




# Detecting the 21 cm Signal from the Epoch of Reionization Using Drift Scans: Correlation of Time-ordered Visibilities

Akash Kumar Patwa  and Shiv Sethi

Raman Research Institute, C. V. Raman Avenue, Sadashivanagar, Bengaluru 560080, India; [akpatwa@rri.res.in](mailto:akpatwa@rri.res.in)

Received 2019 May 7; revised 2019 October 15; accepted 2019 October 23; published 2019 December 10

## Abstract

We present a formalism to extract the HI power spectrum from the epoch of reionization for drift scans using radio interferometers. Our main aim is to determine the coherence timescale of time-ordered visibilities. We compute the two-point correlation function of the HI visibilities measured at different times to address this question. We determine, for a given baseline, the decorrelation of the amplitude and the phase of this complex function. Our analysis uses primary beams of four ongoing and future interferometers—Precision Array for Probing the Epoch of Reionization, Murchison Widefield Array, Hydrogen Epoch of Reionization Array, and Square Kilometre Array (SKA1-Low). We identify physical processes responsible for the decorrelation of the HI signal and isolate their impact by making suitable analytic approximations. The decorrelation timescale of the amplitude of the correlation function lies in the range of 2–20 minutes for baselines of interest for the extraction of the HI signal. The phase of the correlation function can be made small after scaling out an appropriate term, which also causes the coherence timescale of the phase to be longer than the amplitude of the correlation function. We find that our results are insensitive to the input HI power spectrum, and therefore, they are directly applicable to the analysis of the drift scan data. We also apply our formalism to a set of point sources and statistically homogeneous diffuse correlated foregrounds. We find that point sources decorrelate on a timescale much shorter than the HI signal. This provides a novel mechanism to partially mitigate the foregrounds in a drift scan.

*Unified Astronomy Thesaurus concepts:* [Reionization \(1383\)](#); [Early universe \(435\)](#); [Nonclassical interferometry \(1120\)](#); [Cosmology \(343\)](#); [Astronomical techniques \(1684\)](#); [Drift scan imaging \(410\)](#); [Radio interferometry \(1346\)](#); [Astronomical methods \(1043\)](#); [Observational cosmology \(1146\)](#); [Radio astronomy \(1338\)](#); [Interferometric correlation \(807\)](#)

## 1. Introduction

The probe for the end of the cosmic dark age remains an outstanding issue in modern cosmology. From a theoretical consideration, we expect the first luminous objects to appear at a redshift  $z \simeq 30$ . The radiation from these first-light sources ionized and heated the neutral hydrogen (HI) in their neighborhood. As the universe evolved, these ionized regions grew and merged, resulting in a fully ionized universe by  $z \simeq 6$ , as suggested by the measurement of Gunn-Peterson troughs of quasars (Fan et al. 2006). Recent *Planck* results on cosmic microwave background (CMB) temperature and polarization anisotropies fix the reionization epoch at  $z \simeq 7.7$  (Planck Collaboration et al. 2018). The cosmic time between the formation of the first-light sources ( $z \simeq 30$ , the era of cosmic dawn) and the universe becoming fully ionized ( $z \simeq 6$ ) is generally referred to as the epoch of reionization (EoR). Many important astrophysical processes during this era, e.g., the growth and evolution of large-scale structures and the nature of first-light sources, can be best probed using the hyperfine transition of HI. Due to the expansion of the universe, this line redshifts to frequencies 70–200 MHz ( $z \simeq 6$ –20), which can be detected using meter-wave radio telescopes.

Several existing and upcoming radio telescopes aim to detect the fluctuating component of this signal, e.g., radio interferometers—the Murchison Widefield Array (MWA; Bowman et al. 2013; Tingay et al. 2013; Wayth et al. 2018), Low Frequency Array (LOFAR; van Haarlem et al. 2013), Donald C. Backer Precision Array for Probing the Epoch of Reionization (PAPER; Parsons et al. 2014), Hydrogen Epoch of Reionization Array (HERA; DeBoer et al. 2017), and Giant

Metrowave Radio Telescope (GMRT; Paciga et al. 2011). In addition, there are multiple ongoing experiments to detect the global (sky-averaged) HI signal from this era—e.g., Experiment to Detect the Global EoR Signature (EDGES) and Shaped Antenna measurement of the background Radio Spectrum (SARAS; Bowman et al. 2018; Singh et al. 2018).

We focus on the fluctuating component of the HI signal in this paper. There are considerable difficulties in the detection of this signal. Theoretical studies suggest that the strength of this signal is of the order of 10 mK while the foregrounds are brighter than 100 K (for a detailed review, see Furlanetto et al. 2006; Morales & Wyithe 2010; Pritchard & Loeb 2012). These contaminants include diffuse galactic synchrotron, extragalactic point and extended radio sources, supernova remnants, free-free emission, etc. Current experiments can reduce the thermal noise of the system to suitable levels in many hundred hours of integration. The foregrounds can potentially be mitigated by using the fact that the HI signal and its correlations emanate from the three-dimensional large-scale structure at high redshifts. On the other hand, foreground contamination is dominated by spectrally smooth sources. This means that even if foregrounds can mimic the HI signal on the plane of the sky, the third axis, corresponding to the frequency, can be used to distinguish between the two. All ongoing experiments exploit this spectral distinction to isolate the HI signal from foreground contamination (e.g., Parsons & Backer 2009; Parsons et al. 2012).

Using data from ongoing experiments, many pipelines have been developed to analyze the signal (Paciga et al. 2011; Dillon et al. 2015; Beardsley et al. 2016; Choudhuri et al. 2016;

Paul et al. 2016; Trott et al. 2016; Patil et al. 2017). PAPER (Ali et al. 2015) had placed the tightest constraint on the HI power spectrum, but the result has since been retracted (Ali et al. 2018). Their revised upper limit is  $(200 \text{ mK})^2$  at redshift  $z = 8.37$  for  $k \simeq 0.37 \text{ Mpc}^{-1}$  (Kolopanis et al. 2019). The current best upper limits on the HI power spectrum are  $(79.6 \text{ mK})^2$ ,  $k \simeq 0.053h \text{ Mpc}^{-1}$ , and  $z \simeq 10.1$  (LOFAR, Patil et al. 2017) and  $(62.5 \text{ mK})^2$ ,  $k = 0.2h \text{ Mpc}^{-1}$ , and  $z \simeq 7$  (MWA, Barry et al. 2019). More recently, Bowman et al. (2018) reported the detection of an absorption trough of strength 500 mK in the global HI signal in the redshift range  $15 < z < 19$ .

Given the weakness of the HI signal, strong foregrounds, and the requirement of hundreds of hours of integration for detection, one needs extreme stability of the system, precise calibration, and reliable isolation of foregrounds. Drift scans constitute a powerful technique to achieve instrumental stability during an observational run. During such a scan, the primary beam and other instrumental parameters remain unchanged while the sky intensity pattern changes. Two ongoing interferometers, PAPER and HERA, work predominantly in this mode while the others can also acquire data in this mode. Different variants of drift scans have been proposed in the literature:  $m$ -mode analysis (Shaw et al. 2014, 2015, applied to OVRO-LWA data in Eastwood et al. 2018), cross-correlation of the HI signal in time (Paul et al. 2014), drift and shift method (Trott 2014), and fringe-rate method (Parsons et al. 2016, applied to PAPER data). Trott (2014) provided a framework to estimate the uncertainty in the measurement of HI power spectrum based on visibility covariance. Using simulations of visibility covariance, Lanman & Pober (2019) showed that the sample variance can increase up to 20% and 30% on the shortest redundant baselines of HERA and MWA, respectively.

Owing to changing intensity patterns, it is conceptually harder to extract the HI signal from drift scans. As the HI signal is buried beneath instrumental noise, it is imperative that the correct algorithm be applied to retain this sub-dominant component and prevent its loss (e.g., Cheng et al. 2018).

In this paper, we extend the work of Paul et al. (2014) to delay space and, additionally, identify the effects of phase covariance and primary beam size. We also apply our formalism to foregrounds by considering a set of isotropically distributed point sources and statistically homogeneous correlated diffuse emission. We work in both frequency and delay space, the preferred coordinate for separating foregrounds from the HI signal (e.g., Datta et al. 2010; Parsons et al. 2012). Our primary aim is to determine the correlation timescales of time-ordered visibilities of the HI signal in drift scan observations. This information can be used to establish how the HI signal can be extracted from drift scans using the correlation of visibilities measured at different times.

In the next section, we motivate the issue, develop our general formalism, and apply it to the HI signal in frequency and delay space. We use primary beams of PAPER, MWA, HERA, and SKA1-Low for our work. We discuss in detail the analytic approximation of numerical results in this section and Appendix B. In Section 3, we discuss the nature of foregrounds and compute the visibility correlation functions for a set of point sources and diffuse foregrounds. In Section 4, we elaborate on how our formulation can be applied to drift scan

data. We discuss many different approaches to the analysis of data, including comparison with earlier attempts. In the final section, we summarize our main results.

Throughout this paper, we use a spatially flat  $\Lambda$ CDM model with  $H_0 = 100 h \text{ Km s}^{-1} \text{ Mpc}^{-1}$ ,  $h = 0.67$ , and  $\Omega_\Lambda = 0.6911$  (Planck Collaboration et al. 2016).

## 2. HI Visibility Correlation in Drift Scans

The measured visibilities are a function of frequency, baseline, and time. The aim of this section is to determine the correlation structure of visibilities in these domains. In particular, our focus is on the correlation structure of visibilities as a function of time as the intensity pattern changes, for a fixed primary beam, during a drift scan.

This information allows us to average the data in the  $uv$  space with optimal signal to noise and prevent possible HI signal loss. The signal loss could occur if the data are averaged over scales larger than the scales of correlation (see, e.g., Cheng et al. 2018). For instance, the visibilities owing to HI signal are correlated for baselines separated by roughly the inverse of the primary beam, so averaging data over pixels larger than the inverse of the primary beam would result in the loss of HI signal. However, if the data are averaged using pixels much smaller than the correlation scale, then it would result in suboptimal signal to noise.

In this paper, we determine the timescales over which measured visibilities (for a given baseline, etc.) are coherent in time and therefore could be averaged in a drift scan to yield optimal signal to noise without any loss in HI signal. For this purpose, we derive the correlation function of visibilities, arising from the EoR HI signal, measured at two different times in a drift scan.

A pair of antennas of a radio interferometer measures the visibility  $V_\nu$ , which is related to the sky intensity pattern as (Equation (2.21)) of Taylor et al. 1999)

$$V_\nu(u_\nu, v_\nu, w_\nu) = \int \frac{dldm}{n} A_\nu(l, m) I_\nu(l, m) \times \exp[-2\pi i(u_\nu l + v_\nu m + w_\nu(n-1))]. \quad (1)$$

Here,  $\nu$  is the observing frequency.  $(u_\nu, v_\nu, w_\nu)$  are the components of the baseline vector between two antennas measured in units of wavelength.  $(l, m, n)$  define the direction cosine triplet in the sky with  $n = \sqrt{1 - l^2 - m^2}$ .  $A_\nu(l, m)$  is the primary beam power pattern of an antenna element, and  $I_\nu(l, m)$  is the specific intensity pattern in the sky. We further define vectors  $\mathbf{u}_\nu = (u_\nu, v_\nu)$  and  $\boldsymbol{\theta} = (l, m)$ . The intensity pattern owing to the EoR HI gas distribution  $I_\nu(\boldsymbol{\theta})$  can be decomposed into mean and fluctuating components as

$$I_\nu(\boldsymbol{\theta}) = \bar{I}_\nu + \Delta I_\nu(\boldsymbol{\theta}). \quad (2)$$

As an interferometer measures only fluctuating components of the signal, we can write

$$V_\nu(\mathbf{u}_\nu, w_\nu) = \int \frac{d^2\theta}{n} A_\nu(\boldsymbol{\theta}) \Delta I_\nu(\boldsymbol{\theta}) \times \exp[-2\pi i(\mathbf{u}_\nu \cdot \boldsymbol{\theta} + w_\nu(n-1))]. \quad (3)$$

The HI inhomogeneities  $\delta_{\text{HI}}(\mathbf{k})$  arise from various factors such as HI density fluctuations, ionization inhomogeneities, etc. The

fluctuation in the specific intensity  $\Delta I_\nu(\boldsymbol{\theta})$  can be related to the HI density fluctuations in Fourier space,  $\delta_{\text{HI}}(\mathbf{k})$ :

$$\Delta I_\nu(\boldsymbol{\theta}) = \bar{I}_\nu \int \frac{d^3k}{(2\pi)^3} \delta_{\text{HI}}(\mathbf{k}) \exp[i\mathbf{k} \cdot \mathbf{r}]. \quad (4)$$

Here,  $\mathbf{r}$  is the three-dimensional (comoving) position vector and its Fourier conjugate variable is  $\mathbf{k}$ ;  $k$ , the magnitude of the  $\mathbf{k}$  vector, is  $k = |\mathbf{k}| = \sqrt{k_\perp^2 + k_\parallel^2} = \sqrt{k_{\perp 1}^2 + k_{\perp 2}^2 + k_\parallel^2}$ , where  $\mathbf{k}_\perp$  and  $k_\parallel$  are the (comoving) components on the plane of the sky and along the line of sight, respectively. The position vector  $\mathbf{r}$  can be written in terms of the line-of-sight (parallel) and perpendicular components as  $\mathbf{r} = r_\nu \hat{n} + r_\nu \boldsymbol{\theta}$ ;  $r_\nu$  is the comoving distance. Equation (4) reduces to

$$\Delta I_\nu(\boldsymbol{\theta}) = \bar{I}_\nu \int \frac{d^3k}{(2\pi)^3} \delta_{\text{HI}}(\mathbf{k}) \exp[i r_\nu (k_\parallel + \mathbf{k}_\perp \cdot \boldsymbol{\theta})]. \quad (5)$$

As the HI fluctuations are statistically homogeneous, we can define the HI power spectrum  $P_{\text{HI}}(k)$  as<sup>1</sup>

$$\langle \delta_{\text{HI}}(\mathbf{k}) \delta_{\text{HI}}^*(\mathbf{k}') \rangle = (2\pi)^3 \delta^3(\mathbf{k} - \mathbf{k}') P_{\text{HI}}(k). \quad (6)$$

In tracking observations, the primary beam of the telescope follows a particular patch of the sky. In a drift scan, the sky pattern moves with respect to the fixed primary beam. This change of the sky intensity with respect to the fixed phase center introduces a time-dependent phase  $\vartheta(t)$  in the expression of  $\Delta I_\nu(\boldsymbol{\theta})$  in Equation (5), which gives us the fluctuating component of the specific intensity as a function of time:

$$\Delta I_\nu(\boldsymbol{\theta}, t) = \bar{I}_\nu \int \frac{d^3k}{(2\pi)^3} \delta_{\text{HI}}(\mathbf{k}) \exp[i r_\nu (k_\parallel + \mathbf{k}_\perp \cdot (\boldsymbol{\theta} - \boldsymbol{\vartheta}(t)))]. \quad (7)$$

