



# Array element coupling in radio interferometry I: a semi-analytic approach

Alec T. Josaitis<sup>1</sup>,<sup>1</sup>★ Aaron Ewall-Wice<sup>2</sup>,<sup>2</sup>★ Nicolas Fagnoni<sup>1</sup> and Eloy de Lera Acedo<sup>1,3</sup>

<sup>1</sup>*Astrophysics Group - Cavendish Laboratory, University of Cambridge, JJ Thompson Avenue, Cambridge CB3 0HE, UK*

<sup>2</sup>*Astronomy Department, University of California - Berkeley, 425 Campbell Hall, University Dr., Berkeley 94720, USA*

<sup>3</sup>*Kavli Institute for Cosmology, Madingley Road, Cambridge CB3 0HA, UK*

Accepted 2022 March 8. Received 2022 February 9; in original form 2021 October 31

## ABSTRACT

We derive a general formalism for interferometric visibilities, which considers first-order antenna–antenna coupling and assumes steady-state, incident radiation. We simulate such coupling features for non-polarized skies on a compact, redundantly spaced array and present a phenomenological analysis of the coupling features. Contrary to previous studies, we find mutual coupling features manifest themselves at non-zero fringe rates. We compare power-spectrum results for both coupled and non-coupled (noiseless, simulated) data and find coupling effects to be highly dependent on local sidereal time (LST), baseline length, and baseline orientation. For all LSTs, lengths, and orientations, coupling features appear at delays which are outside the foreground ‘wedge’, which has been studied extensively and contains non-coupled astrophysical foreground features. Further, we find that first-order coupling effects threaten our ability to average data from baselines with identical length and orientation. Two filtering strategies are proposed which may mitigate such coupling systematics. The semi-analytic coupling model herein presented may be used to study mutual coupling systematics as a function of LST, baseline length, and baseline orientation. Such a model is not only helpful to the field of 21cm cosmology, but any study involving interferometric measurements, where coupling effects at the level of at least 1 part in  $10^4$  could corrupt the scientific result. Our model may be used to mitigate coupling systematics in existing radio interferometers and to design future arrays where the configuration of array elements inherently mitigates coupling effects at desired LSTs and angular resolutions.

**Key words:** (*cosmology*:) dark ages, reionization, first stars – scattering – techniques: interferometric – techniques: radar astronomy.

## 1 INTRODUCTION

Radio interferometers are key instruments for the fields of cosmology, astrophysics, astronomy, and ionospheric physics. The study of the 21cm emission of the hyperfine splitting of the ground state of neutral hydrogen in the intergalactic medium (IGM) is an active area of research with implications for all such disciplines. Van de Hulst (1945), Madau, Meiksin & Rees (1997), and Wang et al. (2006) describe how a measurement of the 21cm signal probes the era of the formation of the first luminous objects (the ‘Cosmic Dawn’) and the subsequent phase transition of the Universe, during which photons from the first stars ionized the neutral hydrogen of the IGM at  $z \approx 6$  [the ‘Epoch of Reionization’ (EoR)]. Reviews such as Furlanetto, Oh & Briggs (2006), Pritchard & Loeb (2012), Mesinger (2016), and Liu & Shaw (2020) carefully consider various experimental and theoretical considerations regarding a measurement of the 21cm signal, which has since been redshifted to the VHF band.

Considering that the 21cm signal is expected to be many orders of magnitude fainter than astrophysical foregrounds – not to mention the difficulties of terrestrial foregrounds, such as VHF radio frequency interference – it is imperative that any experiment carefully mitigates

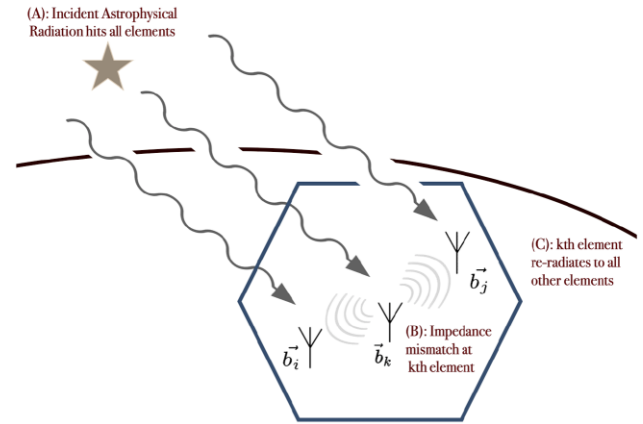
instrumental systematic effects. These systematics result from a variety of sources, such as Faraday rotation due to the ionosphere, impedance mismatches throughout the analogue signal chain, mischaracterization of the primary beam and/or beam sidelobes, among others. Any such systematic effect could distort the spectral structure of either the foregrounds or the EoR signal, making it more difficult to discriminate the latter from the former. Jelić et al. (2010) simulate how Faraday-rotated, polarized galactic emission can affect the calibration of a radio interferometer and contaminate measurements of the 21cm signal. Moore et al. (2013) discuss how this polarized contamination of the intrinsically unpolarized 21cm power spectrum can occur at levels which are several orders of magnitude above the expected 21cm EoR signal. Martinot et al. (2018) describe how such leakage effects are attenuated by at least an order of magnitude due to Faraday rotation from the Earth’s ionosphere, however the modelling and removal of polarized sources required to eliminate polarization leakage exceeds the reasonable capabilities of current 21cm instruments. Not only leakage from polarized foregrounds, but any form of uncalibrated structure in the instrument’s frequency response may be imprinted on the foregrounds and has the potential to leak power into the EoR window at small line-of-sight scales, masking the signal. Zheng et al. (2014) discusses how phase-switching hardware may be used to mitigate circuit-level cross-talk in the analogue signal chains of a 21cm experiment. Hibbard

\* E-mail: [atj28@cam.ac.uk](mailto:atj28@cam.ac.uk) (ATJ); [aaronew@berkeley.edu](mailto:aaronew@berkeley.edu) (AEW)

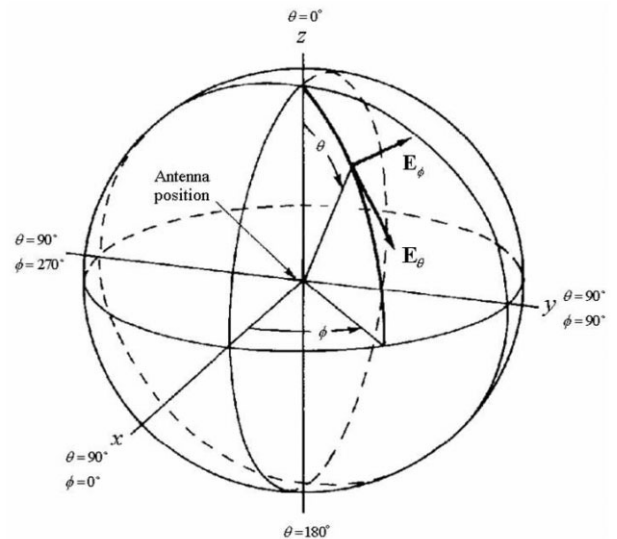
et al. (2020) and Thyagarajan et al. (2016) discuss the challenges that antenna beam chromaticity poses to measurements of both the spatially averaged ('global') 21cm signal and the angular-resolution-dependent (interferometric) 21cm power spectrum, respectively. Visbal, Loeb & Wyithe (2009) analyse how a measurement of the power spectrum across a wide range of redshifts (and, analogously, a wide frequency band) offers the scientific community an understanding of how the EoR and Cosmic Dawn evolve over time, and offers independent and significantly improved constraints on many fundamental, time-independent cosmological parameters which govern the evolution of the early universe. Wideband antenna elements are likely required for such measurements. Unfortunately, Craeye & González-Ovejero (2011) conclude that wider bandwidth antennae require larger element volumes over which to electromagnetically characterize the effective number of degrees of freedom of an antenna element; this scaling challenge makes beam chromaticity inherently more difficult to model through electromagnetic simulation. Furthermore, Craeye & González-Ovejero (2011) discuss how such a wideband system typically permits a greater number of current modes inside the antenna, which yields greater potential for chromatic variation between array elements. The wideband requirements of a 21cm measurement make the challenges of beam chromaticity inherently more difficult to electromagnetically model and constrain during the mechanical construction of a large array.

A conclusive detection of neither the global 21cm signal nor the 21cm power spectrum has yet to occur. Bowman et al. (2018) describe a potential first detection of the global signal by the EDGES collaboration, claiming to have measured an absorption trough centred at 78 MHz. Subsequent examination of the data, such as that performed in Hills et al. (2018), indicates that the EDGES result may be corrupted by unaccounted systematic errors, which result in unphysical parameters for galactic foreground emission, and finds simple, alternative formulations for the signal which equally fit the EDGES data. Sims & Pober (2020) show that a damped sinusoidal systematic is strongly preferred in the EDGES data. More recently, Bevins et al. (2021) identify a similar systematic. First-generation radio interferometers, such as the Precision Array for Probing the Epoch of Reionization (PAPER; Parsons et al. 2012a), the Murchison Widefield Array (Bowman et al. 2013), and the Low Frequency Array (van Haarlem et al. 2013) have made noteworthy upper limits on measurements of the full 21cm power spectrum. Second-generation radio interferometers, such as the Hydrogen Epoch of Reionization Array (HERA; DeBoer et al. 2017) are currently being commissioned, and are designed to have the sensitivity to detect the 21cm power spectrum across a wide range of redshifts.

This work focuses on first-order antenna–antenna coupling in radio interferometers, a form of 'mutual coupling' which affects interferometric measurements of the 21cm power spectrum. Specifically, we derive a semi-analytic model of the interferometric visibility equation which considers this first-order coupling. In this model, incident astrophysical radiation enters all interferometric elements in the array. Each element absorbs most of its incident radiation, although some amount of power is then reflected due to departure from conjugate match at the terminals of the antenna feeds. Thus, each element in the array not only absorbs power from the sky but reradiates power across the array, which is subsequently absorbed by other elements. Being a first-order formalism, we only model the effect of reradiated power being absorbed just once more by all array elements (e.g. sky  $\rightarrow$  antenna  $k \rightarrow$  antenna  $i$ ), not when the scattered power is subsequently absorbed and reradiated again by all elements (e.g. sky  $\rightarrow$  antenna  $k \rightarrow$  antenna  $i \rightarrow$  antenna  $j$ ). Fig. 1 provides a



**Figure 1.** Conceptual outline of first-order array element coupling: Incident astrophysical radiation reaches all elements in the interferometric array. Each element in the array, being imperfectly impedance matched at its feed terminals, will reflect some amount of this incident radiation, and reradiate power across the array, which will be absorbed by another interferometric element. Being a first-order formalism, we only study the effect of this reradiated power being absorbed by other array elements just once (e.g. sky  $\rightarrow$  antenna  $k \rightarrow$  antenna  $i$ ), not when that scattered power is subsequently reradiated by all other elements (e.g. sky  $\rightarrow$  antenna  $k \rightarrow$  antenna  $i \rightarrow$  antenna  $j$ ).



**Figure 2.** Spherical coordinate system, which we adopt as the basis to represent the components of our electric field terms.

simplified, conceptual sketch of our first-order coupling model.

Mutual coupling systematics is a known concern for interferometric 21cm experiments (Parsons et al. 2012b; Chaudhari et al. 2017; Kern et al. 2019). Upper limits from all first-generation interferometers, and current limits from the second-generation HERA (which has yet to achieve its full sensitivity) report excess noise in their power-spectrum measurements with a spectral correlation which is distinct from thermal noise. Analysis of all of these upper limits cite mutual coupling as one unresolved possibility for their excess power (Ali et al. 2015; Beardsley et al. 2016; Ewall-Wice et al. 2016; Mertens et al. 2020; The HERA Collaboration et al. 2022). Ung et al. (2020) discuss the non-trivial relationship between the mutual coupling of low-noise amplifiers and receiver noise temperature,





component of the electric field surface distribution which was reradiated from the  $k^{\text{th}}$  element may contain components from both instrumental polarizations and both incident electric field basis vectors of the  $k^{\text{th}}$  element. Similarly, the voltage response at the terminal of the  $q^{\text{th}}$  pol of the  $j^{\text{th}}$  interferometric element, caused by the  $\hat{\theta}$  basis vector component of the electric field surface distribution which was reradiated from the  $k^{\text{th}}$  element may contain components from both instrumental polarizations and both incident electric field basis vectors of the  $k^{\text{th}}$  element. We formalize this model of antenna-to-antenna coupling in Appendix A by replacing equations (1) and (2) with the voltage responses presented in equations (A10) and (A11), respectively.

The raw data product of our interferometer is the correlation of the voltage responses from any two elements in the array,  $\langle v_i v_j^* \rangle$ , which is referred to as a ‘visibility’,  $V_{ij}$ . If the two elements are the same ( $i = j$ ), the visibility is the product of an autocorrelation. If the two elements are not the same ( $i \neq j$ ), the visibility is the product of a cross-correlation. In either case, the correlation data product is useful because it exploits the spatial coherence of the astrophysical electric field components at two different test points,  $\langle E^\phi(\hat{s}) E^{\theta*}(\hat{s}') \rangle$ . In the next section, we will use equations (A10) and (A11) to calculate the visibility, accounting for first-order antenna–antenna coupling.

## 2.2 General first-order visibility formalism

The voltage responses presented in equations (A10) and (A11) are expressed as functions of feed polarization ( $p, q$ ), interferometric elements ( $i, j$ ), and electric field basis vector components ( $\phi, \theta$ ) which are tangent to the celestial sphere. These expressions can be rewritten in a matrix formalism, to represent all polarizations and basis vector components of the visibility  $\mathbf{V}_{ij}^{0,1}$ , whether to zeroth-order in coupling (no-coupling) or including first-order antenna-to-antenna scattering. Let us define voltage vectors  $\vec{v}_i$  and  $\vec{v}_j$  to characterize, per interferometric element, the induced voltage at the terminal of both feed polarizations.

$$\vec{v}_i = \begin{bmatrix} v_i^p \\ v_i^q \end{bmatrix}, \vec{v}_j = \begin{bmatrix} v_j^p \\ v_j^q \end{bmatrix}. \quad (3)$$

Letting ‘†’ represent the conjugate-transpose, the visibility matrix may be defined as

$$\mathbf{V}_{ij}^{0,1}(v) = \begin{bmatrix} V_{ij}^{pp} & V_{ij}^{pq} \\ V_{ij}^{qp} & V_{ij}^{qq} \end{bmatrix}^{0,1} = \begin{bmatrix} \langle v_i^p v_j^{*p} \rangle & \langle v_i^p v_j^{*q} \rangle \\ \langle v_i^q v_j^{*p} \rangle & \langle v_i^q v_j^{*q} \rangle \end{bmatrix}^{0,1} \quad (4)$$

$$= \langle \vec{v}_i \times \vec{v}_j^* \rangle^{0,1}.$$

Similarly, components of electric field surface distributions (for both incident astrophysical radiation and reradiation for first-order coupling) can be represented as vectors,

$$\vec{\epsilon} = \begin{bmatrix} \epsilon_\phi \\ \epsilon_\theta \end{bmatrix}, \vec{\epsilon}^s = \begin{bmatrix} \epsilon_\phi^s \\ \epsilon_\theta^s \end{bmatrix}. \quad (5)$$

Since the effective height,  $h_p^{\phi}(\hat{s}, \nu)$  is written in terms of feed polarization and electric field basis vectors, it, too, can easily be rewritten in matrix formalism. Following the unit convention in experimental astrophysics (which may differ from optics, or other subdisciplines), we define a ‘Jones Matrix’,  $\mathbf{J}$  with units of length to represent the effective height, where

$$v = \begin{bmatrix} v_p \\ v_q \end{bmatrix} = \mathbf{J} \vec{\epsilon} = \begin{bmatrix} h_p^{\phi} & h_p^{\theta} \\ h_q^{\phi} & h_q^{\theta} \end{bmatrix} \begin{bmatrix} \epsilon_\phi \\ \epsilon_\theta \end{bmatrix}. \quad (6)$$

Using equations (A10) and (A11) for  $\vec{v}_i$  and  $\vec{v}_j$ , respectively, We calculate the first-order visibility terms,  $\langle \vec{v}_i \times \vec{v}_j^* \rangle^1$ , in matrix

formalism.

$$\mathbf{V}_{ij}^1 = \int \int \left( (\mathbf{J}_i(\hat{s}) \vec{\epsilon}(\hat{s}) \times (\mathbf{J}_j(\hat{s}'') \vec{\epsilon}(\hat{s}''))^y \right. \\ \left. \frac{e^{2\pi i \frac{\nu}{c} (R\hat{s} - \vec{b}_i - R\hat{s}'' + \vec{b}_j)}}{|R\hat{s} - \vec{b}_i| |R\hat{s}'' - \vec{b}_j|} R^4 d\Omega d\Omega'' \right) \\ + \sum_{k \neq i} \int \int \left( (\mathbf{J}_i(\hat{b}_{ik}) \vec{\epsilon}_{ki}^s(\hat{s}', \hat{b}_{ki}) \times (\mathbf{J}_j(\hat{s}'') \vec{\epsilon}(\hat{s}''))^y \right. \\ \left. \frac{e^{2\pi i \frac{\nu}{c} (R\hat{s}' - \vec{b}_k - R\hat{s}'' + \vec{b}_j)}}{|R\hat{s}' - \vec{b}_k| |R\hat{s}'' - \vec{b}_j|} R^4 d\Omega' d\Omega'' \right) \\ + \sum_{k \neq j} \int \int \left( (\mathbf{J}_i(\hat{s}) \vec{\epsilon}(\hat{s}) \times (\mathbf{J}_j(\hat{b}_{jk}) \vec{\epsilon}_{kj}^s(\hat{s}', \hat{b}_{kj}))^y \right. \\ \left. \frac{e^{2\pi i \frac{\nu}{c} (R\hat{s} - \vec{b}_i - R\hat{s}' + \vec{b}_k)}}{|R\hat{s} - \vec{b}_i| |R\hat{s}' - \vec{b}_k|} R^4 d\Omega d\Omega' \right). \quad (7)$$

Note that equations (12) and (13) from Kern et al. (2019) are special cases of our equation (7) which only include the reflection from antenna  $j$  to  $i$  (not other elements,  $k$ ). The Kern et al. (2019) coupling model produces autocorrelation coupling terms in visibility, since  $b_{ji} = 0$ , but does not consider cross-correlation terms which are present in equation (7). By Assumption A (stated clearly in Appendix A, we can simplify equation (7), for  $|R\hat{s} - \vec{b}_i| \approx R - \vec{b}_i \cdot \hat{s}$ . This removes all terms of  $R$  from the exponential factors in equation (7). The denominators in the same equation reduce further. For example,  $|R\hat{s} - \vec{b}_i| |R\hat{s}'' - \vec{b}_j| \approx R^2$ , regardless of if we are integrating in the differential direction  $\hat{s}$  or  $\hat{s}''$ .

Finally, we note that the expectation values of each term of equation (7) may be rewritten solely in terms of the expectation value of the electric field surface distributions. For example, in the  $\sum_{k \neq i}$  term:

$$\langle \mathbf{J}_i(\hat{b}_{ik}) \vec{\epsilon}_{ki}^s(\hat{s}', \hat{b}_{ki}) \times (\mathbf{J}_j(\hat{s}'') \vec{\epsilon}(\hat{s}''))^y = \\ \mathbf{J}_i(\hat{b}_{ik}) (\vec{\epsilon}_{ki}^s(\hat{s}', \hat{b}_{ki}) \times \vec{\epsilon}(\hat{s}'')^y) \mathbf{J}_j(\hat{s}'')^y. \quad (8)$$

This latter form will be useful for simplifying equation (7) in the forthcoming two sections.

