Global CD/EoR Signal Detection with a Dense Digital Beamforming Array and Beyond

Junhua Gu\textsuperscript{1*}, Jingying Wang\textsuperscript{2†}, Huanyuan Shan\textsuperscript{2,3}, Qian Zheng\textsuperscript{2}, Quan Guo\textsuperscript{2}, Yan Huang\textsuperscript{1}, Kuanjun Li\textsuperscript{1}, Tianyang Liu\textsuperscript{2,3}

\textsuperscript{1}National Astronomical Observatories, Chinese Academy of Sciences, 20A Datun Road, Beijing 100101, China
\textsuperscript{2}Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, China
\textsuperscript{3}School of Astronomy and Space Science, University of Chinese Academy of Sciences, 19A Yuquan Road, Beijing 100049, China

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

The global neutral hydrogen 21 cm signal extracted from the all-sky averaged radio spectra is one of the signatures of the Cosmic Dawn and Epoch of Reionization (CD/EoR). The frequency-dependency of antenna beam patterns coupled with the strong foreground emission could introduce artificial spectral structures and cause false detection. A digital beamforming array could be potentially employed to form achromatic station beam patterns to overcome this problem. In this work, we discuss the method of forming achromatic beam patterns with a dense regular beamforming array to detect the global CD/EoR signal, covering topics including the array configuration, antenna weight optimization, and error estimation. We also show that based on the equivalence between a beamforming array and an interferometer, most antennas in the array can be removed by canceling redundant baselines. We present an example array design, optimize the antenna weights, and show the final array configuration by canceling redundant baselines. The performance of the example array is evaluated based on a simulation, which provides a positive indication towards the feasibility of detecting the CD/EoR signal using a dense digital beamforming array.

Key words: cosmology: observations – cosmology: dark ages, reionization, first stars – instrumentation: interferometers – methods: observational

1 INTRODUCTION

The Cosmic Dawn and Epoch of Reionization (CD/EoR) are early evolving stages of the universe predicted by today’s cosmology theory. The neutral baryon matter that formed in the recombination epoch was ionized again during CD/EoR, which is expected to be initiated by the first stars (e.g., Furlanetto et al. 2006). Meanwhile, the lacking of direct observational evidence makes CD/EoR a missing link in the evolving process of the universe. This topic has become one of the frontiers of modern observational cosmology and radio astronomy. The central concerns in CD/EoR detection include a series of questions: the details of the first stars, how the neutral hydrogen was ionized, how the halo accretion process modulated the ionization process, etc. The redshift range of the CD/EoR is expected to be about 6 – 27 (e.g., Morales & Wyithe 2010; Madau et al. 1997), during which the 21cm emission line produced by neutral atomic hydrogen is the essential tracer and can be detected in the meter-wave band due to the redshift effect.

There are mainly three paradigms of detection methods for observing the redshifted neutral hydrogen 21cm line as a tracer to the CD/EoR: (1) global averaged signal detection, (2) power spectra measuring, and (3) direct imaging observation. Among these three methods, detecting the global averaged 21cm signal requires the least observation time. Several experiments aiming to detect the sky-averaged CD/EoR neutral hydrogen 21 cm signal are being performed, including LEDA (the radiometer system Price et al. 2018; Bernardi et al. 2016; Spinelli et al. 2021), BIGHORNS (Sokolowski et al. 2015), EDGES (Mozdzen et al. 2016), SCI-HI (Voytek et al. 2014), and REACH (Cumner et al. 2022).

By now, some detection results have been obtained with the EDGES experiment. With the dipole-like antenna used in the EDGES experiment, an absorption feature (i.e., the “dip” feature) of $\delta T_{21} = -500^{+200}_{-500}$ mK was observed (Bowman et al. 2018). This value is much deeper than the prediction of standard cosmology models (e.g., Reis et al. 2021), and different and/or extra mechanisms are required to explain their observational results. Various of models have been proposed (e.g., Barkana 2018; Fialkov et al. 2018; Hirano & Bromm 2018). On the other hand, Singh et al. (2022), based on the data acquired with the SARAS 3 radiometer (Namibissan T. et al. 2021), argues that the extra deep dip found by Bowman et al. (2018) is not evidence for new astrophysics.

* E-mail: jhgu@nao.cas.cn
† E-mail: jywang@shao.ac.cn

© 2022 The Authors
or non-standard cosmology. More detailed cross-checks are required regarding global CD/EoR signal detection.

The frequency-dependency of the antenna beam pattern is one of the key systematics issues. As we have pointed out in our previous work Gu & Wang (2020) and also been mentioned in some other works (e.g., Vedantham et al. 2014; Bernardi et al. 2015; Anstey et al. 2021, 2022; Spinelli et al. 2022), the frequency-dependency of the antenna beam pattern brings extra spectral fluctuations or structures that can significantly bias the global CD/EoR signal detection. The artifacts caused by the frequency-dependent beam are not only determined by the beam pattern itself but also by the brightness temperature angular distribution of the sky. Theoretically, there are at least two methods to solve this problem. One is to mathematically model the antenna beam and the sky brightness temperature angular distribution and disentangle the instrument effect from the measured data (Anstey et al. 2021, 2022); another is to build an instrument, the beam pattern of which is approximately frequency independent in the interested frequency range. Most global CD/EoR detection experiments based on single antennas require the antenna to be optimized for broadband achromatic beam patterns.

Designing a single broadband antenna with frequency-independent beam patterns is challenging. On the other hand, with digital beamforming (DBF) technology, it is possible to make the beam pattern of an array of (usually identical) antennas independent of frequency. This idea was proposed by D’Hulst et al. (2020) and later improved in D’Hulst et al. (2021). The authors described the method of using the beamforming mode of the second station of the Long Wavelength Array (Dowell et al. 2017) to generate frequency-independent beam patterns to detect the global CD/EoR signal. A recent end-to-end simulation work by Price (2022) in the background of the SKA Engineering Development Array v2 (EDA2) reveals that an array with 10⁹ antennas may form beam patterns with sidelobe level ∼−50 dB, so that can be used as an alternative method to measure the global 21 cm signal.

Another series of similar but different methods of measuring the global CD/EoR signal – using a short-spacing interferometer – has been suggested (e.g., Singh et al. 2015; Presley et al. 2015; McKinley et al. 2020). This method utilizes the response of short baselines to the monopole component of the sky brightness temperature distribution.