In Equation (3), we use the expression of  $\Delta I_\nu(\boldsymbol{\theta}, t)$  and expand terms containing  $n$  up to first nonzero order<sup>2</sup> as  $d^2\theta/n \simeq d^2\theta$  and  $w_\nu(n-1) \simeq -(l^2 + m^2)w_\nu/2 = -\theta^2 w_\nu/2$ . This gives us

$$\begin{aligned} V_\nu(\mathbf{u}_\nu, w_\nu, t) &\simeq \bar{I}_\nu \int \frac{d^3k}{(2\pi)^3} \delta_{\text{HI}}(\mathbf{k}) \exp[i r_\nu k_\parallel] \\ &\times \int d^2\theta A_\nu(\boldsymbol{\theta}) \exp\left[-2\pi i \left( \left( \mathbf{u}_\nu - \frac{r_\nu}{2\pi} \mathbf{k}_\perp \right) \cdot \boldsymbol{\theta} \right. \right. \\ &\left. \left. + \frac{r_\nu}{2\pi} \mathbf{k}_\perp \cdot \boldsymbol{\vartheta}(t) - \frac{1}{2} w_\nu \theta^2 \right) \right]. \end{aligned}$$

Next, we compute the two-point visibility correlation function between two different frequencies, baselines, and times:

$$\begin{aligned} &\langle V_\nu(\mathbf{u}_\nu, w_\nu, t) V_{\nu'}^*(\mathbf{u}'_{\nu'}, w'_{\nu'}, t') \rangle \\ &\simeq \bar{I}_\nu \bar{I}_{\nu'} \int \int \frac{d^3k}{(2\pi)^3} \frac{d^3k'}{(2\pi)^3} \langle \delta_{\text{HI}}(\mathbf{k}) \delta_{\text{HI}}^*(\mathbf{k}') \rangle \\ &\times \exp[i(r_\nu k_\parallel - r_{\nu'} k'_\parallel)] \int d^2\theta A_\nu(\boldsymbol{\theta}) \int d^2\theta' A_{\nu'}(\boldsymbol{\theta}') \\ &\times \exp\left[-2\pi i \left( \left( \mathbf{u}_\nu - \frac{r_\nu}{2\pi} \mathbf{k}_\perp \right) \cdot \boldsymbol{\theta} - \left( \mathbf{u}'_{\nu'} - \frac{r_{\nu'}}{2\pi} \mathbf{k}'_\perp \right) \cdot \boldsymbol{\theta}' \right. \right. \\ &\left. \left. + \frac{r_\nu}{2\pi} \mathbf{k}_\perp \cdot \boldsymbol{\vartheta}(t) - \frac{r_{\nu'}}{2\pi} \mathbf{k}'_\perp \cdot \boldsymbol{\vartheta}'(t') - \frac{1}{2} w_\nu \theta^2 + \frac{1}{2} w'_{\nu'} \theta'^2 \right) \right]. \end{aligned} \quad (8)$$

Using Equation (6) in Equation (8) gives the two-point correlation function in terms of the HI power spectrum  $P_{\text{HI}}(k)$ . We first note that the time dependence of Equation (8) occurs as the time difference,  $\Delta t$ , in just one term  $\boldsymbol{\vartheta}'(t') - \boldsymbol{\vartheta}(t) = \Delta\boldsymbol{\vartheta}(\Delta t)$ , which is obtained by dropping the frequency dependence of  $r_\nu$ . This approximation is discussed in detail in the next subsection. Equation (49) is used to express the time-dependent part of the correlation function explicitly in terms of change in the hour angle  $\Delta H$  (for details see Appendix A). This gives us

$$\begin{aligned} &\langle V_\nu(\mathbf{u}_\nu, w_\nu, t) V_{\nu'}^*(\mathbf{u}'_{\nu'}, w'_{\nu'}, t') \rangle \\ &= \bar{I}_\nu \bar{I}_{\nu'} \int \frac{d^3k}{(2\pi)^3} P_{\text{HI}}(k) \exp[ik_\parallel (r_\nu - r_{\nu'})] \\ &\times \exp[i r_\nu k_\perp \cos \phi \Delta H] \\ &\times Q_\nu(\mathbf{k}_\perp, \mathbf{u}_\nu, w_\nu, \Delta H = 0) Q_{\nu'}^*(\mathbf{k}_\perp, \mathbf{u}'_{\nu'}, w'_{\nu'}, \Delta H). \end{aligned} \quad (9)$$

Here,  $\phi$  is the latitude of the telescope, and the Fourier beam (or 2D  $Q$ -integral) is defined as

$$\begin{aligned} Q_\nu(\mathbf{k}_\perp, \mathbf{u}_\nu, w_\nu, \Delta H) &= \int d^2\theta A_\nu(\boldsymbol{\theta}) \\ &\times \exp\left[-2\pi i \left( \mathbf{x}_u \cdot \boldsymbol{\theta} - \frac{1}{2} y \theta^2 \right) \right] \end{aligned} \quad (10)$$

$$\text{with } x_u = u_\nu - \frac{r_\nu}{2\pi} (k_{\perp 1} + k_{\perp 2} \sin \phi \Delta H) \quad (11)$$

$$x_v = v_\nu - \frac{r_\nu}{2\pi} (k_{\perp 2} - k_{\perp 1} \sin \phi \Delta H) \quad (12)$$

$$y = w_\nu + \frac{r_\nu}{2\pi} k_{\perp 1} \cos \phi \Delta H. \quad (13)$$

In this paper, we consider only the zenith drift scan. Nonzenith drift scans can be treated by replacing  $\phi$  with  $\phi + \chi$ , where  $\chi$  is the angle between the latitude of the zenith and the phase center of the observed field (for details, see Appendix A in Paul et al. 2014). This does not impact our main results. Equation (9) can be numerically solved for a given primary beam pattern  $A_\nu(\boldsymbol{\theta})$ . We next discuss the visibility correlation in delay space, the preferred coordinate for analyzing the data.

<sup>1</sup> We also assume here that the HI signal is statistically isotropic, which allows us to write the power spectrum as a function of  $|\mathbf{k}|$ . Statistical isotropy is broken owing to line-of-sight effects such as redshift-space distortion and line-cone anisotropies, which would make the power spectrum depend on the angle between  $\mathbf{k}$  and the line of sight.

<sup>2</sup> As discussed below, we use primary beams corresponding to many ongoing and future radio telescopes for our analysis. For all cases, this approximation holds for the main lobe of the primary beam, which means, as we show later, that our main results are unaffected.

### 2.1. Visibility Correlation in Delay Space

To compute the HI visibility correlation function in delay space ( $\tau$ ), we define

$$V_\tau(\mathbf{u}_0, w_0, t) = \int_{\nu_0-B/2}^{\nu_0+B/2} d\nu V_\nu(\mathbf{u}_\nu, w_\nu, t) \exp[2\pi i t \nu]. \quad (14)$$

Throughout this paper, the subscript “0” under any variable denotes the value of that variable at the central frequency. Throughout this paper, we use  $\nu_0 \simeq 154$  MHz and bandpass  $B \simeq 10$  MHz. Its cross-correlation in delay space can be expressed as

$$\begin{aligned} & \langle V_\tau(\mathbf{u}_0, w_0, t) V_\tau^*(\mathbf{u}'_0, w'_0, t') \rangle \\ &= \iint_{\nu_0-B/2}^{\nu_0+B/2} d\nu d\nu' \langle V_\nu(\mathbf{u}_\nu, w_\nu, t) V_{\nu'}^*(\mathbf{u}'_{\nu'}, w'_{\nu'}, t') \rangle e^{-2\pi i \tau \Delta\nu}. \end{aligned} \quad (15)$$

Equation (15) can be reduced to a more tractable form by making appropriate approximations. We expand frequency-dependent variables in exponents around  $\nu_0$  up to the first order. Thus,  $(r_\nu - r_{\nu'}) \simeq -\dot{r}_0 \Delta\nu$ , denoting  $(dr_\nu/d\nu)_{\nu_0} = \dot{r}_0$ ,  $\nu' - \nu = \Delta\nu$ . To the same order, the approximation made following Equation (9) is also valid. We further approximate  $\mathbf{u}_\nu \simeq \mathbf{u}_0$  and drop the weak frequency dependence of the mean specific intensity and primary beam within the observing bandwidth  $B$ . We discuss the impact of these approximations in Section 2.2. This gives us

$$\begin{aligned} & \langle V_\tau(\mathbf{u}_0, w_0, t) V_\tau^*(\mathbf{u}'_0, w'_0, t') \rangle \\ &= \bar{I}_0^2 \int \frac{d^3k}{(2\pi)^3} P_{\text{HI}}(k) \exp[ir_0 k_{\perp 1} \cos \phi \Delta H] \\ & \quad \times \left( \iint_{\nu_0-B/2}^{\nu_0+B/2} d\nu d\nu' \exp[-i\Delta\nu(k_{\parallel} \dot{r}_0 + 2\pi\tau)] \right) \\ & \quad \times Q_{\nu_0}(\mathbf{k}_{\perp}, \mathbf{u}_0, w_0, \Delta H = 0) Q_{\nu_0}^*(\mathbf{k}_{\perp}, \mathbf{u}'_0, w'_0, \Delta H). \end{aligned} \quad (16)$$

The integrals over  $\nu$  and  $\nu'$  can be solved in two ways. They can be solved by changing the variables from  $(\nu, \nu')$  to  $(x, y)$ .  $x = \nu' - \nu = \Delta\nu$  and  $y = (\nu' + \nu)/2$ . They can also be solved by separating  $\Delta\nu = \nu' - \nu$  and integrating over  $\nu$  and  $\nu'$  individually. The resulting function peaks sharply at  $\tau = -\dot{r}_0 k_{\parallel} / (2\pi)$ . The major contribution to the integral in Equation (16) occurs when  $k_{\parallel} = -2\pi\tau / \dot{r}_0$ , which gives us the well-known correlation scale along the line-of-sight direction (e.g., Paul et al. 2016). We use the  $\delta$ -function approximation for frequency integrals:

$$\begin{aligned} & \iint_{\nu_0-B/2}^{\nu_0+B/2} d\nu d\nu' \exp[-i\Delta\nu(k_{\parallel} \dot{r}_0 + 2\pi\tau)] \\ &= B^2 \text{sinc}^2 \left[ \pi B \left( \tau + \frac{\dot{r}_0}{2\pi} k_{\parallel} \right) \right] \simeq \frac{2\pi B}{|\dot{r}_0|} \delta \left( k_{\parallel} - \frac{2\pi\tau}{|\dot{r}_0|} \right). \end{aligned} \quad (17)$$

This approximation preserves the area under the curve. We note that the delta function approximation used in Equation (17) could break down if  $B$  is small. For  $B = 10$  MHz, which we use in the paper, it is an excellent assumption. For a much smaller  $B$ , the sinc function in the equation can be directly integrated without making any difference in our main results. We denote  $\dot{r}_0 = -|\dot{r}_0|$  because the comoving distance decreases with increasing frequency. Using this in Equation (16), we find, with

$$k_{\parallel} = 2\pi\tau / |\dot{r}_0|,$$

$$\begin{aligned} & \langle V_\tau(\mathbf{u}_0, w_0, t) V_\tau^*(\mathbf{u}'_0, w'_0, t') \rangle \\ & \simeq \bar{I}_0^2 \frac{B}{|\dot{r}_0|} \int \frac{d^2k_{\perp}}{(2\pi)^2} P_{\text{HI}}(k) \exp[ir_0 k_{\perp 1} \cos \phi \Delta H] \\ & \quad \times Q_{\nu_0}(\mathbf{k}_{\perp}, \mathbf{u}_0, w_0, \Delta H = 0) Q_{\nu_0}^*(\mathbf{k}_{\perp}, \mathbf{u}'_0, w'_0, \Delta H). \end{aligned} \quad (18)$$

Here,  $k = \sqrt{k_{\perp 1}^2 + k_{\perp 2}^2 + (2\pi\tau / |\dot{r}_0|)^2}$ . Equation (18) generalizes the results of Paul et al. (2014) to delay space and also accounts for the impact of the  $w$  term. To further simplify Equation (18), we need an expression for the primary beam pattern. We consider four radio interferometers in our analysis.

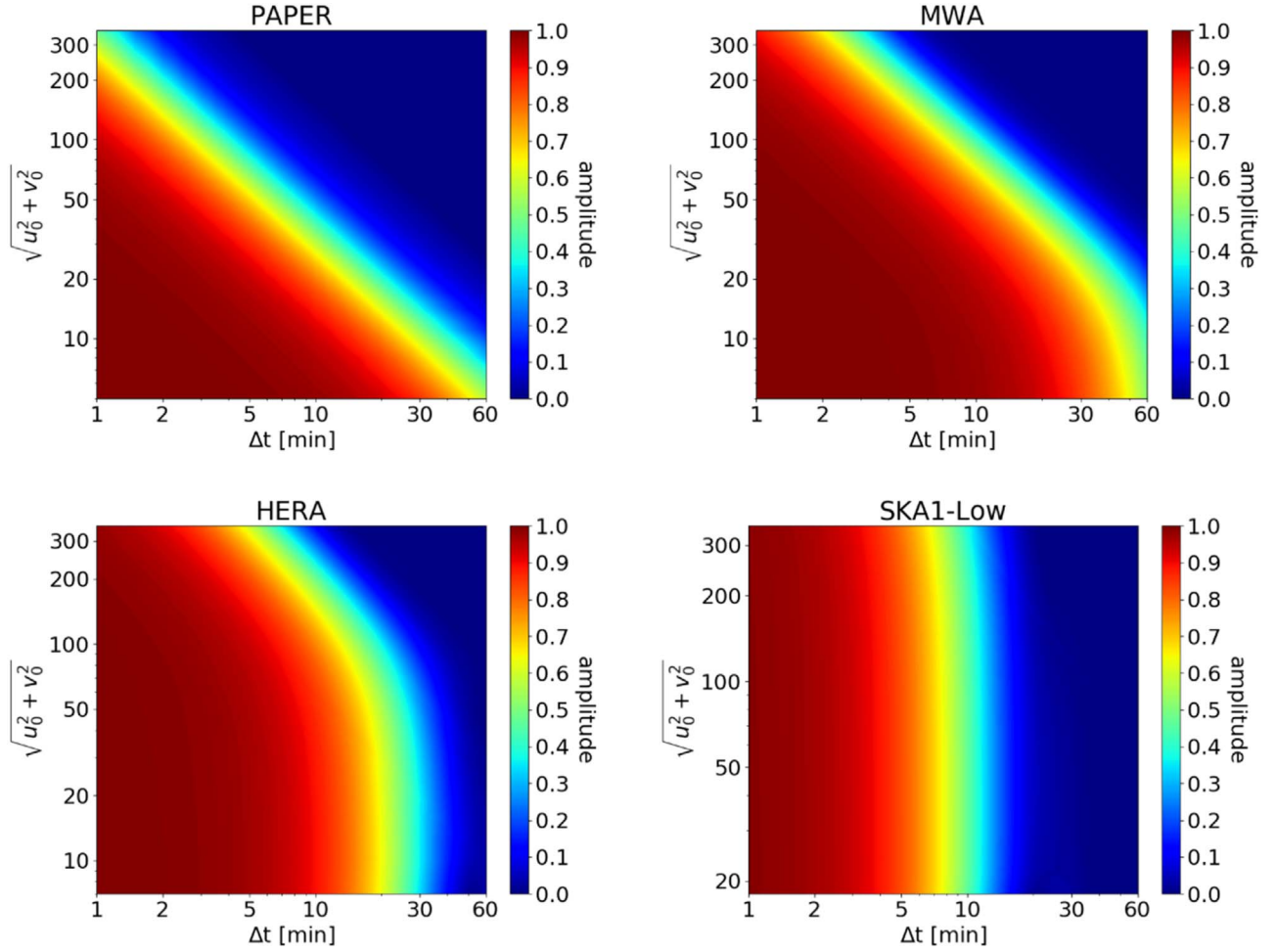
**MWA:** MWA has square-shaped antennas called tiles. Each tile consists of 16 dipoles placed on a mesh and arranged in a  $4 \times 4$  grid at a spacing of roughly 1.1 m. The effective area of a tile  $A_{\text{eff}} = 21.5 \text{ m}^2$  at 150 MHz (Tingay et al. 2013). The square of the absolute value of the 2D Fourier transform of the antenna shape gives the antenna power response. For MWA,  $A_\nu(l, m) = \text{sinc}^2(\pi L_\nu l) \text{sinc}^2(\pi L_\nu m)$ . Here,  $L_\nu = L(\nu / \nu_0)$ ;  $L (\equiv \sqrt{A_{\text{eff}}} / \lambda_0 \simeq 2.4)$  is the length of the square tile in units of central wavelength ( $\lambda_0 \simeq 1.95$  m). Therefore, the 2D primary beam response  $A_\nu(l, m)$  can be represented as a product of two independent 1D patterns,  $A_\nu(l, m) = A_\nu(l) A_\nu(m)$ .