## 2.3 Characterizing the scattered electric field

To simplify (7), we start by examining the scattered electric field terms associated with reradiation of the incident astrophysical radiation. Standard antenna scattering formalisms, such as Jenn & Flokas (1994), characterize the total electric field scattered from the  $p^{\text{th}}$  feed polarization of a single  $k^{\text{th}}$  element in a large array due to impedance mismatch as

$$\vec{E}_{ki}^{s,p}(\hat{s}, \hat{b}_{ki}) = \frac{i\eta_0}{4\lambda R_k} \vec{h}_k^p(\hat{b}_{ki}) (\vec{h}_k^p(\hat{s}) \cdot \vec{E}(\hat{s})) \frac{e^{-2\pi i \frac{\nu}{c} |\vec{b}_{ki}|}}{|\vec{b}_{ki}|} \Gamma_k, \quad (9)$$

where  $\Gamma_k(\nu)$  is the frequency-dependent reflection coefficient of the  $k^{\text{th}}$  feed,  $\eta_0$  is the impedance of free space, and  $R_k$  is the real part of the impedance of the  $k^{\text{th}}$  antenna. This expression assumes that the scattered electric field may be approximated by the far-field limit when it is received by another interferometric element.

The formalism of equation (9) assumes a spherical wave propagation convention of  $e^{-2\pi i \frac{\nu}{c} |\vec{b}_{ki}|}$ . To make this formalism consistent with our chosen convention (see equation A2 in Appendix A), we flip the direction of the complex exponential and integrate over the appropriate electric field surface distribution (equation A3), which involves an integral of the general form  $\int e^{2\pi i \frac{\nu}{c} \vec{s} \cdot \vec{s}} d\vec{s} = \frac{-i}{2\pi} e^{2\pi i \frac{\nu}{c} \vec{s}} + c$ , for some arbitrary constant  $c$ . Using a matrix formalism to accounting

for all instrumental polarizations,

$$\vec{\epsilon}_{ki}^{s,p}(\hat{s}, \hat{b}_{ki}) = \frac{-i\eta_0}{4\lambda R_k} \mathbf{J}_k(\hat{b}_{ki})(\mathbf{J}_k(\hat{s}) \cdot \vec{\epsilon}(\hat{s})) \frac{e^{2\pi i \frac{v}{c} |\hat{b}_{ki}|}}{|\hat{b}_{ki}|} \Gamma_k. \quad (10)$$

It is important to note that our formalism does not consider all reradiation reflections which have occurred in the array since the initial time that an incident wavefront reached the  $i^{\text{th}}$  (equation A10) or  $j^{\text{th}}$  (equation A11) antennae, respectively. Equation (10) only considers how an incident wavefront gets reradiated by each element one time. For our formalism to meaningfully describe the scattering in the array, we make a third assumption, Assumption C that the incident astrophysical radiation changes at time-scales much longer than the time-scale of reflections in the array (or at least those which have an amplitude that could meaningfully induce a terminal voltage above the noise floor of our analogue chain). For most purposes, we consider the sky to not have ‘moved’ over the array during a visibility measurement integrated over the order of minutes, and reflections in the array to have occurred up to thousands of nanoseconds. Thus, the steady-state assumptions of our formalism are reasonable, and give us the opportunity to, in a closed and semi-analytic form, approximate first-order antenna-to-antenna coupling.

Inserting equation (10) into the expectation values of equation (7) as rewritten in the form of equation (8), we notice the per-basis-vector expectation quantities in both coupling terms, namely  $\langle \epsilon^{s,p}(\hat{s}', \hat{b}_{ki}) \epsilon^{*q}(\hat{s}'') \rangle$  and  $\langle \epsilon^{\phi}(\hat{s}) \epsilon^{*,q\theta}(\hat{s}', \hat{b}_{kj}) \rangle$ , can be rewritten in terms of effective heights multiplied by the same expectation found in the zeroth-order term of equation (7), e.g.  $\langle \epsilon^{\phi}(\hat{s}) \epsilon^{*q}(\hat{s}'') \rangle$ . For example, in the  $\sum_{k \neq i}$  term,

$$\begin{aligned} & \mathbf{J}_i(\hat{b}_{ik})(\vec{\epsilon}_{ki}^s(\hat{s}', \hat{b}_{ki}) \times \vec{\epsilon}(\hat{s}'')^y) \mathbf{J}_j(\hat{s}'')^y = \\ & \frac{-i\eta_0}{4\lambda R_k} \frac{e^{2\pi i \frac{v}{c} |\hat{b}_{ki}|}}{|\hat{b}_{ki}|} \Gamma_k \mathbf{J}_i(\hat{b}_{ik}) \mathbf{J}_k(\hat{b}_{ki}) \mathbf{J}_k(\hat{s}') \langle \epsilon^{\phi}(\hat{s}') \epsilon^{*q}(\hat{s}'') \rangle \mathbf{J}_j(\hat{s}'')^y. \end{aligned} \quad (11)$$

## 2.4 Full Stokes, first-order visibility formalism, with summary of key assumptions

We may simplify the zeroth-order terms in the form of equation (11) by making a fourth assumption, Assumption D, that the radiation from astrophysical sources is not spatially coherent. Standard derivations of visibility equations also make this assumption (e.g. Clark 1999), which relies on the Van Cittert–Zernike theorem. Physically, we may interpret Assumption D as saying that the electromagnetic radiation from distant astrophysical sources, at the frequencies over which we calculate visibilities, is the result of random, thermal processes, such as synchrotron radiation. Mathematically, Assumption D asserts that  $\langle \epsilon^{\phi}(\hat{s}) \epsilon^{*q}(\hat{s}') \rangle = 0$ , for  $\hat{s} \neq \hat{s}'$ .

There are three noteworthy simplifications in equation (7) due to Assumption D. First, the effective delay in the exponential terms is based only on the distance between baselines, for example  $\vec{b}_j - \vec{b}_i = \vec{b}_{ij}$ . Second, there is no more radial dependence on  $\vec{R}$  in equation (7). Third, expectation values may be rewritten such that we only need to integrate over one solid angle, not two.

$$\langle \epsilon^{\phi}(\hat{s}) \epsilon^{*q}(\hat{s}') \rangle = \frac{C^{\phi q}(\hat{s}) \delta_D(\hat{s} - \hat{s}')}{|\vec{R}|^2}. \quad (12)$$

The spatial coherence of the electric field may be represented as a ‘coherency matrix’ with the same relationship to electric field surface distributions found in equation (12):

$$\langle \epsilon(\hat{s}) \times \epsilon^y(\hat{s}') \rangle = \frac{\mathbf{C}(\hat{s}) \delta_D(\hat{s} - \hat{s}')}{|\vec{R}|^2}, \quad (13)$$

$$\mathbf{C}(\hat{s}) = \begin{bmatrix} \langle \epsilon_{\phi}(\hat{s}) \epsilon_{\phi}^*(\hat{s}) \rangle & \langle \epsilon_{\phi}(\hat{s}) \epsilon_{\theta}^*(\hat{s}) \rangle \\ \langle \epsilon_{\theta}(\hat{s}) \epsilon_{\phi}^*(\hat{s}) \rangle & \langle \epsilon_{\theta}(\hat{s}) \epsilon_{\theta}^*(\hat{s}) \rangle \end{bmatrix}. \quad (14)$$

Components of the coherency matrix equation (14) are represented in terms of Stokes parameters in various references in the literature. In Thompson, Moran & Swenson (2017),  $\langle |\epsilon_{\phi}|^2 \rangle = \frac{I+Q}{2}$  and  $\langle |\epsilon_{\theta}|^2 \rangle = \frac{I-Q}{2}$ . A full-Stokes representation of the coherency matrix is discussed in Smirnov (2011):

$$\mathbf{C}(\hat{s}) = \begin{bmatrix} I + Q & U + iV \\ U - iV & I - Q \end{bmatrix}. \quad (15)$$

Representing the phase difference between  $\epsilon_{\phi}$  and  $\epsilon_{\theta}$  as  $\delta$ , Smirnov’s formalism is consistent with Thompson and Swenson’s when

$$\begin{aligned} I &= \frac{1}{2}(\langle |\epsilon_{\phi}|^2 \rangle + \langle |\epsilon_{\theta}|^2 \rangle) & Q &= \frac{1}{2}(\langle |\epsilon_{\phi}|^2 \rangle - \langle |\epsilon_{\theta}|^2 \rangle) \\ U &= \langle |\epsilon_{\phi}| |\epsilon_{\theta}| \cos \delta \rangle & V &= \langle |\epsilon_{\phi}| |\epsilon_{\theta}| \sin \delta \rangle \end{aligned} \quad (16)$$

Inserting equations (10) and (12) into equation (7), and recalling the first two simplifications resulting from Assumption D, equation (7) may be rewritten as

$$\begin{aligned} \mathbf{V}_{ij}^1 &= \int \mathbf{J}_i(\hat{s})(\vec{\epsilon}(\hat{s}) \times \vec{\epsilon}^y(\hat{s})) \mathbf{J}_j^y(\hat{s}) e^{2\pi i \frac{v}{c} \vec{b}_{ij} \cdot \hat{s}} d\Omega \\ & - \sum_{k \neq i} \left\{ \frac{i\eta_0 \Gamma_k}{4\lambda R_k |\hat{b}_{ki}|} e^{2\pi i \frac{v}{c} |\hat{b}_{ki}|} \mathbf{J}_i(\hat{b}_{ki}) \mathbf{J}_k(\hat{b}_{ki}) \right. \\ & \left. \left( \int \mathbf{J}_k(\hat{s}) \langle \vec{\epsilon}(\hat{s}) \times \vec{\epsilon}^y(\hat{s}) \rangle \mathbf{J}_j^y(\hat{s}) e^{2\pi i \frac{v}{c} \vec{b}_{kj} \cdot \hat{s}} d\Omega \right) \right\} \\ & + \sum_{k \neq j} \left\{ \left( \int \mathbf{J}_i(\hat{s}) \langle \vec{\epsilon}(\hat{s}) \times \vec{\epsilon}^y(\hat{s}) \rangle \mathbf{J}_k^y(\hat{s}) e^{2\pi i \frac{v}{c} \vec{b}_{ik} \cdot \hat{s}} d\Omega \right) \right. \\ & \left. \frac{i\eta_0 \Gamma_k^*}{4\lambda R_k |\hat{b}_{kj}|} e^{-2\pi i \frac{v}{c} |\hat{b}_{kj}|} \mathbf{J}_k(\hat{b}_{kj}) \mathbf{J}_j(\hat{b}_{jk}) \right\}. \end{aligned} \quad (17)$$

If antenna-to-antenna coupling had not been considered, the visibility equation would only include the first of the three terms in equation (17). We can rewrite this no-coupling term as

$$\begin{aligned} \mathbf{V}_{ij}^0 &= \int \mathbf{J}_i(\hat{s})(\vec{\epsilon}(\hat{s}) \times \vec{\epsilon}^y(\hat{s})) \mathbf{J}_j^y(\hat{s}) e^{2\pi i \frac{v}{c} \vec{b}_{ij} \cdot \hat{s}} d\Omega \\ & = \int \mathbf{J}_i(\hat{s}) \mathbf{C}(\hat{s}) \mathbf{J}_j^y(\hat{s}) e^{2\pi i \frac{v}{c} \vec{b}_{ij} \cdot \hat{s}} d\Omega. \end{aligned} \quad (18)$$

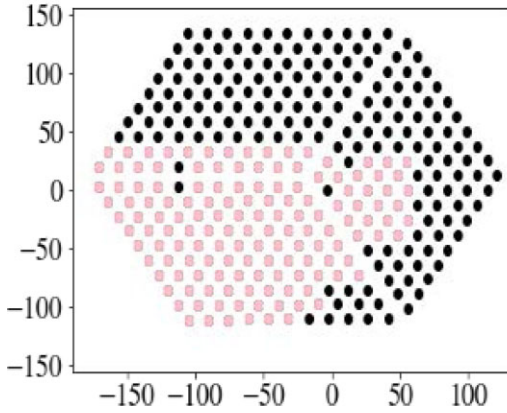
Notice that the second and third (‘coupling’) terms of equation (17) depend on zeroth-order (‘no-coupling’) visibilities of various baselines  $V_{kj}^0$  and  $V_{ik}^0$ . Rewriting equation (17) in terms of zeroth-order visibilities,

$$\begin{aligned} \mathbf{V}_{ij}^1 &= \mathbf{V}_{ij}^0 \\ & + \frac{i\eta_0}{4\lambda} \left( - \sum_{k \neq i} \frac{\Gamma_k}{R_k |\hat{b}_{ki}|} e^{2\pi i \frac{v}{c} |\hat{b}_{ki}|} \mathbf{J}_i(\hat{b}_{ki}) \mathbf{J}_k(\hat{b}_{ki}) \mathbf{V}_{kj}^0 \right. \\ & \left. + \sum_{k \neq j} \frac{\Gamma_k^*}{R_k |\hat{b}_{kj}|} e^{-2\pi i \frac{v}{c} |\hat{b}_{kj}|} \mathbf{V}_{ik}^0 \mathbf{J}_k^y(\hat{b}_{kj}) \mathbf{J}_j^y(\hat{b}_{jk}) \right). \end{aligned} \quad (19)$$

Note that all terms associated with reradiated (mutually coupled) power introduce copies of zeroth-order visibilities. These copies are attenuated by extra beam factors ( $\mathbf{J}\mathbf{J}$  or  $\mathbf{J}^y \mathbf{J}^y$ ), decayed in magnitude by the baseline length, phase-shifted, and then delayed by the time it takes for the reflected radiation to travel the coupled baseline vector. At last, recall that the derivation of equation (19) requires four assumptions, which we previously defined as Assumptions A–D. Assumptions A–B are written in Appendix A. Section 2.3 notes Assumption C, and Assumption D was noted in this section. All four assumptions are summarized in the table below.

Key assumptions	
Assumption	Description
A	$ \vec{R}  \gg  \vec{b}_i $
B	Free space (vacuum) propagation
C	$\Delta t$ (sky change) $\gg \tau$ (delays of scattering)
D	Van Cittert–Zernike theorem

HERA Array Sector: Feeds Positioned on JD 2459122



**Figure 4.** A 143 element subset (pink) of the HERA array, which contains a total of 350 elements. The x-axis (east) and y-axis (north) have units of meters, in arbitrary, local ENU coordinates. This array subset corresponds to all Vivaldi feeds which were positioned on JD 2459 122, a night during which a commissioning observation was recorded by the HERA collaboration.

### 3 SIMULATION OF ZERO-ORDER VISIBILITIES

Armed with a general formalism for first-order array element coupling, let us establish a simulated framework for studying its effects. Specifically, we simulate zeroth- and first-order visibility data for HERA using a simplified version of the formalism. HERA is a 350 element interferometer, built in South Africa’s Karoo Radio Astronomy Reserve. HERA was designed specifically to characterize the evolution of the 21cm signal from cosmic dawn ( $z \approx 30$ ) through the full reionization of the IGM ( $z \approx 6$ ). Each element of the interferometer is a 14 m parabolic dish, capable of observing from 50–250 MHz. For this analysis, we limit our simulations to the frequency range 144–169 MHz. Active electronic receiver components are currently being installed on to HERA’s Vivaldi-style dish feeds; both the receiver system and the feed are characterized in Fagnoni et al. (2021).

While it is beyond the scope of this work to directly compare the results of our simulations to observed HERA data, we none the less choose to study the HERA array as it was configured for a commissioning observation on Julian Date (JD) 2459 122, during which the HERA collaboration recorded data. Only 143 of the 350 Vivaldi feeds were positioned and installed on this JD. Simulating this 143 element subset, rather than the full array, will make it convenient to apply the results of this work to a follow-up analysis which directly compares first-order interferometric visibilities to HERA data. Fig. 4 plots the 143 element array configuration as a function of East / North / Up (ENU) antenna coordinates.

In preparation for the analysis of our simulated HERA visibilities, we confirm that the far-field approximation used in equation (10) is valid across the HERA array. We also analyse simulated HERA beam products and derive how these products, which were calculated using standard electromagnetic simulations software that implicitly accounts for antenna impedance, can be unit-wise compatible with our general formalism, which is explicitly dependent on antenna impedance. See Appendix (B) for our work to reach these conclusions. The HERA beam products were simulated using CST Microwave Studio.<sup>1</sup> Fagnoni et al. (2021) describe in detail the electromagnetic properties of the simulated HERA beam.

#### 3.1 A simplified formalism to simulate

Let us now make two simplifications to the general formalism in order to reduce the computational expense of the simulation over a HERA-like array. First, we only consider an unpolarized astrophysical sky. For now, this consists of the Stokes I component of a combined diffuse (GDSM) and point source (GLEAM) sky. In this case, the coherency matrix greatly simplifies, since the polarized Stokes parameters  $Q = U = V = 0$ . Second, we simulate all beams in the array as having identical electromagnetic properties. Thus,  $\mathbf{J}_i(\hat{s}) = \mathbf{J}_j(\hat{s}) = \mathbf{J}(\hat{s})$ . Further,  $\Gamma_k = \Gamma_i = \Gamma$  and  $R_k = R_i = R_{\text{ant}}$ . With these simplifications, the first-order visibilities of equation (19) simplify to

$$\begin{aligned} \mathbf{V}_{ij}^1 &= \mathbf{V}_{ij}^0 \\ &+ \frac{i\eta_0}{4\lambda} \left( - \sum_{k \neq i} \frac{\Gamma}{R_{\text{ant}} |\vec{b}_{ki}|} e^{2\pi i \frac{\nu}{c} |\vec{b}_{ki}|} \mathbf{J}(\hat{b}_{ik}) \mathbf{J}(\hat{b}_{ki}) \mathbf{V}_{kj}^0 \right. \\ &\left. + \sum_{k \neq j} \frac{\Gamma^*}{R_{\text{ant}} |\vec{b}_{kj}|} e^{-2\pi i \frac{\nu}{c} |\vec{b}_{kj}|} \mathbf{V}_{ik}^0 \mathbf{J}^y(\hat{b}_{kj}) \mathbf{J}^y(\hat{b}_{jk}) \right), \end{aligned} \quad (20)$$

where the zeroth-order visibilities  $\mathbf{V}_{ij}^0$ , formerly expressed as equation (18) now simplify to

$$\mathbf{V}_{ij}^0 = \int I(\hat{s}, \nu) \mathbf{J}(\hat{s}) \mathbf{J}^y(\hat{s}) e^{2\pi i \frac{\nu}{c} \vec{b}_{ij} \cdot \hat{s}} d\Omega. \quad (21)$$

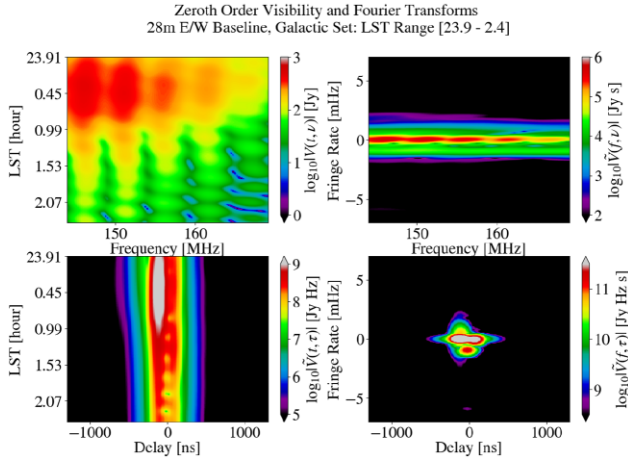
#### 3.2 Zeroth-order visibilities in a simulated HERA-like array

All zeroth-order visibilities presented in this paper are simulated without noise, in the frequency range 144–169 MHz using the formalism of equation (21), as implemented by the Healvis visibility simulator presented in Lanman & Kern (2019). Two sky models are simulated separately. First, a ‘diffuse’ sky model which is derived from the source-subtracted and de-stripped Haslam 408 MHz map presented in Remazeilles et al. (2015). Second, a point source model which includes the GLEAM I catalogue along with point source models of 3C 161, 3C 409, and Cassiopeia A. Multicomponent extended source models for Centaurus A, Hera A, Hydra A, Pictor A, and Virgo A, are also included, as well as an extended source model for Fornax A. For the purposes of this work, the second sky model is called ‘GLEAM I + Extended Sources’; this model is described in detail in Byrne et al. (2022). The zeroth-order visibilities of the two sky models are added together (in the complex plane) to generate a third set of zeroth-order visibility data. This third set shall be dubbed as the ‘Combined Sky’ model, as it contains both diffuse structure and point sources.