Though DBF arrays and interferometers are regarded as different types of instruments, they share a common basic mathematical theory. For example, Ruigrok et al. (2017) pointed out that the auto-correlation of the output of a DBF array can be equivalently replaced by summing the weighted cross-correlation of the outputs of the single antennas that composes the DBF array. Under this treatment, the DBF data acquisition can be performed with a standard correlator. An array with an optimized configuration can form a beam pattern to meet the requirement of global CD/EoR signal detection. Removing redundant baselines can reduce the number of data acquisition channels to save costs. And the number of required antennas can be reduced to several tens.

In this work, we show that a dense regular digital beamforming array, with most of its antennas (∼80%) removed following a particular procedure, can be used to form frequency-independent station beams to detect the global CD/EoR signal.

This paper is organized as follows:

Section 2 describes the design of the digital beamforming array.

Section 3 explains how the number of antennas (so that data acquisition channels) can be significantly reduced based on the mathematical equivalence between the DBF array and the interferometer array.

Section 4 gives an example design and demonstrates the effect of cutting down redundant baselines.

Section 5 presents some aspects that require further study and lists some engineering challenges that must be addressed.

Section 6 presents our conclusions.

In the appendix, we present the deduction of the statistic error estimation.

In the following sections, especially when describing the array configuration, different indexing systems of antenna/baseline are used for different purposes, so we list the indexing-related symbols here beforehand:

- $s$: the size of the regular square grid that is composed of at most $s^2$ antennas, and $s$ is chosen to be an odd number;
- $(p,q)$: antenna index on a regular square grid, i.e., the grid index, and $p = -(s−1)/2 \ldots (s−1)/2$, $q = -(s−1)/2 \ldots (s−1)/2$;
- $i$ and $j$: linear antennas indices, which can be defined as $i(\text{or } j) = [p+ (s−1)/2]s + q + (s−1)/2$ for an antenna with the grid index of $(p,q)$; note that the linear indices do not have to be continuous, and removing an antenna does not affect the linear indices of others;
- $k$: Redundant baselines, a group of equivalent baselines shares the same $k$ index, and $k = 0$ denotes zero-length baselines, corresponding to auto-correlations.

## 2 DESIGNING AN APERTURE ARRAY FOR GLOBAL CD/EOR SIGNAL DETECTION

### 2.1 The Principle of Digital Beamforming

Although DBF is a mature technology, we still give a brief review in the context of global CD/EoR signal detection in this section. A beamforming array or so-called aperture array is composed of a set of single antennas. The output voltages, $v_{\text{ant}}$, are weighted and summed up to produce the station voltage output as

$$v_{\text{sta}}(\nu) = \sum_i w_i v_{\text{ant},i}(\nu),$$

where

$$v_{\text{ant},i}(\nu) = g(\mathbf{n}, \nu)E(\mathbf{n}, \nu) e^{i 2\pi n \cdot x_i / \lambda} dn$$

is the complex voltage output of $i$-th antenna in the station at position $x_i$. $g(\mathbf{n}, \nu)$ is the voltage response to the quasi-monochromatic incident plane wave from direction $\mathbf{n}$ within a narrow band centered on frequency $\nu$, $E$ is the scalar electric field strength at the origin, $\lambda = c/\nu$ is the corresponding wavelength, and $i^2 = -1$. The single antennas in the station are assumed to be identical, just like the treatments in many other works (e.g., Singh et al. 2015; Presley et al. 2015; McKinley et al. 2020). As polarization is not concerned with global CD/EoR signal detection, only scalar theory is
considered here. When necessary, one can easily modify the following deductions for polarization measurements.

The basic idea of measuring the global CD/EoR signal is to acquire the voltage signal output by a single antenna, and if \( \partial g(n, \nu) / \partial \nu = 0 \), then the power spectral density of the antenna output (i.e., the antenna temperature spectrum) is proportional to the sky-averaged radio spectrum, which is modeled to be the sum of the foreground component, the noise component, and the global CD/EoR signal. Theoretically, the foreground component can be easily subtracted by performing a power-law or low-degree log-space polynomial model fitting (e.g., Wang et al. 2006, 2013). However, by performing a power-law or low-degree log-space polynomially, the foreground component can be easily subtracted noise component, and the global CD/EoR signal. Theoretically, given a desired final station beam pattern \( P_{\text{sta}}(n, \nu) \), one can obtain \( w_{p,q}'s \) through an inverse DFT (IDFT).

### 2.2 Determining Antenna Weights

The weights of all the antennas are determined according to the desired station beam pattern \( P_{\text{sta}} \). In the following sections, we set a more strict constraint to the weights as all \( w \)'s are real numbers and

\[
 w_{p,q} = w_{p,-q} = w_{p,q}, 
\]

so that Equation 15 can be rewritten as

\[
 G_n(n, \nu) = \sum_{0 \leq p,q \leq q} w_{p,q} \left[ (2 - \delta_{p,0}) \cos(2\pi d_{\lambda} p n_x) \right. 
\]

\[
 \left. \times (2 - \delta_{q,0}) \cos(2\pi d_{\lambda} q n_y) \right]. 
\]

Without loss of generality, \( w_{p,q}'s \) are normalized by \( w_{0,0} \), i.e., the weight of the central antenna, so that \( w_{0,0} \equiv 1 \). Although values of \( w \)'s can be determined by performing an IDFT, incomplete aperture plane coverage and numerical errors lead the total station beam pattern deviates from the desired one. Practically \( w_{p,q}'s \) are determined by solving an optimization problem

\[
 w = \arg \min_w \int \left| P_{\text{sta}}(n, \nu; w) - \hat{P}_{\text{sat}}(n, \nu) \right|^2 d\nu. 
\]

This optimization problem is solved for each frequency channel independently. Other objective functions may also be used, but we do not perform the test here for conciseness.

As this is a high-dimensional optimization problem, the dimension of which is \((s+1)^2/4-1\), we choose to use the particle swarm optimization (PSO, Bonyadi & Michalewicz 2017) algorithm. The IDFT result is used as an initial guess of the PSO algorithm.