**PAPER, HERA, and SKA1-Low:** individual elements in PAPER, HERA, and SKA1-Low correspond to dishes of diameter 2 m, 14 m, and 35 m, respectively. The beam pattern at a frequency  $\nu$  can be expressed as  $A_\nu = 4|j_1(\pi d_\nu \sqrt{l^2 + m^2}) / (\pi d_\nu \sqrt{l^2 + m^2})|^2$ , where  $j_1(x)$  is the spherical Bessel function and  $d_\nu$  is the diameter of the dish in units of wavelength. Unlike MWA, this primary beam pattern is not separable in  $l$  and  $m$ , or the double integral over angles in Equation (10) cannot be expressed as a product of two separate integrals over  $l$  and  $m$ . We do not consider LOFAR in our analysis as its core primary beam, suitable for EoR studies, is close to SKA1-Low.<sup>3</sup> For MWA and SKA1-Low,  $\phi = -26.7^\circ$ , and for HERA and PAPER,  $\phi = -30.7^\circ$ .

In Figure 1, we show the amplitude of the correlation function (Equation (18)), normalized to unity for  $\Delta t = 0$  as a function of the time difference,  $\Delta t \equiv t' - t$ , in a drift scan. In the figure, we use the HI power spectrum  $P_{\text{HI}}(k)$  given by the simulation of Furlanetto et al. (2006); we discuss the dependence of our results on the input power spectrum below in Section 2.2.1. The figure displays numerical results for different primary beams as a function of baseline length  $|\mathbf{u}_0| = \sqrt{u_0^2 + v_0^2}$  for  $w_0 = 0$  and  $\tau = 0$ . Our numerical results further show that the visibility correlation function in the time domain is nearly independent of  $\tau$ . This is discussed and justified in Appendix B using analytic approximations. Figure 2 complements Figure 1 and allows us to study the change in decorrelation time when the primary beam is changed for a fixed baseline; it will be discussed in detail in the next subsection.

To get analytic insights into the nature of the numerical results displayed in Figures 1 and 2, we consider a separable and symmetric Gaussian beam.

<sup>3</sup> <http://old.astron.nl/radio-observatory/astronomers/lofar-imaging-capabilities-sensitivity/lofar-imaging-capabilities/lofa>



**Figure 1.** Amplitude of the visibility correlation function as a function of  $\Delta t$ , normalized to unity for  $\Delta t = 0$ . The quantity plotted in the figure is  $\langle V_\tau(\mathbf{u}_0, w_0, t) V_\tau^*(\mathbf{u}_0, w_0, t') \rangle / \langle V_\tau(\mathbf{u}_0, w_0, t) V_\tau^*(\mathbf{u}_0, w_0, t) \rangle$  as a function of baseline length  $|\mathbf{u}_0| = \sqrt{u_0^2 + v_0^2}$  and  $\Delta t = t' - t$ , for  $u_0 = v_0$ ,  $w_0 = 0$ , and  $\tau = 0$ . The amplitude of the correlation function decorrelates mainly due to the rotation of the intensity pattern. However, the impact of the traversal of the intensity pattern becomes important for smaller primary beams on small baselines. As seen in the figure, for all baselines for PAPER and large baselines for MWA, HERA, and SKA1-Low, the decorrelation timescales are proportional to  $1/|\mathbf{u}_0|$  and  $1/\sqrt{\Omega}$ . This effect is discussed in Section 2.1.1 (point (b)). On smaller baselines in the MWA, HERA, and SKA1-Low panels, the traversal of the intensity pattern starts dominating the decorrelation. This effect is discussed in Section 2.1.1 (point (a)).

### 2.1.1. Fourier Beam and HI Correlation with Gaussian Beam

The Fourier beam introduced in Equation (10) is the response of the primary beam in the Fourier domain. It has two useful properties which make the computation of the Fourier beam easier. If the primary beam is separable,  $A_\nu(l, m) = A_\nu(l)A_\nu(m)$ , then the Fourier beam is also separable,  $Q_\nu(\mathbf{u}_\nu) = Q_\nu^1(u_\nu)Q_\nu^2(v_\nu)$ , and if the 1D primary beam response,  $A_\nu(l)$ , is an even function, then the 1D Fourier beam,  $Q_\nu^1(u_\nu)$ , satisfies the following relations:

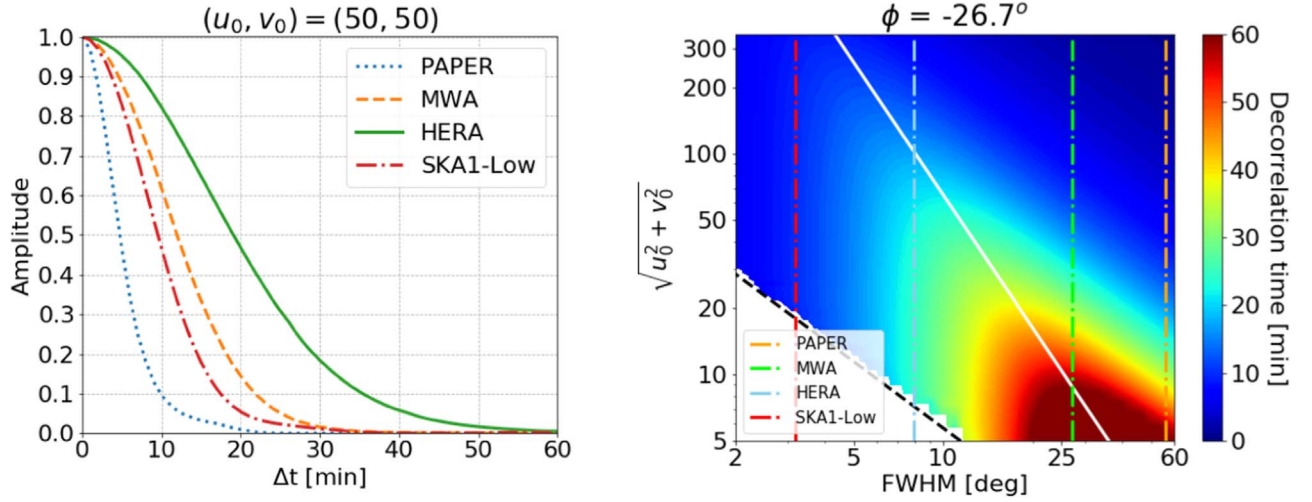
$$\begin{aligned} Q_\nu^1(-x_u, y) &= Q_\nu^1(x_u, y) \\ Q_\nu^1(x_u, -y) &= Q_\nu^{1*}(x_u, y). \end{aligned} \quad (19)$$

The expressions above are also valid for  $Q_\nu^2(v_\nu)$ . This shows that it is sufficient to calculate the Fourier beam for only  $x_u, y \geq 0$ . The variables  $x_u, x_v$ , and  $y$  are defined in Equations (11)–(13).  $x_u$  and  $x_v$  determine the correlation scales in the neighborhood of the Fourier mode,  $2\pi\mathbf{u}_0/r_0$ , at which the  $Q$ -integral receives maximum contribution. The variable  $y$  can be viewed as an effective  $w$  term. We note that when  $y$  is small,  $Q_\nu^1(x_u, y)$  is large but falls very rapidly along  $x_u$ . For larger values of  $y$ ,  $Q_\nu^1(x_u, y)$  is

smaller and goes to zero slowly along  $x_u$ . This behavior can be understood as follows: the effective beam size shrinks for a larger value of the  $w$  term, resulting in a decrease in signal strength but an increase in the correlation scale (e.g., Cornwell et al. 2008; Paul et al. 2016).

The discussion also applies to 2D Fourier beams. The 2D Fourier beam is a function of Fourier coordinates  $x_u, x_v$  and parameter  $y$ . The point  $(x_u, x_v) = (0, 0)$  receives the maximum contribution and picks out Fourier modes,  $k_{\perp 1}, k_{\perp 2}$ . Large beams have smaller Fourier beams, e.g., for PAPER, the Fourier beam is the smallest of all the cases we consider. The width of the Fourier beam decides the range of correlation scales of the HI signal. This range is roughly on the order of  $2/\sqrt{\Omega} \simeq 2d$ , where  $\Omega$  is the primary beam solid angle and  $d$  is the antenna size in units of wavelength. The amplitude of the Fourier beam is more sensitive to  $y$  if the beam is larger (PAPER, MWA).

To gain further analytic insights into the HI correlation function, we use a Gaussian primary beam in our formalism to compute the Fourier beam. For illustration, we choose the Gaussian primary beam of solid angle  $\Omega_{0g}$  at  $\nu_0 = 154.24$  MHz ( $\Omega_{0g} = 0.25/L^2$  roughly matches the MWA primary beam).



**Figure 2.** Left panel: the amplitude of the visibility correlation function is shown as a function of  $\Delta t$  for a fixed baseline for different primary beams. Right panel: the isocontours of the decorrelation time are shown in the primary beam–baseline plane; the decorrelation time is defined as  $\Delta t$  such that the amplitude of the correlation function falls to half its value as compared to  $\Delta t = 0$ . The figure assumes Gaussian beams (Equation (20)) with  $\text{FWHM} = 2\sqrt{\log_e(2)\Omega_{0g}}$ . The region on the bottom left is excluded because the shortest baseline  $\sqrt{u_0^2 + v_0^2} = d_0$ , where  $d_0$  is the primary element of the telescope in units of the central wavelength,  $\lambda_0$ . There could be minor differences between this figure and Figure 1 because we used a fixed telescope latitude  $\phi = -26.7$  for all primary beams. The primary beams of the four interferometers studied in this paper are marked on the figure. The white line demarcates the regions dominated by rotation (above the line) and traversal of intensity pattern (for further discussion, see the text).

This gives us

$$A_{\nu_0}(l, m) = e^{-(l^2+m^2)/\Omega_{0g}}. \quad (20)$$

To compute the Fourier response of a Gaussian beam analytically, we extend the limits of the integral from  $[-1, 1]$  to  $[-\infty, \infty]$ , which is a valid procedure as the integrand falls rapidly outside the support of the primary beam. Using Equation (10), we obtain

$$Q_{\nu_0}(\mathbf{k}_\perp, \mathbf{u}_0, w_0, \Delta H) = \frac{\pi\Omega_{0g}}{1 - i\pi y\Omega_{0g}} \times \exp\left[-\frac{\pi^2\Omega_{0g}(x_u^2 + x_v^2)}{1 - i\pi y\Omega_{0g}}\right]. \quad (21)$$

We assume  $\mathbf{u}_0 = \mathbf{u}'_0$  and  $\mathbf{k}_\perp = (2\pi/r_0)\mathbf{u}_0$  to study the time behavior of the correlation function relevant in a drift scan. The time-dependent part of the visibility correlation function is determined by the product of two Fourier beams separated by drift time  $\Delta H$  in Equation (18). For a Gaussian beam, this product is

$$Q_{\nu_0}(\Delta H = 0)Q_{\nu_0}^*(\Delta H) = \frac{(\pi\Omega_{0g})^2}{(1 - i\pi\Omega_{0g}w_0)(1 + i\pi\Omega_{0g}y)} \times \exp\left[-\frac{\pi^2\Omega_{0g}|\mathbf{u}_0|^2 \sin^2 \phi \Delta H^2}{1 - i\pi y\Omega_{0g}}\right] \quad (22)$$

where only the dependence on the time variable is retained in the left-hand side for brevity. As discussed above,  $y = (w'_0 + u_0 \cos \phi \Delta H)$  acts as an effective  $w$  term. For the zenith drift scan we study in this paper, the  $w$  term is small, so we put  $w_0 = w'_0 = 0$ . We find the amplitude of the product of

the Fourier beams to be

$$\begin{aligned} & |Q_{\nu_0}(\Delta H = 0)Q_{\nu_0}^*(\Delta H)| \\ &= \frac{(\pi\Omega_{0g})^2}{\sqrt{(1 + \pi^2\Omega_{0g}^2 u_0^2 \cos^2 \phi \Delta H^2)}} \\ & \times \exp\left[-\frac{\pi^2\Omega_{0g}|\mathbf{u}_0|^2 \sin^2 \phi \Delta H^2}{1 + \pi^2\Omega_{0g}^2 u_0^2 \cos^2 \phi \Delta H^2}\right]. \end{aligned} \quad (23)$$

Equation (23), along with Equations (18) and (22), allows us to read off several salient features of the visibility correlation function in a drift scan.

Due to the rotation of Earth on its axis, the sources in the sky move with respect to the fixed phase center ( $l = 0, m = 0$ ) of a telescope located at latitude  $\phi$ . The changing intensity pattern is a combination of two motions: rotation around a fixed phase center and the east–west translation of the pattern with respect to the fixed phase center (Equation (47)). In Fourier space, the rotation causes a time-dependent mixing of Fourier modes in the plane of the sky, while the translation introduces a new time-dependent phase that is proportional to  $k_{\perp 1}$ , the component of the Fourier mode in the east–west direction (Equation (49)). In addition to these two effects, which are linear in the angle, we also retain a second-order term that becomes important for large beams (Equations (47) and (49)). The impact of each of these effects on the visibility correlation function is discussed next:

- (a) *Traversal time of coherence scale.* The phase term proportional to  $\exp(ir_0 k_{\perp 1} \cos \phi \Delta H)$  in Equation (18) represents this effect.  $\Delta H \simeq 1/(r_0 k_{\perp 1} \cos \phi)$  is the time over which a coherent feature of linear size  $1/k_{\perp 1}$  is traversed in the east–west direction. As  $r_0 k_{\perp 1} \simeq 2\pi u_0$ ,  $\Delta H \simeq 1/(2\pi u_0 \cos \phi)$  appears to give a rough estimate of the time over which the decorrelation occurs for a given  $u_0$ , the east–west component of the baseline. However, it does not give a reasonable estimate for the decorrelation timescale of the amplitude of the correlation

function as Equation (18) can be multiplied and divided by  $\exp(i2\pi u_0 \cos \phi \Delta H)$ , which allows us to absorb the fastest changing term as the phase term of the correlation function. The correlation timescale of the amplitude of the correlation function depends on the slow phase  $\exp(i(r_0 k_{\perp 1} - 2\pi u_0) \cos \phi \Delta H)$ , whose contribution to the visibility correlation is determined by the primary beam as we discuss below.