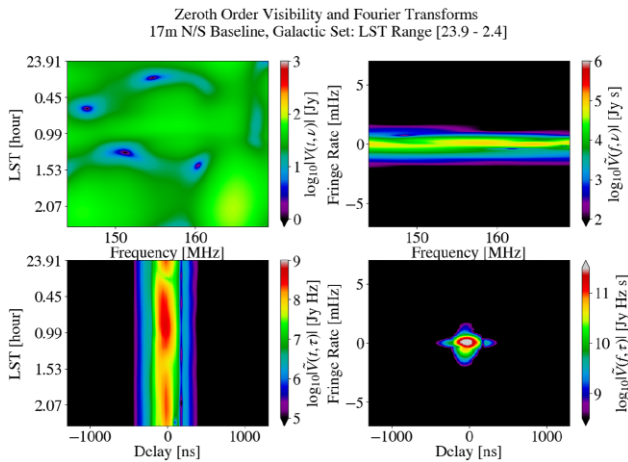
Figs 5 and 6 present the zeroth-order visibility of the combined

<sup>1</sup><https://www.3ds.com/products-services/simulia/products/cst-studio-suite/>





**Figure 5.** Zeroth-order visibility data of the combined (diffuse + point source) sky model, for a 28 m E/W baseline, generated by the healvis simulator. Frequency range: 144–169 MHz. LST range: [23.9–2.4]. Moving clockwise from the top-left panel: Time versus frequency ( $t$  versus  $\nu$ ), fringe rate versus frequency ( $f$  versus  $\nu$ ), fringe rate versus delay ( $f$  versus  $\tau$ ), and time versus delay ( $t$  versus  $\tau$ ).



**Figure 6.** Zeroth-order visibility data of the combined (diffuse + point source) sky model, for a 17 m N/S baseline, generated by the healvis simulator. Frequency range: 144–169 MHz. LST range: [2.4–4.9]. Moving clockwise from the top-left panel: Time versus frequency ( $t$  versus  $\nu$ ), fringe rate versus frequency ( $f$  versus  $\nu$ ), fringe rate versus delay ( $f$  versus  $\tau$ ), and time versus delay ( $t$  versus  $\tau$ ).

sky model, for a 28 m E/W baseline and a 17 m N/S baseline, respectively. All possible Fourier transforms of the zeroth-order visibilities are plotted. Moving clockwise from the top-left panel, the plots for each data set are as follows: time versus frequency ( $t$  versus  $\nu$ ), fringe rate versus frequency ( $f$  versus  $\nu$ ), fringe rate versus delay ( $f$  versus  $\tau$ ), and time versus delay ( $t$  versus  $\tau$ ). The delay spectrum technique, originally explored in Parsons & Backer (2009) and Parsons et al. (2012b) is formed by Fourier transforming visibility data along the frequency axis, and is used extensively in the HERA collaboration to geometrically filter interferometric data; see, for example Kern et al. (2020). The fringe-rate analysis is formed by Fourier transforming visibility data along the time axis and is used extensively to filter interferometric data by temporal variation and sky region (Parsons et al. 2016). In Section 5, we will explain how first-order coupling features evolve as a function of apparent

sky brightness and orientation, and how coupling manifests itself as a function of array position. To do so, we will use the fringe rate versus delay (bottom-right panel) Fourier transform as the main visual representation of coupling features.

## 4 SIMULATION OF FIRST-ORDER VISIBILITIES

To aid in comparing the results of our simulation to data (in a future, companion paper), we choose to simulate the 143 element subset of the array (see Fig. 4) where Vivaldi feeds and active analogue system chains were fully commissioned. Of the hundreds of unique baseline groups which could be derived from this set of 143 elements, for brevity we limit our analysis to visibility data from 28 m E/W baselines and 17 m N/S baselines. These baselines are expected to contribute to HERA’s primary scientific results and are of a convenient length to visually distinguish and differentiate various coupling features. The first-order simulations, which use as input the zeroth-order healvis visibility data, span the same frequency range of 144–169 MHz. Data from two different LST ranges will be studied. First, the LST range of [23.9–2.4], during which the galaxy sets in the western horizon with respect to HERA. Second, the LST range of [2.4–4.9], which captures the transit of Fornax A near the zenith point of HERA. In both LST ranges, coupling effects are measured down to at least 1 part in  $10^4$  compared to peak visibility power. While our analysis focuses on how such coupling effects complicate the detection of the 21cm signal, the first-order coupling model we use to make such visibilities is noteworthy for any experiment using a radio interferometer where internal instrumental systematics at a level of 1 part in  $10^4$  could corrupt scientific results.

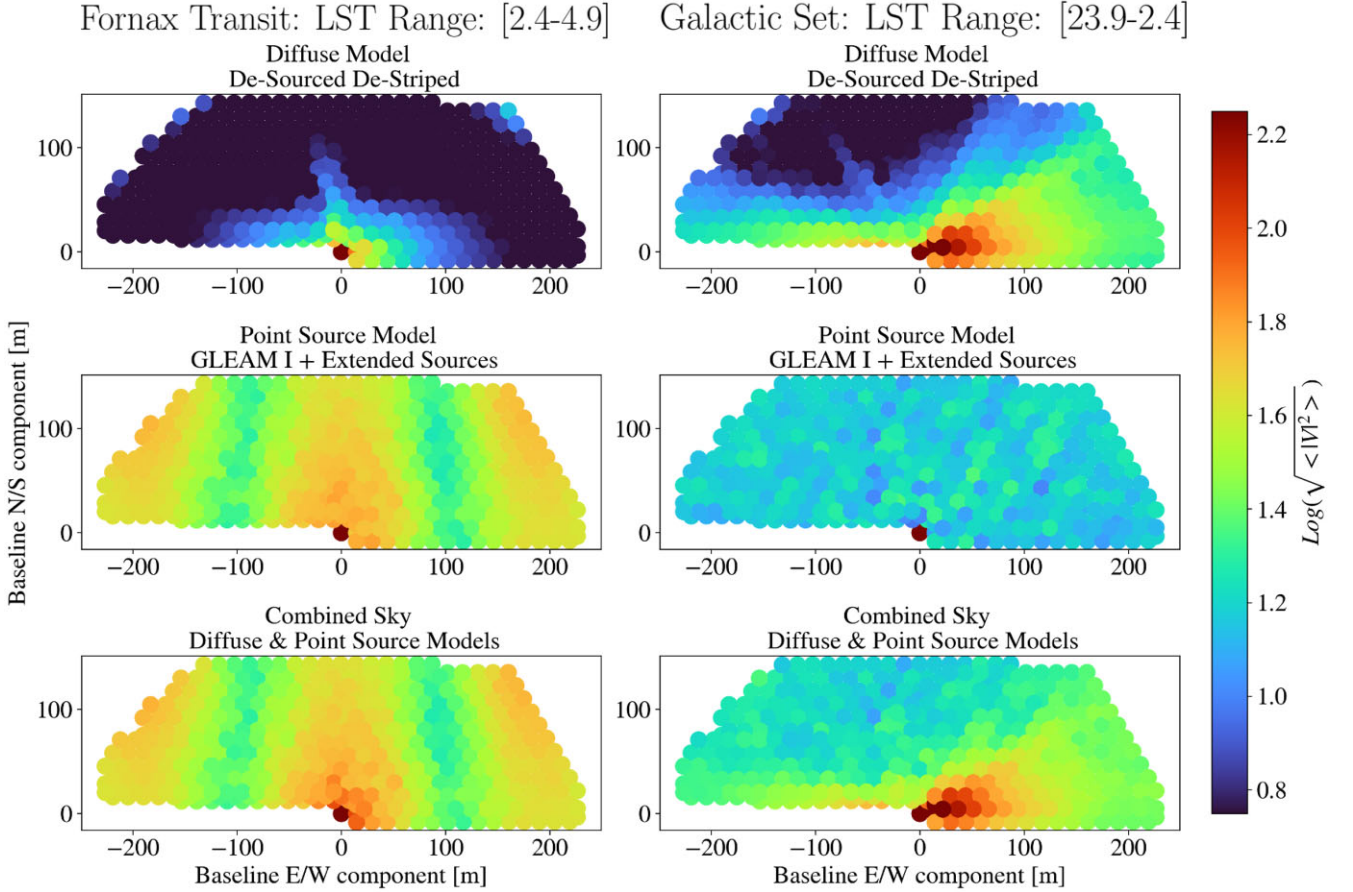
## 5 PHENOMENOLOGY

### 5.1 Using fringe versus delay space to study coupling effects

The structure of the general formalism suggests that fringe rate versus delay is an ideal framework in which we can analyse first-order array element coupling effects. Recall from Section 2.2 that all terms associated with reradiated (mutually coupled) power introduce copies of zeroth-order visibilities. These copies are delayed by the time it takes for the reflected radiation to travel the coupled baseline vectors ( $|\vec{b}_{ki}|$ ,  $|\vec{b}_{kj}|$ , depending on the sum term in equation 20). Note also that, for each sum term in the formalism (either the simplified or general version), the only components which have dot products dependent on sky orientation (e.g.  $\vec{b}_{ij} \cdot \hat{s}$ ) are the zeroth-order visibility terms. In equation (20), none of the other expressions are explicitly dependent on what sky is above the array, and so we infer that any time-dependent feature generated by first-order array element coupling must derive itself from the zeroth-order visibility terms.

As such, we interpret each component in the coupling sum terms of equation (20) as having one explicit fringe term. For each antenna  $k$  included in  $\sum_{k \neq i}$ , we expect coupling features to manifest themselves at the dominant fringe rates of  $\mathbf{V}_{kj}^0$ . Similarly, for  $\sum_{k \neq j}$ , we expect coupling features to manifest themselves at the dominant fringe rates of  $\mathbf{V}_{ik}^0$ . The power of each coupling feature will be proportional to the RMS power of the respective visibilities at the simulated times and frequencies, e.g.  $\sqrt{\langle |\mathbf{V}_{kj}^0|^2 \rangle}$  and  $\sqrt{\langle |\mathbf{V}_{ik}^0|^2 \rangle}$ .

Fig. 7 plots for two LST ranges the RMS power of the visibilities of all possible combinations of  $(k, j)$  and  $(i, k)$  in HERA, as a function of the E/W and N/S component of the baseline. Each colourful dot



**Figure 7.** The RMS of the zeroth-order visibilities of all unique baseline groups in HERA, as a function of the E/W and N/S component of each respective group. Frequency range: [144–169] MHz. In the LST range [23.9–2.4], which corresponds to the galaxy setting in the horizon of HERA, baselines with short N/S components tend to be brighter in the diffuse sky model compared to the point source model, however the point source model dominates when the N/S component is longer. In the LST range [2.4–4.9], which captures Fornax A (a particularly bright point source) transiting the HERA beam, the HERA sky response at most (if not all) baselines is dominated by the point source model. By the method of stationary phase, discussed in Section 6.2, we note that the regions of the sky where the phase of an interferometric visibility is stationary (whether we are discussing the zeroth-order component, or the component associated with antenna–antenna scattering), occurs where the baseline vector aligns with the direction of strongest incident astrophysical radiation at the horizon. When the galaxy is setting, peak RMS power lies in mostly E/W baselines, which is the orientation of the galaxy in the horizon plane during LST range [23.9–2.4]. When Fornax A transits zenith, peak RMS power lies in mostly N/S baselines, which is the orientation of the galaxy in the horizon plane during LST range [2.4–4.9].

in the figures represents a unique baseline group in HERA, and the colour of each dot represents the strength of the RMS visibility. The LST range [23.9–2.4] corresponds to the galaxy setting in the horizon of HERA. In this LST range, baselines with short N/S components tend to be brighter in the diffuse sky model compared to the point source model; however, the point source model dominates when the N/S component is longer. The LST range [2.4–4.9] captures Fornax A, a particularly bright point source, transiting the main lobe of the HERA beam. In this LST range, the HERA sky response at most (if not all) baselines is dominated by the point source model.

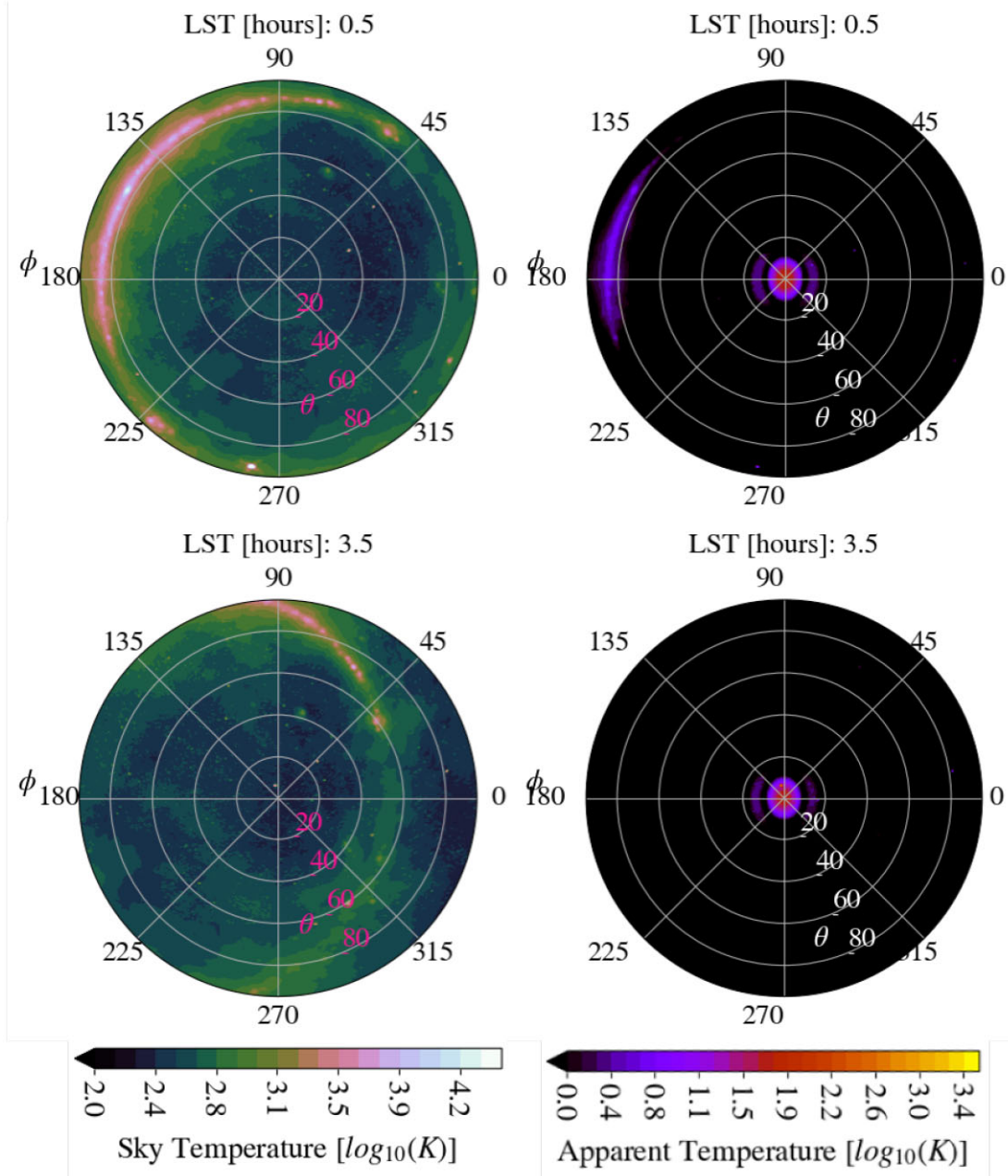
## 5.2 Delay contributions of first-order coupling

Each sum term in equation (20) contains four complex expressions which contribute to the delay at which first-order coupling will manifest itself: First, the constant values (e.g.  $i\eta_0\Gamma$ ). Second, the beam product (e.g.  $J(\hat{b}_{ik})J(\hat{b}_{ki})$ ). Third, delays associated with the copied visibilities (e.g.  $V_{kj}^0$ ). Finally, the complex exponential associated with baseline length (e.g.  $e^{2\pi i \frac{z}{|b_{ki}|}}$ ). Regarding the constants, we

have fit our voltage reflection coefficient measurement using high-order Chebyshev polynomials, thus preserving its spectral shape while ensuring any artefacts in delay space are negligible. We can confirm that all constant values contribute negligible power at non-zero delays.

To gain intuition on the third contribution, Fig. 13 plots the peak-normalized delay spectrum of different healpix beam pixels, caused by the window function applied to the beam products of our first-order visibilities,  $J(\hat{s})J(\hat{s})$  (solid lines) and  $J^y(\hat{s})J^y(\hat{s})$  (dashed lines). Whenever performing a Fourier Transform of a beam product, whether in healvis, or in the first-order coupling algorithms, we use a Blackman–Harris window. Thyagarajan et al. (2016) studied the leakage of foreground power into the time domain associated with windowing in the frequency domain and concluded that the Blackman–Harris window offers a contamination-free dynamic range of at least 1 part in  $10^6$  for voltage signals in the time domain. Thus, foreground leakage caused by the window will be below the expected EoR signal, although to the detriment of measurement sensitivity, which is reduced by approximately 50 percent. The healpix beam pixels plotted in Fig. 13 are at zenith (blue/orange) and





**Figure 8.** Simulations of the combined sky model temperature and apparent temperature (sky multiplied by normalized beam gain) at 150 MHz, rotated into the coordinate system of the simulated HERA beam (see Fig. 2) and plotted as an orthographic projection. At an LST of approximately 0.5, a significant fraction of the apparent brightness (namely the galactic plane) is concentrated near the horizon in the E/W direction. At an LST of approximately 3.5, Fornax A transits zenith, and a significant fraction of the apparent brightness is concentrated near zenith.

