

Detection of cosmic structures using the bispectrum phase. I. Mathematical foundationsNithyanandan Thyagarajan^{1,2,*} and Chris L. Carilli^{1,3}¹*National Radio Astronomy Observatory, Socorro, New Mexico 87801, USA*²*Arizona State University, School of Earth and Space Exploration, Tempe, Arizona 85287, USA*³*Astrophysics Group, Cavendish Laboratory, University of Cambridge, Cambridge CB3 0HE, United Kingdom*

(Received 19 December 2019; accepted 17 May 2020; published 1 July 2020)

Many low-frequency radio interferometers are aiming to detect very faint spectral signatures from structures at cosmological redshifts, particularly of neutral Hydrogen using its characteristic 21 cm spectral line. Due to the very high dynamic range needed to isolate these faint spectral fluctuations from the very bright foregrounds, spectral systematics from the instrument or the analysis, rather than thermal noise, are currently limiting their sensitivity. Failure to achieve a spectral calibration of the instrument with fractional inaccuracy $\lesssim 10^{-5}$ will make the detection of the critical cosmic signal unlikely. The bispectrum phase from interferometric measurements is largely immune to this calibration issue. We present a basis to explore the nature of the bispectrum phase in the limit of small spectral fluctuations. We establish that these fluctuations measure the intrinsic dissimilarity in the transverse structure of the cosmic signal relative to the foregrounds, expressed as rotations in the underlying phase angle. Their magnitude is related to the strength of the cosmic signal relative to the foregrounds. Using a range of sky models, we detail the behavior of the bispectrum phase fluctuations using standard Fourier-domain techniques and find it comparable to existing approaches, with a few key differences. Foreground contamination from *mode-mixing* between the transverse and line-of-sight dimensions is more pronounced than in existing approaches because the bispectrum phase is a product of three individual interferometric phases. The multiplicative coupling of foregrounds in the bispectrum phase fluctuations results in the mixing of foreground signatures with that of the cosmic signal. We briefly outline a variation of this approach to avoid extensive mode-mixing. Despite its limitations, the interpretation of results using the bispectrum phase is possible with forward-modeling. Importantly, it is an independent and a viable alternative to existing approaches.

DOI: [10.1103/PhysRevD.102.022001](https://doi.org/10.1103/PhysRevD.102.022001)**I. INTRODUCTION**

The formation and evolution of large-scale structure in the high-redshift Universe ($1 \lesssim z \lesssim 100$) has been largely underexplored. Probing the early Universe using spectral line tracers on cosmological scales appears promising and could be very rewarding scientifically to both understand these processes and their effects on shaping the astrophysical evolution of the Universe (see e.g., [1–14]). Examples include tomographic mapping of the neutral hydrogen (HI) using the redshifted 21 cm spectral line from the electron spin-flip transition during the Cosmic Dark Ages ($z \gtrsim 30$), the Cosmic Dawn ($15 \lesssim z \lesssim 30$), the Epoch of Reionization (EoR; $6 \lesssim z \lesssim 15$), and the periods when dark energy started becoming significant ($1 \lesssim z \lesssim 3$) and eventually dominant ($z < 1$).

Rapid advances in low radio frequency instrumentation has made it possible for a number of experiments including the Murchison Widefield Array [15–18], the Donald C. Backer Precision Array for Probing the Epoch of

Reionization [19], the Low Frequency Array [20], the Giant Metrewave Radio Telescope EoR experiment [21], the Hydrogen Epoch of Reionization Array [22], the Square Kilometre Array [23], the Canadian Hydrogen Intensity Mapping Experiment [24], and the Hydrogen Intensity and Real-time Analysis Experiment [25] to attempt detecting the cosmic HI structures in these epochs using the redshifted 21 cm spectral line with sufficient sensitivity (see e.g., [26,27]).

With the requirement to isolate very weak spectral signatures in the presence of very bright foreground objects, these experiments are faced with a tremendous challenge of requiring extreme-fidelity spectral calibration [28–31] with fractional inaccuracy typically under $\lesssim 10^{-5}$, failing which the miscalibration will leak sufficient power from the bright foregrounds and contaminate the spectral signatures of the cosmic signal, thereby rendering this critical detection impossible. While advanced calibration methods are being investigated to address the calibration challenge (see e.g., [32–35]), a new and independent approach to detecting the spectral signatures from cosmic structures using the bispectrum phase was presented

*t_nithyanandan@nrao.edu, nithyanandan.t@gmail.com

recently [36], which has the distinct advantage that it is largely impervious to antenna-based calibration and the errors therein. This property has been investigated in detail [37–41]. The bispectrum phase intrinsically measures the symmetry about a point and is invariant to translation [41]. Indeed, the bispectrum phase has been exploited successfully in interferometric imaging experiments where calibration is extremely challenging, such as in deciphering complex structures on stellar surfaces and their surroundings (see [42,43] and references therein), and in the Event Horizon Telescope (EHT) imaging of the shadow of the supermassive black hole at the center of M87 [44–49].

Most of the aforementioned applications of bispectrum phase were used for imaging which is restricted to the transverse plane of the sky. Following up on the idea presented in [36], we explore and exploit new properties of the bispectrum phase while applying it to the spectral (or the frequency) axis, which in cosmological spectral line observations is typically the line-of-sight dimension of the sky. This paper is one in a series of related papers, the others being [36,50–52], and lays the mathematical foundations for our understanding and application of the bispectrum phase approach for spectral line observations where information about the desired signal can be extracted through its distinct spectral signatures. In a companion paper [52], we present the first results from applying this technique to a small sample of data obtained with the HERA telescope.

This paper is organized as follows. In Sec. II, we present a basis using simple complex algebra to simplify the mathematical understanding of phase fluctuations in the limit of small perturbations. We apply this simplification first to phases in the interferometric two-point correlations (or visibilities) in Sec. III. We then extend this formalism to the bispectrum phase in Sec. IV. In Sec. V, we use a range of purely hypothetical to realistic examples to identify and illustrate the relationship and correspondence, and the benefits and limitations, of using our bispectrum phase approach relative to other existing approaches that use visibilities in distinguishing the spectral signatures from the cosmic signal. In Sec. VI, we demonstrate the presence and effects of *mode-mixing* (the coupling of spatial modes along the line-of-sight direction with those in the transverse plane) in our approach similar to that in existing approaches. In the Appendix, we briefly outline an alternate approach that could potentially mitigate contamination from mode-mixing effects. In Sec. VII, we discuss the impact of the intrinsic spectral characteristics of the foregrounds on the spectral signatures of the cosmic spectral line signal and contrast it with existing approaches. We provide a summary of the mathematical formalism of this approach along with its benefits and limitations in Sec. VIII. While we use examples relating to the detection of HI in the intergalactic medium during the EoR, the formalism and the conclusions presented here are

generically applicable to other experiments and science cases as well.

II. LINEAR ORDER FLUCTUATIONS ON COMPLEX VECTORS

Consider complex numbers Z_j , with amplitudes $|Z_j|$ and arguments θ_j in the complex plane, \mathbb{C} , such that $Z_j = |Z_j|e^{i\theta_j}$, for $j = 0, 1, 2, \dots, n$. The real and imaginary parts are denoted by $\Re\{\cdot\}$ and $\Im\{\cdot\}$ respectively. \bar{Z} denotes the complex conjugate of Z .

In this paper, we denote Z_0 as the reference complex vector, and Z_j for $1 \leq j \leq n$ as perturbations over Z_0 . Let $Z_\Sigma = \sum_{j=0}^n Z_j$ denote the resultant complex vector. Throughout this paper, we will often invoke that $|Z_j|/|Z_0| \ll 1$ for $j \geq 1$, so that only linear-order terms in $|Z_j|/|Z_0|$ will be retained while neglecting higher-order terms.

A. Small perturbations from a single cause

For a small perturbation arising from a single cause ($n = 1$), say only one of thermal noise or spectral line fluctuations, $Z_\Sigma = Z_0 + Z_1$. The amplitude and argument of the resultant are

$$|Z_\Sigma|^2 = |Z_0|^2 + |Z_1|^2 + 2\Re\{Z_0\bar{Z}_1\}, \quad (1)$$

$$\theta_\Sigma = \theta_0 + \delta\theta_0, \quad (2)$$

where,

$$\tan \delta\theta_0 = \frac{|Z_1| \sin(\theta_1 - \theta_0)}{|Z_0| + |Z_1| \cos(\theta_1 - \theta_0)}. \quad (3)$$

Assuming the perturbation is small, $|Z_1| \ll |Z_0|$, using Taylor-series expansion and small-angle approximation, and retaining only up to linear-order terms in $|Z_1|/|Z_0|$,

$$|Z_\Sigma| \approx |Z_0| \left(1 + \frac{|Z_1|}{|Z_0|} \cos(\theta_1 - \theta_0) \right) \quad (4)$$

$$= |Z_0| \left(1 + \Re\left\{ \frac{Z_1}{Z_0} \right\} \right), \quad (5)$$

and

$$\tan \delta\theta_0 \approx \delta\theta_0 \approx \frac{|Z_1|}{|Z_0|} \sin(\theta_1 - \theta_0) = \Im\left\{ \frac{Z_1}{Z_0} \right\}. \quad (6)$$

Equation (6) establishes that the fluctuation in the angle of the reference complex vector depends on $|Z_1|/|Z_0|$ and on the relative angle between the two, $\theta_1 - \theta_0$. Hence, if the fluctuation has zero magnitude ($|Z_1| = 0$) or if its angle is exactly aligned with the reference vector ($\theta_1 = \theta_0$), it will cause no perturbation in angle. Thus, perturbations in the underlying phase angle are an intrinsic measure of the dissimilarity of the perturbing signal relative to the

underlying signal, and the magnitude of the phase angle perturbation depends on the ratio between the two.

B. Small perturbations from multiple causes

The relationships can be generalized and extended in case of multiple causes of perturbations such as simultaneous presence of thermal noise and spectral line fluctuations. In the presence of multiple sources of perturbations, the same small-angle and linear-order approximation yields

$$|Z_{\Sigma}| \approx \sum_{j=0}^n |Z_j| \cos(\theta_j - \theta_0) \quad (7)$$

$$= |Z_0| \sum_{j=0}^n \frac{|Z_j|}{|Z_0|} \cos(\theta_j - \theta_0) \quad (8)$$

$$= |Z_0| \sum_{j=0}^n \Re \left\{ \frac{Z_j}{Z_0} \right\} = |Z_0| \Re \left\{ \sum_{j=0}^n \frac{Z_j}{Z_0} \right\}, \quad (9)$$

and

$$\delta\theta_0 \approx \sum_{j=0}^n \frac{|Z_j|}{|Z_0|} \sin(\theta_j - \theta_0) \quad (10)$$

$$= \sum_{j=0}^n \Im \left\{ \frac{Z_j}{Z_0} \right\} = \Im \left\{ \sum_{j=0}^n \frac{Z_j}{Z_0} \right\}. \quad (11)$$

It may be noted that when these perturbations arise from Gaussian noise and are small, wherein $|Z_j|/|Z_0| \ll 1$, the distribution of $\delta\theta_0$ can also be well approximated by a Gaussian distribution [53].

III. SPECTRAL LINE FLUCTUATIONS IN VISIBILITY PHASE

In this section, we develop the mathematical formalism on the interferometric visibility phases that will be later extended to the bispectrum phase. Let $V_p(f)$ denote the spectrum of visibilities of a triad of a baseline (antenna spacing) vectors, \mathbf{b}_p , where, $p = \{1, 2, 3\}$ indexes the baselines comprising the triad. Then, a measured visibility (after calibration) can be written as

$$V_p^m(f) = V_p^T(f) + V_p^N(f), \quad (12)$$

where superscripts m, T, and N denote the measured, true sky, and noise components respectively. The true sky visibility can be further decomposed as

$$V_p^T(f) = V_p^F(f) + V_p^L(f), \quad (13)$$

where F denotes the foregrounds and L denotes the cosmological spectral line signal of interest. Hence,

$$V_p^m(f) = V_p^F(f) + V_p^L(f) + V_p^N(f). \quad (14)$$

In these visibilities, let $\phi_p^m(f)$ be the measured interferometric phase angle, $\phi_p^T(f)$ be the uncorrupted interferometric phase angle relating to the true sky, $\phi_p^F(f)$ be the uncorrupted interferometric phase angle from foregrounds, $\delta\phi_p^L(f)$ be the perturbation caused by cosmic structures of interest probed by the spectral line to the uncorrupted foreground interferometric phase angle, and $\delta\phi_p^N(f)$ be the perturbation to the uncorrupted interferometric phase angle caused by thermal noise. If the perturbations to visibility are small, $|V_p^L(f)| \ll |V_p^F(f)|$ and $|V_p^N(f)| \ll |V_p^F(f)|$, then the approximations in Sec. II B can be employed. Identifying V_p^F , V_p^L , and V_p^N to be respectively Z_0 , Z_1 , and Z_2 in Sec. II B, the perturbations to the visibility phase angle from the cosmic structures and noise are, respectively,

$$\delta\phi_p^L(f) \approx \Im \left\{ \frac{V_p^L(f)}{V_p^F(f)} \right\}, \quad (15)$$

$$\text{and, } \delta\phi_p^N(f) \approx \Im \left\{ \frac{V_p^N(f)}{V_p^F(f)} \right\}. \quad (16)$$

Thus fluctuations in the visibility phase angles are roughly linearly proportional to the fluctuations in the visibilities themselves and inversely proportional to the foreground component of the visibilities under a first-order approximation. Then,

$$e^{i\phi_p^m(f)} = e^{i(\phi_p^F(f) + \delta\phi_p^L(f) + \delta\phi_p^N(f))} \quad (17)$$

$$\approx e^{i\phi_p^F(f)} [1 + i(\delta\phi_p^L(f) + \delta\phi_p^N(f))] \quad (18)$$

$$\approx e^{i\phi_p^F(f)} \left[1 + \frac{1}{2} \left(\frac{V_p^L(f)}{V_p^F(f)} - \frac{\overline{V_p^L(f)}}{\overline{V_p^F(f)}} \right) + \frac{1}{2} \left(\frac{V_p^N(f)}{V_p^F(f)} - \frac{\overline{V_p^N(f)}}{\overline{V_p^F(f)}} \right) \right], \quad (19)$$

where we have used $|\delta\phi_p^L(f)| \ll 1$ and $|\delta\phi_p^N(f)| \ll 1$, and $\Im\{Z\} = (Z - \bar{Z})/2i$, in Eqs. (15) and (16). Thus the visibility phase angle fluctuations $\delta\phi_p^L(f)$ and $\delta\phi_p^N(f)$ are indicators of the spectral line strength to the foreground continuum ratio and the noise to foreground continuum ratio respectively.

If a model for $|V_p^F(f)|$ is available, it can be used to extract partial information about $V_p^L(f)$. Let $|\hat{V}_p^F(f)|$ denote some empirical model of the true sky-based foreground visibility amplitude, $|V_p^F(f)|$. Then, we can reconstruct an estimate of the measured (calibrated) visibility as

$$\begin{aligned}
 \hat{V}_p^m(f) &= |\hat{V}_p^F(f)|e^{i\phi_p^m(f)} = |\hat{V}_p^F(f)|e^{i\phi_p^F(f)} \left[1 + \frac{1}{2} \left(\frac{V_p^L(f)}{V_p^F(f)} - \frac{\overline{V_p^L(f)}}{\overline{V_p^F(f)}} \right) + \frac{1}{2} \left(\frac{V_p^N(f)}{V_p^F(f)} - \frac{\overline{V_p^N(f)}}{\overline{V_p^F(f)}} \right) \right] \\
 &= \hat{V}_p^F(f) \left[1 + \frac{1}{2} \left(\frac{V_p^L(f)}{V_p^F(f)} - \frac{\overline{V_p^L(f)}}{\overline{V_p^F(f)}} \right) + \frac{1}{2} \left(\frac{V_p^N(f)}{V_p^F(f)} - \frac{\overline{V_p^N(f)}}{\overline{V_p^F(f)}} \right) \right] \\
 &= \hat{V}_p^F(f) + \hat{V}_p^L(f) + \hat{V}_p^N(f), \tag{20}
 \end{aligned}$$

where

$$\hat{V}_p^L(f) = \frac{\hat{V}_p^F(f)}{2} \left(\frac{V_p^L(f)}{V_p^F(f)} - \frac{\overline{V_p^L(f)}}{\overline{V_p^F(f)}} \right) \tag{21}$$

$$\text{and, } \hat{V}_p^N(f) = \frac{\hat{V}_p^F(f)}{2} \left(\frac{V_p^N(f)}{V_p^F(f)} - \frac{\overline{V_p^N(f)}}{\overline{V_p^F(f)}} \right). \tag{22}$$

This will not yield a perfect recovery of the fluctuating components because the phase angle fluctuations are related by the imaginary portion of the ratio of the fluctuating component to the foreground and hence only yields a partial recovery, statistically $\sim 50\%$ of the spatial information content. Nevertheless, it can be useful in recovering roughly half the power in the fluctuations.

Note that the use of calibrated interferometric phase angle $\phi_p^m(f)$ is required in Eq. (20) and could not have been substituted by the uncalibrated interferometric phase angle because the latter will result in partial to complete loss of recovery of the visibilities if the measured interferometric phase is uncorrected for the phase corruption of the wavefront introduced by the antenna and the ionosphere. On the other hand, the use of corrected phase carries the burden of having performed extremely accurate calibration, typically with fractional inaccuracy required to be $\lesssim 10^{-5}$. However, the bispectrum phase has the interesting property that it is independent of direction-independent antenna-based calibration and errors therein. The following sections extend this treatment of interferometric phase on visibilities to examine the usefulness of the bispectrum phase while using the raw uncalibrated measurements.