- (b) *Rotation of the intensity pattern.* This effect is captured by the numerator in the Gaussian in Equation (23), which shows that the decorrelation owing to the rotation of the intensity pattern is proportional to  $1/(\Omega_{0g}^{1/2} |\mathbf{u}_0| \sin \phi)$ . This effect, unlike (a), depends on the magnitude of the baseline and not on its east–west component. Equations (10)–(12), along with Equations (47) and (49), allow us to understand this effect. When visibilities at two times are correlated for a given baseline, they respond to different Fourier modes of the HI power spectrum owing to the rotation of the intensity pattern in a drift scan (Equations (10)–(12)). The extent of correlation of the visibilities which receive a contribution from different Fourier modes depends on the primary beam: the smaller the primary beam, the larger the range of Fourier modes that contribute to the correlation. Therefore, the decorrelation time is proportional to  $\Omega_{0g}^{-1/2}$ .
- (c) *Large field of view.* The terms proportional to  $\Omega_{0g}^2$  in Equation (23) (or more generally the terms proportional to  $y$  in Equation (21)) are responsible for this effect. These terms correspond to an effective  $w$  term, a part of which arises from  $w_0$ , and the remainder is the higher-order time-dependent phase in a drift scan. This effect is important when the primary beam or  $w_0$  is large.<sup>4</sup>

We next discuss the relative importance of (a), (b), and (c) in understanding Figures 1 and 2. We first note that (c) does not play an important role in explaining the qualitative features seen in the figures. Its impact is only mildly important for PAPER at the smallest baselines we consider.

For PAPER, the decorrelation time in the figure scales linearly as the inverse of the length of the baseline  $1/|\mathbf{u}_0|$ . Figure 1 shows only the case  $u_0 = v_0$ . We have checked that the behavior seen in the figure is nearly independent of the individual components of the baseline. Also, a comparison of the decorrelation times between PAPER and MWA shows that the decorrelation times scale as  $\Omega_{0g}^{-1/2}$  for baseline  $|\mathbf{u}_0| \gtrsim 25$ . A comparison of these two cases with large baselines  $|\mathbf{u}_0| \gtrsim 150$  for HERA and SKA1-Low also shows the same scaling with the primary beam. This means that (b) is the dominant decorrelation mechanism in all these cases.

For short baselines for MWA, HERA, and SKA1-Low, the behavior is markedly different. If (b) alone determined the decorrelation in these cases, the decorrelation time would be longer as the primary beam is smaller in these two cases, but this behavior is seen only for longer baselines. Therefore, (a) plays an

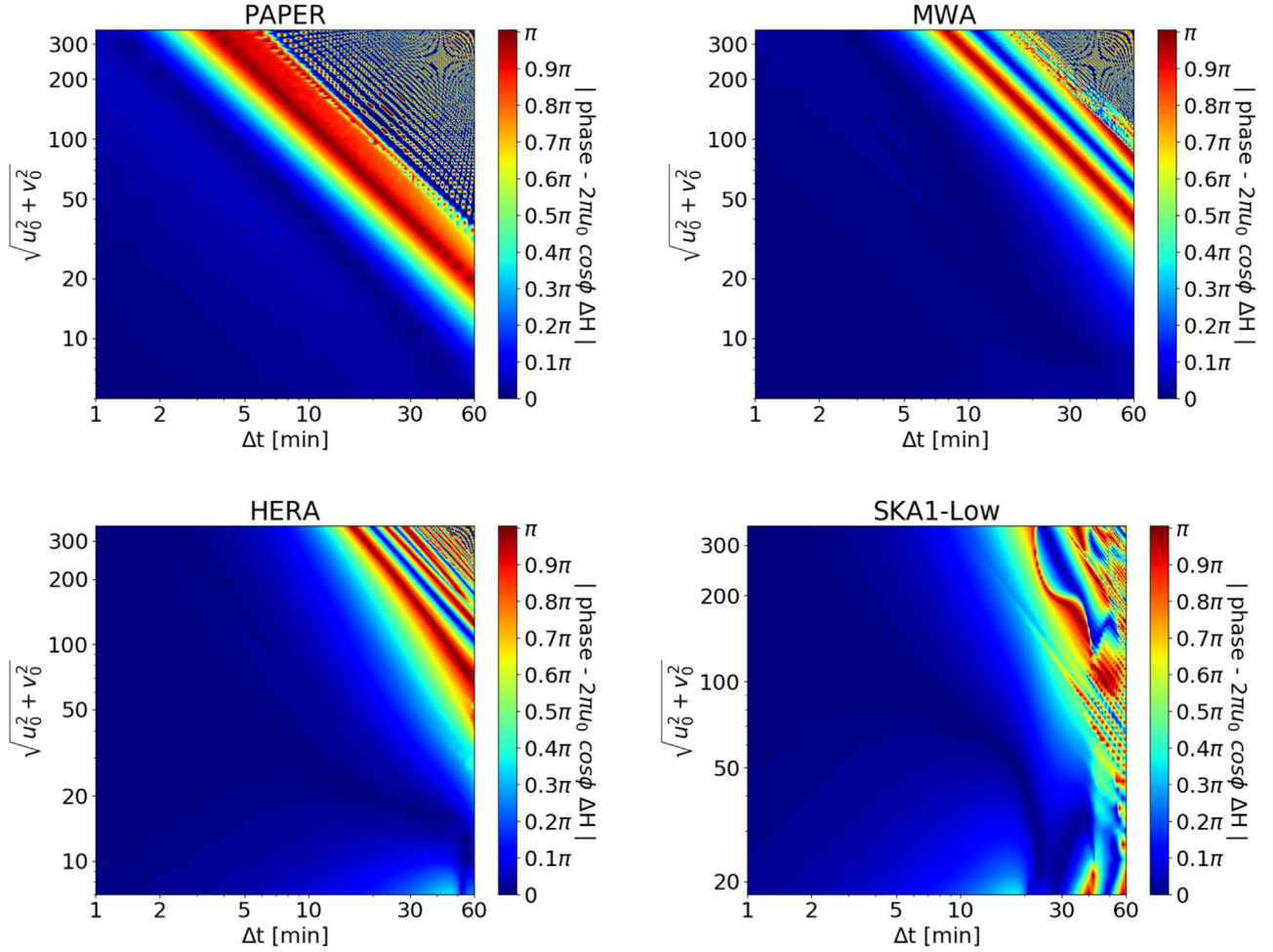
important role in these cases. For large primary beams, (a) is unimportant because the slow phase discussed above is closer to zero, as it gets contribution from a small range of Fourier modes. However, for narrower primary beams, this term gets contribution from a larger range of Fourier modes, which results in cancellation when integration over  $k_{\perp 1}$  is carried out. This results in a reduction of correlation timescale. This effect is more dominant for smaller baselines for the following reason: for a given  $u_0$ , the range of Fourier modes that contribute to the visibility correlation function is  $\Delta k_{\perp 1} \simeq 1/(r_0 \Omega_{0g}^{1/2})$  (i.e., the size of the Fourier beam) centered around  $k_{\perp 1} = 2\pi u_0 / r_0$  (e.g., Equations (10)–(12)). It should be noted that  $\Delta k_{\perp 1}$  is only determined by the size of the primary beam while  $k_{\perp 1}$  scales with the east–west component of the baseline. This implies that for long baselines,  $k_{\perp 1} \gg \Delta k_{\perp 1}$ . In this case, the visibility correlation function is dominated by the contribution of a single Fourier mode, which suppresses the impact of possible cancellations that occur owing to the mixing of Fourier modes, diminishing the impact of (a) for long baselines. However, when  $\Delta k_{\perp 1} \simeq k_{\perp 1}$ , the effect becomes important, and it determines the decorrelation timescale for shorter baselines.

For small baselines and narrower primary beams, both (a) and (b) play an important role so it is worthwhile to investigate the dependence of the decorrelation time on the components of baselines (Figure 1 assumes  $u_0 = v_0$ ). We have checked many different combinations of  $u_0$  and  $v_0$ , and find that the qualitative features of Figure 1 are largely determined by the length of the baseline. But, as discussed below, the phase of the correlation function is dominated by the east–west component of the baseline.

The correlation structure in the primary-beam– $\Delta t$ –baseline space is further explored in Figure 2. In the left panel, we show the amplitude of the correlation function as a function of  $\Delta t$  for a fixed baseline for different primary beams. The right panel shows the isocontours of the decorrelation time in the primary-beam–baseline plane; the decorrelation time is defined as the time difference  $\Delta t$  at which the amplitude of the correlation function falls to half its value at  $\Delta t = 0$ . For each baseline, the decorrelation time reaches a maximum value as a function of the primary beam. Our formalism allows us to understand this general behavior: for a smaller primary beam, the Fourier beam is large, which causes decorrelation owing to mode mixing in the transverse motion of the intensity pattern (point (a)). For a larger primary beam, the rotation of the intensity pattern is responsible for the decorrelation (point (b)). The decorrelation time scales inversely with the baseline length and could reach an hour for the shortest baselines and large primary beams, in agreement with Figure 1. A notable feature of Figure 2 is the alignment of the isocontours of decorrelation time. Its shape is determined by the interplay of decorrelation owing to the rotation and the traversal of the intensity pattern and can be derived analytically.

For large primary beams, the decorrelation time is  $\simeq 1/(|\mathbf{u}_0| \Omega_{0g}^{1/2} |\sin \phi|)$  (point (b), (Equation (23))); the decorrelation profile for large primary beams is seen to follow this function. For small primary beams, the decorrelation time is  $\simeq \Omega_{0g}^{1/2} / \cos \phi$ , nearly independent of the length of the baseline (point (a)). Equating these two expressions gives us  $\Omega_{0g} |\tan \phi| |\mathbf{u}_0| \simeq 1$ . This relation is shown in Figure 2 (white line), and it separates the regions dominated by decorrelation owing to the rotation (above the white line) from the regions in which the translation plays the dominant role. Figure 2 shows the white line adequately captures the essential physics of the separation of the two regions. We note that the large field of view (point (c) above) does not play an important role in our

<sup>4</sup> Throughout our analysis, we assume  $w_0 = 0$ , and we only consider the impact of the time-dependent term. Our assumption would be valid for a zenith drift scan, which we assume, for a near-coplanar interferometric array. Coplanarity is generally a good assumption as our focus for the detection of the HI signal is short baselines, e.g., for MWA  $w_0 \ll |\mathbf{u}|$  for a zenith scan. We can gauge the quantitative impact of nonzero  $w_0$  using Equation (22). The main effect of nonzero  $w_0$  is to yield a smaller effective primary beam (Comwell et al. 2008; Paul et al. 2016, 2014) and to introduce an additional phase in the visibility correlation function (Equation (53)).



**Figure 3.** Absolute value of the phase angle of the visibility correlation function (Equation (24)) as a function of  $\Delta t = t' - t$ . This figure illustrates that the rapidly fluctuating component of the phase of the complex correlation function (Equation (18)) can mostly be removed by multiplying it with  $\exp(-i2\pi u_0 \cos \phi \Delta H)$ . This allows us to determine the timescales for averaging the time-ordered visibilities in drift scans (Section 2.1.2 and 4).

study because of the range of telescope latitudes we consider, which is motivated by the location of radio interferometers studied here. For  $\phi \simeq 90^\circ$ , both translation and large field-of-view effects are negligible, while for  $\phi \simeq 0$ , the impact of rotation is negligible while translation and wide field-of-view effects dominate (Equation (23)).

### 2.1.2. The Phase of Visibility Correlation Function

In the foregoing, we studied the amplitude of the correlation function. As the correlation function (in either frequency or delay space Equation (9) or Equation (15)) is a complex function, we need to know the correlation properties of its phase in addition to complete the analysis.

In Appendix B, we discuss how suitable approximations allow us to discern major contributors to the phase of the correlation function. Equations (52) and (53) show that the phase angle is  $2\pi u_0 \cos \phi \Delta H + \psi_1 + \psi_2$ . The term  $2\pi u_0 \cos \phi \Delta H$  has already been discussed above (point (a) on traversal time of coherence scale). It follows from Equation (53) that both  $\psi_1$  and  $\psi_2$  are small compared to  $2\pi u_0 \cos \phi \Delta H$  as  $\psi_1 \propto \Omega_g$  and  $\psi_2 \propto \Omega_g^2$  for  $\pi^2 \Omega_g^2 y^2 < 1$ .  $\psi_2$  can only be significant when effects arising from a large field of view become important (Equation (53) and discussion on point (c) above), which is not the case for  $w_0 = 0$  and the primary beams we consider in our analysis. The dominant

phase angle  $2\pi u_0 \cos \phi \Delta H$  can be explicitly identified in Equation (52) in this case.

Motivated by our analytic results, we define the phase angle as

$$\psi(\mathbf{u}, t' - t) = \text{Arg}(\exp(-i2\pi u_0 \cos \phi \Delta H) \times \langle V_\tau(\mathbf{u}_0, w_0, t) V_\tau^*(\mathbf{u}_0, w_0, t') \rangle). \quad (24)$$

The multiplication by the additional phase allows for near cancellation of the phase term  $\exp(ik_{\perp 1} r_0 \cos \phi \Delta H)$  in Equation (15) (or a similar term in Equation (9) for correlation in frequency space if  $u_0$  and  $r_0$  are replaced by  $u_\nu$  and  $r_\nu$ , respectively). In Figure 3 we present our numerical results. We notice that the phase angle defined by Equation (24) is small for a wide range of  $\Delta t$ , as suggested by our analytic results. This means, as anticipated, that the phase of the correlation function is nearly  $\exp(i2\pi u_0 \cos \phi \Delta H)$ .<sup>5</sup> The implication of this result for drift scan data analysis will be discussed below.

<sup>5</sup> The origin of this phase can partly be explained by considering a simpler case: a single point source of flux  $F_\nu$  at the phase center. In this case, the visibility  $V_\nu(\mathbf{u}) = F_\nu A_\nu(0)$ , where  $A_\nu(0)$  defines the primary beam response at the phase center,  $l = 0$  and  $m = 0$ . The correlation between visibilities separated by  $\Delta H$  in time in a drift scan is  $V_\nu(\mathbf{u}) V_\nu^*(\mathbf{u}) \simeq F_\nu^2 A_\nu^2(0) \exp(i2\pi u_\nu \cos \phi \Delta H)$ . As discussed in Section 3.1, the same factor scales out of the correlation function for a set of point sources also.



