sky. Nevertheless, it reduces extent-based systematics and requires fewer pixels during computation.

We highlight the change in the 1σ thermal noise level due to the various beam kernel shapes (yellow, Figure 3) given 333 hr of data. Even though the beam kernel with half-power sidelobes and the matched-FWHM Gaussian kernel produce non-optimal errors, their difference from the optimal errors of the Gaussian decomposition beam kernel is negligible. While the beam shapes are quite different from one another past $\sim 15 \lambda$, the majority of the beam volume comes from the center of the beam where they are almost identical. Thus, the 1σ thermal noise level for the matched-FWHM Gaussian kernel is only 8% higher than optimal.

Out of the three kernel shapes, the matched-FWHM Gaussian kernel is preferred during the gridding process due to its ability to reduce systematics while requiring fewer pixels during computation. This does come at the cost of non-optimal errors, but the change is almost negligible. While the framework of optimal map making is extremely useful for reconstructing maps with fully propagated uncertainties, it is more practical to be non-optimal to reduce widefield sidelobe effects in the analysis load.

4.2. Calibration Errors with Misrepresentations of the Instrument

Given the apparent success of compact, non-instrumental $uv$-kernels in the gridding procedure of Section 4.1, it may be tempting to use a similar $uv$-kernel to generate model visibilities for calibration through degridding. However, using a non-instrumental beam is equivalent to a weighting in image space, and thus the model visibilities would no longer represent the instrument, leading to an inaccurate calibration.
expected modeled perfectly results in little-to-no calibration errors in the EoR window as decomposition beam control input beam in degridding. Using the semi-analytic Gaussian Calibration systematic errors caused by misrepresentations of the Figure 4.
The Astrophysical Journal, power sidelobe beam delled sources fl Section 3.3.1. For this experiment, we will also use calibration due to misrepresentations of the instrument, we now measurement discrete Gaussian decomposition has contamination the precludes an EoR (beam fl red, Figure 4). As expected, there are yellow, Figure 4 more at high k-modes. Even though this kernel was preferred in gridding within an optimal map-making framework for its suppression of EoR window spectral contamination and for its compactness, it cannot be used for degridding. This again reiterates that calibration is an instrumental procedure and thus the degridding kernel used must represent the instrument.

However, a further complication arises when the degridding procedure uses our typical hyper-resolved beam like in Section 4.1 instead of the semi-analytic calculation. This fully discrete formulism results in EoR-precluding calibration errors, even with the Gaussian decomposition beam kernel (dashed green, Figure 4). This hyper-resolved beam kernel was at the same resolution as those that were sufficient for gridding. It is apparent from this contamination that the requirements stated in Offringa et al. (2019a) for gridding are not necessarily the same as the requirements for degridding calibration schemes.

More sophisticated calibration schemes will of course do better at reducing spectral contamination into high k-modes. This will depend on the level of flux density modeled, the refinement of the frequency regularization, and the complexity of the instrument in question. For example, if the frequency dependence of an instrument can be accurately represented by a low-order polynomial fit, then most spectral contamination in the EoR window can be avoided (Barry et al. 2016).

Nevertheless, our experiments shown in Figure 4 do not bode well for degridding calibration schemes. Representing the instrumental kernel well enough will prove difficult, especially given its complex and time-varying nature (Joseph et al. 2020; Chokshi et al. 2021). Even if we could capture its response, we would have to greatly increase the resolutions of degridding to levels that we may not be able to achieve. This suggests a paradigm shift in our model generation and calibration procedures.

Thankfully, model visibility generation is not beholden to optimal map making and being able to propagate errors. Thus, degridding does not necessarily have to be used—as long as the result is instrumental model visibilities, any procedure will work in theory. In addition, new calibration schemes are helping to remove the dependence on model visibilities altogether.