IV. SPECTRAL LINE FLUCTUATIONS IN THE BISPECTRUM PHASE

The measured complex bispectrum is written as the product (over index p) of the measured visibilities (may or may not be calibrated):

$$\begin{aligned}
 B_{\nabla}^m(f) &= |B_{\nabla}^m(f)|e^{i\phi_{\nabla}^m(f)} = \prod_{p=1}^3 V_p^m(f) \\
 &= G_{\nabla}(f) \prod_{p=1}^3 [V_p^F(f) + V_p^L(f) + V_p^N(f)], \tag{23}
 \end{aligned}$$

where $G_{\nabla}(f)$ denotes the lumped product of the various direction-independent antenna-based gains associated with the triad, and $|B_{\nabla}^m(f)|$ and $\phi_{\nabla}^m(f)$ denote the amplitude and the phase angle of the measured bispectrum, respectively.

The measured bispectrum phase angle, which is independent of the direction-independent antenna-based gains [37], is

$$\phi_{\nabla}^m(f) = \sum_{p=1}^3 \phi_p^m(f) = \sum_{p=1}^3 [\phi_p^F(f) + \delta\phi_p^L(f) + \delta\phi_p^N(f)] \tag{24}$$

$$= \sum_{p=1}^3 \phi_p^F(f) + \sum_{p=1}^3 \delta\phi_p^L(f) + \sum_{p=1}^3 \delta\phi_p^N(f) \tag{25}$$

$$= \phi_{\nabla}^F(f) + \delta\phi_{\nabla}^L(f) + \delta\phi_{\nabla}^N(f), \tag{26}$$

where $\phi_{\nabla}^F(f) = \sum_{p=1}^3 \phi_p^F(f)$ is the bispectrum phase angle from the foreground structures, $\delta\phi_{\nabla}^L(f) = \sum_{p=1}^3 \delta\phi_p^L(f)$ is the perturbation to the bispectrum phase angle caused by the presence of cosmic structures, and $\delta\phi_{\nabla}^N(f) = \sum_{p=1}^3 \delta\phi_p^N(f)$ is the perturbation to the bispectrum phase angle caused by the presence of thermal noise.

We assume throughout the paper that during an observation the phase center coincides with the pointing center (boresight) of the antenna. For a transit array like HERA, this is typically the zenith. However, the discussion and examples presented here can be generalized to a tracking telescope with arbitrary pointing and phase centers as well.

The following section relates these phase fluctuations to the spatial coherence function corresponding to the cosmic structures and thermal noise. For convenience throughout, we drop the explicit dependence on f unless specified.

A. Bispectrum phase in the limit of small perturbations

The measured bispectrum in Eq. (23) can be expanded as

$$\begin{aligned}
 |B_{\nabla}^m|e^{i\phi_{\nabla}^m} &= \prod_{p=1}^3 V_p^F + V_1^F V_2^F V_3^L + V_1^F V_3^F V_2^L + V_2^F V_3^F V_1^L \\
 &+ V_1^F V_2^F V_3^N + V_1^F V_3^F V_2^N + V_2^F V_3^F V_1^N \\
 &+ \text{higher-order terms in } V^L \text{ and } V^N. \tag{27}
 \end{aligned}$$

Here, we have neglected the gain terms lumped into $G_{\nabla}(f)$ as they multiply across all the terms and are irrelevant for the phase angles and the fluctuations therein.

By keeping terms only up to linear order in V_p^L and V_p^N , we can infer the first-order perturbation to the foreground bispectrum, $B_{\nabla}^F = \prod_{p=1}^3 V_p^F$, arising from cosmic structures and thermal noise, respectively, as

$$B_{\nabla}^L \approx V_1^F V_2^F V_3^L + V_1^F V_3^F V_2^L + V_2^F V_3^F V_1^L, \quad (28)$$

$$\text{and, } B_{\nabla}^N \approx V_1^F V_2^F V_3^N + V_1^F V_3^F V_2^N + V_2^F V_3^F V_1^N. \quad (29)$$

Identifying B_{∇}^F , B_{∇}^L , and B_{∇}^N to be respectively Z_0 , Z_1 , and Z_2 in Sec. II B, the perturbations to the foreground bispectrum phase from the cosmic structures and noise are, respectively,

$$\delta\phi_{\nabla}^L \approx \Im \left\{ \sum_{p=1}^3 \frac{V_p^L}{V_p^F} \right\}, \quad (30)$$

$$\text{and, } \delta\phi_{\nabla}^N \approx \Im \left\{ \sum_{p=1}^3 \frac{V_p^N}{V_p^F} \right\}. \quad (31)$$

This suggests that fluctuations in the bispectrum phase are approximately linearly proportional to the fluctuations in visibilities under a first-order approximation.

Alternatively, from Eqs. (25) and (26),

$$\delta\phi_{\nabla}^L(f) = \sum_{p=1}^3 \delta\phi_p^L(f) \approx \sum_{p=1}^3 \Im \left\{ \frac{V_p^L}{V_p^F} \right\}, \quad (32)$$

and

$$\delta\phi_{\nabla}^N(f) = \sum_{p=1}^3 \delta\phi_p^N(f) \approx \sum_{p=1}^3 \Im \left\{ \frac{V_p^N}{V_p^F} \right\}, \quad (33)$$

which are identical to Eqs. (30) and (31). Effectively, these bispectrum phase fluctuations are a reasonable instrument-independent, and a robust true-sky measure of the dissimilarity of the cosmic structures relative to the foregrounds in the transverse plane of the sky, whose magnitude depends on the ratio between the two.

B. Relation between bispectrum phase fluctuations and visibilities

In the limit of small fluctuations, $|\delta\phi_p^L|, |\delta\phi_p^N| \ll 1$, the bispectrum phase can be expressed using Taylor-series expansion to linear-order terms in $\delta\phi_{\nabla}^L$ and $\delta\phi_{\nabla}^N$ as

$$e^{i\phi_{\nabla}^m} = e^{i(\phi_{\nabla}^F + \delta\phi_{\nabla}^L + \delta\phi_{\nabla}^N)} \quad (34)$$

$$\approx e^{i\phi_{\nabla}^F} [1 + i(\delta\phi_{\nabla}^L + \delta\phi_{\nabla}^N)]. \quad (35)$$

Thus, the perturbed angles, $\delta\phi_{\nabla}^L$ and $\delta\phi_{\nabla}^N$, and also $e^{i\phi_{\nabla}^m}$ contain terms up to linear order in V_p^L/V_p^F and V_p^N/V_p^F . From Eqs. (30) and (31), Eq. (35) can be expressed as

$$e^{i\phi_{\nabla}^m} \approx e^{i\phi_{\nabla}^F} \left[1 + \frac{1}{2} \sum_{p=1}^3 \left(\frac{V_p^L}{V_p^F} - \frac{\overline{V_p^L}}{V_p^F} \right) + \frac{1}{2} \sum_{p=1}^3 \left(\frac{V_p^N}{V_p^F} - \frac{\overline{V_p^N}}{V_p^F} \right) \right]. \quad (36)$$

Hence, the bispectrum phase angle fluctuations, $\delta\phi_{\nabla}^L$ and $\delta\phi_{\nabla}^N$, measure the dissimilarity relative to the foregrounds with magnitudes given by the spectral line to foreground continuum ratio, and noise-to-foreground continuum ratio, respectively. This is very similar to the nature of the visibility phase angle fluctuations.

In principle, we could simply use ϕ_{∇}^m as the mathematical quantity of interest instead of the complex Eulerian representation $e^{i\phi_{\nabla}^m}$, that will be *delay-transformed* as we have in the rest of the paper. However, in practice ϕ_{∇}^m will be noisy and may contain discontinuities at $\pm\pi$ due to phase angle wrapping. Including such discontinuities in a Fourier transform will lead to the classical *ringing* and associated artefacts. The robust removal of such discontinuities especially in the presence of noise and other fluctuations is not straightforward and is a subject of ongoing research (see e.g., [54–58]). Though we do not explore this variant approach in detail in this paper, we present an outline (see Appendix) of the potential advantages using ϕ_{∇}^m could hold in comparison to using $e^{i\phi_{\nabla}^m}$.

Figure 1 corresponds to a hypothetical example that will be discussed in detail later as example(i)(b) in Sec. VA 1 with a slight modification. Briefly, the foreground model is a point source of true flux density 100 Jy, which is $\approx 5^\circ$ off-boresight and has a spectral index $\alpha = 0$ in contrast with the example in Sec. VA 1. The spectral signal comes from a point source at boresight with a cosine-shaped spectrum of characteristic frequency scale, $f_L = 1$ MHz, and amplitude 10 mJy. *Such a model for the foreground, and especially the EoR HI signal, is purely hypothetical and unrealistic.* The antenna spacings correspond to a 50.6 m equilateral triad (antenna layout discussed in Sec. V and illustrated in Fig. 5). The antennas are assumed to be uniformly illuminated dishes of diameter 14 m and have a corresponding *Airy* angular power pattern whose transverse angular structure does not change with frequency. The top subpanel shows the visibility amplitude of the off-boresight foreground point source on “the three differently oriented 50.6 m antenna spacings (red, blue, and black). The off-axis location lowers the apparent flux density as shown due to the power pattern assumed. The fluctuation in amplitude due to the spectral line point source, $\delta|V_p(f)| = |V_p^F(f) + V_p^L(f)| - |V_p^F(f)|$ for the same antenna spacings is shown in the bottom subpanel.

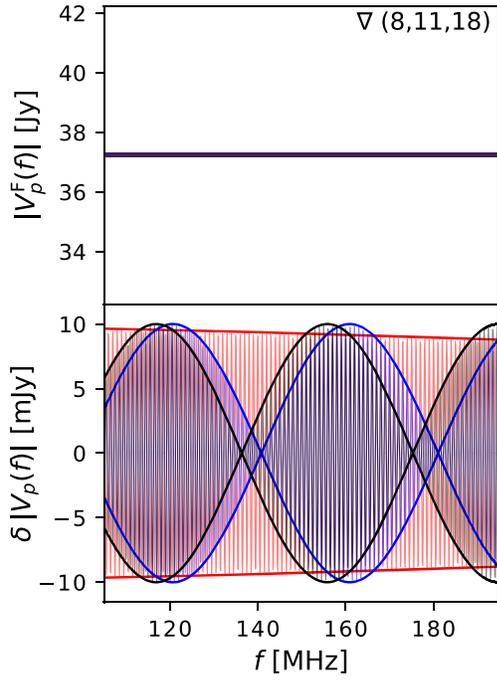


FIG. 1. The amplitude of the visibilities from the foregrounds, $V_p^F(f)$ (top) and the fluctuations therein due to the cosmic H I signal (bottom) given by $\delta|V_p(f)| = |V_p^F(f) + V_p^L(f)| - |V_p^F(f)|$ on three antenna spacings (red, blue, and black) comprising the 50.6 m equilateral triad, $\nabla = (8, 11, 18)$ (refer to the antenna layout in Fig. 5). The foreground model is a point source of true flux density 100 Jy with spectral index, $\alpha = 0$, located off-boresight by $\approx 5^\circ$. The power pattern at this location is the cause of the reduced strength of the apparent foreground visibilities in the top panel. The cosmic H I signal is a purely hypothetical but an unrealistic model consisting of a point source located at boresight and has a cosine-shaped frequency spectrum of characteristic frequency scale, $\delta f_L = 1/\tau_L = 1$ MHz. The envelope of the fluctuations in amplitude obtained using a Hilbert transform is shown to gauge the overall magnitudes of these fluctuations relative to the foreground amplitudes.

The maximum of the envelope of the fluctuations in amplitude is ~ 10 mJy as expected.

Figure 2 shows the actual and predicted (from first-order approximation) values of the fluctuations in the phase angles of the visibilities and the bispectrum for the above example. The visibility phase angle fluctuations predicted after retaining only the first-order perturbations using Eq. (16) for the three antenna spacings (red, blue, and black) are shown in the top subpanel. The middle subpanel shows the actual fluctuations in the bispectrum phase angle (black) and that predicted from Eq. (30) (gray). The envelope of the fluctuations derived using a Hilbert transform shows the overall amplitude of the phase angle fluctuations. The residuals between the predicted and actual bispectrum phase angle fluctuations is shown in the bottom subpanel. The higher-than-linear-order terms are typically $\sim 10^{-4}$ smaller fractionally

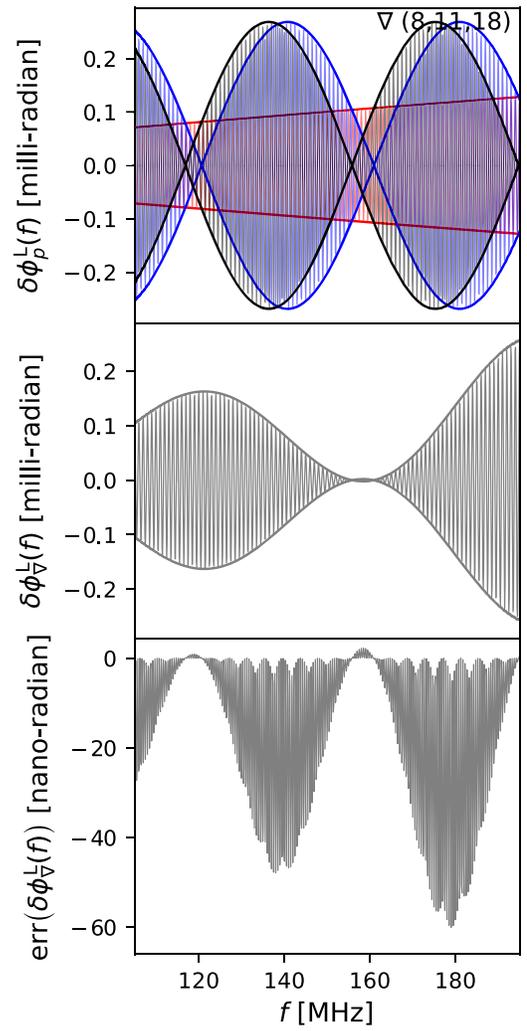


FIG. 2. Predicted and actual values of phase angle fluctuations of the visibility and bispectrum for the sky model considered in Fig. 1. Top: predicted fluctuations in the three visibility phase angles of a 50.6 m equilateral triad in red, blue, and black using Eq. (16). Middle: fluctuations in the actual bispectrum phase angle (black) and the first-order approximation (gray) predicted using Eq. (30). The envelope of these phase angle fluctuations are at a level ~ 0.1 – 0.2 milli-radians, which is similar to and identifiable with the ratio $\sim \delta|V_p(f)|/|V_p^F(f)|$ in Fig. 1. Bottom: bispectrum phase angle residuals from the difference of the first-order prediction and the actual values. These higher-than-linear-order residuals not captured by the first-order approximation are \sim few tens of nano-radians indicating the approximation to linear-order terms has a fractional inaccuracy of only $\lesssim 10^{-4}$ relative to true values.

compared to both the predicted linear-order terms as well as the actual values indicating that the fractional inaccuracy of the linear-order approximation is only $\lesssim 10^{-4}$ relative to true values.

Figures 3 and 4 are the same as Figs. 1 and 2, respectively, but use a spectral index $\alpha = -0.8$ for the foreground point source. This same example is presented in more detail

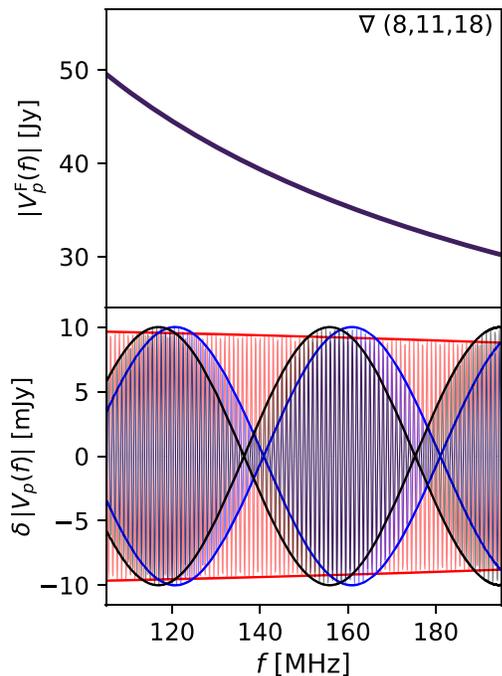


FIG. 3. Same as Fig. 1 but the foreground model has a spectral index, $\alpha = -0.8$.

in Sec. VA 1. The phase angle fluctuations in visibility and bispectrum (Fig. 4) are seen to be correspondingly increased at higher frequencies, and vice versa, relative to that in the previous example illustrated in Fig. 2. This is also in agreement with the predictions in Eqs. (16) and (30) and arises due to the relative decrease of foreground amplitude at higher frequencies and vice versa. The deviation between the predicted and actual values also follows a similar trend where it is higher at higher frequencies and vice versa relative to the previous example. In other words, the first-order approximation is still valid at all frequencies but the prediction is slightly better at lower frequencies than at higher frequencies.