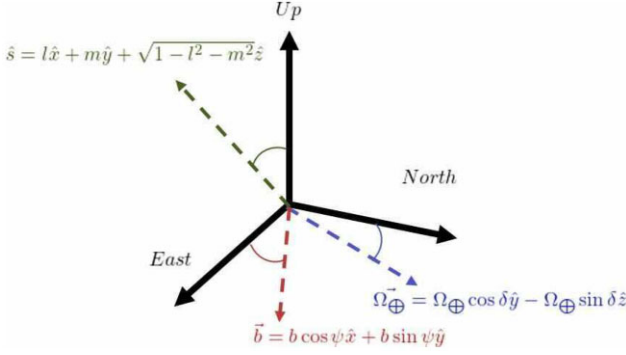
two orthogonal horizon positions (green/red and purple/burgundy, respectively). Power is associated with the windowed beam function at non-negligible delays in both beam products, regardless of pixel location. Nevertheless, for all pixel positions, this window function power is below  $-40$  dB of the peak-normalized delay spectra after  $|\tau| \geq 500$  ns. Since power drops off so quickly at delays greater than 500 ns, regardless of pixel, we approximate the power associated with beam products to be approximately equal across all pixels. Similarly, it can be shown that the Blackman–Harris window function applied to the zeroth-order beam product,  $J(\hat{s})J^y(\hat{s})$ , produces a comparable delay spectrum to those plotted in Fig. 13. Thus, we approximate the ‘delay spectrum associated with the window function of the beam’

as equal across all pixels and all beam products:

$$\tau_{J(\hat{s})J^y(\hat{s})} \approx \tau_{J(\hat{s})J^y(\hat{s})} \approx \tau_{J^y(\hat{s})J^y(\hat{s})} \equiv \tau_{\text{beam}} \approx 500 \text{ ns}. \quad (22)$$

$\tau_{\text{beam}}$  is typically ignored when discussing the delays associated with power in zeroth-order visibilities. However, this non-zero beam delay term cannot be ignored when analysing first-order visibilities. Section 6.2 describes how terms of  $\tau_{\text{beam}}$  manifest themselves in the structure of first-order coupling effects. These beam delay terms are conceptually labelled atop the 2D power-spectrum measurement presented in Fig. 18.

Regarding the fourth delay contribution, we note that zeroth-order visibilities  $V_{kj}^0$  tend to have maximum power at two different delays:



**Figure 9.** Geometry used to compute coordinate expression for fringe rate as a function of sky location and baseline length and location in equation (27).

those associated with the primary beam lobe and those associated with the ‘geometric horizon’ of the baseline ( $|b_{kj}|$ ). The manifestation of foregrounds (the primary source of incident astrophysical radiation) in delay space, as a function of interferometric baseline length, has been studied extensively (see e.g. Datta, Bowman & Carilli 2010; Morales et al. 2012; Parsons et al. 2012b; Thyagarajan et al. 2015). Such studies describe how foregrounds of wideband antenna arrays are contained in a ‘wedge’ of delay versus baseline length (see, for example, the  $V_{ij}^0$  subplots of Figs 16 and 17, and the description of the ‘wedge’ in Fig. 18). In this wedge, power appears at delays approximately equal to zero, regardless of baseline length. Such power comes from incident astrophysical radiation being absorbed in the direction of maximum beam gain. The wedge also contains power at delays which are directly proportional to the baseline length of the measured visibility, with a slope which is inversely proportional to the speed of light. Thyagarajan et al. (2015) describe this ‘pitchfork’ effect at the geometric horizon of the array as fundamental to interferometric visibilities measured by wideband array elements; this is not an instrumental systematic effect, but a fundamental instrument response. As such, the effect can only be mitigated, not entirely removed. Kern et al. (2020) describe how various instrumental effects differ in fringe and delay space compared to the pitchfork effect.

Knowing that zeroth-order visibilities ( $V_{kj}^0$ ) have significant power at two separate delays ( $\tau \approx 0$ ,  $\tau \approx \frac{|b_{kj}|}{c}$ ), and recalling the fourth contribution to total delay (complex exponentials of the form  $e^{2\pi i \frac{v}{c} |b_{ki}|}$ ), we conclude that first-order coupling effects of the  $ij^{th}$  visibility predominately reside at the following delays

$$\tau_{ij}^1 \approx \begin{cases} \tau_{\text{beam}} \rightarrow V_{ij}^0 \text{ ‘wedge’}, \\ \tau_{|b_{ij}|} \rightarrow V_{ij}^0 \text{ ‘wedge’}, \\ \tau_{|b_{ki}|} + \tau_{V_{kj}^0, \text{main lobe}} + \tau_{\text{beam}} \\ \tau_{|b_{ki}|} + \tau_{V_{kj}^0, \text{geometric horizon}} + \tau_{\text{beam}} \end{cases} \quad (23)$$

While zeroth-order visibilities also contain significant power at delays associated with  $\tau_{\text{beam}}$  and  $\tau_{|b_{ij}|}$ , the latter two cases of equation (23) are unique to first-order coupling. The maximum extent in delay of these latter two cases is plotted in yellow in Fig. 18.

### 5.3 Fringe rate as a function of orthographic sky location

In this section, we use the coordinate-free fringe-rate expression from Parsons et al. (2016) to derive a useful analytic formula relating fringe rate to orthographic-projected sky location, baseline length, and baseline orientation. Doing so offers us intuition regarding the fringe rates associated with zeroth-order visibility terms ( $v_{kj}^0$  and  $v_{ij}^0$ ).

According to Parsons et al. (2016),

$$f_r \approx \frac{v}{c} [\hat{s} \cdot (\Omega_{\oplus} \times \mathbf{b})]. \quad (24)$$

This geometry can be translated into the ground-based orthographic coordinate system illustrated in Fig. 9 by aligning the  $\hat{x}$  axis with the E/W direction, the  $\hat{y}$  axis with the N/S direction, and the  $\hat{z}$  axis with the vertical (zenith) direction. If our array is located at a declination of  $\delta$  and our baseline is at an angle of  $\psi$  with respect to the E/W direction, then

$$\Omega_{\oplus} \times \mathbf{b} = b [\hat{x} \sin \delta \sin \psi - \hat{y} \sin \delta \cos \psi + \hat{z} \cos \psi \cos \delta]. \quad (25)$$

Writing  $\hat{s}$  in orthographic coordinates,

$$\hat{s} = \hat{x} \ell + \hat{y} m + \hat{z} \sqrt{1 - \ell^2 - m^2} \quad (26)$$

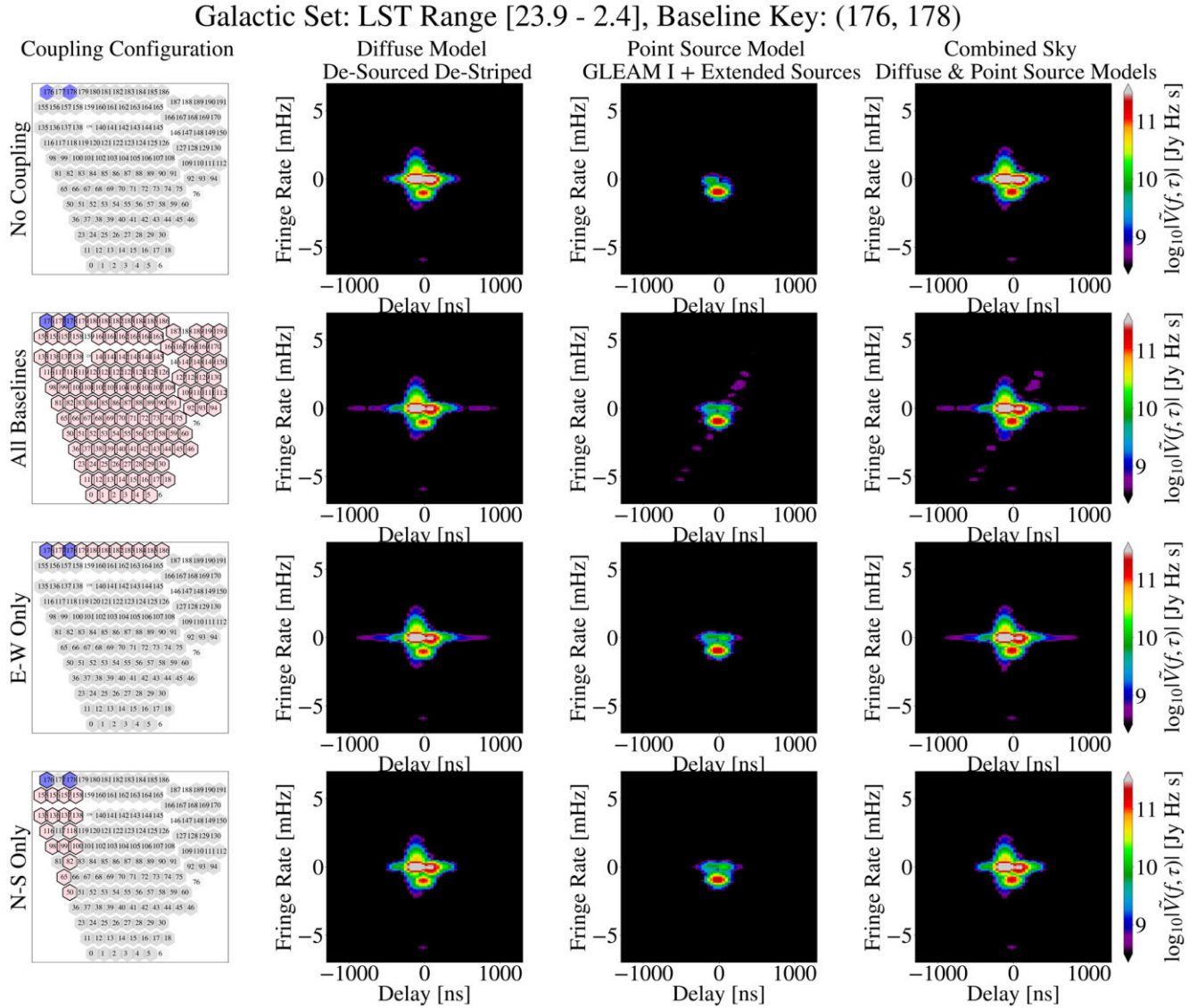
we have

$$f_r(\ell, m) = \frac{bv}{c} \left( \ell \sin \delta \sin \psi - m \sin \delta \cos \psi + \sqrt{1 - \ell^2 - m^2} \cos \psi \cos \delta \right) \quad (27)$$

When the majority of apparent brightness is concentrated at zenith, power is concentrated at fringe rates of  $f_r(0, 0) \approx \frac{bv}{c} \cos \psi \cos \delta$ , which is proportional to the baseline’s projected E/W length ( $b \cos \psi$ ). The same is true when apparent brightness is concentrated near the horizon in the N/S direction. On the other hand, when apparent brightness is concentrated near the horizon, in the E/W direction, then the peak fringe rate is proportional to the N/S projected length of the baseline  $f_r(1, 0) \propto \frac{bv}{c} \sin \psi \cos \delta$ .

### 5.4 Fringe-rate filtering: a potential strategy for mitigating first-order coupling effects

Recall from equation (20) that first-order visibilities consist of a sum of copies of zeroth-order visibilities, and that these visibilities ( $V_{kj}^0$  or  $V_{ik}^0$ , depending on the sum term) are the only fringing component in each respective sum. In the first sum, visibilities will have fringe rates associated with  $b_{kj}$ ,  $k \neq i$ . The fringe terms associated with the main beam of the coupled visibility ( $\frac{|b_{kj}|}{c} \cos \psi \cos \delta$ , or  $\frac{|b_{kj}|}{c} \sin \psi \cos \delta$ ) will equal the fringe term associated with apparent brightness being near zenith of the measured baseline ( $\frac{|b_{ij}|}{c} \cos \psi \cos \delta$ ) when the E/W component of  $|b_{kj}|$  equals the E/W component of  $|b_{ij}|$ . Similarly, for the second sum, the fringe terms associated with the main beam of the coupled visibility ( $\frac{|b_{ik}|}{c} \cos \psi \cos \delta$ , or  $\frac{|b_{ik}|}{c} \sin \psi \cos \delta$ ) will equal the fringe term associated with apparent brightness being near zenith of the measured baseline ( $\frac{|b_{ij}|}{c} \cos \psi \cos \delta$ ) when the E/W component of  $|b_{ik}|$  equals the E/W component of  $|b_{ij}|$ . In our simulated HERA-like array, these relationships are infrequently satisfied. Thus, the fringe rates associated with first-order antenna–antenna coupling will not typically be the same fringe rates associated with the main beam of the measured visibility. Evidence for this can be found in all fringe rate versus delay plots presented in this work (Figs 10 and 11, and all plots in the online supplementary material). To the dynamic range of each figure, the non-fringing and fringing features associated with first-order coupling appear at fringe rates different than the main lobe for all baselines with a non-zero E/W component. This is true for all analysed baseline lengths at all simulated LST ranges. While it is beyond the scope of this work to test filtering strategies for first-order coupling systematics, our conclusion regarding main lobe fringe rates may prove crucial for finding subsets of fringe versus delay space which are not contaminated by mutual coupling.



**Figure 10.** Fringe rate versus delay plot of the first-order visibility of 28 m E/W baseline (176, 178) of the HERA array, plotted in blue, as well all other elements (grey) which were installed on JD 2459 122. Three different array configurations are plotted: ‘All baselines’, ‘E-W only’, and ‘N-S only’. For each array configuration, only antennae which are plotted in pink are included in the first-order coupling calculation (equation 20). All elements of the simulated array configuration use identical beam products ( $\mathbf{J}(\hat{s})$ ,  $\mathbf{J}^*(\hat{s})$ ). LST range: [23.9–2.4], frequency range: [144–169 MHz]. In this LST range, the galaxy sets in the western horizon with respect to HERA. See Fig. 8 for sky brightness as a function of position in the HERA beam.

Let us now broadly categorize coupling features as being non-fringing or having non-zero fringe, and study the two cases separately.

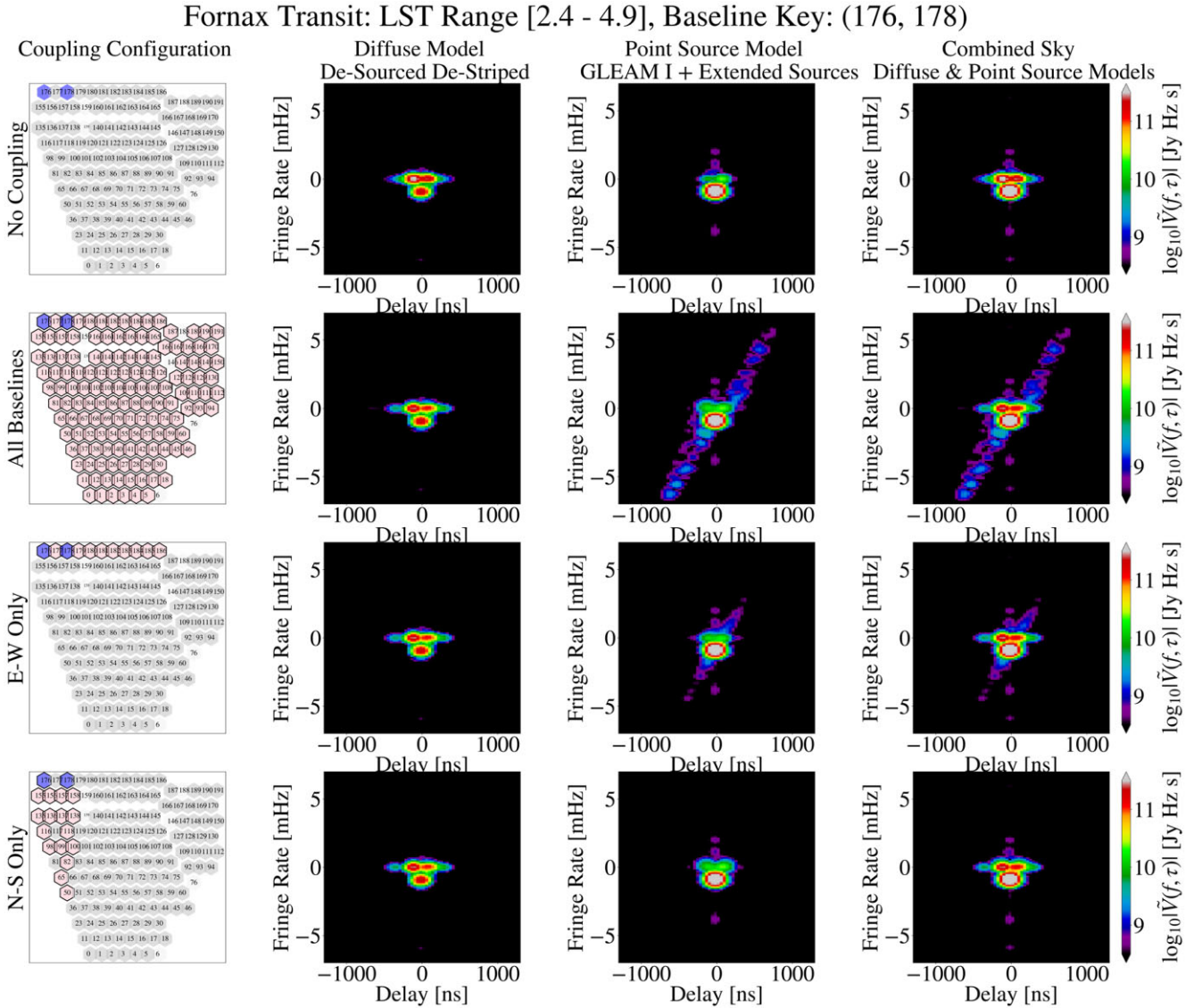
### 5.5 Non-fringing (‘bar’) features

The top row of Fig. 8 shows an LST in our combined sky simulation where a significant portion of the apparent sky brightness is concentrated near the horizon, in the E/W direction. In these figures, the sky has been rotated into an orthographic projection of the spherical coordinate system used for our beam, as presented in Fig. 2. Thus, all plots in Fig. 8 depict the sky as it is ‘seen’ by a HERA beam element as it receives incident astrophysical radiation. At an LST of 0.5 (top row), the galaxy is setting in the western horizon with respect to HERA. Fig. 10 records the first-order coupling ( $\mathbf{V}_{ij}^1$ ) of baseline

(176, 178) – plotted in blue – as calculated by equation (20), in a range of LSTs including 0.5. This coupling term is calculated for all three different sky models (diffuse only, point source only, and the combined sky) for three unique array configurations. For each array configuration, only antennae which are plotted in pink in column 1 are included in the first-order coupling calculation. In addition to the three unique array configurations (rows 2–4), the first row of Fig. 10 plots the zeroth-order visibility of the same sky models (row 1).

The brightness of the galactic plane is a feature which is only present in the diffuse sky model, not the point source model. Therefore, for all array configurations in column 2 of Fig. 10, the apparent brightness is concentrated near the horizon in the E/W direction. As predicted by equation (27), the peak fringe rate of coupling features is proportional to the N/S projected length of the coupled (pink) baselines. For the ‘E/W only’ array configuration (row





**Figure 11.** Fringe rate versus delay plot of the first-order visibility of 28 m E/W baseline (176, 178) of the HERA array, plotted in blue, as well all other elements (grey) which were installed on JD 2459 122. Three different array configurations are plotted: ‘All baselines’, ‘E-W only’, and ‘N-S only’. For each array configuration, only antennae which are included in the first-order coupling calculation (equation 20). All elements of the simulated array configuration use identical beam products ( $\mathbf{J}(\hat{s})$ ,  $\mathbf{J}'(\hat{s})$ ). LST range: [2.4–4.9], frequency range: [144–169 MHz]. In this LST range, Fornax A transits near the zenith point (location of peak gain) of the HERA beam. See Fig. 8 for sky brightness as a function of position in the HERA beam.

3), all coupled baselines have a N/S component of zero with respect to the antennae in baseline (176, 178). Thus, coupling manifests itself at a fringe rate of zero over a broad range of delays, creating a bar-like feature. This bar-like feature is more prominent in the ‘All baselines’ configuration (row 2, column 2) because there are more coupled antennae which have a non-zero E/W baseline component. If a main-beam fringe-rate filtering strategy were adopted, such as that suggested in Section 5.4, our phenomenological understanding of bar-like features suggests we ignore N/S baselines entirely from our analysis if the array is 2D (i.e. not all baselines lie in the N/S direction).