4.3. The Effects of the Horizon

One important aspect of the gridding kernel that has been omitted in this work thus far has been the horizon. Its inclusion into our analytic framework is possible, yet unsustainable on a case-by-case basis for real observational data given that it requires a pixel-by-pixel convolution. Nevertheless, it reveals detriments for widefield telescopes and traditional optimal map-making analysis techniques.

In flat lm-space, the horizon is a circular step function with values of 1 inside a radius of 1, and with values of 0 elsewhere. We multiply this by the beam kernel formulations to generate more realistic representations, such as the left panel of Figure 5. Thus, the analytic Fourier transform is the convolution of the Gaussian decomposition and the circular step function in Fourier space (top panel, Figure 5). We reduce the kernel resolution by 5× in order to perform the analytic convolution

Figure 4. Calibration systematic errors caused by misrepresentations of the control input beam in degridding. Using the semi-analytic Gaussian decomposition beam (see Section 3.3.1) with all the flux density on the sky modeled perfectly results in little-to-no calibration errors in the EoR window as expected (purple). Modifying the degridding kernel to the semi-analytic half-power sidelobe beam (yellow) or the semi-analytic matched-FWHM Gaussian beam (red) results in catastrophic calibration errors. However, even the typical, discrete Gaussian decomposition has contamination the precludes an EoR measurement (dashed green).

In order to quantify the resulting errors that occur in calibration due to misrepresentations of the instrument, we now modify the beam kernel used for degridding. Again, the simulated data visibilities are calculated with a Gaussian decomposition beam in our semi-analytic formulism in Section 3.3.1. For this experiment, we will also use all of the flux density on the sky in calibration to bypass errors arising from frequency-dependent point-spread functions of unmodelled sources (Barry et al. 2016). Any spectral structure in the calibration will thus be solely due to misrepresentations of the beam kernel.

We use a simple calibration procedure: a linear least-squares solution between the data and model visibilities to get the resulting gain parameters of each interferometric element per frequency (Mitchell et al. 2008). In general, this represents the most basic form of calibration and will give us an idea of systematics that will then need to be mitigated through more sophisticated processes like frequency regularization.

As a control, we first test the resulting errors in calibration when the semi-analytic Gaussian decomposition is used. Again, the semi-analytic formulism uses the exact analytic Fourier transform of the beam kernel for all contributing pixels rather than the hyper-resolved lookup table. As expected, there are little to no errors present in calibration, and there are many accessible high k-modes (purple, Figure 4). No flux was subtracted in this test or any of the following experiments in order to compare to this base level of error; otherwise, there is essentially no power left in the control since everything subtracts almost perfectly.

We now modify the degridding beam shape to the semi-analytic beam kernel where all Gaussian components except the main lobe have been reduced by half (yellow, Figure 4). All k-modes are contaminated by calibration errors, and the EoR cannot be measured. This is unfortunate—it means that in order for a simple calibration to work for EoR science, even the sidelobes of the instrumental beam must be represented well.

Modifying the degridding beam shape to the semi-analytic single matched-FWHM Gaussian creates even more calibration error (red, Figure 4). When compared to the semi-analytic half-power sidelobe kernel, this error is between 50% and 200% more at high k-modes. Even though this kernel was preferred in gridding within an optimal map-making framework for its suppression of EoR window spectral contamination and for its compactness, it cannot be used for degridding. This again reiterates that calibration is an instrumental procedure and thus the degridding kernel used must represent the instrument.

Thankfully, model visibility generation is not beholden to optimal map making and being able to propagate errors. Thus, degridding does not necessarily have to be used—as long as the result is instrumental model visibilities, any procedure will work in theory. In addition, new calibration schemes are helping to remove the dependence on model visibilities altogether.

4.3. The Effects of the Horizon

One important aspect of the gridding kernel that has been omitted in this work thus far has been the horizon. Its inclusion into our analytic framework is possible, yet unsustainable on a case-by-case basis for real observational data given that it requires a pixel-by-pixel convolution. Nevertheless, it reveals detriments for widefield telescopes and traditional optimal map-making analysis techniques.