### 3 FROM DBF TO SHORT SPACING INTERFEROMETER

#### 3.1 Forming the Beam with an Equivalent Interferometer

By comparing Equations 4 and 5, one may notice that the measurement of the station auto-correlation signal can be

\[
 \frac{\partial g(n, \nu)}{\partial \nu} \approx 0, 
\]

which will be described in detail later in Section 4.

We assume that the antennas are placed on a regular \( s \times s \) grid as

\[
 x_{p,q} = d(pe_x + qe_y), 
\]

where \( s \) is chosen to be an odd number, the unit vector \( e_x \) and \( e_y \) point to the east and the north direction, respectively, and \( d \) is the antenna spacing. Then we obtain

\[
 G_n(n, \nu) = \sum_{p,q} w_{p,q} \exp[i2\pi d_{\lambda}(pn_x + qn_y)], 
\]

where \( d_{\lambda} \equiv d/\lambda, n_x \equiv n \cdot e_x, \) and \( n_y \equiv n \cdot e_y \). Apparently \( G_n(n, \nu) \) can be represented as the discrete Fourier transform (DFT) to the complex weights on a regular grid. If \( d_{\lambda} > 0.5 \), grating lobes caused by the aliasing effect may appear (e.g., Mailloux 2017).
performed with two methods (e.g., Ruigrok et al. 2017): summing the weighted single antenna voltage signal and then calculating the auto-covariance of the station output, namely

\[ S(\nu) = \left< \sum_i w_i v_{\text{ant},i}(\nu) \right> \left< \sum_j w_j^* v_{\text{ant},j}(\nu) \right> , \]

or summing the weighted auto- and cross-correlations of all antenna pairs, namely

\[ S(\nu) = \sum_{i,j} w_i w_j^* \left< v_{\text{ant},i}(\nu) v_{\text{ant},j}^*(\nu) \right> . \]

The data acquisition of the two methods corresponds to the conventional digital beamforming system and interferometer system, respectively. We name the first method the DBF paradigm (Equation 19) and name the second method the correlation paradigm (Equation 20). Although these two paradigms are equivalent to each other mathematically, the correlation paradigm has an advantage over the DBF paradigm: the weights of the antennas can be determined and applied offline. It gives a second chance to process the data if the instrumental calibration improves after the observation is performed. The correlation paradigm also has a disadvantage: the original time resolution has been lost since the signals acquired have already been integrated, which is fortunately not required in global CD/EoR signal detection.

In the correlation paradigm, the correlation result is only related to the relative position between the two antennas involved. Thus, the number of data acquisition channels can be reduced (to save costs) by canceling redundant baselines. However, one antenna might be in more than one baseline and can be removed only when other non-redundant baselines are unaffected. All the antennas are checked one by one and removed if allowed; until no more antennas can be removed while keeping the baseline coverage unchanged. Note that the final array configuration is affected by the order of the checking and removing procedure, so it is not uniquely determined.

3.2 Statistic Error

According to Equations A19 and A20 in the Appendix, the relative error in the worst case of the correlation paradigm is calculated as

\[ \Delta_{\text{rel}}[S(\nu)]_{\text{max}} = \frac{1}{\sqrt{\Delta \nu \tau}} \sqrt{\frac{1}{N} \sum_{k \neq 0} \left[ \sum_{(l,j) \in \mathcal{B}(k)} w_l w_j^* \right] \left[ \sum_i w_i^2 \right]^{(2)}} (21) \]

\[ \equiv \frac{\eta_{\text{max}}}{\sqrt{\Delta \nu \tau}} . \]

In the next section, we will show that for a set of practical weights, the \( \eta_{\text{max}} \) is \( \sim 1 \) so that the measurement is close to the condition of a single antenna, and the required integration time is around several days.

4 AN EXAMPLE

In this section, we present an example design of the DBF array, optimize the weights, and cut down the redundant baselines to reduce the number of required data acquisition channels. We assume that some facilities of the 21 Centimetre Array (21CMA; Wang et al. 2010; Huang et al. 2016; Zheng et al. 2016) can be directly used in this experiment, so the single antenna design and the array configuration take this into account.

The example shown here is to demonstrate the feasibility of forming frequency-independent beam patterns with an array, not necessarily an optimal design. Both the single antenna design and array configuration can be optimized further. Besides, engineering challenges are not thoroughly handled in this paper, which will be briefly discussed in Section 5.6.

4.1 The Design of Single Antennas in the Array

21CMA is an interferometer working in the 50 – 200 MHz band. Each antenna station of 21CMA is composed of 127 log-periodic antennas optimized for the working band. It is equipped with a digital correlator with 40 data acquisition channels.

We choose to use a modified version of the log-periodic antenna used in the 21CMA array in this example. The length of the longest dipole of the log-periodic antenna is shortened to fulfill the mechanical constraint of the array configuration. The axes of the single antennas are steered to the zenith rather than their original pointings in 21CMA, i.e., the north celestial pole. The beam patterns of the single antenna are calculated with NEC++, a C++ implementation of the Numerical Electromagnetics Code\(^1\) (NEC). We show the beam patterns of 50, 70, 90, and 110 MHz in Figure 1. The NEC code is implemented based on the method of moments (e.g., Davidson 2010). It is worth evaluating if results with higher accuracy can be obtained with other choices, such as the finite element method (FEM) and the finite difference time domain (FDTD). Though log-periodic antennas are wide-band antennas, the single antenna beam pattern in this example is still significantly frequency-dependent. In Section 4.4, we will show that the artificial spectral structure, which is caused by the frequency-dependent beam pattern of a single antenna, may exceed the expected amplitude of the global CD/EoR signal.

4.2 Array Configuration in the DBF Paradigm

The configuration of the DBF array is set to be a 13 × 13 regular grid, i.e., \( s = 13 \), with the spacing to be 1.5 m. In Section 4.5, we will show that \( s = 13 \) ensures a total of 37 data acquisition channels are required after removing all possible redundant baselines; it can fit well into the current 40-channel correlator of the 21CMA. The antennas’ response on the square grid’s boundary may be different from that of the antennas inside, so a set of antennas that are only connected to matching terminal resistors can be placed right outside the grid (Fig. 2).

According to Section 3.1, the DBF array can be equivalent to an interferometer so that the redundant baselines can be canceled. We show the array configuration after removing all possible redundant baselines later (§4.5).