By contrast, all coupled antennae in the ‘N/S only’ array (row 4) of column 2 have a negligible E/W component with respect to (176, 178), and so instead of producing a bar feature, the coupled antennae will manifest themselves in baseline (176, 178) with non-zero fringe.

At the dynamic range of the colour bar of this figure, this is not apparent, both because the gain in the N/S direction for this simulated beam polarization (the N/S oriented dipole of the Vivaldi feed) is significantly lower than the gain of that feed in the E/W direction. Further, Fig. 7 tells us that these fringing N/S baseline contributions will have less RMS power in their visibilities than the RMS of the visibilities which contribute to the ‘E/W’ array configuration. Thus, for this LST range and array configuration, features with non-zero fringe are subordinate to zero-fringe bar features. For the rest of this work, subordinate coupling features are not presented or discussed.

In contrast to the diffuse sky model, the point source sky model (column 3) in Fig. 10 depicts coupling features that dominate at non-zero fringe rates. We note from Fig. 7 that the RMS power of the point source sky at this LST range is fairly homogeneous regardless of baseline length or orientation. Thus, we may treat the apparent

brightness as coming from the pointing centre of the HERA beam, which is zenith. The ‘All baselines’ array configuration of Fig. 10 (row 2, column 3) has a sufficient number of coupled antennae with a non-zero E/W component that we begin to see structure with non-zero fringe rates. Coupling features with non-zero fringe rates are further discussed in the next section.

### 5.6 Fringing (‘cross’) features

The bottom row of Fig. 8 shows an LST in our combined sky simulation where a significant fraction of the apparent brightness is concentrated near zenith. From Fig. 7, we know that the HERA response of most (if not all) baselines in this LST range is dominated by the point source model, not the diffuse model. Fornax A transits the zenith point of the HERA array at an LST of approximately 3.5. Fig. 11 records the first-order coupling ( $\mathbf{V}_{ij}^0$ ), of baseline (176, 178) as calculated by equation (20), over a range of LSTs, including 3.5. The brightness of Fornax A is a feature which is only present in the point source sky model, not the diffuse sky model. Thus, for all array configurations in column 3 of Fig. 11, there is a significantly bright source at zenith. Therefore, per equation (27), the peak fringe rate for first-order coupling features will be proportional to the E/W projected length of the coupled (pink) baselines. For the ‘E/W only’ array configuration (row 3), all coupled antennae have a non-zero E/W component with respect to the antennae in baseline (176, 178) and coupling manifests itself at non-zero fringe rates over a broad range of delays, creating a cross-like feature. The cross feature is weaker in the ‘E/W only’ configuration compared to the ‘All baselines’ configuration because the latter contains more E/W baselines. We note from the ‘N/S only’ configuration that purely N/S baselines do not contribute to the cross feature when the apparent brightness is concentrated at zenith.

The slope of this cross-like coupling feature in fringe versus delay space depends on the relative position of the baseline with respect to all other antennae in the array. In baseline (176, 178), coupled antennae are physically located to the east of the baseline, which, due to the direction of the complex exponential terms in equation (20) makes coupled antennae manifests themselves in the first-order visibility of baseline (176, 178) with positive fringe space at positive delay, and with negative fringe rates at negative delay (see row 2, column 3 of Fig. 11). In the same simulation and array configuration of baseline (128, 130) – row 2, column 3 of Fig. S6 of the online supplemental material – coupled antennae are physically located to the west of the baseline, making the coupling manifest itself with opposite slope compared to baseline (176, 178). By extension, the coupling features of an identical baseline in the middle of the array, for example (162, 164), having coupled antennae on both sides of the observed baseline, will manifest itself along both slopes, making a characteristic ‘X’ feature (see Fig. S2 of the online supplemental material). It is worth clarifying that the ‘X’ feature is purely a consequence of the geometric location of the baseline with respect to the other antennae in the array, not the fact that the baseline happens to be oriented in the E/W direction. Fig. S4 of the online supplemental material depicts a 17 m N/S baseline in the middle of the array, (143, 163), where first-order coupling also manifests itself as an ‘X’.

To the dynamic range of all fringe versus delay plots presented in this work, we note that cross-like coupling systematics do not appear at fringe rates that correspond to the main beam of the measured visibility. This is further evidence that, in a HERA-like redundant array, the E/W component of  $|b_{kj}|$  and  $|b_{ik}|$  rarely equals the E/W component of  $|b_{ij}|$ . Thus, a less-contaminated analysis space, per

baseline group, may reside at fringe rates associated with the main beam of baselines which have a non-zero E/W component. This could be critical to any future filtering strategy used to mitigate antenna–antenna coupling systematics.

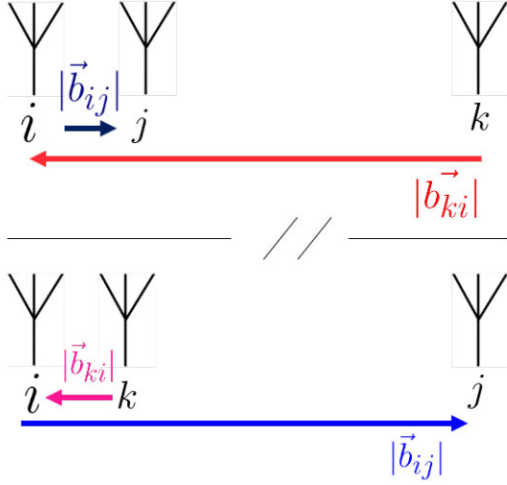
### 5.7 Mutual coupling: a threat to redundant baseline averaging

Three unique 28 m E/W baselines are plotted in the following combinations of two figures: Figs 10 and 11 of this work, and Figs S1 and S2, and S5 and S6 of the online supplemental material. The zeroth-order (row 1, ‘no coupling’) visibilities of all sky models (columns 2–4) of each of these 28 m baselines are identical. As such, without coupling these baselines would be considered ‘redundant’, meaning they could be averaged together in the complex plane without distorting information about incident astrophysical radiation, because their visibilities are identical for all frequencies and times (or, analogously, all delays and fringe rates). However, when we include first-order coupling, these visibilities are no longer identical. First-order coupling makes baselines which were otherwise redundant no longer have identical sky responses. For redundantly configured arrays such as HERA to achieve the measurement sensitivity required to make interferometric measurements of the EoR (the principal scientific goal of the experiment) in any practical duration of observation time, it is imperative, according to analyses such as those presented in Dillon et al. (2020), that visibilities from redundant baselines be able to be redundantly averaged – for at least *some* subset of fringe and delay space. It is beyond the scope of this paper to further address the challenge of non-redundancy caused by array element coupling, but a follow-up study could investigate filtering strategies whereby first-order visibilities of redundant baselines may continue to be identical in certain subsets of fringe and delay space, at least to the level of EoR sensitivity.

## 6 EFFECTS OF FIRST-ORDER COUPLING ON POWER-SPECTRUM MEASUREMENTS

First-order coupling has a significant effect on measurements of the power spectrum, which complicate the fundamental foreground mitigation strategies currently used in 21cm cosmological analyses. We calculate power spectra of our first-order visibility data using the formalism of Parsons et al. (2012b), as normalized using the quadratic estimator strategy of Liu (2020), and instantiated into code using the *hera\_pspec* power-spectrum estimator.<sup>2</sup> To construct the power spectra, we combine information which is redundant both in time and baseline ENU coordinates. Pairs of baselines within each respective redundant baseline group are constructed to estimate a power spectrum. For example, the redundant group of 28 m E/W baselines would contain the following pairs: [(176, 178, ‘xx’), (176, 178, ‘xx’)] and [(162, 164, ‘xx’), (176, 178, ‘xx’)]. For reference, these baselines are plotted for two different LST ranges in Figs 10, 11, and Figs S1 and S2 of the online supplemental material. Power-spectrum measurements which are constructed using ‘auto’ baseline pair combinations contain only those combinations of baseline pairs within a redundant group of one pair with itself, e.g. [(176, 178, ‘xx’), (176, 178, ‘xx’)]. Data sets which are ‘cross’-constructed contain all other possible combinations of baseline pairs in the same redundant group, excluding auto baseline pairs, e.g. [(162, 164, ‘xx’), (176, 178, ‘xx’)], but not [(162, 164, ‘xx’), (162, 164, ‘xx’)].

<sup>2</sup>[https://github.com/HERA-Team/hera\\_pspec](https://github.com/HERA-Team/hera_pspec)

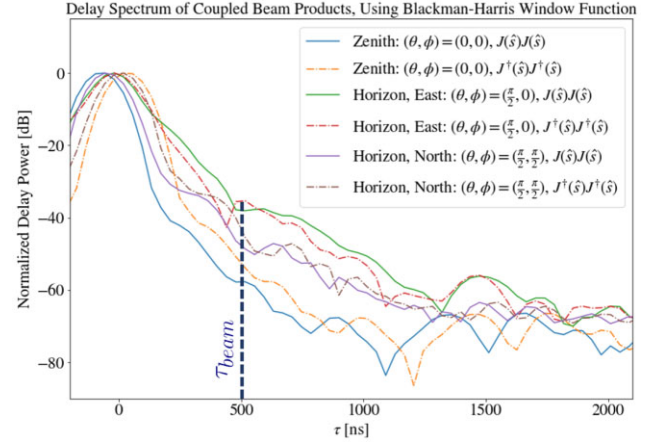


**Figure 12.** Two instances of antenna–antenna coupling terms, which help describe the inverse wedge phenomenon outlined in Fig. 18. Suppose we are analysing the first-order coupling of the  $ij^{\text{th}}$  baseline,  $V_{ij}^1$ , as formalized in equation (20). This visibility will contain copies of all other visibilities, at a delay proportional to the respective baseline distances,  $|b_{ki}|$  or  $|b_{jk}|$ . When the baseline of the measured visibility,  $|b_{ij}|$ , is shortest, we maximize the possible length of  $|b_{ki}|$  and  $|b_{jk}|$ . Similarly, when the baseline of the measured visibility,  $|b_{ij}|$ , is longest, we minimize the possible length of  $|b_{ki}|$ .

Regardless of whether the power spectrum is estimated using auto or cross-baseline pair combinations, we may at our convenience average the data across time (i.e. one power spectrum per time) and/or redundant baseline groups (i.e. one power spectrum per baseline group). We will analyse data which has been averaged along one, and then both, of these axes. First, we study the effects of first-order coupling on autocorrelation data (e.g.  $V_{ii}$  or  $V_{jj}$ ), regardless of if the power-spectrum measurement was constructed using auto or cross-baseline pair combinations. Next, we will average across both time and redundant baseline group, and analyse in one figure the power spectra of all possible baselines in the array, as a function of delay.

### 6.1 Autocorrelation data: a ‘real’ coupling effect

Autocorrelations, which are (to zeroth order) considered as being predominantly real-valued (i.e. having a negligible imaginary component), now contain (to first order) a non-negligible imaginary component. Thus, first-order autocorrelation visibilities are no longer identical in the complex plane. In equation (21), aside from any window function associated with the beam term (which is discussed in Fig. 13 and is typically of a low-enough delay to be negligible at zeroth order), the only explicitly complex component of the visibility equation is the exponential  $e^{2\pi i \frac{b_{ij}}{c} \delta}$ . For an autocorrelation baseline, this term is entirely real because  $b_{ij} = 0$ . However, first-order autocorrelations, per equation (20), have copies of cross-correlations (e.g.  $V_{ki}^0$  which generally have non-zero imaginary components (because  $b_{ki}$  is not generally equal to zero). From Section 5, we know these copied visibilities will manifest themselves in different autocorrelations,  $V_{ii}^1$ , with different amplitudes depending on where in the array the specific antenna is located. Thus, the magnitude of auto visibilities is no longer real-valued and constant across the array:  $\text{imag}(V_{ii}^0) \approx \text{imag}(V_{jj}^0) \approx 0$ ;  $V_{ii}^0 \approx V_{jj}^0 \forall i, j$ . Rather, first-order coupled autocorrelation visibilities are now complex-valued and unequal across the array:  $\text{imag}(V_{ii}^0), \text{imag}(V_{jj}^0) \neq 0$ ;  $V_{ii}^0 \neq V_{jj}^0$ . This non-redundancy is evident in a power-spectrum measurement



**Figure 13.** A peak-normalized delay spectrum of different healpix beam pixels, caused by the window function applied to the beam products of our first-order visibilities,  $J(\hat{s})J(\hat{s})$  (solid lines) and  $J^y(\hat{s})J^y(\hat{s})$  (dashed lines). Window function used: Blackman–Harris. The plotted pixels chosen correspond to zenith (blue/orange) and two orthogonal positions in the horizon plane (green/red and purple/burgundy, respectively). For both beam products, regardless of pixel location, there is power associated with the beam at non-negligible delays. For all pixel positions, power is below  $-40$  dB of the peak-normalized window function power after  $|\tau| \geq 500$  ns. This non-zero delay cannot be ignored when analysing first-order visibilities. One may see this beam delay feature manifest itself in first-order power spectra. See, for example,  $\tau_{\text{beam}}$  in Fig. 18.

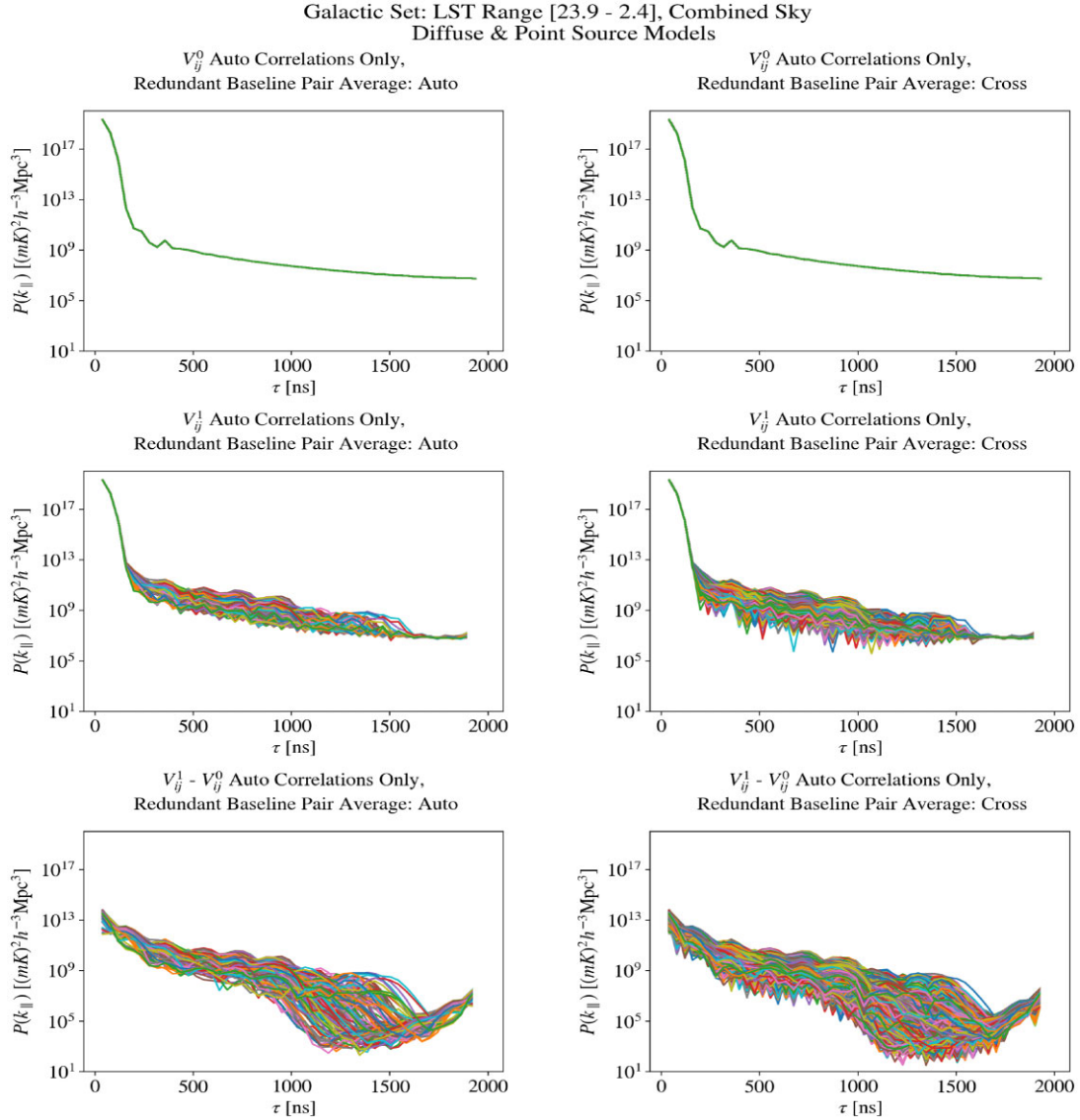
of autocorrelation data, regardless of whether we construct the power spectrum using auto or cross-baseline pairs. Figs 14 and 15 compare the first- and zeroth-order power spectra for the LST ranges of the galaxy setting in the horizon and Fornax A transiting the array, respectively. For the latter LST range, we find significantly less non-redundancy in the power spectrum.

### 6.2 Contamination outside of the wedge: power spectrum measurements as a function of baseline length and delay

Figs 16 and 17 present power spectrum measurements for all possible redundant baseline groups in the array, at two different LST ranges, as a function of baseline length and delay. The top row of plots presents the ‘wedge’ foreground structure, which has been studied extensively in the case of zeroth-order visibility, per the references cited in Section 5.2. The middle rows of Figs 16 and 17 plot the same power-spectrum measurement as the top row, but for first-order visibilities. The bottom row of the figures presents the difference between the first- and zeroth-order visibilities, which is equal to only the sum terms of equation (20). Equivalently, the bottom row of Figs 16 and 17 represents the power and structure uniquely associated with first-order coupling.

It is important to note that the maximum delay of the last two cases of equation (23) can be greater than  $\tau_{|b_{ij}|}$ . Thus, we expect first-order coupling to contaminate delay space beyond the wedge which contains zeroth-order foregrounds. This superwedge contamination is clearly visible in Figs 16 and 17, and is conceptually documented in Fig. 18. The ‘inverse-wedge’ feature presented in these figures depicts the maximum values of the last two delay contributions presented in equation (23). Fig. 12 is a visual aid for explaining how the superwedge features of equation (23) can actually appear at such high delays. Such wedge contamination,





**Figure 14.** Time-averaged (but not baseline-averaged) power-spectrum measurements of autocorrelation data, as a function of delay. The data products used in this plot are a pair of baselines, each belonging to the same redundant group. For example, this redundant group contain autocorrelations, e.g. [(176, 176, 'xx'), (176, 176, 'xx')] or [(162, 162, 'xx'), (178, 178, 'xx')]. Data sets with an 'auto' baseline pair construction only contain baselines paired with itself, e.g. [(176, 176, 'xx'), (176, 176, 'xx')]. Data sets with a 'cross' baseline pair construction contain baseline pairs with all other baseline pairs in the same redundant group, except for itself, e.g. [(162, 162, 'xx'), (178, 178, 'xx')]. Frequency range: [144–169] MHz. LST range: [23.9–2.4], which corresponds to the galaxy setting in the horizon of HERA.