In flat lm-space, the horizon is a circular step function with values of 1 inside a radius of 1, and with values of 0 elsewhere. We multiply this by the beam kernel formulations to generate more realistic representations, such as the left panel of Figure 5. Thus, the analytic Fourier transform is the convolution of the Gaussian decomposition and the circular step function in Fourier space (top panel, Figure 5). We reduce the kernel resolution by 5× in order to perform the analytic convolution
in a reasonable time, which increases the base level of expected error (from blue to yellow, Figure 1).

When the horizon is included, the $uv$-kernel is infinitely extended above floating-point precision due to the interaction of the secondary sidelobes with the horizon edge. While the coordinate distortion caused by the orthogonal projection is an issue (e.g., Brouw 1975), it is not the root cause of this effect. Rather, it is the abrupt transition between beam sensitivity on the sky and an opaque horizon. Regardless of coordinate basis, an abrupt transition creates an infinite response in the Fourier-dual equivalent.

The FEKO simulation used to model the MWA beam does in fact have this abrupt transition since it uses an infinite plane of soil to model the instrumental response (Sutinjo & Hall 2013; Sokolowski et al. 2017). In fact, it is more typical for FEKO simulations to use finite metallic ground planes, not just infinite soil (Bolli et al. 2022), to reduce complexity at the cost of low elevation accuracy (Bolli et al. 2020). Whether the abrupt edge is a real effect or purely simulation-based is unknown. In our simulations, the beam sensitivity at the horizon is 0.3% at its largest, highlighting the severity of this effect in $uv$-space.

This infinite extent is detrimental for gridding/degridding analyses, particularly for optimal map making, because it is impossible to represent the instrumental $uv$-kernel well. In fact, we can only perform limited experiments with this kernel since there is no way to include a ground truth (or simulation thereof) as our control. By gridding all foreground sources, we can probe the expected gridding systematics with the analytic convolution kernel, represented out to $30\,\lambda$ in extent.

The resulting power spectrum for the Gaussian decomposition (solid blue, Figure 6) shows contamination slightly below the EoR level on high $k$-modes. This is unsurprising given the harsh derivatives at the edge of the kernel, even at $+/−15\,\lambda$. While this can be mitigated by windowing or tapering the $uv$-kernel to deviate from an instrumental representation, we cannot predict how non-optimal the resulting errors would be. However, given that the kernel is infinite when there is an abrupt transition in beam sensitivity at the horizon, the kernel used in gridding would have to be finite and thus non-optimal.
in any case. Optimal map making is thus not achievable in the presence of a beam–horizon interaction.

In addition, there is no way to mitigate this contamination for simple calibration without incurring errors reminiscent of those in Figure 4. However, we cannot perform the necessary experiment to constrain requirements of calibration in the presence of the horizon due to the limitations of the in situ simulation framework.

This contamination can be reduced, however, if the instrumental beam kernel has smaller secondary sidelobes. For the Gaussian decomposition with half-power sidelobes, this high $k$-mode power is now lower (dotted blue, Figure 6). If there is little to no beam sensitivity at the horizon, as is the case for the matched FWHM beam, then the contamination traces out the base level of error expected given this simulation resolution (dashed blue, Figure 6).

A further complication arises when FFTs are used to generate the $uv$-kernel. In contrast to the analytic convolution, using FFTs is vastly more efficient and requires significantly less computational load. However, the horizon edge can cause obvious aliasing in the $uv$-domain if the resolution of the image-space beam is not high enough. The base resolution, which is arguably rather small (3000 pixels in the $l$ direction), causes aliasing contamination as seen in the bottom left panel of Figure 5. Increasing the resolution by $10 \times$ (30,000 pixels in the $l$ direction) considerably reduces this contamination (bottom middle panel of Figure 5), and increasing the resolution by $20 \times$ (60,000 pixels in the $l$ direction) essentially reduces the extra contamination completely (bottom right panel of Figure 5).