## 2.2. Approximations and Input Quantities

Our results use an input HI power spectrum, different primary beams, and a set of approximations to transform from frequency to delay space. We discuss the impact of these approximations and input physics on our analysis.

### 2.2.1. Dependence on Input Power Spectrum and the Shape of Primary Beam

The results shown in Figure 1 were derived using the HI power spectrum,  $P(k) \simeq 1/k^n$ , with  $n \simeq 2$ , for a range of scales (Furlanetto et al. 2006). We tested our results with different power-law HI power spectra with spectral indices in the range  $n = 1-3$  and found our results to be insensitive to the input power spectra.

The lack of dependence of the visibility decorrelation time on the input HI power spectrum follows from our analysis. Equations (51) and (52) show that relevant approximations allow us to separate the input power spectrum from the time-dependent part of the correlation function, which means Figure 1 is independent of the HI power spectrum. These equations show that the time dependence of the correlation function is essentially captured by the response of the primary beam in Fourier space. A similar expression was derived in Parsons et al. (2016; their Equation (9)) for cases when the Fourier beam (Equation (10)) has a narrow response (e.g., PAPER).

The only cases not covered by this approximation are small primary beams and small baselines. However, for the limiting cases, we discuss here,  $|\mathbf{u}| \gtrsim 20$  and SKA1-Low primary beam, our numerical results show that the impact of the input HI power spectrum on the decorrelation timescale is negligible.

Our results are insensitive to the shape of the primary beam. We compare our numerical results for instrumental primary beams with a symmetric, separable Gaussian beam by roughly matching  $\Omega_{0g}$  and the main lobe of the instrumental primary beam. We find excellent agreement in explaining the main features of Figures 1–3. Equation (52) adequately explains Figure 1, except for small baselines for HERA and SKA1-Low.

### 2.2.2. Approximations in Transforming from Frequency to Delay Space

Following Equation (15), we discuss various approximations used in making the correlation function in delay space more tractable. In the tracking case, these approximations allow us to find a one-to-one linear relation between the Fourier modes of the HI signal with the variables of radio interferometers (e.g., Paul et al. 2016 and references therein). However, owing to the frequency dependence of the primary beam, the coordinate distance, and the baseline, these commonly used relations are approximate. We assessed the impact of these approximations in Paul et al. (2016) for the tracking case. For a bandwidth  $B = 10$  MHz ( $\nu_0 = 154$  MHz) and MWA primary beam, the error in these relations is less than 5% for  $k_{\parallel} \gtrsim 0.1$  Mpc $^{-1}$ . The modes corresponding to  $k_{\parallel} \lesssim 0.1$  Mpc $^{-1}$  are buried in the foreground wedge and therefore do not play a role in the detection of the HI signal (e.g., Paul et al. 2016). The error increases with bandwidth and primary beam and therefore is expected to be smaller for HERA and SKA1-Low for the same bandwidth. As we also use these approximations in our work to separate the variables on the sky plane from those along the line of sight, we reassess these approximations for a drift scan

and find these errors to be of similar magnitude for the drift scan. As in the tracking case, these approximations allow us to derive the relation between baseline and delay space parameter  $\tau$  and Fourier modes of the HI signal. This simplification allows us to write the frequency-dependent terms in the form expressed in Equation (16).

One outcome of this approximation for drift scans is that the functional form of the decorrelation time shown in Figure 1 is nearly the same in frequency and delay space. Therefore, Figure 1 can be interpreted as displaying the decorrelation time at the center of the bandpass. This assertion is borne out by Equation (51).

Our study is based on the assumption  $\nu_0 \simeq 154$  MHz and  $B \simeq 10$  MHz. It can readily be extended to a different frequency/bandpass by using Equation (51) and/or (52).

We discuss the approximation in transforming from frequency to delay space further with regard to foregrounds and the analysis of drift scan data in later sections (see footnote 6).

It is worthwhile to reiterate the scope of the main approximations we use: (a) for large primary beams and baselines, Equation (51) provides an excellent approximation; (b) for small bandwidths and primary beams, Equation (51) can readily be extended to Equation (52); and (c) for small baselines and primary beams, Equation (51) might not be valid and Equation (18) has to be computed numerically.

## 3. Foregrounds in Drift Scans

In the tracking mode, the foregrounds can be isolated from the HI signal (“EoR window”) by transforming to delay space if the two-dimensional foregrounds are spectrally smooth and therefore their correlation scales differ from the three-dimensional HI signal along the line of sight. However, in tracking mode, we cannot use the difference between correlation properties of foregrounds and the HI signal on the sky plane. In a drift scan, it is possible that the decorrelation time of the HI signal is different from components of foregrounds, which might give us yet another way to mitigate foregrounds.

The aim of this section is to study the decorrelation timescales of two components of foregrounds: the near-isotropic distribution of point sources of flux above 1 Jy and statistically homogeneous and isotropic diffuse foregrounds. In our analysis, the delay space approach continues to be the primary method used to isolate foregrounds from the HI signal, and we therefore present all our results in this space.

### 3.1. Point Sources

In a drift scan, the phase center is held fixed while the intensity pattern changes. The changing intensity pattern owing to a set of point sources can be written as

$$I_{\nu}(\boldsymbol{\theta}, t) = \sum_m F_{\nu}^m \delta^2(\boldsymbol{\theta} - \boldsymbol{\theta}_m(t)). \quad (25)$$

Here,  $F_{\nu}^m$  is the flux of the  $m$ th source and  $\boldsymbol{\theta}_m(t)$  its angular position at time  $t$ . Here, all the angles are measured with respect to the phase center, which is assumed to be fixed at  $\boldsymbol{\theta}_0 = 0$ . The visibility (retaining the  $w$  term) can readily be derived from

the expression above:

$$V_\nu(\mathbf{u}_\nu, w_\nu, t) = \sum_m F_\nu^m A_\nu(\boldsymbol{\theta}_m(t)) \times \exp[-2\pi i(\mathbf{u}_\nu \cdot \boldsymbol{\theta}_m(t) + w_\nu(n_m(t) - 1))]. \quad (26)$$

To discern the main results of this section, we ignore the frequency dependence of source fluxes and primary beam, even though we allow these quantities to be frequency dependent in our simulations.<sup>6</sup> Using Equation (14), the visibility of point sources in delay space is

$$V_\tau(\mathbf{u}_0, w_0, t) \simeq \sum_m F_0^m A_0(\boldsymbol{\theta}_m(t)) B \text{sinc}(\pi B \bar{\tau}^m(t)) e^{2\pi i \nu_0 \bar{\tau}^m(t)} \quad (27)$$

$$\text{where, } \bar{\tau}^m(t) = \tau - \frac{1}{\nu_0}(\mathbf{u}_0 \cdot \boldsymbol{\theta}_m(t) + w_0(n_m(t) - 1)). \quad (28)$$

The correlation function of the visibilities in delay space can be written as

$$\begin{aligned} & \langle V_\tau(\mathbf{u}_0, w_0, t) V_\tau^*(\mathbf{u}_0, w_0, t') \rangle \\ & \simeq B^2 \sum_m \sum_n F_0^m F_0^n A_0(\boldsymbol{\theta}_m(t)) A_0(\boldsymbol{\theta}_n(t')) \\ & \times \text{sinc}(\pi B \bar{\tau}^m(t)) \text{sinc}(\pi B \bar{\tau}^n(t')) e^{2\pi i \nu_0 (\bar{\tau}^m(t) - \bar{\tau}^n(t'))}. \end{aligned} \quad (29)$$

Here, the ensemble average implies averages over all pairs of baselines and times for which  $|\mathbf{u}_0|$  and  $t' - t$  are held fixed. To understand Equation (29), we first consider the tracking case in which source positions are independent of time. In this case, the dominant contribution comes from  $\tau = 2\pi \mathbf{u}_0 \cdot \boldsymbol{\theta}_m / \nu_0$ . This defines the so-called foreground wedge, which is bounded by the maximum value of  $\theta_m$ , which is given approximately by the size of the primary beam. It also follows from the equation that the sum is dominated by terms for which  $m = n$ .

In a drift scan, the source position changes with respect to the primary beam. It means the value of  $\tau$  for which the sum in Equation (29) peaks changes with time. While the broad wedge structure is the same in this case as in the tracking case, given that the dominant contribution comes from sources within the primary beam, the correlation structure becomes more complicated. As  $\boldsymbol{\theta}_n(t') - \boldsymbol{\theta}_m(t)$  remains unchanged during a drift scan, the summation in this case would also generally be dominated by  $m = n$  terms. However, it is possible that a source at one position at a time drifts close to the position of another source at another time. Even though the contribution of this pair could be negligible in tracking mode, it would not be if

<sup>6</sup> We ignore the frequency dependence of the intensity pattern and the primary beam throughout this paper. As we compare our analytic results against simulations in this section, it allows us to verify this assumption more explicitly. We find this assumption to be extremely good for bandwidth  $B \simeq 10$  MHz around a central frequency of  $\nu_0 \simeq 154$  MHz. This approximation can be understood by considering a simpler case: a flat-spectrum source at the phase center. While transforming to delay space, this source receives contribution from only the  $\tau = 0$  mode. If the source is now assumed to have a spectral index, more delay space modes close to  $\tau = 0$  begin to contribute. We find that these modes do not contaminate the EoR window as they lie well within the wedge given the bandwidth and spectral index of interest. The leakage into the EoR window owing to finite bandwidth can be assuaged by using a frequency-space convolving function such as a Blackman–Nuttall window or a Gaussian window we discuss in the section on diffuse foregrounds. The frequency dependence of baselines in phase plays a more important role and is needed to explain the wedge structure for foregrounds (e.g., Paul et al. 2016).

the visibilities are correlated at two different times. The impact of this effect requires details of point-source distribution which we model using a simulation in this paper.

For the case of  $m = n$ , the same source is correlated at two different times. In this case, it follows from Equation (29) that the visibility correlation diminishes as the time separation increases. As the additional time-dependent phase acquired in the drift is proportional to the length of the baseline, the decorrelation timescale is expected to be shorter for longer baselines.

*Point-source simulations:* we generate 15,067 point sources brighter than 1 Jy distributed isotropically on the southern hemisphere (Hopkins et al. 2003). We assume the spectral index of sources to be  $-0.7$ .<sup>7</sup> For this source distribution, we compute the power spectrum in delay space as a function of drift time. In a drift scan, the coordinates of these sources evolve according to Equation (47) with respect to the fixed phase center.

We compute visibilities in delay space for a one-hour drift scan. The visibilities are then correlated in time, and the visibility correlation function is computed by averaging over the number of correlation pairs for which  $t' - t$  and  $|\mathbf{u}_0|$  are held fixed:

$$\begin{aligned} & \langle V_\tau(\mathbf{u}_0, w_0, t) V_\tau^*(\mathbf{u}'_0, w'_0, t') \rangle \\ & = \frac{1}{N_{|\mathbf{u}_0|}} \sum_{|\mathbf{u}_0|} \frac{1}{N_{t'-t}} \sum_{t, t'} V_\tau(\mathbf{u}_0, w_0, t) V_\tau^*(\mathbf{u}'_0, w'_0, t'). \end{aligned} \quad (30)$$

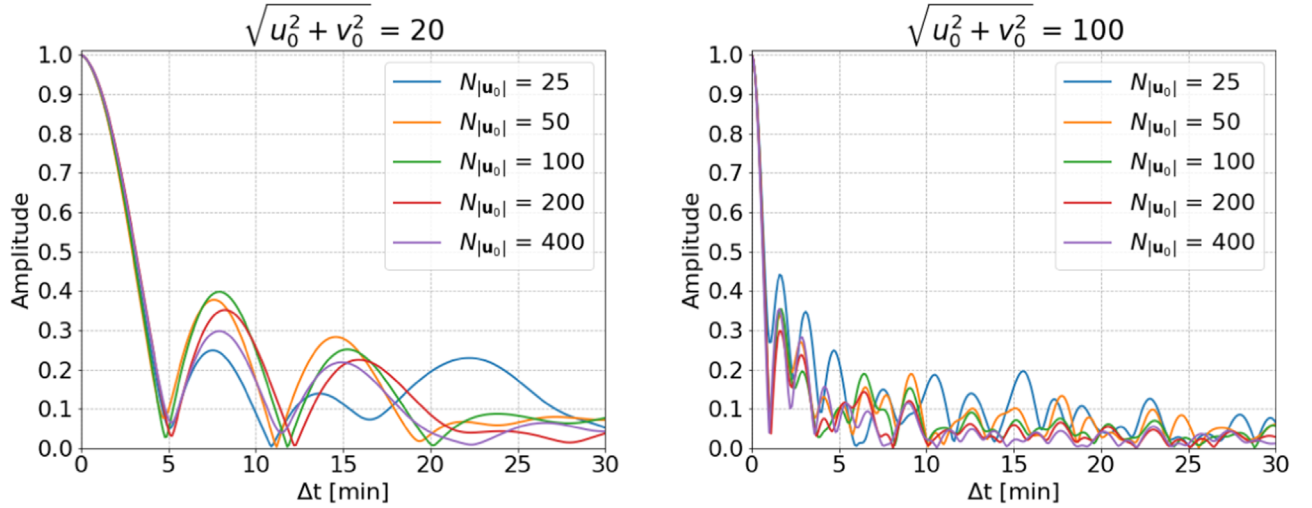
Here,  $N_{|\mathbf{u}_0|}$  and  $N_{t'-t}$  are the number of baseline pairs for fixed  $|\mathbf{u}_0|$  and  $t - t'$ , respectively.

To establish how the amplitude of the visibility correlation behaves as a function of time, baselines, and the number of points over which the average is computed, we choose two representative baselines,  $|\mathbf{u}_0| = 20, 100$ . We carry out averages in a ring of width  $\Delta|\mathbf{u}_0| = 4$ ; each of these rings is populated, randomly and uniformly, with  $N_{|\mathbf{u}_0|} = 25, 50, 100, 200, 400$ .

In Figure 4, the visibility correlation functions are plotted for the two cases using the instrumental parameters of MWA (primary beam and  $\phi$ ) for  $\tau = 0$  and  $w_0 = 0$ . We notice the following: (a) averaging over more baselines causes the correlation function to decorrelate faster when the number of baselines are small but the function converges as the number of baselines is increased, (b) the correlation function decorrelates faster for larger baselines, as anticipated earlier in the section based on the analytic expression, Equation (29), and (c) a comparison between Figures 4 and 1 shows that the decorrelation timescale for the HI signal is much larger than for a set of point sources. For  $|\mathbf{u}_0| = 100$ , the point sources decorrelate to 50% of the peak in less than a minute while this time is nearly 10 minutes for the HI signal.

The structure of the foreground wedge in a drift scan is expected to be similar to the tracking mode; we verify it using analytic estimates and simulations but do not show it here.

<sup>7</sup> Foreground components from both the point sources and diffuse galactic emission are expected to be dominated by synchrotron radiation from the power-law energy distribution of relativistic electrons. The galaxy is optically thin to these photons; therefore, the observed spectrum retains the form of the emitted spectrum, which is featureless. The main mechanism of the absorption of radio photons in the interstellar medium is free-free absorption of thermal and nonthermal electrons. The optical depth of free-free absorption  $\tau = 3.3 \times 10^{-7} (T/10^4)^{-1.35} \nu^{-2.1} \text{EM}$ , where  $\nu$  is in gigahertz and EM, the emission measure, is observationally determined to be  $\text{EM} = 5 \text{ pc cm}^{-3}$  (e.g., Haffner et al. 1999); the optical depth is negligible at frequencies of interest to us.