\(^1\) http://elec.otago.ac.nz/index.php/Necpp
For frequencies higher than 110 MHz, we could not find a set of antennae that would satisfy this requirement.

The profiles of the normalized single antenna beam pattern in the E-plane and the H-plane are shown in Figure 4. We notice that in the high-frequency end (i.e., 110 MHz), sidelobes with level higher than −15 dB appear. In section 4.4, we will show that detecting the global CD/EoR signal is feasible with this level of sidelobes.

### 4.3 Calculating the Weights

The weights are determined by solving the optimization problem Equation 18. The code was mainly run on the ARM64 architecture cluster of the China SKA Regional Centre prototype (An et al. 2022). About 500 core-hours were consumed. Note that a full array configuration composed of $s^2$ antennas is assumed here.

We set the desired station beam pattern to be a Gaussian beam as

$$
\hat{P}_{\text{sta}}(\theta) \propto \exp \left( -\frac{\theta^2}{2\theta_b^2} \right),
$$

where $\theta_b = 15^\circ$, and $\theta$ is the zenith angle. By minimizing the objective function of Equation 18 with the PSO algorithm, we calculate the weight of each antenna in each frequency channel between 50 and 110 MHz, with the channel spacing of 1 MHz, Figure 3 shows the obtained weights of 50 MHz, 70 MHz, 90 MHz, and 110 MHz, respectively; the corresponding profiles of the array beam patterns in the E-plane and the H-plane are shown in Figure 4. We notice that in the high-frequency end (i.e., 110 MHz), sidelobes with level $-40$ dB appears. In section 4.4, we will show that detecting the global 21 cm signal is feasible with this level of sidelobes.

For frequencies higher than 110 MHz, we could not find a set of $w$’s that meet the requirements of global CD/EoR detection. Hopefully, changing the design of the single antennas or optimizing the array configuration (increasing $s$ and/or decreasing $d$) can extend the frequency range.

### 4.4 Evaluating the Influence of the Frequency Dependent Beam

We use the 408 MHz continuous spectra sky map by Remazeilles et al. (2015) (originally created by Haslam et al. 1981) as the sky template and set the spectral index to be $\alpha = -2.7$ to simulate the foreground signal in the interesting frequency range as

$$
T_{\text{sky}}(n, \nu) = T_{408 \text{ MHz}}(n) \left( \frac{\nu}{408 \text{ MHz}} \right)^\alpha.
$$

Note that the spectral index may have a variety of different values in difference literature, most of which are in the range between $-2.75$ (e.g., Irfan et al. 2022) and $-2.5$ (Presley et al. 2015; McKinley et al. 2020).

We assume the instrument is placed at a site with a latitude of 45$^\circ$N. Then we use Equation 12 to evaluate the instant station output $T_{\text{sta}}(\nu)$, calculate the 24 hours average spectrum $\bar{T}_{\text{sta}}(\nu)$, and perform a polynomial fitting to the simulated $(\ln(\nu/\text{MHz}), \ln(T_{\text{sta}}(\nu)/\text{K}))$ data set. Polynomials with different degrees have been tested in previous works (e.g., Bernardi et al. 2016; Singh & Subrahmanyan 2019) to fit the foreground and to mitigate the residual spectral structures caused by the instrumental effects. Most choices of the degrees of the polynomial are between 3 and 7. Here we show the residual of a 3$^{rd}$- and a 7$^{th}$-degree polynomial fitting in Figure 5a and b. If the global CD/EoR absorption feature (i.e., the dip) is around 100 mK, it should be sufficient to detect this feature.

As a comparison, we use the beam pattern of the single antenna to predict the induced antenna temperature $T_{\text{ant}}$ and perform a 3$^{rd}$- and a 7$^{th}$-degree polynomial fittings, the residual of which are shown in Figure 5c and d. The DBF algorithm significantly improves the frequency-independency of the single antennas in the array and makes global CD/EoR signal detection feasible.

We also test recovering the global CD/EoR signal from the simulated data by using a direct Bayesian inference method described in our previous work Gu & Wang (2020). The global CD/EoR signal model evaluating code was implemented according to the work of Mirocha et al. (2012) and Mirocha (2014). The parameters to simulate the global CD/EoR signal...
Figure 3. The weights of four frequency channels, the four pictures are all of the size 13 × 13 pixel², corresponding to the antennas on the 13 × 13 regular grid. The value of each pixel equals the weight of each corresponding antenna.

Figure 4. The normalized array beam pattern profiles in the E-plane and the H-plane.

are modified so that the dip falls within the frequency range of 50 – 110 MHz. We have to admit that the actual condition can hardly be so ideal. We employ the same emcee algorithm (Foreman-Mackey et al. 2013) as in Gu & Wang (2020) to perform the Markov-Chain Monte Carlo sampling to recover the signal. A 7th-degree polynomial in log-space is used to model the foreground and any artificial spectral structures introduced by the uncorrected frequency-dependent station beam pattern. The comparison between the input and the recovered signals is shown in Figure 6. The global CD/EoR signal is recovered in two ways: (1) by calculating the difference between the total signal $T_{total}$ and the foreground model $T_{fg}$, and (2) by directly evaluating the numerical global CD/EoR model. The characters of the input signal have been well recovered. Notably, the parameter space may show degeneracy for some other combinations of the global CD/EoR models and foreground models. More thorough simulation tests may be necessary to ensure the detection of the CD/EoR signal, but beyond the concern of this paper.

4.5 Removing Redundant Baselines

According to Section 3.1, the DBF array can be equivalent to an interferometer by replacing the DBF data acquisition system with a correlator. The number of data acquisition channels can be reduced by removing the redundant baselines. The array configuration after removing all possible redundant baselines is shown in Figure 7. There are 37 antennas left, so the data acquisition can be performed with the 21CMA correlator directly, which has 40 data acquisition channels.

4.6 Error Estimation

The statistic uncertainty is estimated based on the array configuration described in section 4.5, i.e., all possible redundant baselines have been removed.