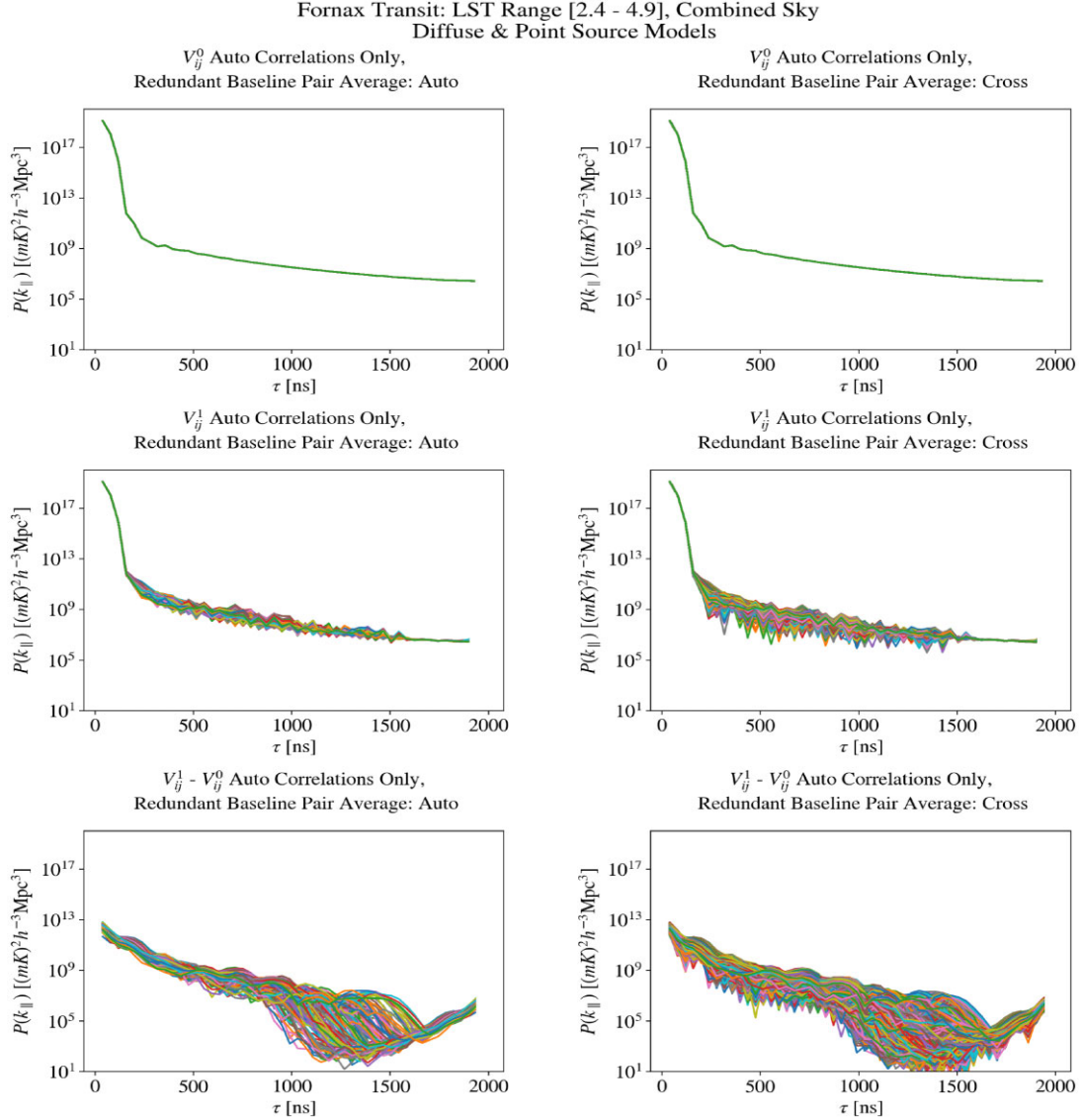
if not mitigated, may pose a critical impasse for a HERA-like interferometer to make a detection of the EoR.

When comparing the first column to the second column in Fig. 16, or also making the same comparison in Fig. 17, we note that the lower-delay 'inverse wedge' feature associated with  $\max(\tau_{|b_{ki}|} + \tau_{V_{kj, \text{main lobe}}}^0 + \tau_{\text{beam}})$ , reduces in power when averaging all cross-baseline pairs, compared to only averaging autobaseline pairs. This is expected behaviour, because intuitively we consider different baselines to receive incident astrophysical radiation with different phases from one another, even if all baselines belong to the same redundant group, by the fact that the baselines are physically located at different points in the array. It is also intuitive that the coupling features in Fig. 16 are brighter than those in Fig. 17 because the apparent brightness of astrophysical radiation at the horizon is brighter in the LST range associated with the setting of the galaxy

than it is in the LST range associated with the transit of Fornax A; see Fig. 8.

### 6.3 Delay-based coupling mitigation strategy

There exist subsets of our analysis space where first-order power spectra are consistent with their zeroth-order counterparts down to at least 10 orders of magnitude in sensitivity. In Figs 16 and 17, this region exists for each baseline group at delays greater than the maximum extent of case 4 of equation (23); this foreground-free, coupling-free region is denoted in Fig. 18 as delays greater than the yellow line. This region could be critical to any future coupling mitigation strategy, although further analysis of simulated data, and comparing said data to observation, is required to confirm whether



**Figure 15.** Time-averaged (but not baseline-averaged) power-spectrum measurements of autocorrelation data, as a function of delay. The data products used in this plot are a pair of baselines, each belonging to the same redundant group. For example, this redundant group contain autocorrelations, e.g. [(176, 176, 'xx')], [(162, 162, 'xx')], [(178, 178, 'xx')]. Data sets with an 'auto' baseline pair construction only contain baselines paired with itself, e.g. [(176, 176, 'xx')], [(176, 176, 'xx')]. Data sets with a 'cross' baseline pair construction contain baseline pairs with all other baseline pairs in the same redundant group, except for itself, e.g. [(162, 162, 'xx')], [(178, 178, 'xx')]. Frequency range: [144–169] MHz. LST range: [2.4–4.9], which captures Fornax A, a particularly bright point source, transiting the HERA beam. In this LST range, the HERA sky response at most (if not all) baselines is dominated by the point source model.

such a strategy will mitigate first-order coupling effects to levels below the expected power of the 21cm EoR signal.

## 7 DISCUSSIONS

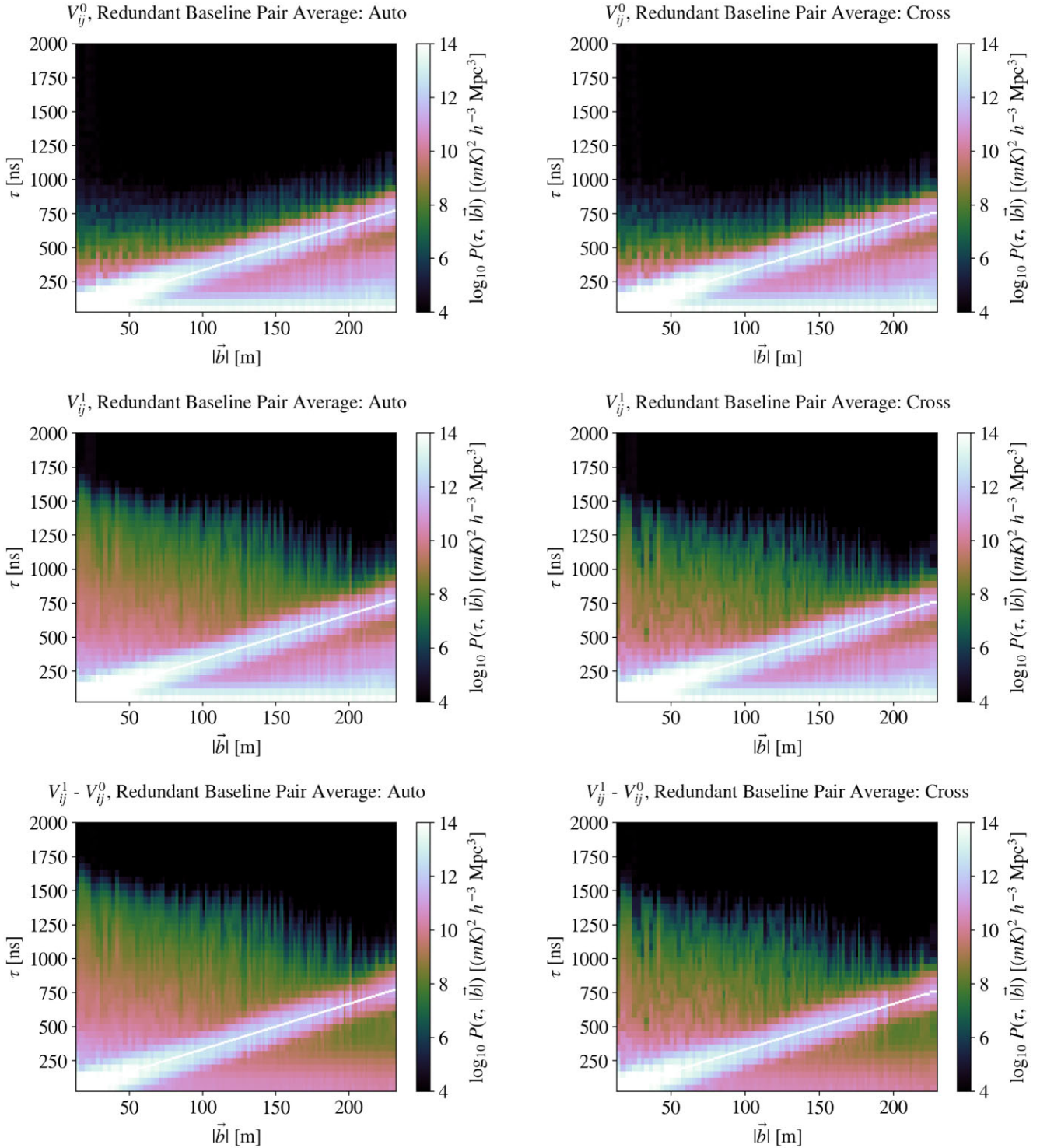
### 7.1 The strengths and weakness of a semi-analytic coupling model, motivating the use of embedded element patterns

The steady-state formalism described above considers the effect of first-order antenna–antenna coupling in interferometric visibilities. Incident astrophysical radiation enters all interferometric elements in the array. Each element absorbs most of its incident radiation, although some amount of power is then reflected due to departure

from conjugate match at the terminals of the antenna feeds. Subsequently, each element in the array reradiates power across the array. By being a first-order formalism, we only model this coupling effect on visibilities when reradiated power is absorbed by just one more element, not when the transmitted power is subsequently absorbed and reradiated by other elements. Such a model produces coupling systematics which extend beyond the wedge in baseline length versus delay space which is thought to contain most zeroth-order astrophysical foregrounds, at both zero and non-zero fringe rates.

The benefit of limiting our analysis to first-order antenna–antenna coupling is that we are able to straightforwardly produce a closed-form expression for coupled visibilities, e.g. equation (19). Upon

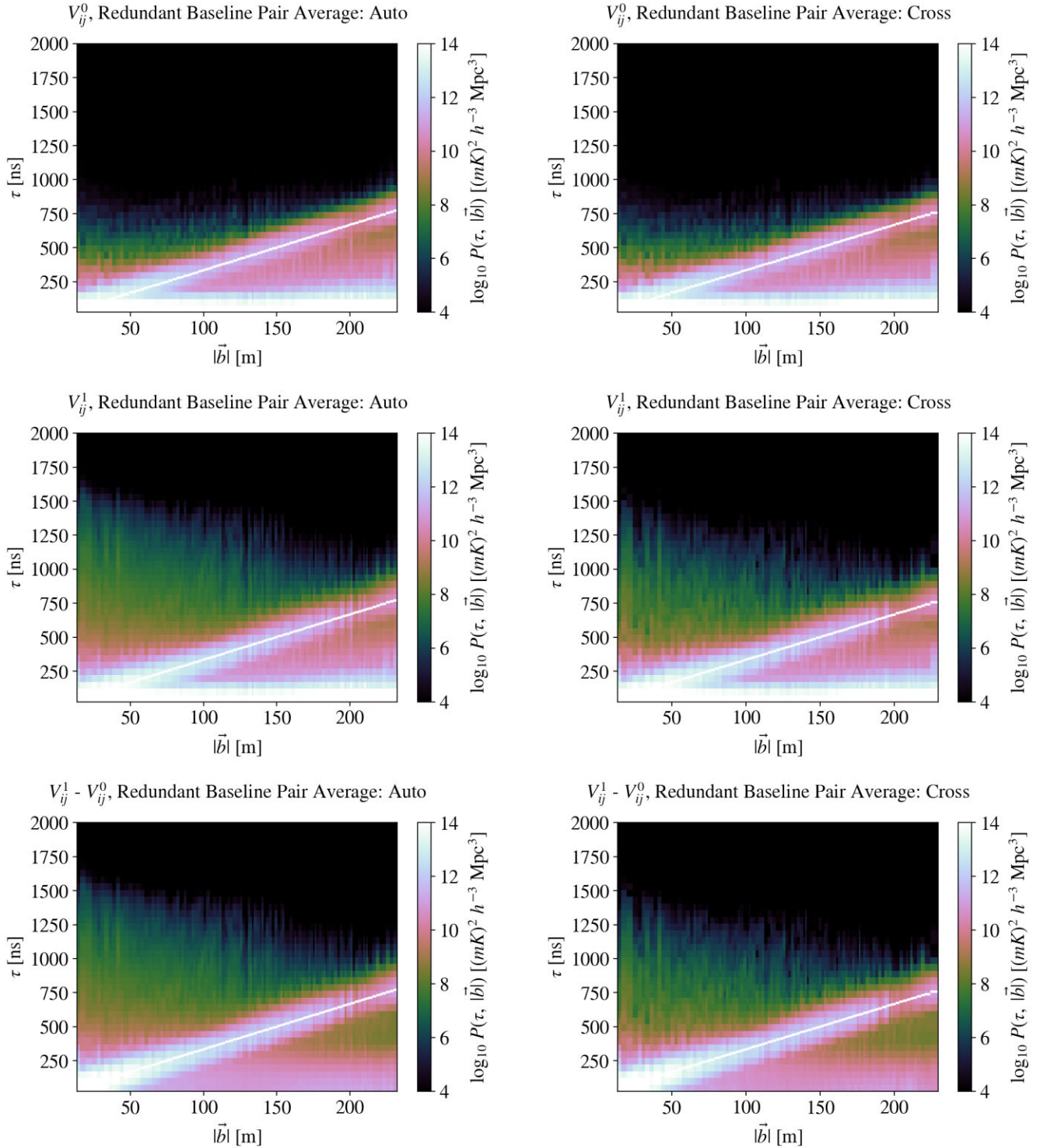
Galactic Set: LST Range [23.9–2.4], Combined Sky  
Diffuse & Point Source Models



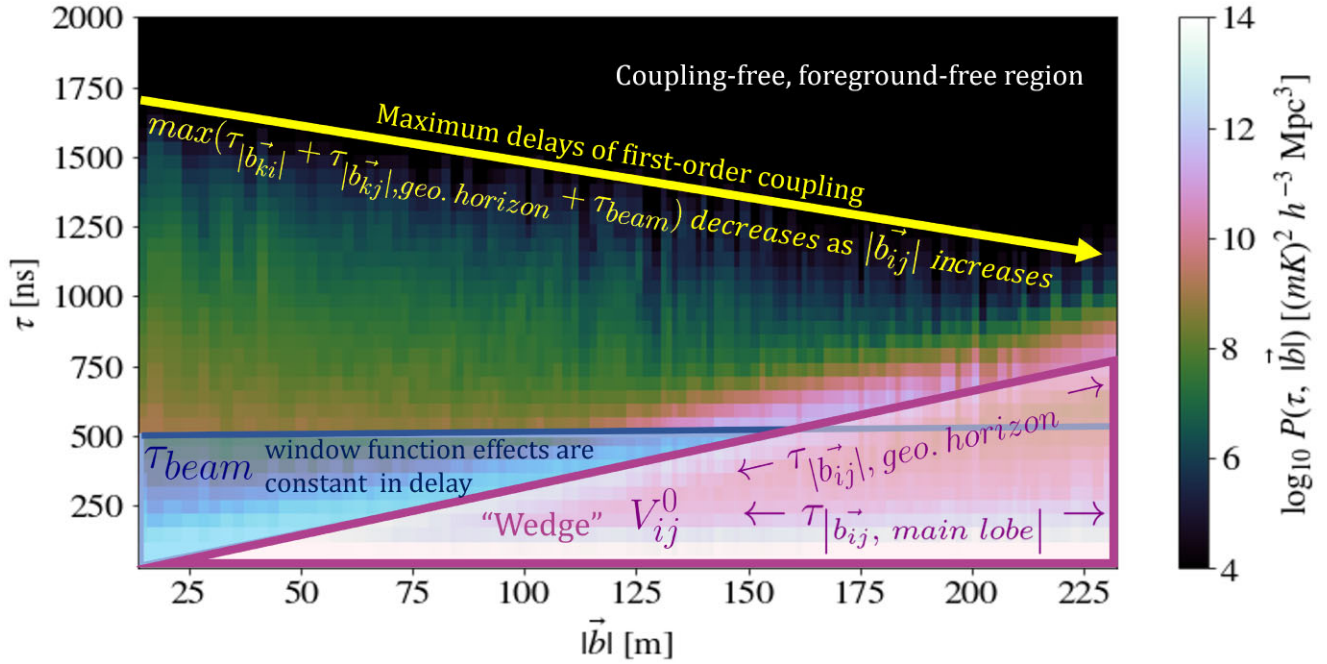
**Figure 16.** Time-averaged, baseline-averaged, power-spectrum measurements as a function of baseline length and delay. The data products used in this plot are a pair of baselines, each belonging to the same redundant group. For example, the redundant group of 28 m E/W baselines would contain the following pairs: [(176, 178, ‘xx’), (176, 178, ‘xx’)] and [(162, 164, ‘xx’), (176, 178, ‘xx’)]. Data sets with a redundant baseline pair average of ‘auto’ are constructed by averaging all baseline pairs in a redundant group with only itself, e.g. [(176, 178, ‘xx’), (176, 178, ‘xx’)]. Data sets with a redundant baseline pair average of ‘cross’ are constructed by averaging a baseline pair with all other baseline pairs in the same redundant group, except for itself, e.g. [(162, 164, ‘xx’), (176, 178, ‘xx’)]. Frequency range: [144–169] MHz. LST range: [23.9–2.4], which corresponds to the galaxy setting in the horizon of HERA.



Fornax Transit: LST Range [2.4-4.9], Combined Sky  
Diffuse & Point Source Models



**Figure 17.** Time-averaged, baseline-averaged, power-spectrum measurements as a function of baseline length and delay. The data products used in this plot are a pair of baselines, each belonging to the same redundant group. For example, the redundant group of 28 m E/W baselines would contain the following pairs: [(176, 178, 'xx'), (176, 178, 'xx')] and [(162, 164, 'xx'), (176, 178, 'xx')]. Data sets with a redundant baseline pair average of 'auto' are constructed by averaging all baseline pairs in a redundant group with only itself, e.g. [(176, 178, 'xx'), (176, 178, 'xx')]. Data sets with a redundant baseline pair average of 'cross' are constructed by averaging a baseline pair with all other baseline pairs in the same redundant group, except for itself, e.g. [(162, 164, 'xx'), (176, 178, 'xx')]. Frequency range: [144–169] MHz. LST range: [2.4–4.9], which captures Fornax A, a particularly bright point source, transiting the HERA beam. In this LST range, the HERA sky response at most (if not all) baselines is dominated by the point source model.