**Figure 4.** The visibility correlation function (Equation (30)) is shown as a function of  $\Delta t$  (normalized to unity for  $\Delta t = 0$ ) for two baselines  $\sqrt{u_0^2 + v_0^2} = 20, 100$  for  $u_0 = v_0$ , for the MWA primary beam and latitude. The visibility correlation function is seen to fall to half its value in a few minutes.

### 3.2. Diffuse Correlated Foregrounds

An important contribution to the foregrounds comes from diffuse galactic emission (DGE), which is correlated on the sky plane; this component of the foregrounds is dominated by optically thin galactic synchrotron emission. The spatial and frequency dependence of this emission is separable if the emission is optically thin, which, as noted above, is a good assumption and is key to the separation of foregrounds from the HI signal. We consider a statistically homogeneous and isotropic component of the diffuse foreground here. This case differs from the HI signal only in different frequency dependencies of the two signals. Therefore, the formulation is similar to the case of the HI signal discussed above.

As we assume the DGE to be statistically homogeneous and isotropic, the two-point function of fluctuations on the plane of the sky in Fourier space could be characterized by a power spectrum  $C_q$  such that  $q = |\mathbf{q}| = \sqrt{q_1^2 + q_2^2}$ , where  $\mathbf{q} = (q_1, q_2)$ , with  $q_1$  and  $q_2$  being the two Fourier components on the sky plane.  $C_q$  can be expressed as

$$\langle I_\nu(\mathbf{q}) I_{\nu'}(\mathbf{q}') \rangle = (2\pi)^2 C_q(\nu, \nu') \delta^2(\mathbf{q} - \mathbf{q}'). \quad (31)$$

For our analysis, we adopt the following form and normalization of  $C_q$ , as appropriate for  $\nu \simeq 150$  (e.g., Ghosh et al. 2012 and references therein):

$$C_q(\nu, \nu') = a_0 \left( \frac{\nu}{\nu_0} \right)^{-\alpha} \left( \frac{\nu'}{\nu_0} \right)^{-\alpha} \left( \frac{q}{1000} \right)^{-\gamma}, \quad (32)$$

where  $\alpha = 0.52$  (Rogers & Bowman 2008) is the spectral index and  $\gamma = 2.34$  (Ghosh et al. 2012) is the index of the spatial power spectrum. The value of  $a_0 = A_0 (2k_B \nu_0^2 / c^2)^2$  is  $237 \text{ Jy}^2$  at  $\nu_0 = 154 \text{ MHz}$ . It rescales the amplitude factor,  $A_0 = 513 \text{ mK}^2$ , given in Ghosh et al. (2012) from  $(\text{mK})^2$  at  $150 \text{ MHz}$  to  $\text{Jy}^2$  at  $\nu_0$ . For a single polarization, this factor should be divided by 4.

Using the formalism used for analyzing the HI signal, it can readily be shown that the visibility correlation function in

frequency space can be related to  $C_q$  as

$$\begin{aligned} & \langle V_\nu(\mathbf{u}_\nu, w_\nu, t) V_{\nu'}^*(\mathbf{u}'_{\nu'}, w'_{\nu'}, t') \rangle \\ &= \int \frac{d^2 q}{(2\pi)^2} C_q(\nu, \nu') e^{iq_1 \cos \phi \Delta H} Q_\nu \\ & \quad \times (\mathbf{q}, \mathbf{u}_\nu, w_\nu, \Delta H = 0) Q_{\nu'}^*(\mathbf{q}, \mathbf{u}'_{\nu'}, w'_{\nu'}, \Delta H), \end{aligned} \quad (33)$$

where the Fourier beam of DGE is

$$\begin{aligned} Q_\nu(\mathbf{q}, \mathbf{u}_\nu, w_\nu, \Delta H) &= \int d^2 \theta A_\nu(\theta) \\ & \quad \times \exp \left[ -2\pi i \left( \mathbf{x}_u \cdot \theta - \frac{1}{2} y \theta^2 \right) \right], \end{aligned} \quad (34)$$

$$\text{with } x_u = u_\nu - \frac{1}{2\pi} (q_1 + q_2 \sin \phi \Delta H), \quad (35)$$

$$x_v = v_\nu - \frac{1}{2\pi} (q_2 - q_1 \sin \phi \Delta H), \quad (36)$$

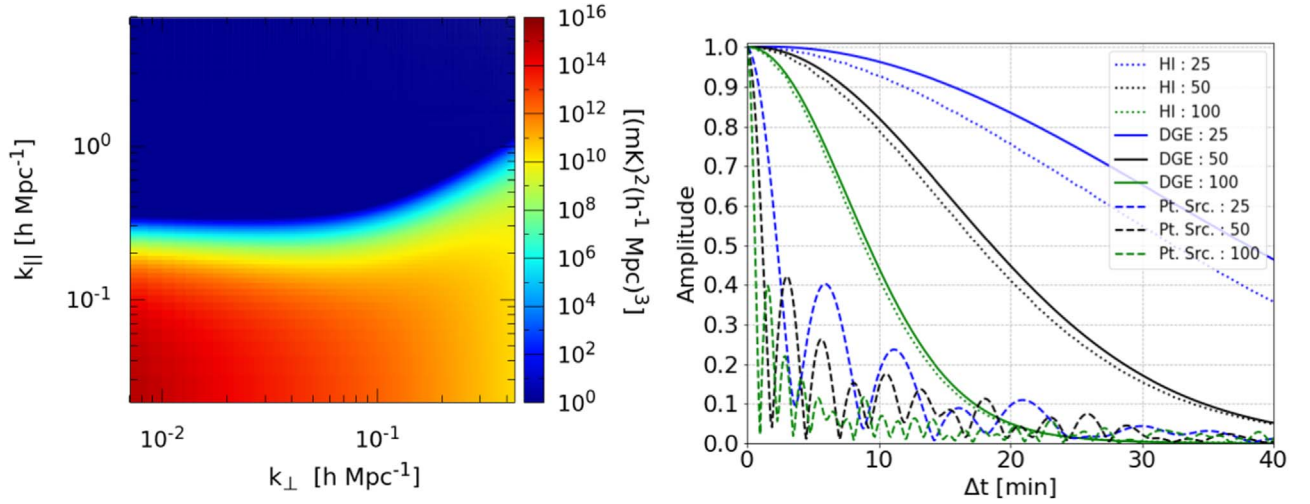
$$y = w_\nu + \frac{1}{2\pi} q_1 \cos \phi \Delta H. \quad (37)$$

In Equation (34), we have used  $Q$ -integrals (or 2D Fourier beam) defined for the HI correlation function (Equation (10)). Comparing Equations (34) and (10), we note the following relation between the Fourier modes of correlated diffuse foregrounds and the HI signal:  $\mathbf{q} \simeq r_0 \mathbf{k}_\perp$ .

As already shown for the HI signal, Equation (34) can be made more tractable by assuming the primary beam to be separable and symmetric. To establish general characteristics of the DGE foreground, we carry out analytical calculations with a symmetric Gaussian beam:  $e^{-(l^2+m^2)/\Omega_g}$ , which allows us to extend the integration limits from  $-\infty$  to  $+\infty$ . Following the HI analysis, we also expand  $n$  to first order. This gives us

$$Q_\nu(\mathbf{q}, \mathbf{u}_\nu, w_\nu, \Delta H) = \pi \Omega'_g \exp[-\pi^2 \Omega'_g (x_u^2 + x_v^2)], \quad (38)$$

where  $\Omega'_g = \Omega_g / (1 - i\pi y \Omega_g)$ . It should be noted that these variables can be read off directly from  $Q$ -integrals defined for the HI signal by putting  $r_0 \mathbf{k}_\perp \simeq \mathbf{q}$ . This shows the equivalence of the HI signal and diffuse foregrounds in the Fourier domain on the plane of the sky.



**Figure 5.** In the left panel, we show the two-dimensional power spectrum of DGE ( $\Delta H = 0$ ) in the  $k_{\parallel} - k_{\perp}$  plane in units of  $(\text{mK})^2 (h^{-1} \text{Mpc})^3$ . The figure assumes  $\nu_0 = 154$  MHz and bandwidth  $B = 10$  MHz. The relation applicable to the HI signal is used to transform from the telescope variables  $(u_0, v_0, \tau)$  to the Fourier modes  $(k_{\perp}, k_{\parallel})$  and to convert the power spectrum to the appropriate units (e.g., Paul et al. 2016). The figure highlights the separation of foregrounds from the EoR window; the bandwidth determines the extent of the flat region parallel to the  $k_{\parallel}$  axis. In the right panel, the visibility correlation function (normalized to unity for  $\Delta t = 0$ ) for DGE is shown for three baselines  $\sqrt{u_0^2 + v_0^2} = 25, 50, 100$  (Equation (40)). We also show the HI and point-source visibility correlation functions for comparison.

We next carry out frequency integrals to transform to delay space. As already discussed in Section 3.1, the main results in the delay space can be obtained by retaining only the frequency dependence of baselines because the foreground wedge in the two-dimensional power spectrum of foregrounds arises largely due to the chromaticity of baselines (e.g., Paul et al. 2016).

The frequency integral can be computed numerically for a finite bandpass. To carry out analytical calculations, the limits of the frequency integral can be extended to infinity. However, under this assumption, the baseline  $(\mathbf{u}_{\nu} = \mathbf{u}_0 \nu / \nu_0)$  also becomes infinity, and the integral does not converge.<sup>8</sup> To correctly pick the relevant scales of the diffuse foregrounds, we apply a Gaussian window function in frequency space ( $\exp(-c_2(\nu - \nu_0)^2)$ ), which allows us to pick the relevant scales within the bandwidth ( $B$ ) of the instrument and also enables us to extend the limits of integration.<sup>9</sup> This gives us

$$\begin{aligned}
 & \tilde{Q}(\mathbf{q}, \mathbf{u}_0, w_0, \Delta H) \\
 &= \int_{\nu_0 - B/2}^{\nu_0 + B/2} d\nu e^{2\pi i \tau \nu} e^{-c_2(\nu - \nu_0)^2} Q_{\nu}(\mathbf{q}, \mathbf{u}_{\nu}, w_{\nu}, \Delta H) \\
 &= \pi \Omega'_g \sqrt{\frac{\pi}{c_1 + c_2}} \exp\left[-\frac{\pi^2 \tau^2}{c_1 + c_2}\right] \\
 & \quad \times \exp\left[2\pi i \tau \nu_0 \left(1 + \frac{c_1}{c_1 + c_2} \frac{1}{|\mathbf{q}_u|} (a_1 + a_2 \sin \phi \Delta H)\right)\right] \\
 & \quad \times \exp\left[-\frac{\Omega'_g}{4} \left(\frac{c_2}{c_1 + c_2} (a_1 + a_2 \sin \phi \Delta H)^2\right.\right. \\
 & \quad \left.\left.+ (a_2 - a_1 \sin \phi \Delta H - |\mathbf{q}_u| \sin \phi \Delta H)^2\right)\right],
 \end{aligned} \tag{39}$$

<sup>8</sup> This highlights the main difference between the HI signal and the two-dimensional diffuse foregrounds. In the former, the frequency integral picks the scale along the line of sight  $k_{\parallel}$  while no such scale exists for diffuse foregrounds.

<sup>9</sup> A similar window (e.g., Blackman–Nuttall window; e.g., Paul et al. 2016) is applied to the data to prevent the leakage of foregrounds from the foreground wedge to the clean EoR window.

where  $c_1 = (|\mathbf{q}_u|/\nu_0)^2 \Omega'_g/4$ ,  $c_2 = 1/(bB^2)$ ,  $\mathbf{q}_u = 2\pi\mathbf{u}_0$ ,  $a_1 = q_1 - 2\pi u_0$ , and  $a_2 = q_2 - 2\pi v_0$ . The parameter  $b$  is a numerical factor that can be tuned to get the desired width of the Gaussian window function. The argument of the factor  $\exp[-2\pi^2 \tau^2 / (c_1 + c_2)]$  in Equation (39) yields the linear relation corresponding to the foreground wedge.

We can read off the correlation scales for diffuse correlation foregrounds from Equation (39). A baseline  $\mathbf{u}_0$  is most sensitive to the Fourier mode  $\mathbf{q}_u$ . As in the case of the HI signal, the decorrelation timescale for a drift scan can be estimated readily by putting  $\mathbf{q} = \mathbf{q}_u$  and simplifying the expression. We finally obtain

$$\begin{aligned}
 & \langle V_{\tau}(\mathbf{u}_0, w_0, t) V_{\tau}^*(\mathbf{u}'_0, w'_0, t') \rangle \\
 &= \iint_{\nu_0 - B/2}^{\nu_0 + B/2} d\nu d\nu' \langle V_{\nu}(\mathbf{u}_{\nu}, w_{\nu}, t) V_{\nu'}^*(\mathbf{u}'_{\nu'}, w'_{\nu'}, t') \rangle \\
 &= \int \frac{d^2 q}{(2\pi)^2} C_q(\nu_0, \nu_0) e^{iq_1 \cos \phi \Delta H} \tilde{Q}(\mathbf{q}, \mathbf{u}_0, \Delta H = 0) \\
 & \quad \times \tilde{Q}^*(\mathbf{q}, \mathbf{u}'_0, \Delta H).
 \end{aligned} \tag{40}$$

Equation (40) gives the general expression for the visibility correlation function in delay space for a drift scan observation. It can be computed by using Equations (32) and (39) in Equation (40). It reduces to the relevant expression for tracking observation for  $\Delta H = 0$ . In Figure 5, we show numerical results obtained from solving Equation (40) for a Gaussian primary beam matched to the main lobe of the MWA primary beam and  $\phi = -26.7^\circ$ . We display the power spectrum in  $k_{\parallel} - k_{\perp}$  plane for  $\Delta H = 0$  and the correlation of diffuse correlated foregrounds as a function of time. Our main conclusions are:

1. Like the point sources, diffuse correlated foregrounds are confined to a wedge and the EoR window is clean for the detection of the HI signal.
2. The diffuse foregrounds decorrelate on timescales comparable to the HI signal. (We note that the difference between the two cases for the shortest baseline is partly because we use the exact MWA beam for the HI case

while we use the Gaussian beam for the diffuse foreground.) This should be contrasted with point-source foregrounds that decorrelate on a much shorter timescale as compared to the HI signal.

#### 4. Analyzing Drift Scan Data

Our study allows us to address the following question: over what time period can the time-ordered visibility data be averaged without diminishing the HI signal? We further seek the optimal signal to noise for the detection of the HI signal. We computed the two-point visibility correlation function to assess the coherence timescale of visibilities. Our results are shown in Figures 1–2 (amplitude of the correlation function as a function of  $\Delta t$ , baseline and primary beam) and 3 (the phase of the complex correlation function). Our study shows that the range of timescales over which time-ordered visibilities can be averaged without the loss of HI signal lies in the range of a few minutes to around 20 minutes.