If a model of the foreground visibilities, \hat{V}_p^F , is available, the spectral line strength from cosmic structures can be approximately estimated from this ratio. We construct the quantity

$$V_{\nabla}(f) = V_{\text{eff}}^F e^{i\phi_{\nabla}^m(f)}, \quad (37)$$

where V_{eff}^F is designed to be an empirical estimate of $\hat{V}_p^F(f)$. Note that the explicit dependence on f has been re-introduced to emphasize that V_{eff}^F is not a function of frequency but only provides an overall amplitude scaling that corresponds to the frequency band of interest. The multiplication by V_{eff}^F that is empirically representative of the effective foreground visibility amplitude from the triads converts the complex bispectrum phase term to an effective flux density. $V_{\nabla}(f)$ has units of Jy but since it is purely

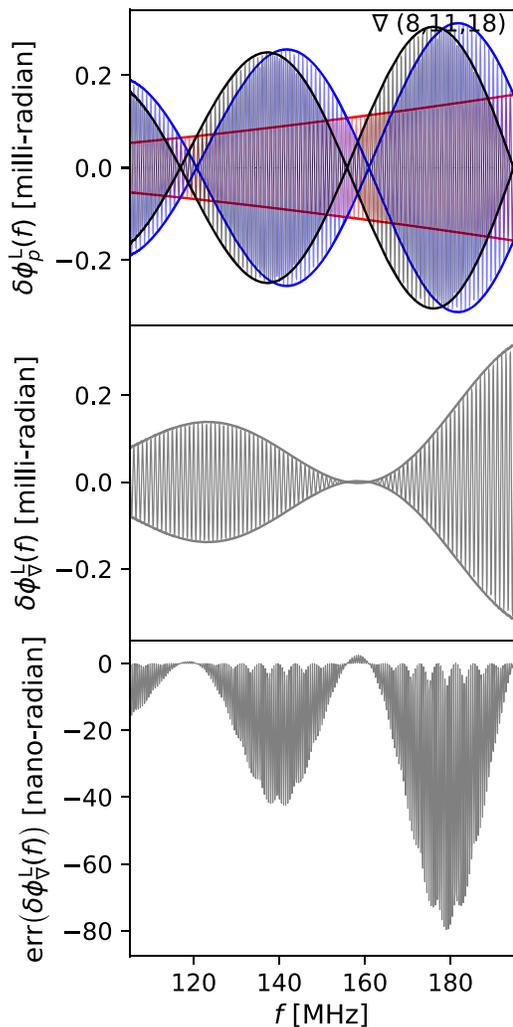


FIG. 4. Same as Fig. 2 but the foreground model has a spectral index, $\alpha = -0.8$ corresponding to that shown in Fig. 3. The visibility and bispectrum phase angle fluctuations are higher at higher frequencies and vice versa when compared to the case with $\alpha = 0$ shown in Fig. 2. This is because the fluctuations in visibility amplitudes remain the same whereas the foreground amplitudes are lower at higher frequencies and vice versa due to the spectral index, $\alpha < 0$. Thus, the deviation between the actual and first-order approximation of the bispectrum phase also becomes higher at higher frequencies and vice versa relative to that when $\alpha = 0$ in Fig. 2. The approximation is still accurate overall, but is slightly better at lower frequencies than at higher frequencies essentially following the ratio $\sim \delta |V_p(f)| / |V_p^F(f)|$.

mathematical rather than a physically valid flux density, we refer to its units by “pseudo Jy” to distinguish it from physical flux density. V_{eff}^F , a scalar, will not be able to fully capture the exact spatial coherence information in $V_p^F(f)$ because the latter consists of independent information on possibly three different spatial modes in addition to containing spectral information. This will introduce an error in the final result but one that appears as a simple scaling error and does not introduce any spectral errors or systematics.

Using Eq. (36),

$$\begin{aligned} V_{\nabla}(f) &= V_{\text{eff}}^F e^{i\phi_{\nabla}^F(f)} \left[1 + \frac{1}{2} \sum_{p=1}^3 \left(\frac{V_p^L(f)}{V_p^F(f)} - \frac{\overline{V_p^L(f)}}{\overline{V_p^F(f)}} \right) \right. \\ &\quad \left. + \frac{1}{2} \sum_{p=1}^3 \left(\frac{V_p^N(f)}{V_p^F(f)} - \frac{\overline{V_p^N(f)}}{\overline{V_p^F(f)}} \right) \right] \\ &= V_{\nabla}^F(f) + V_{\nabla}^L(f) + V_{\nabla}^N(f), \end{aligned} \quad (38)$$

where

$$V_{\nabla}^F(f) = V_{\text{eff}}^F e^{i\phi_{\nabla}^F(f)}, \quad (39)$$

$$V_{\nabla}^L(f) = i\delta\phi_{\nabla}^L(f) V_{\text{eff}}^F e^{i\phi_{\nabla}^F(f)}, \quad (40)$$

and

$$V_{\nabla}^N(f) = i\delta\phi_{\nabla}^N(f) V_{\text{eff}}^F e^{i\phi_{\nabla}^F(f)}. \quad (41)$$

Equation (38) takes a familiar form where the measured visibilities comprise of an additive combination of foregrounds, the cosmological spectral line signal, and measurement noise. For convenience, we define $\gamma_p^F(f) = V_{\text{eff}}^F / V_p^F(f)$. Then,

$$V_{\nabla}^L(f) = \frac{1}{2} e^{i\phi_{\nabla}^F(f)} \sum_{p=1}^3 [\gamma_p^F(f) V_p^L(f) - \overline{\gamma_p^F(f) V_p^L(f)}], \quad (42)$$

$$V_{\nabla}^N(f) = \frac{1}{2} e^{i\phi_{\nabla}^F(f)} \sum_{p=1}^3 [\gamma_p^F(f) V_p^N(f) - \overline{\gamma_p^F(f) V_p^N(f)}]. \quad (43)$$

However, differing from the standard approach, Eqs. (42) and (43) show that the effective visibilities denoting contributions of the spectral line signal and noise to the bispectrum phase are now weighted by the foreground spectra. This is expected because the bispectrum phase is a measure of the ratio of the fluctuating signal to the foregrounds as noted earlier. It must also be noted that $e^{i\phi_{\nabla}^F(f)}$ and $\gamma_p^F(f)$ (with $|\gamma_p^F(f)| \sim 1$) are expected to exhibit only slow spectral variations. Thus, the excess spectral variance from rapid fluctuations such as from the cosmic line signal will still be distinguishable as will be demonstrated later through examples.

The process of determining V_{∇} , as presented here, is empirical and has the following reasoning. We note that the variance due to cosmic line signal and noise fluctuations is approximately the sum of the variances in the fluctuations in the individual interferometric phases. Assuming that the cosmic line signal strength and the noise rms measured on each of the baselines forming the triad do not differ significantly between the baselines, the fluctuations are inversely dependent on the foreground visibility measured

on the respective baselines. The weakest visibility amplitude among the baselines in the triad will induce the maximum fluctuations which will dominate the overall budget of fluctuations in the measured bispectrum phase. Therefore, we obtain V_{eff}^F by averaging in inverse quadrature as

$$(V_{\text{eff}}^F)^{-2} = \sum_{p=1}^3 |\hat{V}_p^F|^{-2} \quad (44)$$

where

$$\hat{V}_p^F = \frac{\int W(f) \hat{V}_p^F(f) df}{\int W(f) df}. \quad (45)$$

$\hat{V}_p^F(f)$ denotes a reliable visibility model (obtained either through calibration or modeling), which is then averaged over the frequency sub-band of interest, with the same optional spectral window weighting, $W(f)$, that may get used in the further processing as described below. It must be noted that the choice of V_{eff}^F above is not entirely rigorous and could be substituted with any other reasonable estimate.

It must be emphasized that the model or calibrated $\hat{V}_p^F(f)$ does not need to be fractionally as accurate as $\sim 10^{-5}$, for example, as in other standard approaches for detecting faint spectral lines. It is simply used to obtain an average scalar to scale the bispectrum phase to be in the same units as flux density. This procedure does not introduce any potential spectral artefact except for an overall uniform but minor error in the scaling because the choice of the scalar may not have been rigorous.

C. Delay spectrum of the bispectrum phase

Since the context of this paper is the detection of distinctive spectral features, we employ the delay spectrum technique [59,60], which is essentially a Fourier domain method for spectral discrimination. We define the delay transform, which is simply a Fourier transform, of a complex-valued spectrum, $Z(f)$, as

$$\tilde{Z}(\tau) = \int Z(f) e^{i2\pi f\tau} df. \quad (46)$$

Consider the delay transform of $V_{\nabla}^L(f)$, denoted by $\tilde{V}_{\nabla}^L(\tau)$. From Eq. (42), it can be seen that $\tilde{V}_{\nabla}^L(\tau)$ is formed from a convolution of $\tilde{\mathcal{E}}_{\nabla}^F(\tau)$, $\tilde{\gamma}_p^F(\tau)$, and $\tilde{V}_p^L(\tau)$, which are the delay-domain duals of $e^{i\phi_{\nabla}^F(f)}$, $\gamma_p^F(f)$, and $V_p^L(f)$, respectively. $\tilde{V}_p^L(\tau)$ contains the structural information about the cosmic spectral line signal. However, because of the weighting from foregrounds, it gets convolved by the predominantly smooth spectral structure response from the foregrounds within the sub-band in which the delay transform is computed,

TABLE I. Description of symbols.

Symbol	Description	Section	
Z, \bar{Z}	A complex number and its conjugate	II	
p	Index of antenna spacing (baseline) vector	III	
\mathbf{b}_p	Antenna spacing (baseline) vector indexed by p	III	
f	Frequency	III	
τ	Delay, the Fourier dual of frequency, f	IV C	
κ_{\parallel}	Pseudo line-of-sight wavenumber modes corresponding to τ	IV D	
$V_p^T(f)$	True sky visibility spectrum on \mathbf{b}_p	III	
$V_p^m(f)$	Measured visibility spectrum on \mathbf{b}_p	III	
$V_p^F(f)$	$\tilde{V}_p^F(\tau)$	True foreground visibility spectrum and its delay transform on \mathbf{b}_p	III, IV C
$V_p^L(f)$	$\tilde{V}_p^L(\tau)$	True spectral line visibility spectrum and its delay transform on \mathbf{b}_p	III, IV C
$V_p^N(f)$	$\tilde{V}_p^N(\tau)$	Noise visibility spectrum and its delay transform on \mathbf{b}_p	III, IV C
$\phi_p^m(f)$	Spectrum of interferometric phase angle on measured visibility, $V_p^m(f)$	III	
$\phi_p^T(f)$	Spectrum of interferometric phase angle on true visibility, $V_p^T(f)$	III	
$\phi_p^F(f)$	Spectrum of interferometric phase angle on foreground visibility, $V_p^F(f)$	III	
$\delta\phi_p^L(f)$	Spectrum of perturbed interferometric phase angle due to $V_p^L(f)$	III	
$\delta\phi_p^N(f)$	Spectrum of perturbed interferometric phase angle due to $V_p^N(f)$	III	
$\hat{V}_p^F(f)$	Model of foreground visibility spectrum, $V_p^F(f)$	III	
$\hat{V}_p^m(f)$	Estimate of measured visibility spectrum, $V_p^m(f)$, using $\phi_p^m(f)$ and $\hat{V}_p^F(f)$	III	
$\hat{V}_p^L(f)$	Estimate of spectrum of spectral line visibility, $V_p^L(f)$	III	
$\hat{V}_p^N(f)$	Estimate of noise visibility spectrum, $V_p^N(f)$	III	
$B_{\nabla}^m(f)$	Spectrum of measured visibility bispectrum	IV	
$B_{\nabla}^F(f)$	Spectrum of foreground visibility bispectrum	IVA	
$B_{\nabla}^L(f)$	Spectrum of perturbation to $B_{\nabla}^F(f)$ due to the cosmic spectral line signal	IVA	
$B_{\nabla}^N(f)$	Spectrum of perturbation to $B_{\nabla}^F(f)$ due to noise	IVA	
$\phi_{\nabla}^m(f)$	Spectrum of phase angle on measured bispectrum, $B_{\nabla}^m(f)$	IV	
$\phi_{\nabla}^F(f)$	Spectrum of phase angle on foreground bispectrum, $B_{\nabla}^F(f)$	IV	
$\delta\phi_{\nabla}^L(f)$	Spectrum of perturbation to $\phi_{\nabla}^F(f)$ due to the cosmic spectral line signal	IV	
$\delta\phi_{\nabla}^N(f)$	Spectrum of perturbation to $\phi_{\nabla}^F(f)$ due to noise	IV	
V_{eff}^F	A scalar estimate of $\hat{V}_p^F(f)$ obtained empirically over the sub-band	IV B	
$V_{\nabla}(f)$	$\tilde{V}_{\nabla}(\tau)$	Representation of $e^{i\phi_{\nabla}^m(f)}$ in flux density units	IV B, IV C
$V_{\nabla}^F(f)$	$\tilde{V}_{\nabla}^F(\tau)$	Representation of $e^{i\phi_{\nabla}^F(f)}$ in flux density units and its delay transform	IV B, IV C
$V_{\nabla}^L(f)$	$\tilde{V}_{\nabla}^L(\tau)$	Representation of $e^{i\delta\phi_{\nabla}^L(f)}$ in flux density units and its delay transform	IV B, IV C
$V_{\nabla}^N(f)$	$\tilde{V}_{\nabla}^N(\tau)$	Representation of $e^{i\delta\phi_{\nabla}^N(f)}$ in flux density units and its delay transform	IV B, IV C
$W(f)$	$\tilde{W}(\tau)$	Spectral window function and its delay transform	IV C
$\gamma_p^F(f)$	$\tilde{\gamma}_p^F(\tau)$	Normalized true foreground response in visibility and its delay transform on \mathbf{b}_p	IV B, IV C
$e^{i\phi_{\nabla}^E(f)}$	$\tilde{\mathcal{E}}_{\nabla}^E(\tau)$	Complex Eulerian representation of $\phi_{\nabla}^E(f)$ and its delay transform	IV, IV C
$V_{\nabla}(f)W(f)$	$\tilde{\Psi}_{\nabla}(\tau)$	Windowed $V_{\nabla}(f)$ and its delay transform	IV C
$P_{\nabla}(\kappa_{\parallel})$		Delay-domain power spectrum of windowed bispectrum phase, $V_{\nabla}(f)W(f)$	IV D

$$\begin{aligned} \tilde{V}_{\nabla}^L(\tau) &= \frac{1}{2} \tilde{\mathcal{E}}_{\nabla}^F(\tau) * \sum_{p=1}^3 [\tilde{\gamma}_p^F(\tau) * \tilde{V}_p^L(\tau) \\ &\quad - \tilde{\gamma}_p^F(-\tau) * \tilde{V}_p^L(-\tau)], \end{aligned} \quad (47)$$

$$\begin{aligned} \tilde{\Psi}_{\nabla}(\tau) &= \tilde{V}_{\nabla}(\tau) * \tilde{W}(\tau) = \int V_{\nabla}(f) W(f) e^{i2\pi f\tau} df, \\ &= \tilde{W}(\tau) * [\tilde{V}_{\nabla}^F(\tau) + \tilde{V}_{\nabla}^L(\tau) + \tilde{V}_{\nabla}^N(\tau)], \end{aligned} \quad (48)$$

which can be further expanded as

$$\begin{aligned} \tilde{\Psi}_{\nabla}(\tau) &= \tilde{W}(\tau) * \tilde{\mathcal{E}}_{\nabla}^F(\tau) \\ &\quad * \left\{ V_{\text{eff}}^F \delta(\tau) + \frac{1}{2} \sum_{p=1}^3 \left[\tilde{\gamma}_p^F(\tau) * \left(\tilde{V}_p^L(\tau) + \tilde{V}_p^N(\tau) \right) \right. \right. \\ &\quad \left. \left. - \tilde{\gamma}_p^F(-\tau) * \left(\tilde{V}_p^L(-\tau) + \tilde{V}_p^N(-\tau) \right) \right] \right\} \end{aligned} \quad (49)$$

where $W(f)$ is an optional spectral window weighting usually chosen to control the quality of the delay spectrum [27,61] and has an effective bandwidth, ΔB . $\tilde{W}(\tau)$ is its delay-domain dual. $\delta(\tau)$ is a *delta* function at $\tau = 0$ in the *delay* domain. $\tilde{\Psi}_{\nabla}(\tau)$ has units of “pseudo Jy Hz” for reasons explained earlier.

D. Delay power spectrum

We obtain the analogous power spectrum of the bispectrum phase in the delay domain as [59,62]

$$P_{\nabla}(\kappa_{\parallel}) \equiv |\tilde{\Psi}_{\nabla}(\tau)|^2 \left(\frac{A_e}{\lambda^2 \Delta B} \right) \left(\frac{D^2 \Delta D}{\Delta B} \right) \left(\frac{\lambda^2}{2k_B} \right)^2, \quad (50)$$

with

$$\kappa_{\parallel} \equiv \frac{2\pi\tau f_r H_0 E(z)}{c(1+z)^2}, \quad (51)$$

where A_e is the effective area of the antenna, ΔB is the effective bandwidth, λ is the wavelength of the band center, k_B is the Boltzmann constant, c is the speed of light in vacuum, f_r is the rest-frame frequency of the cosmic spectral line signal, z is the redshift, $D \equiv D(z)$ is the transverse comoving distance, and ΔD is the comoving depth along the line of sight corresponding to ΔB at redshift z . H_0 , h , and $E(z) \equiv [\Omega_M(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda]^{1/2}$ are standard terms in cosmology. In this paper, we use cosmological parameters from [63] with $H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}$. $P_{\nabla}(\kappa_{\parallel})$ is in units of “pseudo $\text{mK}^2 (\text{Mpc}/h)^3$.” Note that we use κ_{\parallel} to explicitly distinguish it from the line-of-sight wavenumber k_{\parallel} as the two are very similarly defined mathematically but are not exactly related because the origin of fluctuations in the bispectrum phase are not

identical to those in standard visibilities. κ_{\parallel} has units of “pseudo $h \text{ Mpc}^{-1}$.”