Aliasing in the $uv$-kernel via the FFT can contribute to the power spectra. While easily fixed with image-space filtering, this cannot be done for optimal map-making formulisms. The kernel created with the base resolution FFT contributes approximately an order of magnitude higher contamination on high $k$-modes for the beams with sidelobes (solid and dotted orange, Figure 6). Of course, when there is little to no beam sensitivity at the horizon, as is the case for the FWHM Gaussian beam, then there is no extra contamination (dashed orange, Figure 6). Fortunately, the higher-resolution FFT kernels presented in Figure 5 do not add significant contamination to the power spectra and are omitted for clarity. The level of contamination expected given a $uv$-kernel of 30 $\lambda$ can be reached with approximately $5 \times$ image-space resolution.

Thus, not only is $uv$-kernel resolution important as explored in Offringa et al. (2019a), but image-space resolution can also contribute to contamination if not properly handled within an optimal map-making framework.

Regardless of whether or not the Fourier transform is performed analytically or with an FFT, a beam–horizon interaction leads to a kernel that is infinitely extended. Therefore, it is essentially impossible to represent the instrumental beam well. Thus, if optimal map-making and/or simple calibration procedures are to be used for EoR studies, then the instrument must have little to no beam–horizon interaction.

5. Discussion

The SKA era promises to push the boundaries of our current understanding of the universe, and with it our current analysis methodologies. The EoR science case is an extreme example of required precision and accuracy, in both analysis and measurement, and is thus a benchmark toward the future success of the SKA. Studies have already begun to re-evaluate traditional radio astronomy analysis in the context of the EoR.

In particular, traditional radio astronomy imaging techniques rely on a spreading function, or kernel, in pixelated $uv$-space. Extreme resolution is required to reduce spectral structure caused by the movement of baselines convolved with the $uv$-kernel (Offringa et al. 2019a). However, other aspects of the kernel, including its extent, shape, and horizon, also have an important role in reconstructed power spectra analyses. This work extends that of Offringa et al. (2019a) to explore these other aspects and how they propagate to calibration, subtraction, and our final measurements, particularly for optimal map making.

To test these effects, we built a semi-analytic framework within a current, working EoR pipeline for the MWA. We introduced a Gaussian decomposition of the image-space beam in directional cosine space which allows us to analytically transform and easily modify the gridding shape. The base level of error expected from this framework is $10^{13} \text{mK}^2$ smaller than the foregrounds, allowing for an unhindered exploration of the aforementioned effects.

Our first experiment quantifies the effect of choosing a gridding kernel shape in optimal map making. This choice is a balance between reducing systematics caused by the edge derivative, increasing the number of pixels required for each gridded visibility, and increasing uncertainties due to non-optimal beam shapes.

For an MWA-style Gaussian decomposition of the beam, the $uv$-kernel must extend out to at least 0.001%–0.0001% of the maximum value in order to measure the EoR at high $k$. Decreasing the sidelobes to half the power does not change this result drastically. However, a kernel with no sidelobes has a much more compact $uv$-kernel, which results in lower high-$k$ systematics for a given kernel size. Even though this is a misrepresentation of the input beam kernel shape, it does not affect high-$k$ modes in the power spectrum, which bodes well for foreground avoidance techniques using gridding. In addition, this only increases uncertainties by roughly 8% compared to the optimal case.

Interestingly, all kernel shapes perform poorly at small $k$-modes where foregrounds are prevalent, even the Gaussian decomposition of the beam. The resolution of the Gaussian decomposition of the beam did not match the input, which was semi-analytic in nature, and thus the direct foreground subtraction performed poorly. Given this result and those of Offringa et al. (2019a), simple foreground removal techniques will most likely not reveal the EoR within foreground-dominated regions given current computing resources, and thus more advanced methods will most likely need to be employed.

Our second experiment further supports this point. We use a traditional calibration scheme that calculates a gain parameter for every frequency channel using a linear least-squares reduction between input visibilities and a generated model. Even with 100% of the foregrounds in the model visibilities, the base level of calibration error is orders of magnitude higher than the EoR signal, again due to the difference between the semi-analytic input visibilities and the high-resolution, yet discrete degridded model visibilities.