We calculate the $\eta_{\text{max}}$ using Equation A20 and the weights obtained in Section 4.3. The result is shown in Figure 8. The 50 MHz channel has the largest $\eta_{\text{max}}$, so we use this value to estimate the statistic error for a ten-day observation as an example. By using Equation A19, and let $\eta_{\text{max}} \approx 2.5$, $\tau = 10$ days, $\Delta \nu = 1$ MHz, the relative statistic error of $S(\nu)$ is calculated to be

$$\Delta_{\text{rel}}[S(\nu)]_{\max} = 2.7 \times 10^{-6}. \quad (25)$$

Assuming the absorption feature of the global CD/EoR signal is about 100 mK, the sky-averaged Galactic foreground is $10^4$ K in the worst case; the required relative statistic error should be less than $10^{-5}$. So the above array design should be sufficient to constrain the global CD/EoR signal if the instrument calibration error is well controlled.
Figure 5. The residual of the 3rd-degree (left) and the 7th-degree (right) polynomial fitting to the \((\ln(\nu/\text{MHz}), \ln(\bar{T}_{\text{obs}}(\nu)/\text{K}))\) data set that is simulated based on the array (upper) beam pattern obtained in Section 4.3, and the single antenna (lower) beam pattern described in Section 4.1. Note that the y-axis units are in mK and K for the array and the single antenna data, respectively.

Figure 6. A comparison of the input and recovered signal from the simulated data assuming the array to have a configuration as is described in Section 4.2.

Figure 7. The array configuration after removing all possible redundant baselines.

5 DISCUSSION

5.1 Do We Really Need an Array?

As shown in Section 4.4, a well-designed array does help suppress the frequency-dependency of the beam pattern. But antenna arrays are usually much more expensive and complicated than single antenna systems.

On the single antenna side, the frequency-dependent beam pattern can be suppressed by narrowing down the working frequency range, i.e., dividing the whole frequency range into several sub-bands and building single antennas for each of them. However, if the character bandwidth of the global CD/EoR signal features (e.g., the dip feature) happens to be broader than the sub-bands, the detection can be more difficult.

Therefore, whether to use an array to detect the global CD/EoR signal depends on the compromise among several factors, including the cost, the desired instrument performance, and the technical capacity.
5.2 An Non-Square Array

Since the acquired antenna temperature spectra are averaged within a period longer than 24 hours, the station beam pattern profiles in the east-west direction are integrated out eventually, and only the frequency-dependency of the profiles in the north-south direction matters; it is not strictly true due to the array is not located on the equator but is still acceptable in our condition. Thus it is not necessary to restrict the array configuration to be a full $s \times s$ square array. The antenna number along the east-west direction can be reduced to 1 so that we could have a $1 \times s$ linear array deployed along the local meridian. And still, with the procedure described in Section 3, the number of antennas can be reduced while the baseline coverage is kept unchanged. With this method, the frequency-dependency of the 24-hour averaged station beam patterns can hopefully be better suppressed with the same number of data acquisition channels.

5.3 Forming Achromatic Beam with SKA LFAA Station?

The future SKA Low-Frequency Aperture Array (LFAA) antenna station utilizes DBF technology to form the station beam, making it possible to perform global CD/EoR signal detection. Price (2022) simulated the end-to-end drift-scan observation of the radio sky at 50–100 MHz using a zenith-phased array in the background of the EDA2 array, a precursor instrument for SKA. They concluded that an EDA2-like array with $O(10^5)$ antennas can form station beam patterns with side lobe level $< -50$ dB, so that can be used to detect the global CD/EoR signal. The huge required number of antennas makes it hard to implement. The simulation result of our example design described in Section 4 suggests that a properly set desired station beam pattern, a carefully chosen array configuration, and a set of optimized antenna weights may enable the detection of global CD/EoR signal with a small number of antennas.

According to the SKA system requirement specification\(^2\), 256 antennas (SKA1-SYS_REQ 3339) will be randomly distributed within a circular region with a diameter about 35 m (SKA1-SYS_REQ 2140). If this configuration is chosen, the average distance between antennas is $\sim 1.94$ m, which is larger than the half wavelength of 78 MHz. This could make it hard to form desired beam pattern at frequency $> 78$ MHz. A similar reason may cause the EDA2-like array to need $O(10^5)$ antennas to detect the global CD/EoR signal. Nevertheless, it is worth evaluating the performance of the future LFAA station by using the antenna weight optimizing procedure described in Section 2.2; we will leave this in future work.

5.4 A Possible Method to Suppress the Grating Lobes

Limited by the geometric dimension of single antennas, the requirement of $d_s \leq 0.5$ cannot always be met. In our example array (Section 4), $d_s$ is limited by the length (123.6 cm) of the longest dipole of the log-periodic antenna.

A possible strategy is rotating the antenna about the vertical axis by $45^\circ$. It will allow us to place the single antennas closer to build a denser array; the grating lobes can be further suppressed. However, this will bring two shortages: (1) when the array becomes denser, the coupling and crosstalk between neighboring antennas will be more severe, and thus it becomes harder to predict the beam pattern of the single antennas, and (2) rotating the antennas by $45^\circ$ will destroy the symmetry of the single antenna beam pattern, then the symmetry constraints over the $w_{p,q}$’s described as Equation 16 no longer work if we still require the final array beams to be symmetric (as is described by Equation 23). Then the number of free parameters of the optimization problem to solve $w$’s (i.e., Equation 18) will increase to $s^2 - 1$ (almost multiplied by 4), and thus the optimization problem will be harder to solve.

Note that only making the array denser might not improve the overall quality of the formed beams due to the equivalent aperture becoming smaller. A compensatory method is increasing the number of antennas, i.e., enlarging the parameter $s$. This will require a larger number of data acquisition channels. This effort requires further evaluation to value its necessity.

5.5 Suggestions about Antenna Design and Spacing

The final station beam pattern is comprehensively determined by (1) array configuration, (2) antenna weights $w_{p,q}$, and (3) the beam pattern of single antennas, while most efforts of this work are focused on the first two factors. The single antennas should have a beam pattern as independent from the frequency as possible. However, this design target may be limited by other factors. For example, considering that the antennas are rather close to each other, the dimensions of the single antennas should not be too large to fit into the array, and this factor puts some constraints on the single antenna design.