**Figure 18.** A conceptual outline of the different features of a wedge power spectrum containing first-order visibilities, as calculated using the simplified version of our coupling formalism, presented in equation (20). Section 5.2 describes how first-order coupling features not only manifest themselves at delays corresponding to zeroth-order visibilities (e.g.  $\tau_{|b_{ij}|}$  and  $\tau_{\text{beam}}$ ), but at much longer delays, e.g.  $\max(\tau_{|b_{ki}|} + \tau_{|b_{kj}|, \text{geo. horizon}} + \tau_{\text{beam}})$ . This latter sum extends to delays beyond the ‘wedge’ which has been extensively studied in the literature and contains astrophysical foreground features in visibilities that do not consider first-order coupling. From equation (20), we note that first-order visibilities consist of copies of zeroth-order visibilities ( $V_{kj}^0$  and  $V_{ik}^0$ ) which are translated to the delays presented in the last two cases of equation (23). Superwedge features associated with both the main lobe and pitchfork effect at the geometric horizon of the copied visibilities manifest themselves at delays which are inversely proportional to the copied baseline lengths, giving these first-order coupling features a distinct ‘inverse wedge’ shape (yellow). For each baseline length, delays greater than the yellow line are free from first-order coupling effects. Fig. 12 offers insight as to why the relationship between baseline length and maximum delay of first-order coupling is inversely proportional. Fig. 13 and Section 5.2 describe the origin of  $\tau_{\text{beam}}$ .

deriving a closed-form expression, Section 3.2 of this work lays the framework for a simulation of zeroth-order visibilities which are coupled to first order.

While our first-order phenomenological analysis (Section 5) benefits from the fact that it can easily be simulated, we warn the reader that first-order antenna–antenna coupling effects cannot generally be considered the dominant mutual coupling effect. For example, note that the coupling terms in equation (19) are attenuated by baseline lengths ( $|\hat{b}_{ki}|$  and  $|\hat{b}_{kj}|$ ) and voltage reflection coefficients ( $\Gamma_k < 0, \forall k$ ). Depending on how conjugate mismatched a feed is to its terminals (essentially, by comparing the ratio of  $\Gamma_k$  to  $|\hat{b}_{ki}|$  and  $|\hat{b}_{kj}|$ ), a second-order coupling event of a short baseline could be stronger in magnitude than a first-order coupling event from a longer baseline. It is beyond the scope of this document to analyse how common a HERA-like array experiences second-order coupling effects which are stronger in magnitude than the first-order effects derived above.

To account for all orders of antenna–antenna coupling in the general visibility formalism, one could calculate the active array element pattern (also referred to as the embedded element pattern) of each element in the array. This data product, originally considered in Hannan (1964), is analogous to the ‘effective height’ terms in equations (1) and (2). In an active array element pattern, *all* orders of antenna coupling (not just first-order coupling) are considered by modifying the gain of an array element as a function of azimuth, zenith, and frequency. Assuming we had the computational resources to calculate these embedded element patterns (whether using com-

mercial electromagnetic simulation software, such as CST, or by means of a proprietary software), the sums present in equations (A10) and (A11) would not need to be explicitly considered, as they would be assumed into the embedded element pattern. In essence, the embedded element pattern would replace  $h_i^{px}(\hat{s})$  and  $h_j^{qy}(\hat{s})$ . It is beyond the scope of this work to derive a general visibility formalism which uses embedded element patterns, although we note that the concept of using embedded element patterns has previously been applied in other radio interferometric applications, such as Sutinjo et al. (2017).

## 7.2 Mutual coupling: a thread to redundant baseline averaging and to wedge foreground containment

Comparing Figs 10 and 11 to other baselines of the respective LSTs, which are found in the online supplementary material, it is clear that baselines with the same ENU coordinates experience first-order coupling differently; this is a function of where each specific baseline is located in the array. This non-redundancy in a baseline group is summarized in the power-spectrum measurements of Figs 14 and 15. Redundant arrays such as HERA (or its pathfinder project, PAPER), were commissioned with the assumption that data from baseline groups with identical ENU coordinates could be averaged together. Any introduction of non-redundancy in such a baseline group threatens the observation strategy, power-spectrum sensitivity requirements, and overall analysis pipeline of such an array.

Furthermore, the general analysis approach of interferometric arrays has focused on discriminating the EoR signal from spectrally smooth astrophysical foregrounds which are many orders of magnitude brighter. One prevailing, and reasonably conservative strategy for such foreground mitigation, discussed primarily in Parsons et al. (2012b) and Thyagarajan et al. (2015) is ‘avoidance’: One seeks to measure the EoR at delays and interferometric baseline lengths which are entirely distinct from the delays and baseline lengths associated with foregrounds. Such strategies discuss a ‘wedge’ of foreground space, such as those presented in the top rows of Figs 16 and 17. The ability to entirely avoid foreground contamination is threatened by first-order coupling features which extend beyond the zeroth-order wedge, and are conceptually described in Fig. 18. Other analyses, such as Kern et al. (2019), discuss filtering strategies for non-fringing coupling features, but such strategies are not sufficient for mitigating coupling features with non-zero fringe, such as those characterized in this work.

### 7.3 Array designs and observation strategies which mitigate coupling effects

First-order array element coupling effects are a function of LST, baseline length, and baseline orientation. Since it is a function of all three effects, a general strategy for filtering mutual coupling is not trivial and beyond the scope of this analysis. Nevertheless, the presented analysis motivates how our semi-analytic coupling model may be used by the reader to mitigate coupling systematics in existing radio interferometers and to design future arrays where the configuration of array elements inherently mitigates coupling effects at desired LSTs and/or angular resolutions on the sky.

While it is beyond the scope of this work to fully discuss how one could position the elements of an interferometric array in order to mitigate fringing and non-fringing coupling effects, we briefly outline a mitigation strategy for the observation of only one particular LST range. Suppose, for example, that an experiment requires an interferometric array to only observe the galactic centre. In this LST range, the apparent brightness of the sky in any array element would be at zenith. By equation (27), the principle first-order coupling features will be concentrated at fringe rates of  $f_r(0, 0) \approx \frac{bv}{c} \cos \alpha \cos \delta$ , which is proportional to a baseline’s projected E/W length ( $b \cos \alpha$ ). At this desired LST range, coupling effects would be entirely avoided if all baselines used in power-spectrum measurements were configured in a N/S orientation (e.g. row 4, column 3 of Fig. 11 and Figs S2, S6, and S8 of the online supplemental material).

Similarly, an observer could mitigate first-order coupling effects at a particular angular resolution on the sky (i.e. for one or a few interferometric baseline lengths) if he chose to only make measurements at LSTs where the apparent brightness of the sky in any array element were at zenith.

## 8 CONCLUSION

We derive a general, fully polarized, semi-analytic formalism for interferometric visibilities, which considers how impedance mismatch at the feed terminals of each element in an array will cause incident astrophysical radiation to reradiate and couple into all other elements in the array. Our model assumes steady-state, incident radiation. Since it is a first-order model, we only capture the effect of reradiation being absorbed by all other elements in the array just once (e.g. sky  $\rightarrow$  Antenna k  $\rightarrow$  Antenna i), not when reradiated power is subsequently absorbed and reradiated to all

elements multiple times (e.g. sky  $\rightarrow$  Antenna k  $\rightarrow$  Antenna i  $\rightarrow$  Antenna j). We simulate first-order coupling features on a HERA-like redundant array using non-polarized skies with diffuse and point-source emission; a phenomenological analysis of the coupling features is presented. Contrary to previous studies, we find mutual coupling features manifest themselves at non-zero fringe rates. We compare power-spectrum results for both coupled and non-coupled (noiseless, simulated) data and find coupling effects to be highly dependent on LST, baseline length, and baseline orientation. For all LSTs, lengths, and orientations, coupling features appear at delays which are outside the foreground ‘wedge’ containing non-coupled astrophysical foreground features. Further, we find that first-order coupling effects threaten our ability to average data from baselines with identical length and orientation. We propose two filtering strategies, one in delay space and the other in fringe-rate space, which may mitigate such coupling systematics. The semi-analytic coupling model herein presented may be used to study mutual coupling systematics as a function of LST, baseline length, and baseline orientation. Such a model is not only helpful to the field of 21cm cosmology, but any study involving interferometric measurements, where coupling effects at the level of at least 1 part in  $10^4$  could corrupt the scientific result. Our model may be used to mitigate coupling systematics in existing radio interferometers and to design future arrays where the configuration of array elements inherently mitigates coupling effects at desired LSTs and angular resolutions.

## ACKNOWLEDGEMENTS

We thank James Aguirre, Jacqueline Hewitt, Nicholas Kern, Daniel Jacobs, and Vincent MacKay for discussions which directly contributed to this work. Our compliments to Bang Nhan and Honggeun Kim for the high-precision resimulation of Nicolas Fagnoni’s CST models of the HERA Vivaldi feed, which were used to construct all beam products analysed in this paper. This material is based upon work supported by the National Science Foundation under grant nos. 1636646 and 1836019 and institutional support from the HERA collaboration partners. This research is funded in part by the Gordon and Betty Moore Foundation. HERA is hosted by the South African Radio Astronomy Observatory, which is a facility of the National Research Foundation, an agency of the Department of Science and Innovation.

## SOFTWARE AND DATA AVAILABILITY

A demonstration of how to apply the first-order coupling formalism to interferometric visibility data may be found in the first author’s public Github repository.<sup>3</sup> The healvis visibility simulator may be found in Lanman & Kern (2019).<sup>4</sup> Simulated visibilities (both zero and first order) are output as uvh5 data objects, which may be conveniently read and accessed via pyuvdata (Hazelton et al. 2017).<sup>5</sup> Running this analysis on simulated HERA-like data products was made substantially more convenient using the ‘i/o’ functions and data containers defined in the hera\_cal repository.<sup>6</sup> We use the hera\_pspec power spectrum estimator to calculate all power spectra herein

<sup>3</sup><https://github.com/alphabetangojuliett/CoupledRadioInterferometer>

<sup>4</sup><https://github.com/rasg-affiliates/healvis>

<sup>5</sup><https://github.com/RadioAstronomySoftwareGroup/pyuvdata>

<sup>6</sup>[https://github.com/HERA-Team/hera\\_cal](https://github.com/HERA-Team/hera_cal)



presented.<sup>7</sup> Additional plots of coupled and non-coupled visibilities, as a function of fringe rate versus delay, may be found in the online supplemental material. Simulated data of the electromagnetic properties of HERA, including CST beam products and reflection coefficients, may be requested via email from the authors.

## REFERENCES

- Ali Z. S. et al., 2015, *ApJ*, 809, 61
- Beardsley A. P. et al., 2016, *ApJ*, 833, 102
- Bevins H. T. J., Handley W. J., Fialkov A., de Lera Acedo E., Greenhill L. J., Price D. C., 2021, *MNRAS*, 502, 4405
- Bowman J. D. et al., 2013, *Publ. Astron. Soc. Aust.*, 30, e031
- Bowman J. D., Rogers A. E. E., Monsalve R. A., Mozden T. J., Mahesh N., 2018, *Nature*, 555, 67
- Byrne R., Morales M. F., Hazelton B., Sullivan I., Barry N., Lynch C., Line J. L. B., Jacobs D. C., 2022, *MNRAS*, 510, 2011
- Chaudhari S. C., Gupta Y., Kumar A., Shinde N. D., Gupta S., Vishwakarma A., 2017, *J. Astron. Instrum.*, 6, 1641017
- Clark B. G., 1999, in Taylor G. B., Carilli C. L., Perley R. A., eds, ASP Conf. Ser. Vol. 180, Synthesis Imaging in Radio Astronomy II. Astron. Soc. Pac., San Francisco, p. 1
- Craeye C., González-Ovejero D., 2011, *Radio Sci.*, 46, RS2012
- Datta A., Bowman J. D., Carilli C. L., 2010, *ApJ*, 724, 526
- DeBoer D. R. et al., 2017, *PASP*, 129, 045001
- Dillon J. S. et al., 2020, *MNRAS*, 499, 5840
- Ewall-Wice A. et al., 2016, *MNRAS*, 460, 4320
- Fagnoni N. et al., 2021, *MNRAS*, 500, 1232
- Furlanetto S. R., Oh S. P., Briggs F. H., 2006, *Phys. Rep.*, 433, 181
- Hannan P., 1964, *Proc. IEEE*, 12, 423
- Hazelton B. J., Jacobs D. C., Pober J. C., Beardsley A. P., 2017, *J. Open Source Soft.*, 2, 140
- Hibbard J., Tauscher K., Rapetti D., Burns J., 2020, American Astronomical Society Meeting Abstracts, 236, 340.01
- Hills R., Kulkarni G., Meerburg P. D., Puchwein E., 2018, *Nature*, 564, E32
- Jelić V., Zaroubi S., Labropoulos P., Bernardi G., de Bruyn A. G., Koopmans L. V. E., 2010, *MNRAS*, 409, 1647
- Jenn D. C., Flokas V., 1994, *Proc. IEEE*, 3, 1870
- Kern N. S., Parsons A. R., Dillon J. S., Lanman A. E., Fagnoni N., de Lera Acedo E., 2019, *ApJ*, 884, 105
- Kern N. S. et al., 2020, *ApJ*, 888, 70
- Lanman A. E., Kern N., 2019, Astrophysics Source Code Library, record ascl:1907.002
- Liu A., 2020, <http://reionization.org/science/memos/>, Last accessed 1 April 2022
- Liu A., Shaw J. R., 2020, *PASP*, 132, 062001
- Madau P., Meiksin A., Rees M. J., 1997, *ApJ*, 475, 429
- Martinot Z. E., Aguirre J. E., Kohn S. A., Washington I. Q., 2018, *ApJ*, 869, 79
- Mertens F. G. et al., 2020, *MNRAS*, 493, 1662
- Mesinger A., 2016, Astrophysics and Space Science Library, Vol. 423, Understanding the Epoch of Cosmic Reionization: Challenges and Progress. Springer, Berlin
- Moore D. F., Aguirre J. E., Parsons A. R., Jacobs D. C., Pober J. C., 2013, *ApJ*, 769, 154
- Morales M. F., Hazelton B., Sullivan I., Beardsley A., 2012, *ApJ*, 752, 137
- Orfanidis S. J., 2013, Electromagnetic Waves and Antenna. Rutgers University, NJ, USA
- Parsons A. R., Backer D. C., 2009, *AJ*, 138, 219
- Parsons A., Pober J., McQuinn M., Jacobs D., Aguirre J., 2012a, *ApJ*, 753, 81
- Parsons A. R., Pober J. C., Aguirre J. E., Carilli C. L., Jacobs D. C., Moore D. F., 2012b, *ApJ*, 756, 165
- Parsons A. R., Liu A., Ali Z. S., Cheng C., 2016, *ApJ*, 820, 51
- Pritchard J. R., Loeb A., 2012, *Rep. Prog. Phys.*, 75, 086901
- Remazeilles M., Dickinson C., Banday A. J., Bigot-Sazy M. A., Ghosh T., 2015, *MNRAS*, 451, 4311
- Sims P. H., Pober J. C., 2020, *MNRAS*, 492, 22
- Smirnov O. M., 2011, *A&A*, 527, A106
- Sutinjo A. T., Ung D. C. X., Colegate T. M., Wayth R. B., Hall P. J., de Lera Acedo E., 2017, *Proc. IEEE*, 65, 3967
- The HERA Collaboration et al., 2022, *ApJ*, 925, 221
- Thompson A. R., Moran J. M., Swenson George W. J., 2017, Interferometry and Synthesis in Radio Astronomy, 3rd edn. Springer, Berlin
- Thyagarajan N. et al., 2015, *ApJ*, 804, 14
- Thyagarajan N., Parsons A. R., DeBoer D. R., Bowman J. D., Ewall-Wice A. M., Neben A. R., Patra N., 2016, *ApJ*, 825, 9
- Ung D. C. X., Sokolowski M., Sutinjo A. T., Davidson D. B., 2020, *Proc. IEEE*, 68, 5395
- Van de Hulst H. C., 1945, Ned. Tijd. Natuurkunde, 11, 210
- van Haarlem M. P. et al., 2013, *A&A*, 556, A2
- Visbal E., Loeb A., Wyithe S., 2009, *J. Cosmol. Astropart. Phys.*, 2009
- Wang X., Tegmark M., Santos M. G., Knox L., 2006, *ApJ*, 650, 529
- Zheng H. et al., 2014, *MNRAS*, 445, 1084

## SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

**Figure S1.** Fringe rate versus delay plot of the first-order visibility of 28m E/W baseline (162, 164) of the HERA array, plotted in green, as well all other elements (grey) which were installed on JD 2459122.

**Figure S2.** Fringe rate versus delay plot of the first-order visibility of baseline (162, 164) of the HERA array, plotted in green, as well all other elements (grey) which were installed on JD 2459122.

**Figure S3.** Fringe rate versus delay plot of the first-order visibility of 17m N/S baseline (143, 163) of the HERA array, plotted in hot pink, as well all other elements (grey) which were installed on JD 2459122.

**Figure S4.** Fringe rate versus delay plot of the first-order visibility of 17m N/S baseline (143, 163) of the HERA array, plotted in hot pink, as well all other elements (grey) which were installed on JD 2459122.

**Figure S5.** Fringe rate versus delay plot of the first-order visibility of 28m E/W baseline (128, 130) of the HERA array, plotted in amber, as well all other elements (grey) which were installed on JD 2459122.

**Figure S6.** Fringe rate versus delay plot of the first-order visibility of 28m E/W baseline (128, 130) of the HERA array, plotted in amber, as well all other elements (grey) which were installed on JD 2459122.

**Figure S7.** Fringe rate versus delay plot of the first-order visibility of 17m N/S baseline (135, 155) of the HERA array, plotted in hot pink, as well all other elements (grey) which were installed on JD 2459122.

**Figure S8.** Fringe rate versus delay plot of the first-order visibility of 17m N/S baseline (135, 155) of the HERA array, plotted in hot pink, as well all other elements (grey) which were installed on JD 2459122.

**Figure S9.** Fringe rate versus delay plot of the first-order visibility of a baseline with both E/W and N/S components (100, 135) of the HERA array, plotted in hot pink, as well all other elements (grey) which were installed on JD 2459122.

**Figure S10.** Fringe rate versus delay plot of the first-order visibility of a baseline with both E/W and N/S components (100, 135) of the HERA array, plotted in hot pink, as well all other elements (grey) which were installed on JD 2459122.

**Figure S11.** Fringe rate versus delay plot of the first-order visibility of a baseline with both E/W and N/S components (109, 150) of the

<sup>7</sup>[https://github.com/HERA-Team/hera\\_pspec](https://github.com/HERA-Team/hera_pspec)

HERA array, plotted in hot pink, as well all other elements (grey) which were installed on JD 2459122.