In a scenario that includes noise, the noise bias can be avoided by estimating the delay cross-power spectrum by replacing $|\tilde{\Psi}_{\nabla}(\tau)|^2$ in Eq. (50) with $\Re\{\tilde{\Psi}_{\nabla}(\tau)\tilde{\Psi}_{\nabla}^{\prime}(\tau)\}$, where, $\tilde{\Psi}_{\nabla}^{\prime}(\tau)$ is another independent realization of $\tilde{\Psi}_{\nabla}(\tau)$. The cross term, $\Re\{\tilde{\Psi}_{\nabla}(\tau)\tilde{\Psi}_{\nabla}^{\prime}(\tau)\}$, serves the purpose of removing the noise bias or systematics in case of nonredundancy, etc. assuming the cosmic signal component remains fully correlated in both $\tilde{\Psi}_{\nabla}(\tau)$ and $\tilde{\Psi}_{\nabla}^{\prime}(\tau)$. In a noiseless and an ideal scenario, the cross-power spectrum reduces to the auto-power spectrum given in Eq. (50).

Table I lists the symbols most relevant in this formalism along with their descriptions. Column 1 contains the symbols. Subcolumns in column 1 separated by a vertical delimiter denote Fourier domain duals of each other with the first subcolumn representing the spectral (frequency) domain, and the second its Fourier dual in the delay domain. The second column provides a brief description of the symbol(s). The third column points to the section in the text in which the symbol is primarily introduced.

V. DEMONSTRATION WITH EXAMPLES OF SKY MODELS

For purposes of demonstration, we assume ideal cases without noise or other systematics. Only foregrounds and fluctuations from the cosmic spectral line signal will be considered. Hence, the noise terms can be ignored, and the use of auto-power spectrum will suffice. We consider the redshifted 21 cm signal from HI from the EoR as our target cosmic signal, and foregrounds in the corresponding 100–200 MHz frequency band. Note that some of the examples where the foregrounds and the HI signal from the EoR, especially the latter, are modeled as point sources are purely hypothetical and unrealistic. The purpose of such examples is to progressively build an intuition for the behavior of the bispectrum phase from simple to intermediate scenarios, eventually culminating in a more realistic example towards the end.

Four examples of sky models are considered, which are hereafter denoted as (i), (ii), (iii), and (iv). In all these examples, the boresight points to RA(J2000) = $05^{\text{h}}32^{\text{m}}39^{\text{s}}32$, Dec(J2000) = $-30^{\circ}44'05''1$. In example (i), the foreground model is a point source of strength $V^F(f) = A(\hat{s}_F) V^F(f/f_F)^{\alpha} e^{-i2\pi(f/c)\mathbf{b}_p \cdot \hat{s}_F}$ with spectral index α , and a pivot frequency for reference f_F , from a location \hat{s}_F where the angular power pattern of the antenna is given by $A(\hat{s})$. The cosmic signal is also a point source of strength $V^L(f) = A(\hat{s}_L) V^L \cos(2\pi f\tau_L + \theta_0) e^{-i2\pi(f/c)\mathbf{b}_p \cdot \hat{s}_L}$ which is a cosine-shaped spectral ripple of characteristic frequency scale $\delta f_L = 1/\tau_L$, and an arbitrary angle θ_0 , at location \hat{s}_L . Both the foreground and the spectral line signal appear as point sources in the transverse plane with $|V^L(f)| \ll |V^F(f)|$. Specifically, we adopt the values $V^F = 100 \text{ Jy}$,

TABLE II. Description of examples of sky models.

Example	Sky model	
	Foregrounds	Cosmic H I signal
(i)(a)	Point source ^a at boresight ^b of strength 100 Jy at 150 MHz and $\alpha = -0.8$	Point source ^a at boresight with cosine-shaped spectrum of amplitude 10 mJy and characteristic frequency scale of 1 MHz
(i)(b)	Same as example (i)(a) ^a but $\approx 5^\circ$ off-boresight	Same as example (i)(a) ^a
(ii)	Objects from the GLEAM catalog within 15° of boresight	Same as example (i)(a) ^a
(iii)	Same as example (i)(a) ^a	FAINT GALAXIES model from 21cmFAST simulations centered on boresight
(iv)	Same as example (ii)	Same as example (iii)

^aA point source model for the foregrounds and the EoR H I signal, especially the latter, is unrealistic and purely hypothetical.

^bBoresight points to RA(J2000) = $05^{\text{h}}32^{\text{m}}39^{\text{s}}32$, Dec(J2000) = $-30^\circ44'05''1$.

$f_{\text{F}} = 150$ MHz, $\alpha = -0.8$, $V^{\text{L}} = 10$ mJy, and $\delta f_{\text{L}} = 1/\tau_{\text{L}} = 1$ MHz. This example is further subdivided into two cases: (a) the location of the foreground and the cosmic signal point sources are colocated, $\hat{s}_{\text{F}} = \hat{s}_{\text{L}}$, and exactly at boresight, and (b) the foreground and the cosmic signal locations are not colocated, $\hat{s}_{\text{F}} \neq \hat{s}_{\text{L}}$, \hat{s}_{L} points to the boresight, and \hat{s}_{F} points to $\approx 5^\circ$ off-boresight.

In example (ii), the foreground consists of objects from the GLEAM catalog [64] within a circle of 30° diameter around boresight. The cosmic H I signal is modeled as a point source at boresight with a cosine-shaped spectrum, the same as in the previous example. In example (iii), the foreground consists of a point source of flux density 100 Jy at 150 MHz, located at boresight and spectral index $\alpha = -0.8$ for the foreground model [as in example (i)(a)]. As our fiducial EoR H I model, we use the FAINT GALAXIES model [65,66] publicly available¹ from 21cmFAST simulations [67] centered on boresight. In example (iv), the foreground model consists of objects from the GLEAM catalog of radio sources [as in example (ii)] and the fiducial 21cmFAST EoR H I model from example (iii).

The examples are briefly summarized in Table II. We reiterate that the hypothetical models of the foregrounds and the cosmic spectral line signal, especially the latter, being point sources in examples (i), (ii), and (iii), are unrealistic. However, having a point source for the sky model results in the vanishing of the bispectrum phase angle which serves as a useful point of reference. And the cosine-shaped spectrum having a single characteristic frequency scale (correspondingly an impulse in the Fourier domain) serves the very useful purpose of understanding the response (also referred to as the impulse response or the transfer function) of the bispectrum phase

statistic towards a single impulse input. Example (iv) presents a realistic realization of both foregrounds and the cosmic EoR H I signal.

Figure 5 shows the antenna layout used in simulating visibilities, which is shown in local eastward and northward coordinates along the x and y axes respectively. The array is assumed to be coplanar and located at a latitude of $-30^\circ.7224$ and a longitude of $+21^\circ.4278$. The circles denote dish-

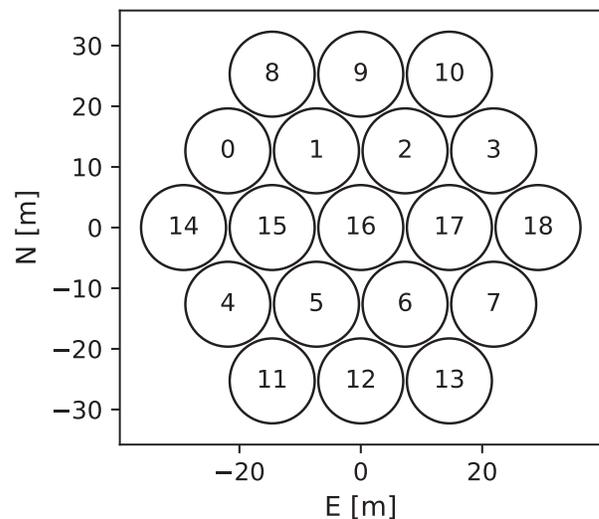


FIG. 5. Redundantly spaced antenna layout used in the simulations. They are assumed to be located at a latitude of $-30^\circ.7224$ and a longitude of $+21^\circ.4278$. The circles denote dish-shaped antennas, each of diameter 14 m. The numerals denote the antenna numbering. The x and y axes denote the local eastward and northward coordinates respectively. The array is assumed to be coplanar. The shortest spacing between antennas is 14.6 m. The specific triad classes chosen in this study are the 14.6 m [for example, $\nabla = (0, 1, 8)$] and 50.6 m equilateral triads [for example, $\nabla = (8, 11, 18)$].

¹<http://homepage.sns.it/mesinger/EOS.html>.

shaped antennas each of diameter 14 m. The numerals denote the antenna numbering. The shortest antenna spacing is 14.6 m. The two classes of antenna triads frequently used in this paper are the 14.6 m and 50.6 m equilateral triads. Specific triads in each class are $\nabla = (0, 1, 8)$ and $\nabla = (8, 11, 18)$ respectively.

Though the antenna placements are redundant, neither the mathematical formalism nor the results derived in this paper assume or require such a redundancy. The power patterns are also assumed to be identical between all the antennas with a uniform circular illumination of the aperture corresponding to an *Airy* angular power pattern. In order to clearly isolate the findings reported in this paper from the spectral characteristics of the power pattern, we have further assumed that the power patterns are achromatic. The angular structure of the power pattern is identical at all frequencies in the 100–200 MHz band and is derived from the analytical expression evaluated at 150 MHz.

We choose our spectral window function, $W(f)$ to be the “modified” Blackman-Harris window [61] with an effective bandwidth of $\Delta B = 42$ MHz centered at 150 MHz. Although this choice of ΔB may include significant evolution of the properties of the cosmic signal within the sub-band, our aim in this paper is to demonstrate the spectral properties of bispectrum phase with highest resolution in Fourier space, $\delta\tau_w = 1/\Delta B \approx 0.024 \mu\text{s}$. This choice of ΔB is also the maximum that can fit inside the 100–200 MHz band without abrupt truncation of the window function at the edges. The effective area of the antennas is chosen to be $A_e = 100 \text{ m}^2$ and is assumed to remain constant across the sub-band. Visibilities of the sky models in these examples were simulated using the Precision Radio Interferometry Simulator [PRISim²; [68]].

$$\begin{aligned} \tilde{V}_{\nabla}^L(\tau) &= \frac{A(\hat{s}_L)}{A(\hat{s}_F)} V^L \tilde{W}(\tau) * \tilde{\gamma}^F(\tau) e^{i\theta_0} * \frac{1}{2} [\delta(\tau - \tau_L) + \delta(\tau + \tau_L)] * \frac{1}{2i} \sum_{p=1}^3 [\delta(\tau - \tau_p) - \delta(\tau + \tau_p)] \\ &= \frac{A(\hat{s}_L)}{A(\hat{s}_F)} \frac{V^L \tilde{W}(\tau) * \tilde{\gamma}^F(\tau) e^{i\theta_0}}{4i} * \sum_{p=1}^3 [\delta(\tau - (\tau_p + \tau_L)) + \delta(\tau - (\tau_p - \tau_L)) - \delta(\tau + (\tau_p + \tau_L)) - \delta(\tau + (\tau_p - \tau_L))]. \end{aligned} \quad (53)$$

Without loss of generality, we can choose $\theta_0 = 0$.

Example (i)(a): Spectral line signal transversally collocated with the foreground.— When the cosmic signal and the foreground object both modeled as point sources are collocated, $\hat{s}_F = \hat{s}_L = \hat{s}_0$, then $\delta\phi_p^L(f) = 0$, and thus $\delta\phi_{\nabla}^L(f) = 0$. This can be qualitatively reasoned as follows. Regardless of the spectral structure, in any given frequency channel, the sky appears as a point source in the transverse sky plane. Therefore, $\phi_{\nabla}^E(f) = \delta\phi_{\nabla}^L(f) = 0$. This can be understood mathematically as well. The visibilities from

²PRISim is publicly available for use under the MIT license at <https://github.com/nithyanandan/PRISim>.

A. Frequency-domain view of the sky models

The spectra of phases in the visibilities and the bispectrum are investigated to gain an understanding of their behavior in the frequency domain for each of the sky model examples.

1. Example (i): Unresolved foreground and unresolved spectral line signal

From Eq. (15), the perturbation in the interferometric phase angle is given by

$$\begin{aligned} \delta\phi_p^L(f) &= \frac{A(\hat{s}_L) V^L \cos(2\pi f \tau_L + \theta_0)}{A(\hat{s}_F) V^F \left(\frac{f}{f_F}\right)^\alpha} \\ &\quad \times \Im \{ e^{-i2\pi \mathbf{b}_p \cdot (\hat{s}_L - \hat{s}_F)} \} \\ &= \frac{A(\hat{s}_L) V^L \cos(2\pi f \tau_L + \theta_0)}{A(\hat{s}_F) V^F \left(\frac{f}{f_F}\right)^\alpha} \\ &\quad \times \sin \left(2\pi \frac{f}{c} \mathbf{b}_p \cdot (\hat{s}_L - \hat{s}_F) \right). \end{aligned} \quad (52)$$

We define $\tau_p(\hat{s}_L, \hat{s}_F) \equiv \tau_p \equiv \mathbf{b}_p \cdot (\hat{s}_L - \hat{s}_F)/c$.

For a point source foreground, regardless of α , $\phi_{\nabla}^E(f) = 0$, $e^{i\phi_{\nabla}^E(f)} = 1$, and thus, $\tilde{\mathcal{E}}_{\nabla}^F(\tau) = \delta(\tau)$. With $V_{\text{eff}}^F = V^F$, we get $\gamma^F(f) = (f/f_F)^{-\alpha}$ which is a smooth function of frequency. Thus $\tilde{\gamma}^F(\tau)$, the delay-domain dual of $\gamma^F(f)$, is a sharply peaked function in τ . From Eq. (47), the delay spectrum of the spectral line fluctuations in the bispectrum phase is given by

the foregrounds and the cosmic spectral line signal are such that $\Im \{ V_p^L(f)/V_p^F(f) \} = 0$. Therefore, from Eq. (30), $\delta\phi_{\nabla}^L(f) = 0$. This will also be true even without the linear-order approximation. Thus, even though the spectral structures are very different between the foreground and the hypothetical cosmic spectral line model, the spectral fluctuations from the latter will be indistinguishable from the foregrounds in $\phi_{\nabla}(f)$ because of the perfect relative symmetry in the transverse sky structure between the foregrounds and the cosmic signal (equivalent to the vanishing of the imaginary part in the visibility ratios).

Figure 6(a) shows for example (i)(a) the amplitude of the visibilities on the three baselines comprising the 50.6 m

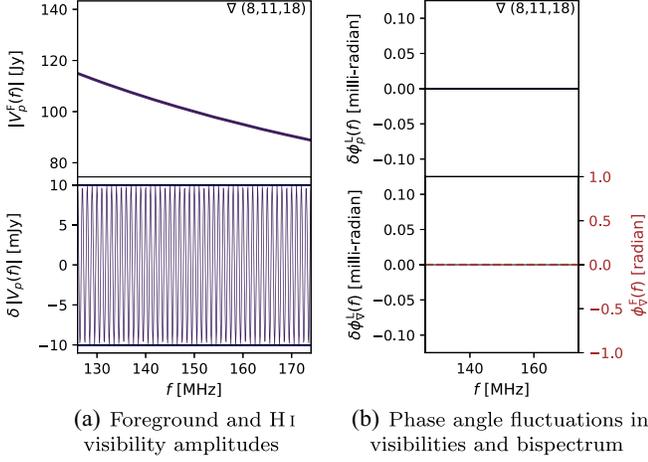


FIG. 6. Left: foreground visibility amplitudes for a point source with a spectral index, $\alpha = -0.8$ (top) and the fluctuations therein (bottom), $\delta|V_p(f)| = |V_p^F(f) + V_p^H(f)| - |V_p^F(f)|$, for example (i)(a) measured on three antenna spacings (red, blue, and black) comprising the 50.6 m equilateral triad. Right: perturbations due to the fluctuating H I spectrum in the three visibility phase angles, $\delta\phi_p^L(f)$, in red, blue, and black (top) and the bispectrum phase angle (bottom) with the foreground component $\phi_V^F(f) = 0$ (dashed brown curve to be read off the y axis placed on the right) and the fluctuations caused by the cosmic H I component $\delta\phi_V^L(f)$ (solid gray, y axis on the left). Phase angle fluctuations in both the visibilities and the bispectrum are absent because of the colocation and symmetry in the transverse structure of the cosmic H I signal relative to the point source foreground model. The frequency ranges in the x -axis are restricted only to enhance readability.

equilateral triad due to the foreground (top), and the perturbations in visibility amplitudes (bottom) obtained as $\delta|V_p(f)| = |V_p^F(f) + V_p^H(f)| - |V_p^F(f)|$ due to the hypothetical H I spectral line signal. Figure 6(b) shows the fluctuations in the phase angles of the visibility on the three antenna spacings, $\delta\phi_p^L(f)$ (top) and the bispectrum, $\delta\phi_V^L(f)$ (bottom). They are both identically zero as expected even without any linear-order approximation.