About the detailed antenna beam pattern shape, as we mentioned in Section 2.1, the array should have a $d_s \leq 0.5$; otherwise, grating lobes may appear. Limited by the dimension of single antennas, this requirement cannot always be fulfilled, especially at the high-frequency end. If the single

\(^2\) https://www.skao.int/sites/default/files/documents/d3-SKA-TEL-SKO-0000008-Rev11_SKA1SystemRequirementSpecification_0.pdf

MNRAS 000, 1–11 (2022)
antenna beam pattern can concentrate to the zenith direction, the grating lobes arise by a $d_1 > 0.5$, which usually have large zenith angles, can be suppressed.

5.6 Engineering Challenges

5.6.1 An Array in the Real World

In previous computations and most other works (e.g., Singh et al. 2015; Presley et al. 2015; McKinley et al. 2020; Dilullo et al. 2020), the single antennas in an array are assumed to be identical, but in the real world, this assumption does not always hold. As we have stated in our previous work Gu & Wang (2020), any gain calibration error $> 10^{-4}$ may significantly bias the detection result or make the detection completely impossible. Similarly, any difference between the single antennas in electronic/mechanical aspects might introduce errors comparable to or larger than the target CD/EoR signal.

The above computations assume that the single antenna beam pattern $|g(n,\nu)|^2$ can be precisely measured beforehand, but for antennas working in the meter-wave band, this kind of measurement is challenging. Calibrating the beam pattern with a bright radio source does not work in the case of low-frequency single antennas because of their small effect area and wide beam. Neither is it easy to perform a precise anechoic chamber measurement because (1) the absorption material used to build meter-wave anechoic chamber is usually not good enough to emulate the free space, and (2) ambient facilities (e.g., other antennas, the ground plane) around the antenna being calibrated has to be considered, which is hard to be fit into a common microwave darkroom. The recent development of drone-based antenna measurement technology (e.g., Paonessa et al. 2015; Chang et al. 2015) may be used to perform such kind of calibration; however, the accuracy is to be validated. The tolerance of inaccuracy in the single antenna beam pattern measurement will be discussed in our future work.

For an array composed of a set of closely placed antennas, like the one we described in Section 4, the coupling or crosstalk between neighboring antennas is an issue that cannot be ignored. When designing single antennas, this effect should be taken into account. When performing an actual beam pattern measurement, the single antenna under test should also be placed inside the array.

5.6.2 The Challenges in the Analog Front-end

The antenna temperature induced by the foreground emission is four orders higher than that induced by the CD/EoR signal. This extreme contrast raises a strict requirement for the analog front-end. The analog front-end is composed of filters and amplifiers, which are expected to be almost identical, stable enough, and well-calibrated (better than the level of $10^{-4}$ dB).

Besides the active analog devices, the co-axis cables connecting the antennas and the data acquisition systems must be cut precisely. Considering the extreme requirements of the system calibration (e.g., Gu & Wang 2020), an error that is only less than the wavelength is far from sufficient. How sensitive the result is to the cable length error requires further evaluation in future works.

5.6.3 Self-Generated and External Weak RFIs

Different from single antenna experiments such as EDGES (Bowman & Rogers 2010), BIGHORNs (Sokolski et al. 2015), SCI-HI (Voytek et al. 2014), SARAS (Patra et al. 2013), and PRI$^2$M (Philip et al. 2019), the data acquisition system for an array that is composed of tens of antennas is much more complicated, as the complexity scales with the square of the number of antennas. The self-generated radio frequency interference (RFI) can be well controlled in single antenna experiments; on the other hand, the high power consumption of the data acquisition system of an array makes it hard to build the shielding system.

Unlike strong radio RFIs, which can be easily identified and excluded, weak RFIs below the detection threshold are hard to handle (e.g., see Section 4.4 of Wang et al. 2021), especially for the global EoR detection experiments that have given up sky imaging. Such RFIs might remain in the flagged data and show the appearance of spectral features that are hard to be explained, which may mislead the CD/EoR signal detection. Further discussion about the RFIs is beyond the scope of this paper; we will leave this in future work.

6 CONCLUSIONS

This work introduces a method to form achromatic beam patterns with a dense regular DBF array to detect the all-sky averaged 21 cm signal from the Cosmic Dawn and Epoch of Reionization. Section 2.2 gives a procedure to optimize the antenna weights to form the desired station beam pattern. The equivalence between a DBF array and an interferometer reduces the number of necessary antennas in the DBF array.

An example array design with antennas arranged in a regular $13 \times 13$ square grid is given. We calculate the antenna weights to form the desired station beam pattern and evaluate the performance of global CD/EoR signal detection. A simulation aiming to evaluate the performance of the example array provides a positive indication towards the feasibility of using the dense regular DBF array to detect the global CD/EoR signal. The final array configuration after removing all possible redundant baselines is shown in Figure 7.

We also give some general suggestions about the array design and briefly discuss the engineering challenges that must be overcome in practical experiments.

ACKNOWLEDGEMENTS

This work is supported by National SKA Program of China No. 2020SKA0110200 and 2020SKA0110100. We acknowledge the supports from NSFC under 11973070 and Key Research Program of Frontier Sciences, CAS, Grant No. ZDBS-LY-7013.

This work used resources of China SKA Regional Centre prototype funded by the National Key R&D Programme of China (2018YFA0404603).

We would also like to thank Marta Spinelli for her comments on the final manuscript.
DATA AVAILABILITY

The code and data underlying this article will be shared on reasonable request to the corresponding authors.