Motivated by our theoretical analysis, we define the quantity

$$\mathcal{C}_\tau(\mathbf{u}_0, w_0, t' - t) \equiv \exp(-i2\pi u_0 \cos \phi \Delta H) \times \langle V_\tau(\mathbf{u}_0, w_0, t) V_\tau^*(\mathbf{u}_0, w_0, t') \rangle. \quad (41)$$

Notice that  $\mathcal{C}_\tau(\mathbf{u}_0, w_0, t' - t) = \mathcal{C}_\tau^*(\mathbf{u}_0, w_0, t - t')$ . Our analysis shows that the complex number  $\mathcal{C}_\tau(\mathbf{u}_0, w_0, t' - t)$  is dominated by its real component with a phase that remains small over the coherence timescale of the amplitude (Figures 3 and 2). Our aim is to extract  $\mathcal{C}_\tau(\mathbf{u}_0, w_0, t' - t)$  from the data and then suitably weight it to extract the HI signal, optimally and without the loss of HI signal.<sup>10</sup> We discuss two possible ways to extract the HI signal. The first is based on averaging the visibilities before computing the correlation function.

We consider visibilities measured with time resolution  $\Delta H$  ( $\Delta H$  is assumed to be much smaller than the coherence scale of visibilities for any baseline of interest to us, e.g.,  $\Delta H = 10$  s). Let us denote the measured visibilities as  $V_n$ , where  $n$  corresponds to the time stamp; each visibility is a function of baseline and either  $\nu$  or  $\tau$ . As noted above, we could use data in either frequency or delay space. For the discussion here, we consider delay space and express all quantities as functions of  $\nu_0$ . For brevity, we only retain the time dependence of measured visibilities. We define

$$\mathcal{V} = \sum_{n=1}^N \exp(i2\pi u_0 \cos \phi \Delta H n) V_n. \quad (42)$$

The total time over which the visibilities are averaged  $T = N \Delta H$  should be small enough such that the signal decorrelation is negligible (Figure 1). For instance, we could choose  $N$  such that the decorrelation is 0.9, which corresponds roughly to 10 minutes for MWA for  $\sqrt{u_0^2 + \nu_0^2} \simeq 20$ . It also follows that if the

<sup>10</sup> To prevent HI signal loss, the simplest way to extract the HI signal from drift scans would be to not use the coherence of visibilities in time. Assuming visibilities are measured with time resolution much shorter than the coherence timescale, visibilities with identical time stamps can be squared (after averaging over redundant baselines) to compute the power spectrum. This gives an unbiased estimator of the HI signal. However, in such a procedure, visibilities measured at two different times are treated as uncorrelated, which results in an estimator with higher noise compared to what is achievable using further information regarding coherence of visibilities in time. If the time resolution of visibilities is around 10 s and the coherence time is around 10 minutes, then the noise rms of the visibility correlation is higher by roughly the square root of the ratio of these two times.

visibilities are averaged for a period much longer than the correlation scale of the signal, there would be serious loss of the HI signal. Even though we define  $\mathcal{V}$  for a single baseline  $\mathbf{u}_0$ , it can also be obtained by averaging visibilities over all redundant baselines. The correlation function that extracts the HI signal  $|\langle V_\tau(\mathbf{u}_0, w_0, t) V_\tau^*(\mathbf{u}_0, w_0, t) \rangle|$  then is

$$\mathcal{C}_{\text{HI}} \simeq \frac{1}{N^2} \mathcal{V} \mathcal{V}^*. \quad (43)$$

Notice that  $\mathcal{C}_{\text{HI}}$  is nearly the same as the expression in Equation (41) in this case. A longer stream of data of length  $K \gg N$  can be divided into time slices of  $N \Delta H$ . The correlation function can be estimated for each slice using this method (coherent averaging as the number of pairs is  $\simeq N^2$ ) and then averaged further over different time slices (incoherent averaging over  $K/N$  slices).  $\mathcal{C}_{\text{HI}}$  is also optimal as the noise rms is nearly the same for each pair of correlated visibilities. We note that the HI signal is mostly contained in the real part of this resulting function, as the phase angle is small for timescales over which the visibilities are averaged (Figure 3).

A much better method to utilize the functional form shown in Figure 1 is to use the estimator

$$\mathcal{C}_{\text{HI}} \simeq \frac{1}{N^2} \sum_{n'} \sum_n \exp(-i2\pi u_0 \cos \phi \Delta H (n' - n)) \times V_n V_{n'}^* g^{-1}(n' - n). \quad (44)$$

Here,  $g(n' - n)$  corresponds to the time decorrelation function shown in Figure 1; by construction,  $g(n' - n)$  is real,  $g(n - n) = 1$ , and  $g(n' - n) = g(n - n')$ . The difference between this approach and the first method is that visibilities are correlated first and then averaged. This yields the same final expression as the first method if  $g(n' - n)$  is applied for a suitable time interval such that it is close to unity. A distinct advantage of this method is that we could only retain cross-correlations such that  $n' \neq n$ , which allows us to avoid self-correlation or noise bias; the total number of cross-correlations are  $\simeq N^2/2$  in this case. This estimator is unbiased with respect to the detection of the HI signal but does not minimize noise rms. The following estimator is both unbiased and optimal:

$$\mathcal{C}_{\text{HI}} = \frac{\sum_{n'} \sum_n \exp(-i2\pi u_0 \cos \phi \Delta H (n' - n)) V_n V_{n'}^* g(n' - n)}{\sum_{n'} \sum_n g^2(n' - n)}. \quad (45)$$

The estimator is unbiased for any choice of  $g(n' - n)$ . However, when using this estimator, small values of  $g(n' - n)$  (e.g.,  $g(n' - n) < 0.3$ ) should be avoided to prevent averaging over very noisy visibility pairs. As in the first method, the real part of this function dominates the HI signal.

The amplitude of  $\mathcal{C}_{\text{HI}}$  for both the proposed estimators extracts the visibility correlation function at equal time,  $\langle V_\tau(\mathbf{u}_0, w_0, t) V_\tau^*(\mathbf{u}_0, w_0, t) \rangle$ , which is real. The estimation of HI power spectrum from this function has been extensively studied in the analysis of EoR tracking data (e.g., Paul et al. 2016).

Our method has similarities with other approaches proposed to analyze the drift scan data. In Parsons et al. (2016), the

fringe-rate filters have been applied on the visibility data. We apply a similar filter to reduce the rapid oscillations of the phase of the correlation function. We note that the filter applied in Parsons et al. (2016) takes into account all the components of Earth’s rotation (Equation (49)). In our analysis, we identify the different roles played by these components. We show how the components responsible for the rotation and translation of the intensity pattern cause the decorrelation of the amplitude of the correlation function while the component that gives rise to the translation dominates the phase of the correlation function. In  $m$ -mode analysis (Shaw et al. 2014, 2015), the intensity pattern is expanded using spherical harmonics and the time variation of the intensity pattern is solely owing to the change in the azimuthal angle  $\phi$ . This time variation can then be Fourier transformed to extract  $m$ -modes of the data. The filter we apply in Equation (42) corresponds to a similar process. Equation (42) can be viewed as a Fourier transform in which a single mode is extracted for a time window of the duration given roughly by the decorrelation time of the amplitude of the correlation function. Our analysis shows that such a procedure, directly applied on measured visibilities, can extract the relevant information of the HI signal.

#### 4.1. Impact on Foregrounds

The measured visibilities are a linear sum of the HI signal, foregrounds, and noise, which are uncorrelated with each other. In this paper, we also compute the timescale of the decorrelation of a set of point sources and statistically homogeneous and isotropic diffuse foregrounds. Does our method allow us to mitigate foregrounds?

First, we notice that the phase factor  $\exp(-i2\pi \cos \phi u_0 \Delta H)$  we apply to curtail rapid oscillations of the correlation function of the HI signal has the same form for the foregrounds (Equations (33) and (29)). Hence, it does not play a role in separating foregrounds from the HI signal.

However, the decorrelation timescale of point sources is smaller than the HI signal. In this case, the following situation is possible: two visibilities separated in time are correlated such that the HI component is fully extracted ( $g(n' - n) = 1$ ) but the point-source component is uncorrelated. This means that there would be partial decorrelation of this component of foregrounds when either of the two methods discussed above are used to extract the HI signal. But this argument does not apply to diffuse foregrounds.

Therefore, it is possible to partly reduce the level of foregrounds in a drift scan but the primary method of separating foregrounds from the HI signal remains the transformation to delay space, as in a tracking observation.

### 5. Summary and Conclusion

In this paper, we address the following question: over what timescales are time-ordered visibilities coherent in a drift scan for the EoR HI signal, set of point sources, and diffuse correlated foregrounds? This is an extension of our earlier work (Paul et al. 2014) and has similarities with other approaches in the literature (Shaw et al. 2014; Parsons et al. 2016). Our main theoretical tool is the complex two-point correlation function of visibilities measured at different times. We consider the primary beams of PAPER, MWA, HERA, and SKA1-Low for our analysis. Our main results can be summarized as follows:

1. Figure 1 shows the amplitude of the correlation function of HI visibilities in time for four interferometers. The correlation timescales vary from a few minutes to nearly 20 minutes for the cases considered. We identify the three most important factors that cause decorrelation: (a) traversal time across a coherent feature, (b) rotation of sky intensity pattern, and (c) large field of view.
2. The time variation of the phase of the HI correlation function is dominated by a filter function that is determinable in terms of measurable quantities (component of east–west baseline, latitude of the telescope, etc.). This filter function can be absorbed into an overall phase. The phase angle of the resultant function is small, which means the complex correlation function is dominated by its real part. The phase angle remains small over the coherence timescale of the amplitude of the correlation function (Figure 3).
3. Our results are valid in both frequency and delay space and are insensitive to the input HI power spectrum. By implication, they are directly applicable to the analysis of EoR drift scan data.
4. The nature of foregrounds in a drift scan is different from the tracking mode owing to the time dependence of the sky intensity pattern. We consider two components of foregrounds for our analysis: the set of point sources and statistically homogeneous diffuse correlated emission. The decorrelation timescales for these components are displayed in Figures 4 and 5. The point sources decorrelate faster than the HI signal. This provides a novel way to partly mitigate foregrounds using only information on the sky plane. However, the diffuse foreground decorrelation timescale is comparable to that of the HI signal, and the contamination from this component cannot be removed in a drift scan on the sky plane. By implication, the delay space formalism remains the principal method for isolating foregrounds from the HI signal (Figure 5).

We discuss in detail how our formalism can be used to extract the HI signal from the drift scan data. We argue that many different approaches might be possible for the lossless retrieval of the HI signal while optimizing the noise. In the future, we hope to apply our formalism to publicly available drift scan data.

### Appendix A Coordinate Transformation

Here we discuss the sky coordinate system  $(l, m, n)$  in terms of  $(\delta, \phi, H)$  with  $\delta, \phi, H$  representing the declination, the terrestrial latitude of the telescope, and the hour angle, respectively. From Equation (A4.7) of Christiansen & Hoegbom (1969),

$$\begin{aligned} l &= \cos \delta \sin H \\ m &= \cos \delta \cos H \sin \phi - \sin \delta \cos \phi \\ n &= \cos \delta \cos H \cos \phi + \sin \delta \sin \phi. \end{aligned} \quad (46)$$

In a drift scan, the primary beam remains unchanged with respect to a fixed phase center chosen to be  $l = m = 0$ . The coordinates of intensity pattern  $(l, m, n)$  change with time, in

the first order in  $\Delta H$ , as

$$\begin{aligned}\Delta l &= (m \sin \phi + n \cos \phi) \Delta H \\ \Delta m &= -l \sin \phi \Delta H \\ \Delta n &= -l \cos \phi \Delta H.\end{aligned}\quad (47)$$

The change in hour angle,  $\Delta H$ , can be expressed in terms of radians as

$$\Delta H [\text{in rad}] = \frac{\pi}{12} \frac{\Delta t [\text{in minutes}]}{60}.\quad (48)$$

We use Equation (47) to express the time-dependent part of Equation (9) explicitly in terms of change in hour angle  $\Delta H$ . Equation (48) can be used to express  $\Delta H$  in terms of drift time  $\Delta t$  for a zenith scan:

$$\begin{aligned}-\frac{r_0}{2\pi} \mathbf{k}_\perp \cdot \Delta \vartheta (\Delta t) &= -\frac{r_0}{2\pi} (k_{\perp 1} \Delta l + k_{\perp 2} \Delta m) \\ &= -\frac{r_0}{2\pi} (k_{\perp 1} (m \sin \phi + n \cos \phi) \Delta H - k_{\perp 2} l \sin \phi \Delta H) \\ &\simeq -\frac{r_0}{2\pi} (k_{\perp 1} \cos \phi \Delta H + (-l k_{\perp 2} + m k_{\perp 1}) \sin \phi \Delta H) \\ &\quad + \frac{1}{2} (l^2 + m^2) \frac{r_0}{2\pi} k_{\perp 1} \cos \phi \Delta H.\end{aligned}\quad (49)$$

We use the flat-sky approximation  $n \simeq 1 - \frac{1}{2}(l^2 + m^2)$  in writing Equation (49).