Although this example is unrealistic and the spectral structure of H I fluctuations does not manifest at all in the bispectrum phase in a manner useful towards its detection, it nevertheless reveals an important property of the spectrum of bispectrum phase fluctuations. If the transverse structure of the fluctuating signal is in perfect relative symmetry with respect to the underlying foreground transverse structures, then regardless of their inherently distinct spectral structures, no spectral signatures from the fluctuating cosmic spectral line signal will manifest in the bispectrum phase.

Next, the frequency-domain behavior of the sky model examples (i)(b), (ii), (iii), and (iv) are discussed below in detail and illustrated collectively in Figs. 7 and 8, which characterize the amplitudes and phases, respectively.

Example (i)(b): Spectral line signal transversally displaced from the foreground.— In contrast to example (i)(a) in Sec. VA 1, usually $\exists \mathbf{b}_p$ such that $\mathbf{b}_p \cdot (\hat{\mathbf{s}}_L - \hat{\mathbf{s}}_F) = c\tau_p \neq 0$. Under such conditions, Eqs. (52) and (53) yield a nonzero response, at a minimum of 4 and possibly up to 12 sharply peaked distinct delays (four for each p) at $\tau = \pm(\tau_p - \tau_L)$ and $\tau = \pm(\tau_p + \tau_L)$. The delay spectrum of spectral line visibility can be expressed as delta functions at $\tau = \frac{\mathbf{b}_p \cdot \hat{\mathbf{s}}_L}{c} \pm \tau_L$:

$$\begin{aligned} \tilde{V}_p^L(\tau) &= \frac{A(\hat{\mathbf{s}}_L)V^L e^{i\theta_0}}{2} [\delta(\tau - \tau_L) + \delta(\tau + \tau_L)] \\ &\quad * \delta\left(\tau - \frac{\mathbf{b}_p \cdot \hat{\mathbf{s}}_L}{c}\right) * \tilde{W}(\tau) \\ &= \frac{A(\hat{\mathbf{s}}_L)V^L e^{i\theta_0}}{2} \left[\tilde{W}\left(\tau - \left(\frac{\mathbf{b}_p \cdot \hat{\mathbf{s}}_L}{c} + \tau_L\right)\right) \right. \\ &\quad \left. + \tilde{W}\left(\tau - \left(\frac{\mathbf{b}_p \cdot \hat{\mathbf{s}}_L}{c} - \tau_L\right)\right) \right]. \end{aligned} \quad (54)$$

Mathematically, the delay spectrum of the bispectrum phase using the linear approximation is expected to differ from the standard delay spectrum in the following ways:

- (1) The former is sharply peaked but *broader than the latter* because of the convolution with the delay response of the foreground spectrum term, $\tilde{\gamma}^F(\tau)$. This broadening behavior is more clearly illustrated in other examples that follow.
- (2) The former peaks at four distinct delays corresponding for each antenna spacing because the bispectrum phase fluctuations are a product of the cosine-shaped spectrum with the sine of the phase angle from the displacement of the cosmic H I source relative to the foreground object, whereas the latter has only a pair of peaks for each visibility from the cosine-shaped spectrum.
- (3) The spatial position (and the spatial structure, in general) of the foreground influences where the delays are in the bispectrum phase delay spectrum. This reaffirms that the bispectrum phase fluctuations are a measure of the relative transverse-plane asymmetry (or dissimilarity) between the cosmic fluctuations and the dominant foregrounds, whose magnitude depends on the ratio between the two as predicted to first order by Eqs. (30) and (32).

Figure 7(a) shows that the positional displacement between the H I and the foreground models causes a slow variation in the envelope of the visibility fluctuations with a maximum amplitude of 10 mJy and fastest spectral variations are on scales of $\delta f_L = 1/\tau_L = 1$ MHz. Figure 8(a) shows the fluctuations in the phase angles of the visibility (top subpanel) and the bispectrum (bottom subpanel). In contrast with example (i)(a) [see Fig. 6(a)], the phase angle fluctuations in both the visibilities and bispectrum phase are nonzero *because of the relative asymmetry in the*

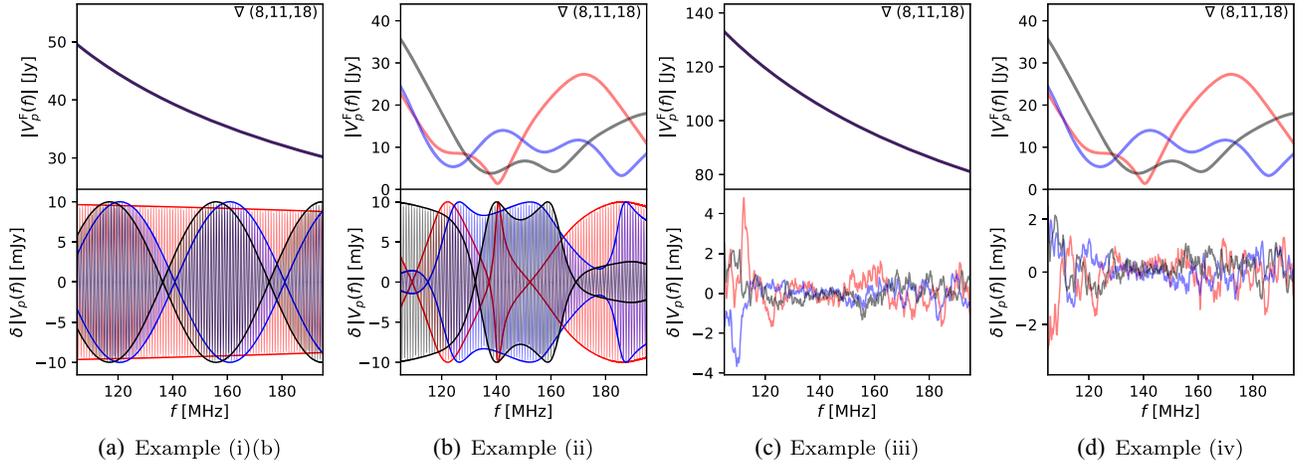


FIG. 7. Frequency spectra of amplitudes of visibilities due to foregrounds (top subpanels) and fluctuations therein caused by the cosmic HI spectral line signal (bottom subpanels) for the examples specified. The 50.6 m equilateral triad used in these panels is specified at the top of the top subpanels. The three visibilities are shown in red, blue, and black. (a) Example (i)(b): the foreground visibility amplitudes are smaller than in Fig. 6(a) by a factor equal to the power pattern of the Airy disk at its angular separation from boresight. The visibility amplitude fluctuations are seen to have an envelope of amplitude ~ 10 mJy with the fastest spectral variation on scales of $\delta f_L = 1/\tau_L = 1$ MHz. The slower variation of the envelope is determined by the location of the cosmic HI relative to the foreground point source. (b) Example (ii): the foreground visibilities obtained from the GLEAM catalog have a richer spectra owing to the wide-field distribution of foregrounds in the transverse direction, and yet exhibit smooth spectra. The broadband changes in the amplitude of the fluctuations are also due to the wide-field spatial distribution of the foreground objects. (c) Example (iii): the point source foreground has an extremely smooth spectrum, while the realistic HI model exhibits rich spectral fluctuations on a wide range of scales. (d) Example (iv): although the foregrounds from the GLEAM catalog show a rich foreground structure, they are still much smoother compared to the cosmic HI spectral line signal obtained from the 21cmFAST simulations.

transverse structure between the foreground and the HI model, and increase in amplitude towards higher frequencies due to the lowered foreground visibility amplitudes caused by $\alpha < 0$.

2. Example (ii): Realistic foreground and unresolved spectral line signal

Here, we consider a foreground model determined by the GLEAM catalog [64]. Visibilities were modeled for sources in the GLEAM catalog within a circle of 30° diameter around boresight. The cosmic HI signal is modeled as the same unrealistic point source at boresight with a cosine-shaped spectrum as in the previous example.

Figure 7(b) shows the amplitude of the foreground visibilities and the fluctuations in the amplitude due to the HI fluctuations. The fastest spectral fluctuations are on scales of $\delta f_L = 1/\tau_L = 1$ MHz while the broadband fluctuations in the amplitude are determined by the locations of the various GLEAM catalog objects relative to the HI point source at boresight. Figure 8(b) shows the phase angle fluctuations in the visibility (top) and the bispectrum (bottom). The broadband changes in the amplitude of the phase angle fluctuations are due to the changing amplitudes of the foreground visibility spectra. Although the foregrounds have richer spectral structure they are still relatively smooth.

3. Example (iii): Unresolved foreground and fiducial cosmic spectral line signal

This example consists of a point source of flux density 100 Jy at 150 MHz, located at boresight and spectral index $\alpha = -0.8$ for the foreground model. As our fiducial EoR HI model, we use the FAINT GALAXIES model [65,66] from 21cmFAST simulations [67] centered on boresight.

Figure 7(c) shows the foreground visibility amplitude (top subpanel) and the fluctuations in the amplitude caused by the cosmic HI fluctuations (bottom subpanel). Figure 8(c) shows the phase angle fluctuations in the visibilities (top subpanel) and the bispectrum (bottom subpanel). The phase angle fluctuations span a wide range of frequency scales. They are seen to be at a level $\lesssim 10^{-4}$ radians (which depends on the level $\lesssim 10^{-4}$ radians (which depends on the foreground visibility amplitudes) that is consistent with Eqs. (30) and (32), and have coherent structures on frequency scales that approximately correspond to those in Fig. 7(c).

4. Example (iv): Realistic foreground and fiducial cosmic spectral line signal

We consider the GLEAM catalog of radio sources for our foreground model and the fiducial 21cmFAST EoR HI model from the example above. This represents a realistic realization of both foregrounds and the cosmic EoR HI signal. Figure 7(d) shows the foreground visibility amplitudes and the fluctuations therein for a 50.6 m equilateral

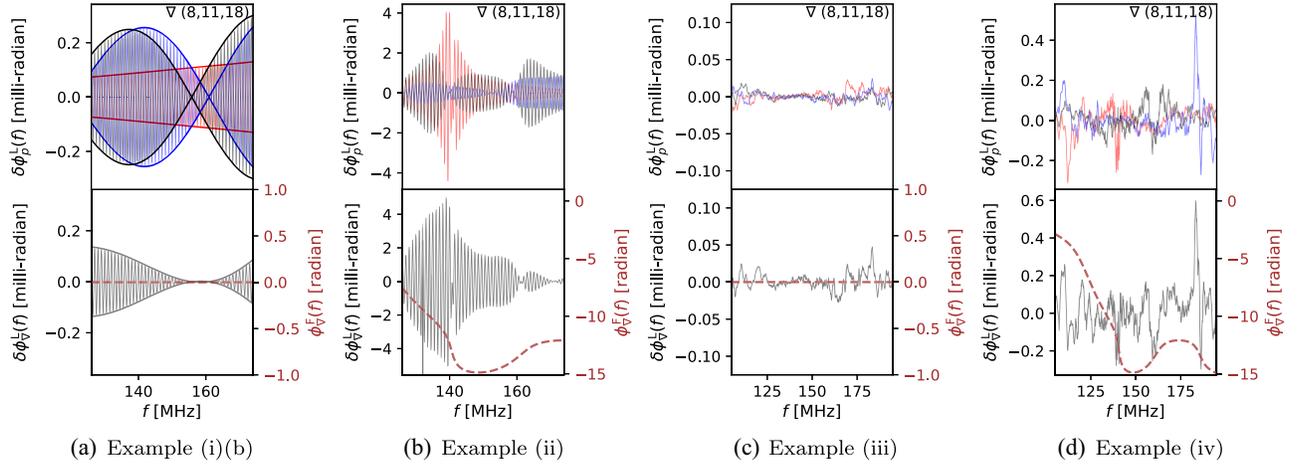


FIG. 8. Frequency spectra of the phase angle fluctuations on the visibilities (top subpanels) and the bispectrum due to the cosmic H I spectral line fluctuations. The top subpanel shows the nonvanishing visibility phase fluctuations on the three antenna spacings (red, blue, and black). The bottom subpanels show the component of bispectrum phase from the foreground (dashed brown curve to be read off the y axis placed on the right) and the bispectrum phase angle fluctuations (gray curves). The 50.6 m equilateral triad used in these panels is specified at the top of the top subpanels. (a) Example (i)(b): As the foreground visibility amplitude decreases at higher frequencies due to the spectral index [see Fig. 7(a)], the amplitude of the phase angle fluctuations increases. Both these phase angle fluctuations are nonzero unlike when the cosmic H I and foreground models were perfectly colocated with respect to each other in example (i)(a) [see Fig. 6(b)]. (b) Example (ii): The foreground component of the bispectrum phase angle is nonzero and has smooth spectral structure which is unwrapped to remove discontinuities at odd multiples of $\pm\pi$. The envelope of phase fluctuations in both the visibilities (top subpanel) and the bispectrum (bottom subpanel) has amplitudes inversely proportional to the foreground visibility amplitudes [see Fig. 7(b)]. (c) Example (iii): The point source foreground yields $\phi_p^F(f) = 0$ (dashed brown curve with its y axis placed on the right). There is a significant correspondence between the shape and scale of these phase angle fluctuations in visibilities and the bispectrum to those in the visibility amplitudes in the bottom subpanel of Fig. 7(c). (d) Example (iv): The cosmic H I spectral line signal from the 21cmFAST simulations shows a lot more spectral structure in the phase angle fluctuations of the visibilities and the bispectrum relative to the smooth spectral structure from the GLEAM foregrounds. There is a significant correspondence between the phase angle fluctuations shown here and the amplitude fluctuations in Fig. 7(d). The sharp spikes in the spectra of phase angle fluctuations are generally regions with low foreground amplitudes (for example, near 140 and 187 MHz). The frequency ranges in the x-axis are restricted only to enhance readability.

triad. Figure 8(d) shows the phase angle fluctuations of the visibilities and the bispectrum. The sharp spikes, for example at ≈ 186 MHz, are due to low foreground amplitudes and thus reaffirm that they are a function of the ratio of the H I fluctuations to the foreground strength. Although the foreground model from the GLEAM catalog exhibits a relatively rich spectral structure, they are still much smoother compared to the cosmic H I spectral line fluctuations from the EoR obtained using 21cmFAST simulations. Redshifted 21 cm interferometer experiments aim to detect this spectral distinction.

B. Delay- (Fourier-)domain view of the sky models

The Fourier-domain view into the visibilities and the bispectrum phases of these examples are explored using their delay spectra.

1. Example (i): Unresolved foreground and unresolved spectral line signal

Spectral line signal transversally colocated with the foreground.— As discussed in Sec. VA 1, despite the

cosmic spectral line signal having a cosine-shaped spectral structure, its signatures are completely absent in the bispectrum phase angle. Therefore, the delay spectrum of the bispectrum phase considered in example (i)(a) is not expected to show any signatures in the Fourier domain as well.

The delay spectra of the rest of the sky model examples (i)(b), (ii), (iii), and (iv) are discussed below in detail and illustrated in Fig. 9.