APPENDIX A: STATISTIC ERROR OF THE CORRELATION PARADIGM

Based on the Equation 20, given that all w’s are real numbers and under the constraint of Equation 16, the measured value \( S(\nu) \) can be rewritten as

\[
S(\nu) = \sum_i w_i^2 < |v_{ant,i}|^2 > + 2 \sum_{i<j} w_i w_j < \Re(v_{ant,i}v_{ant,j}) > \\
= \sum_i w_i^2 < \Re(v_{ant,i})^2 > \\
+ \sum_i w_i^2 < \Im(v_{ant,i})^2 > \\
+ 2 \sum_{i<j} w_i w_j < \Re(v_{ant,i})\Re(v_{ant,j}) > \\
+ 2 \sum_{i<j} w_i w_j < \Im(v_{ant,i})\Im(v_{ant,j}) > \\
= \sum_{i,j} w_i w_j (< \Re(v_{ant,i})\Re(v_{ant,j}) > + < \Im(v_{ant,i})\Im(v_{ant,j}) >),
\]

(A1)

where \( \Re(\cdot) \) and \( \Im(\cdot) \) denote the real and imaginary parts of a complex number, respectively. Note that a certain pair of antennas say \((i_1, j_1)\) could have been removed so that \(< \Re(v_{ant,i})\Re(v_{ant,j}) > \) and \(< \Im(v_{ant,i})\Im(v_{ant,j}) > \) could be missing. However, the removing antenna strategy described in Section 3.1 ensures that there is always at least one equivalent baseline remaining in the array. No matter whether one baseline is missing, its correlation value is always replaced by the average values obtained from the still remaining equivalent baselines as

\[
\Re(v_{ant,i,j}) = C^r(k) \equiv \frac{1}{r(k)} \sum_{i,j \in b(k)} \Re(v_{ant,i})\Re(v_{ant,j})
\]

(A4)

\[
\Im(v_{ant,i,j}) = C^i(k) \equiv \frac{1}{r(k)} \sum_{i,j \in b(k)} \Im(v_{ant,i})\Im(v_{ant,j}),
\]

(A5)

where \( k \) is the index of the baseline from the \( i_1 \)-th antenna to the \( j_1 \)-th antenna, \( b(k) \) is a function mapping the \( k \)-th baseline to the set of still existing corresponding antenna pairs \((i, j)\) in the regular grid, \( r(k) \) is the number of members in the set \( b(k) \). Correspondingly we also define a function \( B(k) \), mapping \( k \)-th baseline to the set of all possible antenna pairs \((i, j)\), whether they are removed or not.

Then

\[
S(\nu) = \sum_k [(< C^r(k) > + < C^i(k) >) \sum_{i,j \in B(k)} w_i w_j].
\]

(A6)

In a certain frequency channel with a bandwidth of \( \Delta \nu \), centered on frequency \( \nu \), the variance of the real (and also imaginary) part of the complex voltage follows a normal distribution with the variance of

\[
\text{Var}[\Re(v_{ant,i})] = \frac{1}{2} Rk_B T_{ant} \Delta \nu,
\]

(A7)

and the expected value of 0, \( R \) is the system impedance, \( k_B \) is the Boltzmann constant, and \( T_{ant} \) is the antenna temperature. In the meter-wave band, the system temperature is dominated by the Milky Way foreground, so we do not distinguish between the antenna and system temperatures.

Given that \( \Re(v_{ant,i})'s, \Im(v_{ant,i})'s, \Re(v_{ant,j})'s, \) and \( \Im(v_{ant,j})'s \) follow identical normal distributions and \( \rho_{i,j} \) is the correlation coefficient between \( \Re(v_{ant,i}) \) and \( \Re(v_{ant,j}) \) (and also between \( \Im(v_{ant,i}) \) and \( \Im(v_{ant,j}) \)), \( |\rho_{i,j}| < 1 \), and \( \rho_{i,i} = 1 \). So the variance

\[
\text{Var}[\Re(v_{ant,i})\Re(v_{ant,j})] = \frac{1 + \rho_{i,j}^2}{4} (Rk_B T_{ant} \Delta \nu)^2,
\]

(A8)

and

\[
\text{Var}[\Im(v_{ant,i})\Im(v_{ant,j})] = \text{Var}[\Re(v_{ant,i})\Re(v_{ant,j})].
\]

(A9)

Then

\[
\text{Var}[C^r(k)] = \text{Var}[C^i(k)] = \frac{1 + \rho_{i,j}^2}{4r(k)} (Rk_B T_{ant} \Delta \nu)^2
\]

(A10)

During a period of integration time \( \tau \), \( C^r(k) \) and \( C^i(k) \) are measured \( \Delta \nu \tau \) times respectively. Thus

\[
\text{Var} [< C^r(k) > + < C^i(k) >] = \frac{1 + \rho_{i,j}^2}{2r(k)\Delta \nu \tau} (Rk_B T_{ant} \Delta \nu)^2
\]

(A11)

\[
\text{Var}[S(\nu)] = \frac{(Rk_B T_{ant} \Delta \nu)^2}{\Delta \nu \tau} \sum_k \frac{1 + \rho_{i,j}^2}{2r(k)} \left( \sum_{i,j \in B(k)} w_i w_j \right)^2
\]

\[
= \frac{(Rk_B T_{ant} \Delta \nu)^2}{\Delta \nu \tau} \left[ \frac{1}{N} \left( \sum_i w_i^2 \right)^2 + \sum_{k \neq 0} \frac{1 + \rho_{i,j}^2}{2r(k)} \left( \sum_{i,j \in B(k)} w_i w_j \right)^2 \right],
\]

(A12)

and the standard deviation of \( S(\nu) \) is

\[
\sigma[S(\nu)] = \frac{Rk_B T_{ant} \Delta \nu}{\sqrt{\Delta \nu \tau}} \times \sqrt{\frac{1}{N} \left( \sum_i w_i^2 \right)^2 + \sum_{k \neq 0} \frac{1 + \rho_{i,j}^2}{2r(k)} \left( \sum_{i,j \in B(k)} w_i w_j \right)^2}. \]

(A13)

The expected value of \( S(\nu) \) is dominated by auto-correlations so that it can be approximated as

\[
E[S(\nu)] \approx \sum_{i,j} w_i w_j \delta_{i,j} (< \Re(v_{ant,i})\Re(v_{ant,j}) >
\]

\[
+ < \Im(v_{ant,i})\Im(v_{ant,j}) >) = \sum_{i} w_i^2(< \Re(v_{ant,i}) >^2 + < \Im(v_{ant,i}) >^2) = Rk_B T_{ant} \Delta \nu \sum_{i} w_i^2.
\]

(A14)

(A15)

Then we can calculate the relative uncertainty as

\[
\Delta_{rel}[S(\nu)] \equiv \frac{\sigma[S(\nu)]}{E[S(\nu)]} \times \sqrt{\frac{1}{N} \sum_{k \neq 0} \frac{1 + \rho_{i,j}^2}{2r(k)} \left( \sum_{i,j \in B(k)} w_i w_j \right)^2}}.
\]

(A17)
in the worst case
\[ \Delta_{\text{rel}}[S(\nu)]_{\text{max}} = \frac{1}{\sqrt{\Delta \nu \tau}} \sqrt{\frac{1}{N} + \sum_{k \neq 0} \frac{1}{\tau(i,j)} \left( \frac{1}{\langle \sum_{i,j} w_i w_j \rangle} \right)^2}. \]

We define
\[ \eta_{\text{max}} \equiv \sqrt{\frac{1}{N} + \sum_{k \neq 0} \frac{1}{\tau(i,j)} \left( \frac{1}{\langle \sum_{i,j} w_i w_j \rangle} \right)^2}, \]

(A20)
to represent the ratio of the relative uncertainty of the array output to that of a single antenna in the worst case.