We also explore the consequences of misrepresentations of the beam kernel used to generate the input visibilities during
the calibration process. An incorrect beam shape drastically increases calibration error such that an EoR measurement is precluded. This is true for both the single Gaussian and the Gaussian decomposition with reduced sidelobes, indicating the importance of correctly modeled sidelobes in calibration. Instrumental beams are vital in calibration procedures to help reduce the base level of error in sky-based, traditional calibration.

Including the horizon in our framework introduces even more potential challenges. If there is beam sensitivity which interacts with the horizon, even on a small but non-floating-point precision level, then the resulting uv-kernel is infinitely extended above machine precision. This cannot be represented well in a discrete, finite analysis, and it will always have high edge derivatives. This generates systematics at about the EoR signal at high k-modes in the power spectrum, but this can be easily mitigated by choosing a non-optimal compact kernel. In any case, optimal map making is not an option when there is a beam–horizon interaction because the infinite kernel cannot be represented. However, this poses a potential problem for traditional calibration procedures where the instrumental kernel is not compact.

The inclusion of the horizon in optimal map making frameworks also complicates the use of FFT algorithms, which are discrete and thus cause aliasing in the presence of a large derivative. Compared to the analytic convolution, the FFT of the image-space beam can generate systematics up to an order of magnitude greater. This can be mitigated by increasing the resolution on the image-space beam. However, it should be noted that the limitations of the in situ simulation framework cannot probe quantitative requirements in this case, as the true uv-kernel is infinitely extended and cannot be represented.

The simulations presented in the work indicate that gridding is sufficient for foreground avoidance with reconstructed power spectra (Section 4.1), but that it quickly fails when used in other contexts. Traditional calibration requires degridding resolutions which we cannot obtain (Section 4.2). In addition, inclusion of a horizon for widefield instrument creation creates infinite uv-kernels, even when beam sensitivity is as low as 0.3% at the horizon (Section 4.3). Given this evidence, it is becoming clear that gridding is best used only as a histogramming step, and that using gridding/degridding to calibrate or perform truly optimal map making is relatively unobtainable for EoR reconstructed power spectra studies using widefield instrumentation.

This work, in addition to Offringa et al. (2019a), indicates that advancements beyond traditional radio astronomy imaging methods are needed in order to measure the EoR. Progress over the past decade has started to decouple EoR analysis methodologies from the narrowfield assumptions that built the foundations of radio astronomy. While it is still possible to use techniques, namely gridding/degridding, complications arise due to the widefield nature of most EoR interferometers and the spectral requirement for calibration of power spectra analyses, especially within an optimal map-making framework. By using an in situ simulation, we explored various aspects of the gridding/degridding kernel in this work, and planned further work will assess mitigation strategies in observational data.

The authors would like to thank Cathryn Trott, Bart Pindor, Randall Wayth, Daniel Mitchell, and Miguel Morales for their informative discussions, and Chris Riseley for his advice. N.B. acknowledges the support of the Forrest Research Foundation, under a postdoctoral research fellowship. This research was supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013. This work was supported by resources awarded under Astronomy Australia LTDs merit allocation scheme on the OzSTAR national facility at Swinburne University of Technology. OzSTAR is funded by Swinburne University of Technology and the National Collaborative Research Infrastructure Strategy (NCRIS). The International Centre for Radio Astronomy Research (ICRAR) is a Joint Venture of Curtin University and The University of Western Australia, funded by the Western Australian State government.