Spectral line signal transversally displaced from foreground.— Figure 9(a) shows the delay spectra of the visibilities (left subpanel) and the bispectrum phase (right subpanel) for example (i)(b). The region shaded in yellow in the left subpanel denotes the modes expected to be contaminated by the foregrounds, namely, the *foreground wedge* [27,28,60–62,69–81], whose boundaries are determined by the horizon delay limit, $\tau_{p,h} = |\mathbf{b}_p|/c$. Since we use equilateral triads in these examples, $\tau_{p,h} = |\mathbf{b}_p|/c$ is identical for all p . The foreground delay spectra (solid lines) are seen to be the result of a convolution of a delta function at $\tau = 0$ with the spectral window function's delay-domain response, $\tilde{W}(\tau)$. The location of the peak

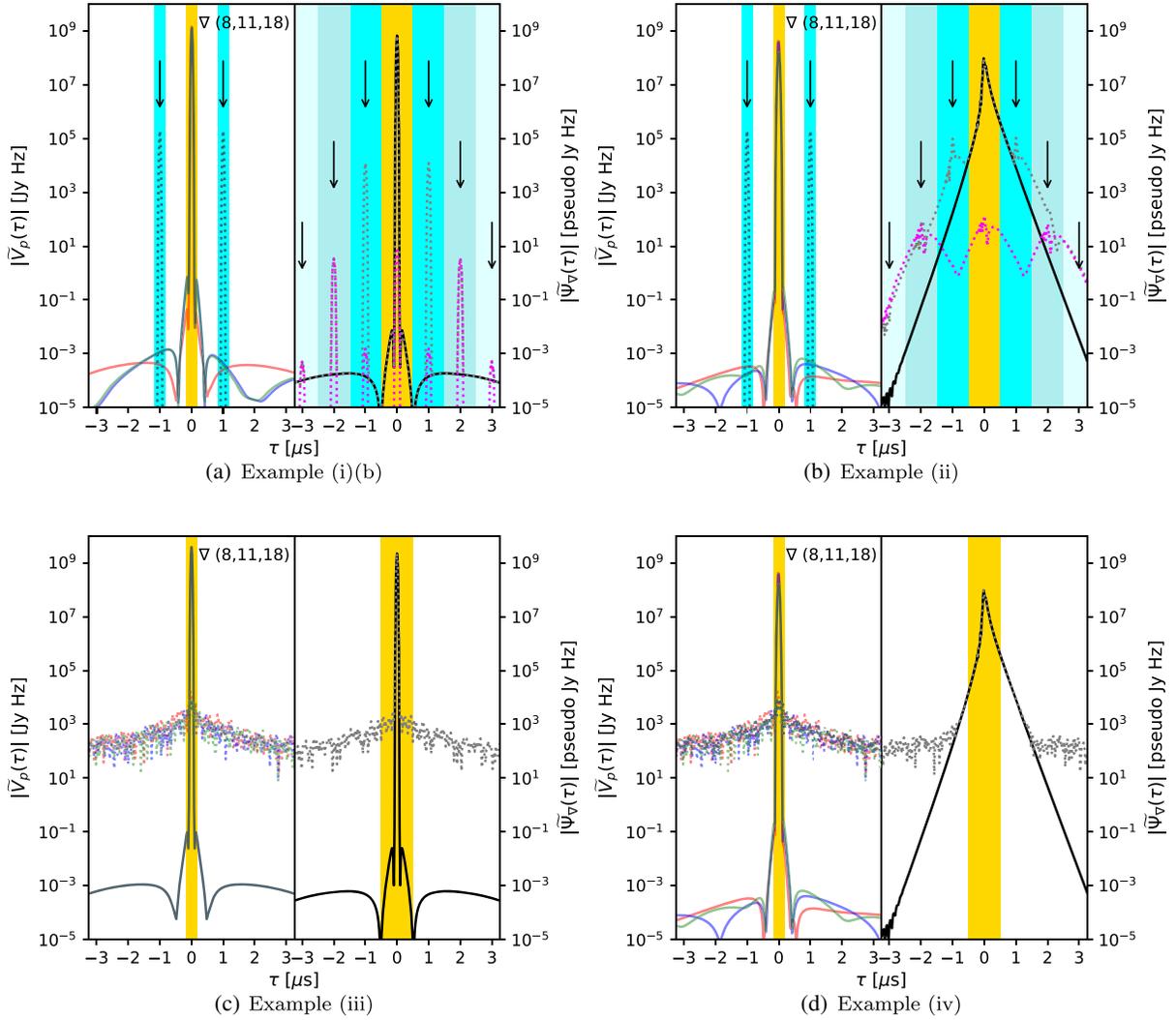


FIG. 9. Standard delay spectrum of the three visibilities, $\tilde{V}_p(\tau)$ (red-, blue-, and green-colored curves in the left subpanels) and the bispectrum phase delay spectrum, $\tilde{\Psi}_\nabla(\tau)$ (right subpanel) for the specified examples. The foreground and the cosmic HI signatures in both subpanels are shown in solid and dotted curves, respectively. The 50.6 m equilateral triad used in these panels is specified at the top of the left subpanels. The yellow-shaded region denotes the foreground-dominated modes (*foreground wedge*) in the delay spectra of the visibilities (left subpanels) and the bispectrum phase (right subpanels). The latter is wider due to triple convolution resulting from the multiplicative combination of the three visibility phases. In examples (i)(b) and (ii) that contain a point source cosmic HI signal with a cosine-shaped spectrum, the downward arrows at $\tau = \pm\tau_L$ where the principal harmonic mode of the cosmic HI signal is expected. The other downward arrows at $\tau = \pm 1 \mu\text{s}$, $\tau = \pm 2 \mu\text{s}$, $\tau = \pm 3 \mu\text{s}$ in the right subpanels denote the higher-order harmonics. Even higher-order harmonics are expected to be present *but* at negligible levels and are not shown. The pink dotted curves in examples (i)(b) and (ii) denote the difference between the delay spectra of the actual and the linear-order approximation of the bispectrum phase indicating that practically all the power in the second and third harmonics at $\tau = \pm 2 \mu\text{s}$ and $\tau = \pm 3 \mu\text{s}$ is entirely absent in the linear-order approximation. The second- and third-order terms also contribute at $\tau = 0 \mu\text{s}$ and $\tau = \pm 1 \mu\text{s}$ respectively but are much smaller (by 10 and 7 orders of magnitude respectively) than that from the linear-order approximation at these harmonics. The cyan region denotes the possible range of offsets (same width as the foreground wedge) that the expected delay mode the HI signal could be subject to, i.e., $(\tau_L - \tau_{p,h}) \leq |\tau_{p,L}^{\text{obs}}| \leq (\tau_L + \tau_{p,h})$ instead of being precisely centered at $\tau_{p,L}^{\text{obs}} = \pm\tau_L$. The detailed analysis of these results for each of the examples is presented in the corresponding section in the text. In general, the foreground-dominated modes are wider in the bispectrum phase relative to the visibility delay spectra. Nevertheless, the cosmic spectral line signal is still detectable with a similar dynamic range and shape in the higher-order Fourier modes of the bispectrum phase compared to that in standard visibilities.

of the foreground visibility delay spectra are displaced from $\tau = 0 \mu\text{s}$ by a small amount $\delta\tau_p = \mathbf{b}_p \cdot \hat{\mathbf{s}}_F/c \lesssim 0.013 \mu\text{s}$. Since this displacement is smaller than the resolution or the

width of the response of the window function ($\delta\tau_p < \delta\tau_w$), the displacement is not discernible. The HI spectrum which is cosine shaped in this hypothetical example manifests as

two delta functions at $\tau = \pm\tau_L = \pm 1 \mu\text{s}$ (expected at the locations of the downward arrows). Because the delay-domain response depends both on the transverse location and the frequency spectrum of the signal, the location of the delta functions of the cosine-shaped HI spectrum in general could be subject to a delay offset that could be as large as the horizon delay limit, $\tau_{p,h}$. Thus, in general, depending on the location of the HI signal, the delta functions corresponding to the cosmic HI signal could be located anywhere in the cyan-shaded regions with $(\tau_L - \tau_{p,h}) \leq |\tau_{p,L}^{\text{obs}}| \leq (\tau_L + \tau_{p,h})$ instead of being precisely centered at $\tau_{p,L}^{\text{obs}} = \pm\tau_L$.

The foreground component of the delay spectrum of the bispectrum phase, $\tilde{\Psi}_{\nabla}(\tau)$, in the right subpanel of Fig. 9(a) is very similar in magnitude, shape, and dynamic range to its counterpart in the standard delay spectrum, $\tilde{V}_p(\tau)$ in the left subpanel. The general expectation for the foreground-contaminated modes is that they will be wider because the term, $\tilde{\mathcal{E}}_{\nabla}^{\text{F}}(\tau)$, is derived from the convolution of the delay transforms of the three visibility phase terms that appear as a product in the frequency spectrum. This triple product of visibility phase terms will, in general, widen the analogous foreground wedge through the aforementioned convolution. Hence, the foreground wedge in the delay spectrum of the bispectrum phase is determined by the sum of the three horizon limits in delay (Fourier) domain, $|\tau_{\nabla,h}| \leq \sum_p \tau_{p,h}$, and is shown by the yellow-shaded region. The three shades of cyan denote the expected delay-offset locations of the HI signal from the first- and higher-order harmonics of the cosine-shaped spectrum. The bright cyan shade denotes the range of delay offset around the expected location of the first harmonic centered around $(n\tau_L - \sum_p \tau_{p,h}) \leq |\tau_{\nabla,L}^{\text{obs}}| \leq (n\tau_L + \sum_p \tau_{p,h})$ with $n = 1$ and primarily arises from the linear-order terms discussed in Sec. IV A.

The actual bispectrum phase will contain higher-order perturbations $\sim \cos^n(2\pi f\tau_L)$, and will not be captured by the linear-order approximation. The medium and pale shades of cyan regions correspond to the second and third harmonics ($n = 2, 3$ respectively) expected from the second- and third-order perturbations respectively. The second-order terms are expected to contribute to both the zeroth and second harmonics since $\cos^2\theta \sim 1 + \cos 2\theta$ (yellow and medium-cyan regions respectively) while the third-order terms will contribute to the first and third harmonics since $\cos^3\theta \sim 3\cos\theta + \cos 3\theta$ (dark- and pale-cyan regions respectively). There will be harmonics of even higher orders which have not been shown here because of their rapidly diminishing strengths. The width of each of these regions is the same as that of the yellow-shaded region.

Unlike when the HI and foreground models are colocated and symmetric relative to each other in the transverse plane, the presence of nonzero phase angle fluctuations manifests prominently in the delay spectrum of the bispectrum phase (right subpanel) as sharply peaked functions (gray dotted lines) at $\tau \simeq \pm\tau_L = \pm 1 \mu\text{s}$ with small displacements around

these locations $\sim \pm\tau_p$, where $\tau_p \lesssim 0.013 \mu\text{s}$. These displacements are also indiscernible because $\delta\tau_p < \delta\tau_w$ and hence the 12 delta functions predicted in Eq. (53) have blended into two sharply peaked functions one each on either side of $\tau = 0 \mu\text{s}$. Besides the principal first harmonics at $\tau \simeq \pm\tau_L$, sharp peaks are also seen as indicated by the downward arrows at $\tau \simeq \pm n\tau_L$ with $n = 1, 2, \dots$ and $n = 1$ being the principal (or first) harmonic. The pink dotted curve shows the difference between the delay spectrum of the actual bispectrum phase and that obtained with the linear-order approximation. The second and third harmonic components that arise from second- and third-order terms in the expansion of the bispectrum phase are lower relative to the first harmonic by factors $\sim 10^{-4}$ and $\sim 10^{-7}$ respectively which agree well with the fractional inaccuracy of the first-order prediction illustrated earlier in Sec. IV A. The second- and third-order terms also contribute to the zeroth and first harmonics that are not entirely represented by the linear-order approximation as indicated by the pink dotted curves around $\tau \simeq 0 \mu\text{s}$ and $\tau \simeq \pm\tau_L$ respectively, but these are negligible contributions (fractionally $\simeq 10^{-8}$) compared to the linear-order or actual values of the bispectrum phase delay spectra. There will be even higher-order harmonics in the actual bispectrum phase angles missed by the linear-order expansion but these contributions are expected to be even more increasingly negligible and are not shown.

2. Example (ii): Realistic foreground and unresolved spectral line signal

Relative to the standard visibility delay spectra, $\tilde{V}_p(\tau)$ in Fig. 9(a), the visibility delay spectra for the GLEAM foreground model [left subpanel in Fig. 9(b)] appear wider and fill the foreground wedge (central yellow region) as expected. The delay spectrum of the foreground component in the bispectrum phase [right subpanel of Fig. 9(b)] is much wider filling the correspondingly wider yellow central region due to $\tilde{\mathcal{E}}_{\nabla}^{\text{F}}(\tau)$ which is formed by the triple convolution of the delay spectra of the visibility phase terms arising from the foregrounds as predicted by Eqs. (47) and (49). The cosmic HI signatures (gray dotted curves) are centered on the expected delays $\tau = \tau_L$ but as in the previous case, there are fainter copies at higher delay harmonics. As seen earlier, the higher-order harmonics that are not fully represented in the linear-order approximation contribute negligibly (by few to many orders of magnitude) to the zeroth and first harmonics (pink dotted curves).

The most notable observation is that the cosmic HI signatures appear to have the shape of the foreground signatures indicating they resulted from a convolution of a delta function with the foreground terms $\tilde{\mathcal{E}}_{\nabla}^{\text{F}}(\tau)$ and $\tilde{\gamma}_p(\tau)$ as detailed in Eqs. (47) and (49). The previous example also had these effects but the foreground spectral signatures were not as rich to be clearly visible as in the present example.

3. Example (iii): Unresolved foreground and fiducial cosmic spectral line signal

The left subpanel of Fig. 9(c) shows the standard visibility delay spectrum for the baselines comprising the 50.6 m equilateral triad (left subpanel) of the point source foreground and the fiducial EoR H I model from the 21cmFAST simulations. On the right subpanel, the delay spectrum of the corresponding bispectrum phase is shown—the foregrounds in solid black, and the combined foregrounds and EoR H I fluctuations in dotted gray. Because of the simple spectral structure of the foregrounds, the delay spectra from either approach look remarkably similar to each other in their overall characteristics.

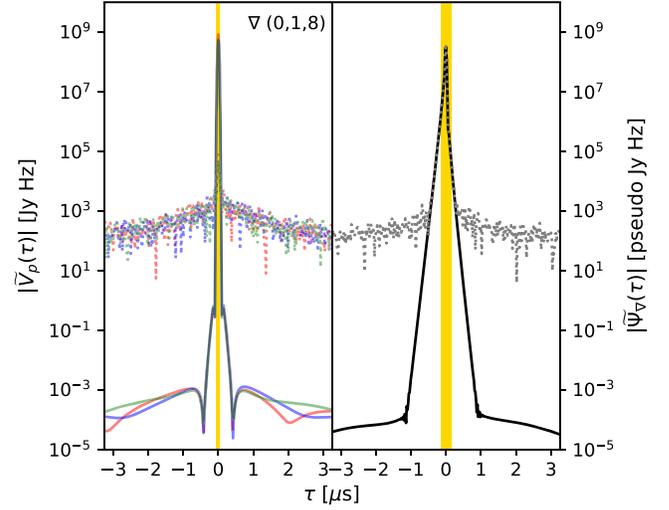
4. Example (iv): Realistic foreground and fiducial cosmic spectral line signal

Figure 9(d) shows the delay spectra of the visibilities and the bispectrum phase. The foreground (solid red, green, and blue curves) and the fluctuating H I components in the standard delay spectrum (left subpanel) are found to occupy the foreground wedge (yellow shaded region) and extend beyond into the EoR window respectively. The delay spectrum of the bispectrum phase (right subpanel) shows the foreground component (solid black curve) significantly wider due to the presence of significant spectral modes in the $\tilde{\mathcal{E}}_{\nabla}^F(\tau)$ term arising out of the GLEAM foregrounds. This is in significant contrast to example (iii) [see Fig. 9(c)] because the visibilities from the GLEAM foregrounds intrinsically contain more spectral structure than a single point source. The extent of this widening decreases with decreasing antenna spacings in the triad as will be shown in Sec. VI and thus will be less severe for a 14.6 m equilateral triad. The H I component (dotted gray) does separate from the foregrounds at a level and shape similar to that in the standard delay spectrum. However, because the foreground component is significantly wider, the number of modes in which the H I is detectable in the bispectrum phase approach is reduced to $|\tau| \gtrsim 1 \mu\text{s}$ relative to $|\tau| \gtrsim 0.12 \mu\text{s}$ in the standard delay spectrum of the visibilities.

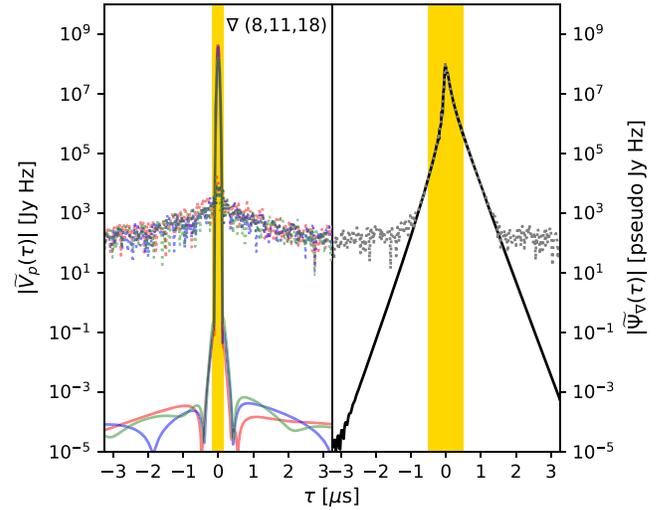
In summary, the foregrounds occupy a larger range of inner spectral modes and hence the range of detectable cosmic spectral line signal modes are reduced but it is still significantly detectable in the remainder of the higher-order spectral modes. An alternate approach using the bispectrum phase angle is briefly outlined in the Appendix that could potentially avoid the widening of foreground contamination into larger spectral modes significantly.

VI. MODE-MIXING IN BISPECTRUM PHASE

Mode-mixing in the context of spectral line experiments in the presence of foregrounds refers to the dependence of the line-of-sight spatial modes on the transverse spatial modes [27,28,60–62,69–81] and is now commonly referred to as the foreground wedge. To examine this effect in our



(a) Delay spectra of a 14.6 m equilateral triad



(b) Delay spectra of a 50.6 m equilateral triad

FIG. 10. Top: delay spectra of a 14.6 m equilateral triad. The left subpanel shows delay spectra of the three visibilities (red, green, and blue) comprising the triad for the foregrounds (solid curves) and the EoR H I signal (dotted). The right subpanel shows the delay spectra of the bispectrum phase with foregrounds only (solid black curve), and with the EoR H I fluctuations also present (gray dotted curve). Bottom: same as Fig. 10(a) (top) but for a 50.6 m equilateral triad, and thus identical to Fig. 9(d). Both the delay spectra of the visibilities (left subpanel) and the bispectrum phase (right subpanel) are wider in the case of the 50.6 m equilateral relative to the 14.6 m equilateral triad. This indicates that in the bispectrum phase, the transverse foreground modes also contaminate the line-of-sight foreground modes, as is the case in a standard delay spectrum approach using visibilities. Although cosmic signal-dominated modes are still accessible with a 50.6 m equilateral triad, they are much more accessible with a 14.6 m triad where the foreground contamination is much more tightly restricted.

bispectrum phase approach, we consider the 14.6 and 50.6 m equilateral triads. The former samples lower order transverse spatial modes relative to the latter. In both cases, the foreground model is drawn from the GLEAM catalog, and the EoR H I model from the fiducial 21cmFAST simulation of the EoR.