REFERENCES

An T., Wu X., Lao B., Guo S., Xu Z., Lv W., Zhang Y., Zhang Z., 2022, Science China Physics, Mechanics & Astronomy, 65, 1
Anstey D., de Lera Acedo E., Handley W., 2021, MNRAS, 506, 2041
Anstey D., Cummer J., de Lera Acedo E., Handley W., 2022, MNRAS, 509, 4679
Barkana R., 2018, Nature, 555, 71
Bernardi G., McQuinn M., Greenhill L. J., 2015, ApJ, 799, 90
Bernardi G., et al., 2016, MNRAS, 461, 2847
Bonyadi M. R., Michalewicz Z., 2017, Evolutionary Computation, 25, 1
Bowman J. D., Rogers A. E. E., 2010, Nature, 468, 796
Bowman J. D., Rogers A. E. E., Monsalve R. A., Mozdzen T. J., Mahesh N., 2018, Nature, 555, 67
Chang C., Monstein C., Refregier A., Amara A., Glauser A., Casura S., 2015, PASP, 127, 1131
Cummer J., et al., 2022, Journal of Astronomical Instrumentation, 11, 2250001
Davidson D. B., 2010, Computational electromagnetics for RF and microwave engineering, Cambridge University Press
Dilullo C., Taylor G. B., Dowell J., 2020, Journal of Astronomical Instrumentation, 9, 2050008
Dilullo C., Dowell J., Taylor G. B., 2021, Journal of Astronomical Instrumentation, 10, 2150015
Dowell J., Taylor G. B., Schinzel F. K., Kassim N. E., Stovall K., 2017, MNRAS, 469, 4537
Fialkov A., Barkana R., Cohen A., 2018, Phys. Rev. Lett., 121, 011101
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Furlanetto S. R., Oh S. P., Briggs F. H., 2006, Phys. Rep., 433, 181
Gu J., Wang J., 2020, MNRAS, 492, 4080
Haslam C. G. T., Klein U., Salter C. J., Stoffel H., Wilson W. E., Cleary M. N., Cooke D. J., Thomasson P., 1981, A&A, 100, 209
Hirano S., Bromm V., 2018, MNRAS, 480, L85
Huang Y., Wu X.-P., Zheng Q., Gu J.-H., Xu H., 2016, Research in Astronomy and Astrophysics, 16, 36
Irfan M. O., et al., 2022, MNRAS, 509, 4923
Madau P., Meiksin A., Rees M. J., 1997, ApJ, 475, 429
Mailoux R. J., 2017, Phased array antenna handbook. Artech house
McKinley B., Trott C. M., Sokolowski M., Wayth R. B., Sutinjo A., Patra N., Nambissan T. J., Ung D. C. X., 2020, MNRAS, 499, 52
Mirocha J., 2014, MNRAS, 443, 1211
Mirocha J., Skory S., Burns J. O., Wise J. H., 2012, ApJ, 756, 94
Morales M. F., Wyithe J. S. B., 2010, ARA&A, 48, 127
Mozdzen T. J., Bowman J. D., Monsalve R. A., Rogers A. E. E., 2016, MNRAS, 455, 3890
Nambissan T. J., et al., 2021, arXiv e-prints, p. arXiv:2104.01756
Paonesa F., et al., 2015, in 2015 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting. pp 1372–1373
Patra N., Subrahmanyan R., Raghunathan A., Udaya Shankar N., 2013, Experimental Astronomy, 36, 319
Philip L., et al., 2019, Journal of Astronomical Instrumentation, 8, 1950004
Presley M. E., Liu A., Parsons A. R., 2015, ApJ, 809, 18
Price D. C., 2022, arXiv e-prints, p. arXiv:2210.04693
Price D. C., et al., 2018, MNRAS, 478, 4193
Reis I., Fialkov A., Barkana R., 2021, MNRAS, 506, 5479
Remazeilles M., Dickinson C., Banday A. J., Bigot-Sazy M. A., Ghosh T., 2015, MNRAS, 451, 4311
Ruigrok E., Gibbons S., Wapenaar K., 2017, Journal of Seismology, 21, 495
Singh S., Subrahmanyan R., 2019, ApJ, 880, 26
Singh S., Subrahmanyan R., Udaya Shankar N., Raghunathan A., 2015, ApJ, 815, 88
Singh S., et al., 2022, Nature Astronomy, 6, 607
Sokolowski M., et al., 2015, Publ. Astron. Soc. Australia, 32, e004
Spinelli M., Bernardi G., Garsden H., Greenhill L. J., Fialkov A., Dowell J., Price D. C., 2021, MNRAS, 505, 1575
Spinelli M., Kyriakou G., Bernardi G., Bolli P., Greenhill L. J., Fialkov A., Garsden H., 2022, MNRAS, 515, 1580
Vedantham H. K., Koopmans L. V. E., de Bruyn A. G., Wijnholds S. J., Ciardi B., Brentjens M. A., 2014, MNRAS, 437, 1056
Voytek T. C., Natarajan A., Jiuregui García J. M., Peterson J. B., López-Cruz O., 2014, ApJ, 782, L9
Wang X., Tagmark M., Santos M. G., Knox L., 2006, ApJ, 650, 529
Wang J., et al., 2010, ApJ, 723, 620
Wang J., et al., 2013, ApJ, 763, 90
Wang J., et al., 2021, MNRAS, 505, 3698
Zheng Q., Wu X.-P., Johnston-Hollitt M., Gu J.-h., Xu H., 2016, ApJ, 832, 190

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

MNRAS 000, 1–11 (2022)