**ORCID iDs**

Nichole Barry https://orcid.org/0000-0003-2064-6979

Aman Chokshi https://orcid.org/0000-0003-1130-6390

**References**

Ali, Z. S., Parsons, A. R., Zheng, H., et al. 2015, *ApJ*, 809, 61

Barry, N., Beardsley, A. P., Byrne, R., et al. 2019a, *PASA*, 36, e026

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

Barry, N., Wilensky, M., Trott, C. M., et al. 2019b, *ApJ*, 884, 1

Beardsley, A. P., Hazelton, B. J., Sullivan, I. S., et al. 2016, *ApJ*, 833, 102

Bolli, P., Bereczi, M., Ninni, P. D., Labate, M. G., & Viron, G. 2020, in 14th European Conf. on Antennas and Propagation (EuCAP) (Copenhagen, Denmark: IEEE), 1

Bolli, P., Davidson, D. B., Bereczi, M., et al. 2022, *JATIS*, 8, 011017

Brouw, W. 1975, in Methods in Computational Physics: Advances in Research and Applications, ed. B. Alder, S. Fernbach, & M. Rotenberg (Amsterdam: Elsevier), 131

Byrne, R., Morales, M. F., Hazelton, B., et al. 2019, *ApJ*, 875, 70

Byrne, R., Morales, M. F., Hazelton, B., et al. 2022, *MNRAS*, 510, 2011

Carozzi, T. D. 2015, *MNRAS Lett.*, 451, L6

Carozzi, T. D., & Woon, G. 2009, *MNRAS*, 395, 1558

Carroll, P. A., Line, J., Morales, M. F., et al. 2016, *MNRAS*, 461, 4151

Cheng, C., Parsons, A. R., Kolopanis, M., et al. 2018, *ApJ*, 868, 26

Chokshi, A., Line, J. L. B., Barry, N., et al. 2021, *MNRAS*, 502, 1990

Choudhuri, S., Bharadwaj, S., Roy, N., Ghosh, A., & Ali, S. S. 2016, *MNRAS*, 459, 151

Cooley, J. W., & Tukey, J. W. 1965, *Math. Comput.*, 19, 297

Cornwall, T., & Perley, R. 1992, *A&A*, 261, 353

Cornwell, T. J., Golap, K., & Bhatnagar, S. 2008, *ISTSP*, 2, 647

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

Dillon, J. S., Neben, A. R., Hewitt, J. N., et al. 2015a, *PhRvD*, 91, 123011

Dillon, J. S., Tagmark, M., Liu, A., et al. 2015b, *PhRvD*, 91, 023002

Furlanetto, S. R. 2019, The Cosmic 21 cm Revolution, 2514–3433 (Bristol: IOP Publishing), 3

Hamaker, J. P., Bregman, J. D., & Sault, R. J. 1996, A&AS, 117, 137

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

Hurley-Walker, N., Callingham, J. R., Hancock, P. J., et al. 2017, *MNRAS*, 464, 1146

Jelic, V., Zorouhi, S., Labropoulos, P., et al. 2008, *MNRAS*, 389, 1319

Joseph, R. C., Trott, C. M., Wayth, R. B., & Nasirudin, A. 2020, *MNRAS*, 492, 2017

Kerrigan, J. R., Pober, J. C., Ali, Z. S., et al. 2018, *ApJ*, 864, 131

Kolopanis, M., Jacobs, D. C., Cheng, C., et al. 2019, *ApJ*, 883, 133

Li, W., Poher, J. C., Barry, N., et al. 2019, *ApJ*, 887, 141

Li, W., Poher, J. C., Hazelton, B. J., et al. 2018, *ApJ*, 863, 170

Liu, A., Parsons, A. R., & Trott, C. M. 2014, *PhRvD*, 90, 023018

Lomb, N. R. 1976, *Ap&SS*, 39, 447

Mellema, G., Koopmans, L. V. E., Abdalla, F. A., et al. 2008, *ExA*, 36, 235

Mirocha, J. 2019, The Cosmic 21 cm Revolution, 2514–3433 (Bristol: IOP Publishing), 2

Mitchell, D. A., Greenhill, L. J., Wayth, R. B., et al. 2008, *ISTSP*, 2, 707

Morales, M. F., Beardsley, A., Poher, J., et al. 2019, *MNRAS*, 483, 2207

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