Figure 10(a) shows the delay spectra of the three (red, blue, green) visibilities (left subpanel) and of the bispectrum phase (right subpanel) corresponding to the 14.6 m equilateral triad. The solid curves correspond to the foreground component while the dotted curves represent the case when the cosmic H I fluctuations are present. Figure 10(b) is the same but for a 50.6 m equilateral triad and is identical to Fig. 9(d). In both the delay spectra (left and right subpanels), the foreground component widens for the 50.6 m equilateral triad relative to the 14.6 m equilateral triad. This conclusively proves that the delay spectrum of the bispectrum phase is also subject to mode-mixing effects in general, wherein the transverse spatial modes contaminate the line-of-sight spatial modes, similar to the standard delay

spectrum. However, the foreground wedge is still limited in extent and the cosmic signal is detectable in the line-of-sight modes even for the 50.6 m equilateral triad. The foreground contamination is found to be much more limited and a wider range of cosmic signal-dominated modes are accessible with the usage of smaller triads such as the 14.6 m equilateral triad. The Appendix presents an outline of a variant to this approach using the bispectrum phase angle $\phi_{\nabla}^m(f)$ instead of $e^{i\phi_{\nabla}^m(f)}$ which is expected to be not as susceptible to mode-mixing as the latter is.

VII. IMPACT OF FOREGROUND SPECTRAL CHARACTERISTICS

Here, we compare the effects of the foreground spectral characteristics, such as the spectral index, on the delay spectra of the bispectrum phase and that of the visibilities. The example is similar to that in Sec. V A 1. It consists of a 100 Jy point-source foreground model $\approx 5^\circ$ off-boresight and an unrealistic point-source H I model at boresight of

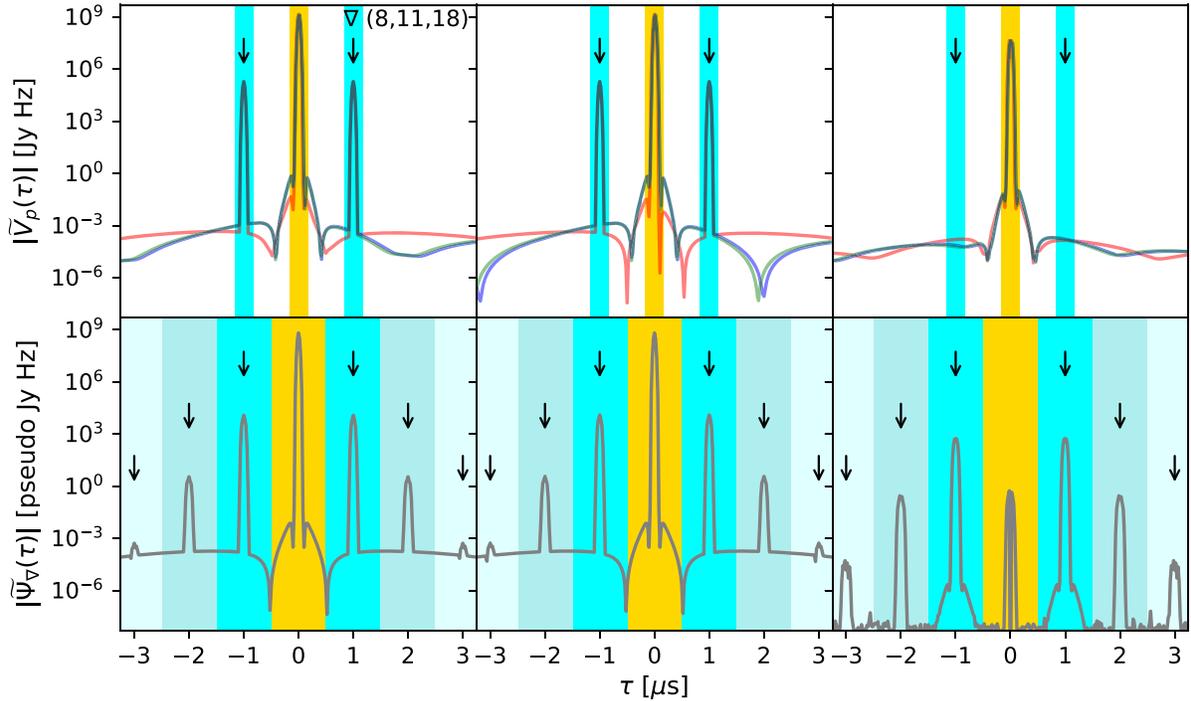


FIG. 11. Top panels correspond to the delay spectrum of the visibilities on a 50.6 m equilateral triad. Bottom panels correspond to the delay spectrum of the bispectrum phase on the same triad. The sky model used is similar to the hypothetical and unrealistic example (i)(b) with minor differences. The left and middle panels in both rows are derived using $\alpha = -0.8$ and $\alpha = 0$ respectively for the point source foreground. The right panels show the absolute value of the difference in the two delay spectra from these sky models. The downward arrows in the top panel indicate the expected location of the cosmic H I signal at $\tau = \pm\tau_L$, whereas the downward arrows in the bottom panels show the location of the n th harmonics $\tau = \pm n\tau_L$ as discussed in previous examples. The difference of the visibility delay spectra (top right subpanel) shows a residual that is reduced in magnitude but is purely foreground-based due to the differences in the spectral index of the two foreground models used and a complete absence of the H I signatures at $\tau = \pm\tau_L$. The difference between the delay spectra of the bispectrum phase using the different spectral indices in the foreground models is shown in the bottom right subpanel. In contrast with the differenced visibility delay spectra (top-right subpanel), the H I signatures do not vanish and the residuals have foregroundlike signatures even at the harmonics where the H I signatures are expected. This implies there is a mixing of the spectral characteristics of the foregrounds into the H I signatures.

amplitude 10 mJy with a cosine-shaped spectrum of characteristic scale $\delta f_L = 1/\tau_L = 1$ MHz. The 50.6 m equilateral triad is used.

The left panels of Fig. 11 use a spectral index, $\alpha = -0.8$, for the foreground model while the middle panels use $\alpha = 0$. The right panels denote the absolute value of the difference between the delay spectra in the middle and the left panels. The top and bottom panels apply to the delay spectrum of the visibilities and the bispectrum phase respectively. In case of the latter, since different scalings to obtain the “pseudo” flux densities may have been applied, we normalized their peaks to be equal before the differencing. The downward arrows indicate where the HI signatures are expected as sharply peaked functions. In the case of the bispectrum phase, the multiple downward arrows on each side of $\tau = 0 \mu\text{s}$ denote the different harmonics of the $\tau_L = 1/\delta f_L = 1 \mu\text{s}$ spectral mode as discussed in the examples above.

In the case of the standard visibility delay spectrum (top panels), the HI signatures are completely absent in the difference (right subpanel) and the peak of the foreground component has reduced by more than an order of magnitude. The difference still has a finite width around $\tau = 0$ which can be attributed to the spectral index being different between the foreground models. This shows that the HI signatures were not affected by the spectral characteristics of the foreground component and resulted in a perfect subtraction *because of their additive behavior*. The residual purely arises from the spectral characteristics of the two foreground models.

In the case of the delay spectrum of the bispectrum phase (bottom panels), the differencing reduces the amplitude of the HI signatures at the indicated harmonics but they do not vanish (right subpanel). Around each of these harmonics, the convolving effect of the foreground delay spectrum shapes can be seen. The fact that the HI signatures at the harmonics do not vanish and the residuals at these harmonics contain the foregroundlike signatures support the findings of Eq. (49) wherein the spectral characteristics of the foregrounds are mixed with those of the cosmic HI signal *multiplicatively*.

VIII. SUMMARY

Numerous low-frequency radio interferometric measurements are underway to detect the large-scale distribution of baryons in the early Universe. This includes the detection of HI using its 21 cm spectral line signature from high redshifts such as that from the Dark Ages, the Cosmic Dawn, the Epoch of Reionization, and the dark energy-dominated epoch in the Universe. These are expected to be extremely faint spectral signatures where the uncertainties are likely to be dominated by systematic uncertainties (especially of a spectral nature) from the instrument compounded by overwhelmingly bright and undesirable foreground emission from the Galaxy and extragalactic

objects, rather than thermal noise in the detectors. One of the key challenges is the high-accuracy spectral calibration of the instrument which is typically required to have a fractional inaccuracy $\lesssim 10^{-5}$. The use of the bispectrum phase, which is independent of direction-independent antenna-based calibration and errors therein, has been presented as a viable alternative to statistically detect the presence of spectral line fluctuations. In this paper, we lay the foundational steps toward understanding the bispectrum phase in the context of the detection of faint cosmic spectral line fluctuations and examine its potential benefits and limitations.

The principal quantity investigated here for detection—the bispectrum phase—intrinsically measures the asymmetry (or dissimilarity) of the spatial distribution of the cosmic spectral line signal relative to the foregrounds in the transverse sky plane and is expressed as a rotation or fluctuation of the dominant phase angle from the bright foregrounds. In this paper, we focus on using the bispectrum phase to distinguish the faint cosmic spectral line fluctuations from the bright but spectrally smooth foregrounds along the spectral dimension, or equivalently along the line of sight. In the limit of small spectral fluctuations relative to the foregrounds, an approximate correspondence has been established between the approaches using the standard spatial coherence (visibilities) and that using the bispectrum phase. Specifically, the exact mathematical description for the spectral fluctuations in phase angles of both the visibilities and bispectrum have been established using a linear-order approximation (purely for analytical tractability) as being related to the ratio of the strength of the fluctuating signal to that of the foregrounds. Thus, existing Fourier domain techniques (e.g., delay transform) can be readily employed to isolate these fluctuating spectral signatures of the cosmic signal.

We have demonstrated that the linear-order approximation is a valid and useful handle to understand the bispectrum phase in the context of detecting faint spectral line fluctuations from cosmic structures. Although the linear-order approximation neglects the effects from the higher-order perturbation terms which manifest in actual measurements at higher harmonic modes in the delay spectrum, they are found to be negligible, typically by at least a few orders of magnitude.

Except in virtually impossible scenarios where there is perfect symmetry of structure in the transverse sky plane between the source of spectral fluctuations and the foregrounds, the delay spectrum of the bispectrum phase corresponds well with that from the visibilities especially in aspects such as the foreground peak, the magnitude and shape of spectral line signatures, and the dynamic range between the two. This is confirmed using a wide variety of examples which ranged from simple point source models for foregrounds with zero or nonzero spectral index placed at boresight or off-boresight and a hypothetical and

unrealistic point source HI signal with a cosine-shaped spectrum to a realistic wide-area model of the foregrounds using the GLEAM catalog and a fiducial EoR model from 21cmFAST simulations.

In the nearly impossible scenario that the transverse portion of the structures sourcing the spectral line fluctuations are perfectly symmetric relative to the foregrounds, the fluctuations in the bispectrum phase vanish and are undetectable even though there are clear spectral structures in the visibilities that will be detected in a standard delay spectrum. One of the key limitations of this approach stems from the fact that in the bispectrum phase, the foreground component contains a triple-product of the three interferometric visibility phase terms and this leads to a triply convolved and a wider response in the delay spectrum leading to higher levels of contamination in the low-order spectral modes thereby affecting detectability of the cosmic spectral line signal in these modes. *In spite of this, the cosmic signal is still detectable on a wide range of Fourier modes of the bispectrum phase.* A slightly modified approach using the bispectrum phase angles is also briefly presented in the Appendix that could potentially avoid this disadvantage to a significant extent. Further, since the bispectrum phase angle fluctuations depend on the foregrounds which are coupled multiplicatively, rather than additively, the spectral line signatures in a delay spectrum appear convolved with the delay response of the foreground spectral characteristics. Therefore, a straightforward interpretation of the spectral signatures seen in the delay spectrum is difficult and requires either a deconvolution approach to decouple the foreground effects or a detailed forward-modeling.

Despite the limitations, the bispectrum phase approach is an intrinsic measure of the dissimilarity between the cosmic and the contaminating foreground structures and appears to be a viable, independent, and powerful tool to detect faint cosmic spectral line signatures in experiments where calibration of the instrument without corrupting the signatures of the cosmic signal is challenging. In a companion paper [52], we present the first results from applying this technique to a small sample of data from the HERA telescope.

ACKNOWLEDGMENTS

N.T. is a Jansky Fellow of the National Radio Astronomy Observatory. We acknowledge softwares including PYTHON, NUMPY, SCIPY, ASTROPY, MATPLOTLIB that made the numerical computations and the creation of figures presented in this manuscript possible. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

APPENDIX: AN ALTERNATE APPROACH WITH BISPECTRUM PHASE ANGLES

We briefly outline an alternate approach wherein the bispectrum phase angle, $\phi_{\nabla}^m(f)$, is used as the primary physical quantity of interest in deriving the results rather than using the complex Eulerian version, $e^{i\phi_{\nabla}^m(f)}$, as was done in the main text. We assume that the measured bispectrum phase angles are unwrapped accurately without introducing artefacts. In analogy to the complex Eulerian counterparts in Sec. IV A, we define

$$\begin{aligned} V_{\nabla}(f) &= V_{\text{eff}}^F \phi_{\nabla}^m(f) \\ &= V_{\text{eff}}^F [\phi_{\nabla}^F(f) + \delta\phi_{\nabla}^L(f) + \delta\phi_{\nabla}^N(f)] \end{aligned} \quad (\text{A1})$$

$$\begin{aligned} &= V_{\text{eff}}^F \left[\phi_{\nabla}^F(f) + \frac{1}{2i} \sum_{p=1}^3 \left(\frac{V_p^L(f)}{V_p^F(f)} - \frac{\overline{V_p^L(f)}}{\overline{V_p^F(f)}} \right) \right. \\ &\quad \left. + \frac{1}{2i} \sum_{p=1}^3 \left(\frac{V_p^N(f)}{V_p^F(f)} - \frac{\overline{V_p^N(f)}}{\overline{V_p^F(f)}} \right) \right] \\ &= V_{\nabla}^F(f) + V_{\nabla}^L(f) + V_{\nabla}^N(f), \end{aligned} \quad (\text{A2})$$

where

$$V_{\nabla}^F(f) = \phi_{\nabla}^F(f) V_{\text{eff}}^F, \quad (\text{A3})$$

$$V_{\nabla}^L(f) = \delta\phi_{\nabla}^L(f) V_{\text{eff}}^F, \quad (\text{A4})$$

$$\text{and } V_{\nabla}^N(f) = \delta\phi_{\nabla}^N(f) V_{\text{eff}}^F. \quad (\text{A5})$$

Combining the equations above,

$$V_{\nabla}^L(f) = \frac{1}{2i} \sum_{p=1}^3 [\gamma_p^F(f) V_p^L(f) - \overline{\gamma_p^F(f)} \overline{V_p^L(f)}], \quad (\text{A6})$$

$$V_{\nabla}^N(f) = \frac{1}{2i} \sum_{p=1}^3 [\gamma_p^F(f) V_p^N(f) - \overline{\gamma_p^F(f)} \overline{V_p^N(f)}]. \quad (\text{A7})$$

Comparing Eq. (A6) above with its analog, Eq. (42), we notice that the key difference (disregarding some constants of proportionality) is the absence of the foreground bispectrum phase term, $e^{i\phi_{\nabla}^F(f)}$, in Eq. (A6). Since this is the term that is a triple-product of the visibility phases in the triad that leads to a triple-convolution in the delay spectrum resulting in a significant widening of the foreground wedge discussed in Secs. V and VI, its absence in this alternate approach with bispectrum phase angles could potentially mitigate the contamination from mode-mixing presented earlier to a substantial extent. This will be explored in detail in future work.