### Appendix B Further Simplification of the Visibility Correlation Function

In this appendix, we discuss how the visibility correlation function can be further simplified for large primary beams and long baselines. This allows us to discern several generic properties of the correlation function. We start with the HI visibility correlation function in frequency space (Equation (9)):

$$\begin{aligned}\langle V_\nu(\mathbf{u}_\nu, w_\nu, t) V_{\nu'}^*(\mathbf{u}'_{\nu'}, w'_{\nu'}, t') \rangle \\ &= \bar{I}_\nu \bar{I}_{\nu'} \int \frac{d^3 k}{(2\pi)^3} P_{\text{HI}}(k) e^{ik_\parallel \dot{r}_0 |\Delta \nu|} e^{i r_\nu k_{\perp 1} \cos \phi \Delta H} \\ &\quad \times Q_\nu(\mathbf{k}_\perp, \mathbf{u}_\nu, w_\nu, \Delta H = 0) Q_{\nu'}^*(\mathbf{k}_\perp, \mathbf{u}'_{\nu'}, w'_{\nu'}, \Delta H).\end{aligned}$$

The Fourier beam can be expressed as (Equation (10))

$$\begin{aligned}Q_\nu(\mathbf{k}_\perp, \mathbf{u}_\nu, w_\nu, \Delta H) &= \int d^2 \theta A_\nu(\boldsymbol{\theta}) \\ &\quad \times \exp \left[ -2\pi i \left( \mathbf{x}_u \cdot \boldsymbol{\theta} - \frac{1}{2} y \theta^2 \right) \right],\end{aligned}\quad (50)$$

with

$$\begin{aligned}x_u &= u_\nu - \frac{r_\nu}{2\pi} (k_{\perp 1} + k_{\perp 2} \sin \phi \Delta H) \\ x_v &= v_\nu - \frac{r_\nu}{2\pi} (k_{\perp 2} - k_{\perp 1} \sin \phi \Delta H) \\ y &= w_\nu + \frac{r_\nu}{2\pi} k_{\perp 1} \cos \phi \Delta H.\end{aligned}$$

We consider a Gaussian beam  $A(l, m) = e^{-(l^2+m^2)/\Omega_g}$  to compute the Fourier beam:

$$\begin{aligned}Q_\nu(\mathbf{k}_\perp, \mathbf{u}_\nu, w_\nu, \Delta H) &= Q(x_u, x_v, y) = \frac{\pi \Omega_g}{1 - i\pi y \Omega_g} \\ &\quad \times \exp \left[ -\frac{\pi^2 \Omega_g (x_u^2 + x_v^2)}{1 - i\pi y \Omega_g} \right].\end{aligned}$$

For  $\Omega'_g \equiv \Omega_g / (1 - i\pi y \Omega_g)$

$$\begin{aligned}Q_\nu(\mathbf{k}_\perp, \mathbf{u}_\nu, w_\nu, \Delta H) &= Q(x_u, x_v, y) \\ &= \pi \Omega'_g \times \exp[-\pi^2 \Omega'_g (x_u^2 + x_v^2)].\end{aligned}$$

If  $\Omega_g$  is large, e.g., PAPER or MWA beams, we can use the  $\delta$ -function approximation to solve  $Q_\nu(\mathbf{k}_\perp, \mathbf{u}_\nu, w_\nu, \Delta H = 0)$ , which gives us

$$\begin{aligned}Q_\nu(\mathbf{k}_\perp, \mathbf{u}_\nu, w_\nu, \Delta H = 0) &= \delta \left( u_\nu - \frac{r_\nu}{2\pi} k_{\perp 1} \right) \delta \left( v_\nu - \frac{r_\nu}{2\pi} k_{\perp 2} \right) \\ Q_\nu(\mathbf{k}_\perp, \mathbf{u}_\nu, w_\nu, \Delta H = 0) &= \left( \frac{2\pi}{r_\nu} \right)^2 \delta^2 \left( \mathbf{k}_\perp - \frac{2\pi}{r_\nu} \mathbf{u}_\nu \right).\end{aligned}$$

This allows us to express the HI visibility correlation function in frequency space as

$$\begin{aligned}\langle V_\nu(\mathbf{u}_\nu, w_\nu, t) V_{\nu'}^*(\mathbf{u}'_{\nu'}, w'_{\nu'}, t') \rangle \\ &= \frac{\bar{I}_\nu \bar{I}_{\nu'}}{r_\nu^2} e^{2\pi i u_\nu \cos \phi \Delta H} Q_{\nu'}^*(\mathbf{k}_\perp, \mathbf{u}'_{\nu'}, w'_{\nu'}, \Delta H) \\ &\quad \times \int \frac{dk_\parallel}{2\pi} P_{\text{HI}}(k) e^{ik_\parallel \dot{r}_0 |\Delta \nu|}.\end{aligned}\quad (51)$$

In the previous equation, we have used  $\mathbf{k}_\perp = 2\pi \mathbf{u}_\nu / r_\nu$ . Equation (51) gives an excellent approximation for MWA and PAPER, and for HERA and SKA1-Low for long baselines in frequency space. This can be readily be computed at any frequency and explains the features seen in Figure 1.

We can extend our analysis to the HI visibility correlation function in delay space (Equation (15)):

$$\begin{aligned}\langle V_\tau(\mathbf{u}_0, w_0, t) V_\tau^*(\mathbf{u}'_0, w'_0, t') \rangle \\ &= \iint_{\nu_0 - B/2}^{\nu_0 + B/2} d\nu d\nu' \langle V_\nu(\mathbf{u}_\nu, w_\nu, t) V_{\nu'}^*(\mathbf{u}'_{\nu'}, w'_{\nu'}, t') \rangle e^{-2\pi i \tau \Delta \nu}.\end{aligned}$$

Here,  $B$  is the observational bandwidth. We make the same approximations discussed in Section 2.1, which gives us

$$\begin{aligned}\langle V_\tau(\mathbf{u}_0, w_0, t) V_\tau^*(\mathbf{u}'_0, w'_0, t') \rangle \\ &= \frac{\bar{I}_0^2}{r_0^2} e^{2\pi i u_0 \cos \phi \Delta H} Q_{\nu_0}^*(\mathbf{k}_\perp, \mathbf{u}'_0, w'_0, \Delta H) \\ &\quad \times \int \frac{dk_\parallel}{2\pi} P_{\text{HI}}(k) \iint_{\nu_0 - B/2}^{\nu_0 + B/2} d\nu d\nu' e^{i \Delta \nu (k_\parallel \dot{r}_0 - 2\pi \tau)} \\ &\langle V_\tau(\mathbf{u}_0, w_0, t) V_\tau^*(\mathbf{u}'_0, w'_0, t') \rangle \\ &\simeq \frac{\bar{I}_0^2}{r_0^2} e^{2\pi i u_0 \cos \phi \Delta H} Q_{\nu_0}^*(\mathbf{k}_\perp, \mathbf{u}'_0, w'_0, \Delta H) \\ &\quad \times \int \frac{dk_\parallel}{2\pi} P_{\text{HI}}(k) \frac{2\pi B}{|\dot{r}_0|} \delta \left( k_\parallel - \frac{2\pi \tau}{|\dot{r}_0|} \right).\end{aligned}$$

In deriving this equation, we use the following result from Section 2.1:

$$\iint_{\nu_0-B/2}^{\nu_0+B/2} d\nu d\nu' e^{i\Delta\nu(k_{\parallel}\dot{r}_0-2\pi\tau)} \\ = B^2 \text{sinc}^2 \left[ \pi B \left( \tau - \frac{|\dot{r}_0|}{2\pi} k_{\parallel} \right) \right] \simeq \frac{2\pi B}{|\dot{r}_0|} \delta \left( k_{\parallel} - \frac{2\pi\tau}{|\dot{r}_0|} \right).$$

The H I signal is strongly correlated when  $|\mathbf{u}_0 - \mathbf{u}'_0| \lesssim 2/\Omega_g^{1/2}$ , which allows us to use  $\mathbf{u}'_0 \approx \mathbf{u}_0$ . This gives us

$$\langle V_{\tau}(\mathbf{u}_0, w_0, t) V_{\tau}^*(\mathbf{u}'_0, w'_0, t') \rangle \\ \simeq \frac{\bar{I}_0^2 B}{r_0^2 |\dot{r}_0|} e^{2\pi i u_0 \cos \phi \Delta H} Q_{\nu_0}^*(\mathbf{k}_{\perp}, \mathbf{u}_0, w'_0, \Delta H) P_{\text{HI}}(k), \quad (52)$$

where  $k = \sqrt{(2\pi\tau/|\dot{r}_0|)^2 + (2\pi u_0/r_0)^2 + (2\pi\nu_0/r_0)^2}$ . Though Equation (52) was derived using a Gaussian beam, it is in excellent agreement with the numerical results for MWA and PAPER, and for HERA and SKA1-Low for longer baselines ( $|\mathbf{u}| \gtrsim 150$ ) shown in Figure 1. Equation (52) also shows that the decorrelation time is expected to be nearly independent of the delay parameter  $\tau$ .

We next give explicit forms of the amplitude and the phase of the Fourier beam. We have

$$Q_{\nu}(\mathbf{k}_{\perp}, \mathbf{u}_{\nu}, w_{\nu}, \Delta H) \\ = Q(x_u, x_v, y) = \frac{\pi\Omega_g}{1 - i\pi y\Omega_g} \exp \left[ -\frac{\pi^2\Omega_g(x_u^2 + x_v^2)}{1 - i\pi y\Omega_g} \right],$$

where  $x_u^2 + x_v^2 = |\mathbf{u}_{\nu}|^2 \sin^2 \phi \Delta H^2$  and  $y = w_{\nu} + u_{\nu} \cos \phi \Delta H$ . Then,

$$Q_{\nu}(\mathbf{k}_{\perp}, \mathbf{u}_{\nu}, w_{\nu}, \Delta H) = \pi z_1 z_2 \\ = \pi a_1 e^{i\psi_1} a_2 e^{i\psi_2} = \pi a_1 a_2 e^{i(\psi_1 + \psi_2)}$$

$$\text{Amp}[Q_{\nu}(\mathbf{k}_{\perp}, \mathbf{u}_{\nu}, w_{\nu}, \Delta H)] = \pi a_1 a_2$$

$$\text{Arg}[Q_{\nu}(\mathbf{k}_{\perp}, \mathbf{u}_{\nu}, w_{\nu}, \Delta H)] = \psi_1 + \psi_2$$

$$z_1 = a_1 e^{i\psi_1} = \frac{\Omega_g}{1 - i\pi y\Omega_g} \\ z_2 = a_2 e^{i\psi_2} = \exp \left[ -\frac{\pi^2\Omega_g(x_u^2 + x_v^2)}{1 - i\pi y\Omega_g} \right].$$

On solving  $a_1$ ,  $\psi_1$ ,  $a_2$ ,  $\psi_2$  in terms of known quantities, we find

$$a_1 = \frac{\Omega_g}{\sqrt{1 + \pi^2\Omega_g^2 y^2}} \\ \psi_1 = \arctan(\pi\Omega_g y) \\ a_2 = \exp[-\pi^2(x_u^2 + x_v^2)a_1 \cos \psi_1] \\ = \exp \left[ -\pi^2(x_u^2 + x_v^2) \frac{\Omega_g}{1 + \pi^2\Omega_g^2 y^2} \right] \\ \psi_2 = -\pi^2(x_u^2 + x_v^2)a_1 \sin \psi_1 \\ = -\pi^2(x_u^2 + x_v^2) \frac{\Omega_g}{1 + \pi^2\Omega_g^2 y^2} (\pi\Omega_g y).$$

Hence,

$$\text{Amp}[Q_{\nu}(\mathbf{k}_{\perp}, \mathbf{u}_{\nu}, w_{\nu}, \Delta H)] = \pi a_1 a_2 \\ = \frac{\pi\Omega_g}{\sqrt{1 + \pi^2\Omega_g^2 y^2}} \exp \left[ -\pi^2(x_u^2 + x_v^2) \frac{\Omega_g}{1 + \pi^2\Omega_g^2 y^2} \right] \\ \text{Arg}[Q_{\nu}(\mathbf{k}_{\perp}, \mathbf{u}_{\nu}, w_{\nu}, \Delta H)] = \psi_1 + \psi_2 \\ = \arctan(\pi\Omega_g y) - \pi^2(x_u^2 + x_v^2) \frac{\Omega_g}{1 + \pi^2\Omega_g^2 y^2} (\pi\Omega_g y). \quad (53)$$

The total phase acquired by the H I visibility correlation function is  $2\pi u_0 \cos \phi \Delta H + \psi_1 + \psi_2$ .

## ORCID iDs

Akash Kumar Patwa  <https://orcid.org/0000-0002-6216-2430>

## References

- Ali, Z. S., Parsons, A. R., Zheng, H., et al. 2015, *ApJ*, 809, 61  
 Ali, Z. S., Parsons, A. R., Zheng, H., et al. 2018, *ApJ*, 863, 201  
 Barry, N., Wilensky, M., Trott, C. M., et al. 2019, *MNRAS*, 484, 1  
 Beardsley, A. P., Hazelton, B. J., Sullivan, I. S., et al. 2016, *ApJ*, 833, 102  
 Bowman, J. D., Cairns, I., Kaplan, D. L., et al. 2013, *PASA*, 30, 31  
 Bowman, J. D., Rogers, A. E. E., Monsalve, R. A., Mozdzien, T. J., & Mahesh, N. 2018, *Natur*, 555, 67  
 Cheng, C., Parsons, A. R., Kolopanis, M., et al. 2018, *ApJ*, 868, 26  
 Choudhuri, S., Bharadwaj, S., Chatterjee, S., et al. 2016, *MNRAS*, 463, 4093  
 Christiansen, W. N., & Hoegbom, J. A. 1969, *Radiotelescopes* (Cambridge: Cambridge Univ. Press)  
 Cornwell, T. J., Golap, K., & Bhatnagar, S. 2008, *ISTSP*, 2, 647  
 Datta, A., Bowman, J. D., & Carilli, C. L. 2010, *ApJ*, 724, 526  
 DeBoer, D. R., Parsons, A. R., Aguirre, J. E., et al. 2017, *PASP*, 129, 045001  
 Dillon, J. S., Neben, A. R., Hewitt, J. N., et al. 2015, *PhRvD*, 91, 123011  
 Eastwood, M. W., Anderson, M. M., Monroe, R. M., et al. 2018, *AJ*, 156, 32  
 Fan, X., Strauss, M. A., Becker, R. H., et al. 2006, *AJ*, 132, 117  
 Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, *PhR*, 433, 181  
 Ghosh, A., Prasad, J., Bharadwaj, S., Ali, S. S., & Chengalur, J. N. 2012, *MNRAS*, 426, 3295  
 Haffner, L. M., Reynolds, R. J., & Tufte, S. L. 1999, *ApJ*, 523, 223  
 Hopkins, A. M., Afonso, J., Chan, B., et al. 2003, *AJ*, 125, 465  
 Kolopanis, M., Jacobs, D. C., Cheng, C., et al. 2019, *ApJ*, 883, 133  
 Lanman, A. E., & Pober, J. C. 2019, *MNRAS*, 487, 5840  
 Morales, M. F., & Wyithe, J. S. B. 2010, *ARA&A*, 48, 127  
 Paciga, G., Chang, T.-C., Gupta, Y., et al. 2011, *MNRAS*, 413, 1174  
 Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, *A&A*, 594, A13  
 Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2018, arXiv:1807.06209  
 Parsons, A. R., & Backer, D. C. 2009, *AJ*, 138, 219  
 Parsons, A. R., Liu, A., Aguirre, J. E., et al. 2014, *ApJ*, 788, 106  
 Parsons, A. R., Liu, A., Ali, Z. S., & Cheng, C. 2016, *ApJ*, 820, 51  
 Parsons, A. R., Pober, J. C., Aguirre, J. E., et al. 2012, *ApJ*, 756, 165  
 Patil, A. H., Yatawatta, S., Koopmans, L. V. E., et al. 2017, *ApJ*, 838, 65  
 Paul, S., Sethi, S. K., Morales, M. F., et al. 2016, *ApJ*, 833, 213  
 Paul, S., Sethi, S. K., Subrahmanyam, R., et al. 2014, *ApJ*, 793, 28  
 Pritchard, J. R., & Loeb, A. 2012, *RPPh*, 75, 086901  
 Rogers, A. E. E., & Bowman, J. D. 2008, *AJ*, 136, 641  
 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, *PhRvD*, 91, 083514  
 Singh, S., Subrahmanyam, R., Shankar, N. U., et al. 2018, *ExA*, 45, 269  
 Taylor, G. B., Carilli, C. L., & Perley, R. A. (ed.) 1999, in ASP Conf. Ser. 180, Synthesis Imaging in Radio Astronomy II (San Francisco, CA: ASP)  
 Tingay, S. J., Goeke, R., Bowman, J. D., et al. 2013, *PASA*, 30, 7  
 Trott, C. M. 2014, *PASA*, 31, e026  
 Trott, C. M., Pindor, B., Procopio, P., et al. 2016, *ApJ*, 818, 139  
 van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, *A&A*, 556, A2  
 Wayth, R. B., Tingay, S. J., Trott, C. M., et al. 2018, *PASA*, 35, e033