**Figure S12.** Fringe rate versus delay plot of the first-order visibility of a baseline with both E/W and N/S components (109, 150) of the HERA array, plotted in hot pink, as well all other elements (grey) which were installed on JD 2459122.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

## APPENDIX A: MOTIVATING THE STANDARD VISIBILITY EQUATION

Before establishing a visibility equation, we begin by discussing the propagation of radiation from a single astrophysical source to an interferometric array on Earth. By the linearity of Maxwell's equations, we may superimpose the electric fields at all measurement locations,  $b_i$ , by the source point. The propagation of this electric field from the source to a test point can be characterized by some function  $P(\vec{R}, \vec{b}_i)$ :

$$\int \int \int P(\vec{R}, \vec{b}_i) \vec{E}(\vec{R}) dxdydz. \quad (A1)$$

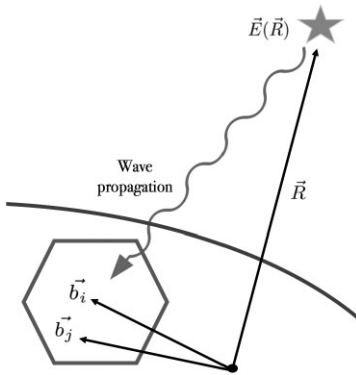
We now make our first assumption, Assumption A, which is both logical and also serves as the basis for standard visibility derivations, such as Clark (1999). Namely, we assume that the astrophysical sources are very far away from the array, so that  $|\vec{R}| \gg |\vec{b}_i|$ . In this case, all we can hope to characterize is the surface brightness of a source, and information about brightness along the radial axis is lost. We imagine a ‘celestial sphere’ of radius  $|\vec{R}|$ , and focus on integrating the electric field surface distribution,  $\vec{E}(\vec{R})$ .

We now introduce Assumption B, that the space within the celestial sphere is empty. Thus, by Huygen's Principle for spherical wave propagation, the propagation function of equation (A1) is

$$P(\vec{R}, \vec{b}_i) = e^{2\pi i \frac{v}{c} |\vec{R} - \vec{b}_i|}. \quad (A2)$$

Thus, the electric field at interferometric element  $i$ , as measured by integrating over surface area elements  $ds$ , is

$$\vec{E}(\vec{R}, \vec{b}_i) = \int \vec{E}(\vec{R}) \frac{e^{2\pi i \frac{v}{c} |\vec{R} - \vec{b}_i|}}{|\vec{R} - \vec{b}_i|} ds. \quad (A3)$$



**Figure A1.** Electric field propagation from astrophysical source to two elements of an interferometric array. Both the source and the interferometric elements in the array are referenced to an arbitrary ground point.

We now imagine that each interferometric element consists of an electromagnetically sensitive structure, such as a feed, or a dish with a feed near the dish's optical focus. The structure has two orthogonal feed polarizations,  $p$  and  $q$ . Assuming that the electric field of the astrophysical sources propagates at a frequency which is in the electromagnetically sensitive band of each feed polarization, then the open-circuit voltage of the  $p$ -oriented feed of the  $i^{th}$  element is

$$v_i^p = \vec{h}_i^p \cdot \vec{E}. \quad (A4)$$

Let us represent our array's interferometric elements in terms of local, Cartesian, ENU coordinates,  $(x, y, \text{ and } z)$  and the electric field components in terms of orthogonal spherical coordinates  $(r, \theta, \text{ and } \phi)$ , where (in radians)  $0 \leq \theta \leq \pi$  and  $0 \leq \phi \leq 2\pi$ . Recall that  $\hat{\theta}$  and  $\hat{\phi}$  are tangent to the celestial sphere. Fig. 2 graphically represents this system, whose relationship to Cartesian coordinates is

$$\hat{r} = \cos(\phi) \sin(\theta) \hat{x} + \sin(\phi) \sin(\theta) \hat{y} + \cos(\theta) \hat{z}, \quad (A5)$$

$$\hat{\theta} = \cos(\phi) \cos(\theta) \hat{x} + \sin(\phi) \cos(\theta) \hat{y} - \sin(\theta) \hat{z}, \quad (A6)$$

$$\hat{\phi} = -\sin(\phi) \hat{x} + \cos(\phi) \hat{y}. \quad (A7)$$

Noting that our differential surface area element  $ds$  points in the radial direction,  $\hat{s} = \frac{\vec{R}}{|\vec{R}|}$ , we can integrate over a differential solid angle,  $d\Omega$ , where  $ds = R^2 d\Omega$ . Thus, the voltage response at the terminal of the  $p^{th}$  pol of the  $i^{th}$  interferometric element, caused by the surface distribution of the  $\hat{\phi}$  basis vector component of the electric field is

$$v_i^{p\phi} = \int h_i^{p\phi}(\hat{s}) \varepsilon^\phi(\hat{s}) \frac{e^{2\pi i \frac{v}{c} |\vec{R}\hat{s} - \vec{b}_i|}}{|\vec{R}\hat{s} - \vec{b}_i|} R^2 d\Omega. \quad (A8)$$

Similarly, the voltage response at the terminal of the  $q^{th}$  pol of the  $j^{th}$  interferometric element, caused by the surface distribution of the  $\hat{\theta}$  basis vector component of the electric field is

$$v_j^{q\theta} = \int h_j^{q\theta}(\hat{s}') \varepsilon^\theta(\hat{s}') \frac{e^{2\pi i \frac{v}{c} |\vec{R}\hat{s}' - \vec{b}_j|}}{|\vec{R}\hat{s}' - \vec{b}_j|} R^2 d\Omega'. \quad (A9)$$

The above voltage responses do not consider reflections of incident radiation due to impedance mismatch at each feed. This work considers first-order reradiation of the astrophysical electric field, which arrives at the  $i^{th}$  and  $j^{th}$  elements after being previously reradiated by only one other element  $k$  (and not more than one element). Because the array may be located generally on the geographic (2D) plane, not simply in a (1D) line, the scattered electric field (from the  $k^{th}$  element), upon arrival to any instrumental polarization of all other elements, can be a function of both instrumental polarizations of the radiating ( $k^{th}$ ) element. Consequently, the scattered electric field can be a function of both incident astrophysical electric field basis vectors. We may generally define the electric field surface distribution which is reradiated by a  $k^{th}$  element, as received by the  $i^{th}$  element, as  $\varepsilon^{s\phi, p\theta}(\hat{s}, \hat{b}_{ki})$ . The first superscript (e.g. ‘ $s\phi$ ’) refers to  $\hat{\phi}$  component of what is received by the  $i^{th}$  element from the radiating  $k^{th}$  element. After the comma in the superscript (e.g.  $p\theta$ ), we denote the incident electric field basis vector (e.g.  $\hat{\theta}$ ) contribution to each instrument polarization (e.g.  $p$ ) of the radiating  $k^{th}$  element. A similar, general expression is used for the electric field distribution which is scattered from the  $k^{th}$  element to the  $j^{th}$  element, e.g.  $\varepsilon^{s\theta, q\phi}(\hat{s}, \hat{b}_{kj})$ .

To summarize, the voltage response at the terminal of the  $p^{th}$  pol of the  $i^{th}$  interferometric element, caused by the  $\hat{\phi}$  basis vector component of the electric field distribution which was reradiated from the  $k^{th}$  element may contain components from both instrumental polarizations and both incident electric field basis vectors of the  $k^{th}$

element. Similarly, the voltage response at the terminal of the  $q^{th}$  pol of the  $j^{th}$  interferometric element, caused by the  $\hat{\theta}$  basis vector component of the electric field distribution which was reradiated from the  $k^{th}$  element may contain components from both instrumental polarizations and both incident electric field basis vectors of the  $k^{th}$  element. We formalize this model of antenna-to-antenna coupling by rewriting equations (A8) and (A9):

$$v_i^{p\phi} = \int h_i^{p\phi}(\hat{s}) \varepsilon^\phi(\hat{s}) \frac{e^{2\pi i \frac{v}{c} |R\hat{s} - \vec{b}_i|}}{|R\hat{s} - \vec{b}_i|} R^2 d\Omega \\ + \sum_{k \neq i} \int \frac{e^{2\pi i \frac{v}{c} |R\hat{s} - \vec{b}_k|}}{|R\hat{s} - \vec{b}_k|} \left[ \varepsilon^{s\phi, p\phi}(\hat{s}', \hat{b}_{ki}) + \varepsilon^{s\phi, p\theta}(\hat{s}', \hat{b}_{ki}) \right. \\ \left. + \varepsilon^{s\theta, q\phi}(\hat{s}', \hat{b}_{ki}) + \varepsilon^{s\theta, q\theta}(\hat{s}', \hat{b}_{ki}) \right] h_i^{p\phi}(\hat{b}_{ik}) R^2 d\Omega. \quad (A10)$$

$$v_j^{q\theta} = \int h_j^{q\theta}(\hat{s}) \varepsilon^\theta(\hat{s}) \frac{e^{2\pi i \frac{v}{c} |R\hat{s} - \vec{b}_j|}}{|R\hat{s} - \vec{b}_j|} R^2 d\Omega \\ + \sum_{k \neq j} \int \frac{e^{2\pi i \frac{v}{c} |R\hat{s} - \vec{b}_k|}}{|R\hat{s} - \vec{b}_k|} \left[ \varepsilon^{s\phi, p\phi}(\hat{s}', \hat{s}'_{kj}) + \varepsilon^{s\phi, p\theta}(\hat{s}', \hat{s}'_{kj}) \right. \\ \left. + \varepsilon^{s\theta, q\phi}(\hat{s}', \hat{s}'_{kj}) + \varepsilon^{s\theta, q\theta}(\hat{s}', \hat{s}'_{kj}) \right] h_j^{q\theta}(\hat{s}'_{jk}) R^2 d\Omega. \quad (A11)$$

## APPENDIX B: COMPATIBILITY OF HERA BEAM WITH ASSUMPTIONS IN THE GENERAL FORMALISM

### B1 Satisfaction of far-field limit in HERA's horizon plane

To simulate antenna-to-antenna scattering in a HERA-like array using the general formalism, we must first establish that the far-field approximation used in equation (10) is valid. In HERA, each of the Vivaldi feeds (which contain the feed terminals) is positioned above the rims of each dish. HERA is a drift-scan array with the peak beam gain of all feeds oriented directly towards zenith. Conventionally, one would consider the far field limit of radiation coming from zenith. This is guaranteed for incident astrophysical radiation by Assumption A. However, in our coupling formalism, the scattered radiation comes exclusively from the horizon plane (e.g.  $\hat{b}_{ik}$ ,  $\hat{b}_{kj}$ ). Thus, we must ensure that far field limit is not only satisfied for incident astrophysical radiation (coming from zenith), but is also satisfied for scattered radiation along the horizon direction, emitted at distances as short as 14 m. It can be shown by analysing CST electromagnetic simulations of the beam that, in the horizon plane, the beam gain of a Vivaldi feed with and without the surrounding dish are nearly identical. Our result implies that, in the horizon plane (but not in the zenith direction), the feed alone will be the dominant radiator of a scattered electric field due to impedance mismatch at the feed terminals.

Let us now ensure that this feed-to-feed radiation satisfies the far-field limit. We consider the far-field limit to be valid at distances greater than the Fraunhofer distance,  $d_F$ , where

$$d_F = \frac{2D^2}{\lambda}. \quad (B1)$$

HERA's Vivaldi feeds have a diameter approximately equal to  $D \approx 2$  m. HERA operates in the range of 50–250 MHz, corresponding to  $\lambda \in [1.2, 6]$  m, separated by a minimum of 14 m. Thus, the far-field limit is practically satisfied for the reradiated electric field of all  $\hat{b}_{ki}$  baseline directions considered in equation (19), allowing us to apply the general formalism to HERA.

### B2 Relating electromagnetic (EM) simulation (CST) outputs to formalism

Using the definition of our Jones matrix in terms of effective height (equation 6), we calculate

$$\mathbf{J}(\hat{s})\mathbf{J}^y(\hat{s}) = \begin{bmatrix} |h_p^\phi(\hat{s})|^2 + |h_p^\theta(\hat{s})|^2 & h_p^\phi h_q^{*\phi}(\hat{s}) + h_p^\theta h_q^{*\theta}(\hat{s}) \\ h_q^\phi h_p^{*\phi}(\hat{s}) + h_q^\theta h_p^{*\theta}(\hat{s}) & |h_q^\phi(\hat{s})|^2 + |h_q^\theta(\hat{s})|^2 \end{bmatrix}. \quad (B2)$$

Typically, this beam term is normalized so that the power at some specific, peak point is 1. For example, to normalize such that the power (per polarization  $p$ ) at zenith is 1, we stipulate that  $1 = |\mathbf{J}_p(\hat{s}_z)|^2$ , where  $\hat{s}_z$  is the unit vector pointing in the zenith direction. When we simulate the beam of a HERA interferometric element in CST (whether it be just the Vivaldi feed, or we also include the dish) there are two primary data products: First, a unitless directivity beam,  $D_p(\hat{s})$ , which may similarly be normalized at zenith. Second, an electric field beam,  $\mathbf{E}(\hat{s})$ , to be discussed below. Both beams use the spherical coordinate basis vectors represented in Fig. 2.

The coupling terms in the first-order visibility formalism contain an additional multiplication of Jones matrices, this time pointing in the direction of the baseline being considered. Namely,  $\mathbf{J}(\hat{b}_{ik})\mathbf{J}(\hat{b}_{ki})$  and  $\mathbf{J}^y(\hat{b}_{kj})\mathbf{J}^y(\hat{b}_{jk})$ . These terms do not simplify nicely into power beams, because they are not conjugate pairs. Rather, we have the multiplication of two complex numbers where none of the numbers, or both of the numbers, are conjugated. Physically, we interpret this behaviour as reradiation due to the antenna impedance at the feed terminals not being conjugate-matched to the rest of the analogue chain. Mathematically, this is why we see ‘double-dagger’ or ‘no-dagger’ terms multiplied by  $\Gamma$  (the voltage reflection coefficient) in equation (19).

It can be shown using orthographic projections of the HERA beams that, in the CST electric field basis vector convention, the  $\hat{\phi}$  component of the electric field remains unchanged regardless of whether the baseline vector points towards the radiating or receiving element,  $\mathbf{J}^\phi(\hat{b}_{ik}) = \mathbf{J}^\phi(\hat{b}_{ki})$ . It may similarly be shown that, when the baseline vectors reside in the horizon plane ( $\theta \approx \frac{\pi}{2}$ ), the  $\hat{\theta}$  component of the beam flips sign:  $\mathbf{J}^{\theta \approx \frac{\pi}{2}}(\hat{b}_{ik}) \approx -\mathbf{J}^{\theta \approx \frac{\pi}{2}}(\hat{b}_{ki})$ . Incorporating this sign flip for the  $\hat{\theta}$  basis vector components, the Jones matrices in the coupling terms may be explicitly written as

$$\mathbf{J}(\hat{b}_{ik})\mathbf{J}(\hat{b}_{ki}) = \begin{bmatrix} h_p^\phi h_p^\phi(\hat{b}_{ki}) - h_q^\phi h_p^\theta(\hat{b}_{ki}) & h_p^\theta h_q^\theta(\hat{b}_{ki}) - h_p^\theta h_p^\phi(\hat{b}_{ki}) \\ h_p^\phi h_q^\phi(\hat{b}_{ki}) - h_q^\phi h_q^\theta(\hat{b}_{ki}) & h_q^\theta h_q^\theta(\hat{b}_{ki}) - h_p^\theta h_q^\phi(\hat{b}_{ki}) \end{bmatrix}, \quad (B3)$$

$$\mathbf{J}^y(\hat{b}_{kj})\mathbf{J}^y(\hat{b}_{jk}) = \begin{bmatrix} h_p^{*\phi} h_p^{*\phi}(\hat{b}_{kj}) - h_q^{*\phi} h_p^{*\theta}(\hat{b}_{kj}) & h_q^{*\phi} h_p^{*\phi}(\hat{b}_{kj}) - h_q^{*\phi} h_q^{*\theta}(\hat{b}_{kj}) \\ h_p^{*\phi} h_q^{*\phi}(\hat{b}_{kj}) - h_p^{*\theta} h_p^{*\phi}(\hat{b}_{kj}) & h_q^{*\phi} h_q^{*\theta}(\hat{b}_{kj}) - h_p^{*\theta} h_q^{*\phi}(\hat{b}_{kj}) \end{bmatrix}. \quad (B4)$$

We must now relate these Jones matrix terms to the CST electric field and directivity data products, which contain, for all directions, the far-field electric field response in units ( $\frac{V}{m}$ ), at a reference distance 1 m away from the Vivaldi feed terminals. To be clear, the basis vectors output by CST for electric field and directivity data products are only ( $\hat{\phi}$ ) and ( $\hat{\theta}$ ), because ( $\hat{r}$ ) is fixed at 1 m for all points. For both feed polarizations, ( $p, q$ ), the CST-calculated electric field beam may be considered as a 2x2 matrix:

$$\mathbf{E}(\hat{s}) = \begin{bmatrix} E_p^\phi & E_p^\theta \\ E_q^\phi & E_q^\theta \end{bmatrix}. \quad (B5)$$

Standard antenna textbook references, such as Section 16.5 of Orfanidis (2013), relate per-polarization components of the effective height matrix,  $h_p(\hat{s})$ , to the per-polarization gain of the antenna,



$G_p(\hat{s})$ , where both  $\vec{h}$  and  $G_p$  are functions of  $\theta$  and  $\phi$ :

$$\frac{\eta_0 |\vec{h}_p(\hat{s})|^2}{4R_{\text{ant}}} = \frac{\lambda^2 G_p(\hat{s})}{4\pi}. \quad (\text{B6})$$

In our case, we seek to relate effective height to directivity, rather than gain, because directivity is another data product which is output by CST. It is well understood that gain is related to directivity by the relationship  $G_p(\hat{s}) = e_{\text{rad}} D_p(\hat{s})$ , where  $e_{\text{rad}}$  is the radiation efficiency. Plugging this relationship into equation (B6) and re-ordering terms,

$$\sqrt{\frac{\eta_0}{R_{\text{ant}}}} |\vec{h}_p(\hat{s})| = \sqrt{\frac{\lambda^2 e_{\text{rad}} D_p(\hat{s})}{\pi}}. \quad (\text{B7})$$

To be useful to the coupling formalism we use to make simulated visibilities, which is presented in equation (20), we must relate the direction-dependent effective height terms of equations (B3) and (B4) to directivity – in the direction of feed-to-feed scattering, which is the horizon plane ( $\hat{s} = \hat{b}_{ki}$ ).

$$\hat{b}_{ki} = \frac{\vec{h}_p(\hat{b}_{ki})}{|\vec{h}_p(\hat{b}_{ki})|} = \frac{\vec{E}^p(\hat{b}_{ki})}{|\vec{E}^p(\hat{b}_{ki})|}. \quad (\text{B8})$$

Multiplying both sides of equation (B7) by equation (B8), we derive the non-general result

$$\sqrt{\frac{\eta_0}{R_{\text{ant}}}} \vec{h}_p(\hat{b}_{ki}) = \frac{\vec{E}^p(\hat{b}_{ki})}{|\vec{E}^p(\hat{b}_{ki})|} \sqrt{\frac{\lambda^2 e_{\text{rad}} D_p(\hat{b}_{ki})}{\pi}}, \quad (\text{B9})$$

where  $|E_p(\hat{b}_{ki})| = \sqrt{|E_p^\phi(\hat{b}_{ki})|^2 + |E_p^\theta(\hat{b}_{ki})|^2}$ .

Recall that the CST far-field electric field response,  $\vec{E}^p(\hat{b}_{ki})$ , is calculated at a reference distance of 1 m away from the feed terminals. When we plug this reference value into our impedance-normalized Jones matrix terms, e.g.  $\sqrt{\frac{\eta_0}{R_{\text{ant}}}} h_p^\phi(\hat{b}_{ki})$  or  $\sqrt{\frac{\eta_0}{R_{\text{ant}}}} h_q^\theta(\hat{b}_{ki})$ , as calculated by equation (B9), all factors of  $R_{\text{ant}}$  and  $\eta_0$  cancel in the coupling formalism as presented in equation (20). Thus, while the coupling formalism is written in terms of antenna impedance, the CST data products used in our simulations do not need to be corrected for antenna impedance.

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.