- [1] R. A. Sunyaev and Y. B. Zeldovich, Formation of clusters of Galaxies: Protocluster fragmentation and intergalactic gas heating, *Astron. Astrophys.* **20**, 189 (1972), <https://ui.adsabs.harvard.edu/abs/1972A%26A....20..189S/abstract>.
- [2] D. Scott and M. J. Rees, The 21-cm line at high redshift: A diagnostic for the origin of large scale structure, *Mon. Not. R. Astron. Soc.* **247**, 510 (1990), <https://ui.adsabs.harvard.edu/abs/1990MNRAS.247..510S/abstract>.
- [3] N. Y. Gnedin and J. P. Ostriker, Reionization of the universe and the early production of metals, *Astrophys. J.* **486**, 581 (1997).
- [4] P. Madau, A. Meiksin, and M. J. Rees, 21 centimeter tomography of the intergalactic medium at high redshift, *Astrophys. J.* **475**, 429 (1997).
- [5] P. A. Shaver, R. A. Windhorst, P. Madau, and A. G. de Bruyn, Can the reionization epoch be detected as a global signature in the cosmic background?, *Astron. Astrophys.* **345**, 380 (1999), <https://ui.adsabs.harvard.edu/abs/1999A%26A...345..380S/abstract>.
- [6] P. Tozzi, P. Madau, A. Meiksin, and M. J. Rees, Radio signatures of H I at high redshift: Mapping the end of the “Dark Ages”, *Astrophys. J.* **528**, 597 (2000).
- [7] I. T. Iliev, P. R. Shapiro, A. Ferrara, and H. Martel, On the direct detectability of the cosmic dark ages: 21 centimeter emission from Minihalos, *Astrophys. J. Lett.* **572**, L123 (2002).
- [8] X. Fan, V. K. Narayanan, M. A. Strauss, R. L. White, R. H. Becker, L. Pentericci, and H.-W. Rix, Evolution of the ionizing background and the epoch of reionization from the spectra of $z \sim 6$ quasars, *Astron. J.* **123**, 1247 (2002).
- [9] N. Y. Gnedin and P. A. Shaver, Redshifted 21 centimeter emission from the pre-reionization era. I. Mean signal and linear fluctuations, *Astrophys. J.* **608**, 611 (2004).
- [10] X. Fan, C. L. Carilli, and B. Keating, Observational constraints on cosmic reionization, *Annu. Rev. Astron. Astrophys.* **44**, 415 (2006).
- [11] S. R. Furlanetto, S. P. Oh, and F. H. Briggs, Cosmology at low frequencies: The 21 cm transition and the high-redshift universe, *Phys. Rep.* **433**, 181 (2006).
- [12] R. Barkana and A. Loeb, The physics and early history of the intergalactic medium, *Rep. Prog. Phys.* **70**, 627 (2007).
- [13] M. F. Morales and J. S. B. Wyithe, Reionization and cosmology with 21-cm fluctuations, *Annu. Rev. Astron. Astrophys.* **48**, 127 (2010).
- [14] J. R. Pritchard and A. Loeb, 21 cm cosmology in the 21st century, *Rep. Prog. Phys.* **75**, 086901 (2012).
- [15] C. J. Lonsdale *et al.*, The Murchison widefield array: Design overview, *Proc. IEEE* **97**, 1497 (2009).
- [16] J. D. Bowman *et al.*, Science with the Murchison widefield array, *Pub. Astron. Soc. Aust.* **30**, e031 (2013).
- [17] S. J. Tingay *et al.*, The Murchison widefield array: The square kilometre array precursor at low radio frequencies, *Pub. Astron. Soc. Aust.* **30**, E007 (2013).
- [18] A. P. Beardsley *et al.*, Science with the Murchison widefield array: Phase I results and phase II opportunities, *Pub. Astron. Soc. Aust.* **36**, e050 (2019).
- [19] A. R. Parsons, D. C. Backer, G. S. Foster, M. C. H. Wright, R. F. Bradley, N. E. Gugliucci, C. R. Parashare, E. E. Benoit, J. E. Aguirre, D. C. Jacobs, C. L. Carilli, D. Herne, M. J. Lynch, J. R. Manley, and D. J. Werthimer, The precision array for probing the epoch of re-ionization: Eight station results, *Astron. J.* **139**, 1468 (2010).
- [20] M. P. van Haarlem *et al.*, LOFAR: The Low-Frequency Array, *Astron. Astrophys.* **556**, A2 (2013).
- [21] G. Paciga, J. G. Albert, K. Bandura, T.-C. Chang, Y. Gupta, C. Hirata, J. Odegova, U.-L. Pen, J. B. Peterson, J. Roy, J. R. Shaw, K. Sigurdson, and T. Voytek, A simulation-calibrated limit on the H I power spectrum from the GMRT epoch of reionization experiment, *Mon. Not. R. Astron. Soc.* **433**, 639 (2013).
- [22] D. R. DeBoer *et al.*, Hydrogen Epoch of Reionization Array (HERA), *Pub. Astron. Soc. Aust.* **129**, 045001 (2017).
- [23] G. Mellema *et al.*, Reionization and the cosmic dawn with the square kilometre array, *Exp. Astron.* **36**, 235 (2013).
- [24] K. Bandura *et al.*, Canadian hydrogen intensity mapping experiment (CHIME) pathfinder, *Proc. SPIE Int. Soc. Opt. Eng.* **9145**, 22 (2014).
- [25] L. B. Newburgh *et al.*, HIRAX: A probe of dark energy and radio transients, *Proc. SPIE Int. Soc. Opt. Eng.* **9906**, 96065X (2016).
- [26] A. P. Beardsley *et al.*, The EoR sensitivity of the Murchison widefield array, *Mon. Not. R. Astron. Soc.* **429**, L5 (2013).
- [27] N. Thyagarajan *et al.*, A study of fundamental limitations to statistical detection of redshifted H I from the epoch of reionization, *Astrophys. J.* **776**, 6 (2013).
- [28] A. Datta, J. D. Bowman, and C. L. Carilli, Bright source subtraction requirements for redshifted 21 cm measurements, *Astrophys. J.* **724**, 526 (2010).
- [29] N. Barry, B. Hazelton, I. Sullivan, M. F. Morales, and J. C. Pober, Calibration requirements for detecting the 21 cm epoch of reionization power spectrum and implications for the SKA, *Mon. Not. R. Astron. Soc.* **461**, 3135 (2016).
- [30] C. M. Trott and R. B. Wayth, Spectral calibration requirements of radio interferometers for epoch of reionisation science with the SKA, *Pub. Astron. Soc. Aust.* **33**, e019 (2016).
- [31] A. H. Patil *et al.*, Upper limits on the 21 cm epoch of reionization power spectrum from one night with LOFAR, *Astrophys. J.* **838**, 65 (2017).
- [32] A. Ewall-Wice, J. S. Dillon, A. Liu, and J. Hewitt, The impact of modelling errors on interferometer calibration for 21 cm power spectra, *Mon. Not. R. Astron. Soc.* **470**, 1849 (2017).
- [33] J. L. Sievers, Calibration of quasi-redundant interferometers, [arXiv:1701.01860](https://arxiv.org/abs/1701.01860).
- [34] J. S. Dillon, S. A. Kohn, A. R. Parsons, J. E. Aguirre, Z. S. Ali, G. Bernardi, N. S. Kern, W. Li, A. Liu, C. D. Nunhokee, and J. C. Pober, Polarized redundant-baseline calibration for 21 cm cosmology without adding spectral structure, *Mon. Not. R. Astron. Soc.* **477**, 5670 (2018).
- [35] N. Orosz, J. S. Dillon, A. Ewall-Wice, A. R. Parsons, and N. Thyagarajan, Mitigating the effects of antenna-to-antenna variation on redundant-baseline calibration for 21 cm cosmology, *Mon. Not. R. Astron. Soc.* **487**, 537 (2019).
- [36] N. Thyagarajan, C. L. Carilli, and B. Nikolic, Detecting Cosmic Reionization Using the Bispectrum Phase, *Phys. Rev. Lett.* **120**, 251301 (2018).
- [37] R. C. Jennison, A phase sensitive interferometer technique for the measurement of the Fourier transforms of spatial

- brightness distributions of small angular extent, *Mon. Not. R. Astron. Soc.* **118**, 276 (1958).
- [38] S. R. Kulkarni, Self-noise in interferometers: Radio and infrared, *Astron. J.* **98**, 1112 (1989).
- [39] *Synthesis Imaging in Radio Astronomy II*, edited by G. B. Taylor, C. L. Carilli, and R. A. Perley, Astronomical Society of the Pacific Conference Series, Vol. 180 (Astronomical Society of the Pacific, 1999), http://www.aspbbooks.org/a/volumes/table_of_contents/?book_id=292.
- [40] A. R. Thompson, J. M. Moran, and G. W. Swenson Jr., in *Interferometry and Synthesis in Radio Astronomy*, 2nd ed., edited by A. Richard Thompson, James M. Moran, and George W. Swenson Jr. (Wiley, New York, 2001), p. 692.
- [41] J. D. Monnier, Phases in interferometry, *New A Rev.* **51**, 604 (2007).
- [42] J. D. Monnier, J.-P. Berger, R. Millan-Gabet, W. A. Traub, F. P. Schloerb, E. Pedretti, M. Benisty, N. P. Carleton, P. Hagenauer, P. Kern, P. Labeye, M. G. Lacasse, F. Malbet, K. Perraut, M. Pearlman, and M. Zhao, Few skewed disks found in first closure-phase survey of Herbig Ae/Be stars, *Astrophys. J.* **647**, 444 (2006).
- [43] J. D. Monnier, M. Zhao, E. Pedretti, N. Thureau, M. Ireland, P. Muirhead, J. P. Berger, R. Millan-Gabet, G. Van Belle, T. ten Brummelaar, H. McAlister, S. Ridgway, N. Turner, L. Sturmman, J. Sturmman, and D. Berger, Imaging the surface of Altair, *Science* **317**, 342 (2007).
- [44] K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay, A.-K. Baczko, D. Ball, M. Baloković, J. Barrett *et al.* (Event Horizon Telescope Collaboration), First M87 Event Horizon Telescope results. I. The shadow of the supermassive black hole, *Astrophys. J. Lett.* **875**, L1 (2019).
- [45] K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay, A.-K. Baczko, D. Ball, M. Baloković, J. Barrett *et al.* (Event Horizon Telescope Collaboration), First M87 Event Horizon Telescope results. II. Array and instrumentation, *Astrophys. J. Lett.* **875**, L2 (2019).
- [46] K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay, A.-K. Baczko, D. Ball, M. Baloković, J. Barrett *et al.* (Event Horizon Telescope Collaboration), First M87 Event Horizon Telescope results. III. Data processing and calibration, *Astrophys. J. Lett.* **875**, L3 (2019).
- [47] K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay, A.-K. Baczko, D. Ball, M. Baloković, J. Barrett *et al.* (Event Horizon Telescope Collaboration), First M87 Event Horizon Telescope results. IV. Imaging the central supermassive black hole, *Astrophys. J. Lett.* **875**, L4 (2019).
- [48] K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay, A.-K. Baczko, D. Ball, M. Baloković, J. Barrett *et al.* (Event Horizon Telescope Collaboration), First M87 Event Horizon Telescope results. V. Physical origin of the asymmetric ring, *Astrophys. J. Lett.* **875**, L5 (2019).
- [49] K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay, A.-K. Baczko, D. Ball, M. Baloković, J. Barrett *et al.*, First M87 Event Horizon Telescope results. VI. The shadow and mass of the central black hole, *Astrophys. J. Lett.* **875**, L6 (2019).
- [50] C. L. Carilli, B. Nikolic, N. Thyagarayan, and K. Gale-Sides, H I 21-cm cosmology and the bispectrum: Closure diagnostics in massively redundant interferometric arrays, *Radio Sci.* **53**, 845 (2018).
- [51] C. L. Carilli *et al.*, Imaging and modeling data from the hydrogen epoch of reionization array, *Astrophys. J. Supp. Ser.* **247**, 67 (2020).
- [52] N. Thyagarajan, C. Carilli, B. Nikolic, J. Kent, A. Mesinger, N. Kern, G. Bernardi, and S. Matika (HERA Collaboration), following paper, Detection of cosmic structures using bispectrum phase. II. First results from application to cosmic reionization using the hydrogen epoch of reionization array, *Phys. Rev. D* **102**, 022002 (2020).
- [53] P. C. Crane and P. J. Napier, Sensitivity, in *Synthesis Imaging in Radio Astronomy*, Astronomical Society of the Pacific Conference Series, Vol. 6, edited by R. A. Perley, F. R. Schwab, and A. H. Bridle (Astronomical Society of the Pacific, 1989), p. 139, http://www.aspbbooks.org/a/volumes/article_details/?paper_id=5980.
- [54] R. M. Goldstein, H. A. Zebker, and C. L. Werner, Satellite radar interferometry: Two-dimensional phase unwrapping, *Radio Sci.* **23**, 713 (1988).
- [55] M. A. Schofield and Y. Zhu, Fast phase unwrapping algorithm for interferometric applications, *Opt. Lett.* **28**, 1194 (2003).
- [56] V. V. Volkov and Y. Zhu, Deterministic phase unwrapping in the presence of noise, *Opt. Lett.* **28**, 2156 (2003).
- [57] V. Katkovnik, J. Astola, and K. Egiazarian, Phase local approximation (PhaseLA) technique for phase unwrap from noisy data, *IEEE Trans. Image Process.* **17**, 833 (2008).
- [58] I. Shevkunov, A new phase unwrapping method, *J. Phys. Conf. Ser.* **737**, 012065 (2016).
- [59] A. Parsons, J. Pober, M. McQuinn, D. Jacobs, and J. Aguirre, A sensitivity and array-configuration study for measuring the power spectrum of 21 cm emission from reionization, *Astrophys. J.* **753**, 81 (2012).
- [60] A. R. Parsons, J. C. Pober, J. E. Aguirre, C. L. Carilli, D. C. Jacobs, and D. F. Moore, A per-baseline, delay-spectrum technique for accessing the 21 cm cosmic reionization signature, *Astrophys. J.* **756**, 165 (2012).
- [61] N. Thyagarajan, A. R. Parsons, D. R. DeBoer, J. D. Bowman, A. M. Ewall-Wice, A. R. Neben, and N. Patra, Effects of antenna beam chromaticity on redshifted 21 cm power spectrum and implications for hydrogen epoch of reionization array, *Astrophys. J.* **825**, 9 (2016).
- [62] N. Thyagarajan *et al.*, Foregrounds in wide-field redshifted 21 cm power spectra, *Astrophys. J.* **804**, 14 (2015).
- [63] P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett *et al.* (Planck Collaboration), Planck 2015 results. XIII. Cosmological parameters, *Astron. Astrophys.* **594**, A13 (2016).
- [64] N. Hurley-Walker *et al.*, Galactic and Extragalactic All-Sky Murchison widefield array (GLEAM) survey. I. A low-frequency extragalactic catalogue, *Mon. Not. R. Astron. Soc.* **464**, 1146 (2017).
- [65] A. Mesinger, B. Greig, and E. Sobacchi, The evolution of 21 cm structure (EOS): Public, large-scale simulations of cosmic dawn and reionization, *Mon. Not. R. Astron. Soc.* **459**, 2342 (2016).
- [66] B. Greig and A. Mesinger, Simultaneously constraining the astrophysics of reionization and the epoch of heating with 21CMMC, *Mon. Not. R. Astron. Soc.* **472**, 2651 (2017).

- [67] A. Mesinger, S. Furlanetto, and R. Cen, 21CMFAST: A fast, seminumerical simulation of the high-redshift 21-cm signal, *Mon. Not. R. Astron. Soc.* **411**, 955 (2011).
- [68] N. Thyagarajan, S. Harish, M. Kolopanis, S. Murray, and D. Jacobs, PRISim v2.2.1 (Zenodo, 2020), <https://doi.org/10.5281/zenodo.3892099>.
- [69] J. D. Bowman, M. F. Morales, and J. N. Hewitt, Foreground contamination in interferometric measurements of the redshifted 21 cm power spectrum, *Astrophys. J.* **695**, 183 (2009).
- [70] A. Liu, M. Tegmark, J. Bowman, J. Hewitt, and M. Zaldarriaga, An improved method for 21-cm foreground removal, *Mon. Not. R. Astron. Soc.* **398**, 401 (2009).
- [71] A. Liu, A. R. Parsons, and C. M. Trott, Epoch of reionization window. I. Mathematical formalism, *Phys. Rev. D* **90**, 023018 (2014).
- [72] A. Liu, A. R. Parsons, and C. M. Trott, Epoch of reionization window. II. Statistical methods for foreground wedge reduction, *Phys. Rev. D* **90**, 023019 (2014).
- [73] A. Liu and M. Tegmark, A method for 21 cm power spectrum estimation in the presence of foregrounds, *Phys. Rev. D* **83**, 103006 (2011).
- [74] A. Ghosh, J. Prasad, S. Bharadwaj, S. S. Ali, and J. N. Chengalur, Characterizing foreground for redshifted 21 cm radiation: 150 MHz giant metrowave radio telescope observations, *Mon. Not. R. Astron. Soc.* **426**, 3295 (2012).
- [75] M. F. Morales, B. Hazelton, I. Sullivan, and A. Beardsley, Four fundamental foreground power spectrum shapes for 21 cm cosmology observations, *Astrophys. J.* **752**, 137 (2012).
- [76] C. M. Trott, R. B. Wayth, and S. J. Tingay, The impact of point-source subtraction residuals on 21 cm epoch of reionization estimation, *Astrophys. J.* **757**, 101 (2012).
- [77] H. Vedantham, N. Udaya Shankar, and R. Subrahmanyam, Imaging the epoch of reionization: Limitations from foreground confusion and imaging algorithms, *Astrophys. J.* **745**, 176 (2012).
- [78] J. S. Dillon, A. Liu, and M. Tegmark, A fast method for power spectrum and foreground analysis for 21 cm cosmology, *Phys. Rev. D* **87**, 043005 (2013).
- [79] J. C. Pober, A. R. Parsons, J. E. Aguirre, Z. Ali, R. F. Bradley, C. L. Carilli, D. DeBoer, M. Dexter, N. E. Gugliucci, D. C. Jacobs, P. J. Klima, D. MacMahon, J. Manley, D. F. Moore, I. I. Stefan, and W. P. Walbrugh, Opening the 21 cm epoch of reionization window: Measurements of foreground isolation with PAPER, *Astrophys. J. Lett.* **768**, L36 (2013).
- [80] J. S. Dillon, A. Liu, C. L. Williams, J. N. Hewitt, M. Tegmark, E. H. Morgan, A. M. Levine, M. F. Morales, S. J. Tingay, G. Bernardi, J. D. Bowman, F. H. Briggs, R. C. Cappallo, D. Emrich, D. A. Mitchell, D. Oberoi, T. Prabu, R. Wayth, and R. L. Webster, Overcoming real-world obstacles in 21 cm power spectrum estimation: A method demonstration and results from early Murchison widefield array data, *Phys. Rev. D* **89**, 023002 (2014).
- [81] N. Thyagarajan *et al.*, Confirmation of wide-field signatures in redshifted 21 cm power spectra, *Astrophys. J. Lett.* **807**, L28 (2015).