


Generalized redundant calibration of radio interferometers

Prakruth Adari¹ and Anže Slosar²

Brookhaven National Laboratory, Upton, New York 11973, USA and Physics and Astronomy Department, Stony Brook University, Stony Brook, New York 11794, USA

 (Received 30 July 2021; revised 5 July 2022; accepted 19 July 2022; published 5 August 2022)

Redundant calibration is a technique in radio astronomy that allows calibration of radio arrays whose antennas lie on a lattice by exploiting the fact that redundant baselines should see the same sky signal. Because the number of measured visibilities scales quadratically with the number of antennas but the number of unknowns describing the individual antenna responses and the available information about the sky scales only linearly with the array size, the problem is always overconstrained as long as the array is big and dense enough. This is true even for nonlattice array configurations. In this work we study a generalized algorithm in which a per-antenna gain is replaced with a number of gains. We show that it can successfully fit data from an approximately redundant array on square lattice with pointing and geometry errors, but that the model's parameters are difficult to link to the quantities of interest. We discuss the parametrization, limitations, and possible extensions of this algorithm.

DOI: [10.1103/PhysRevD.106.043006](https://doi.org/10.1103/PhysRevD.106.043006)

I. INTRODUCTION

The 21-cm emission from neutral hydrogen is promising to transform our understanding of the universe across the ages: from low-redshift observations of neutral hydrogen in galaxies through the epoch of reionization all the way to the dark ages at redshift $z \sim 100$ in the future.

The field received a major boost when it was realized that developments in computing and rf technology allow telescopes to be built almost entirely in software. A number of experiments were born, some already operating, such as Canadian Hydrogen Intensity Mapping Experiment (CHIME) [1], Tianlai [2], Hydrogen Era of Reionization Array (HERA) [3], Murchison Widefield Array [4], Precision Array for Probing the Epoch of Reionization [5], and Giant Metrewave Radio Telescope [6]; some are under construction, including Hydrogen Intensity and Real-Time Analysis eXperiment (HIRAX) [7], BAO from Integrated Neutral Gas Observations [8], Square Kilometer Array-low [9], and Canadian Hydrogen Observatory and Radio-transient Detector [10] and some proposals for future very large facilities, such as Packed Ultra-wideband Mapping Array [11].

Large interferometric radio arrays require a large number of calibration parameters. It was soon noticed that arrays of indistinguishable elements on regular lattice possess strong redundancy. The total number of possible pairs of antennas, corresponding with the total number of measured visibilities, scales with the square of the number of elements in the array. On the other hand the number of unique baselines (given by the number of unique separation vectors between all possible pairs of antennas) scales only linearly with the number of

elements. Since all baselines made of identical elements and spanning the same distance vector should measure the same signal, we can use this to back out calibration factors of individual antennas, without ever knowing anything about the actual sky signal. The solution is unique up to intrinsic degeneracies of the system corresponding to the overall shift, scaling in amplitude, and translation of the sky signal (i.e., applying a phase gradient across the uv-plane). This calibration procedure is known as redundant calibration and has been worked in detail in [12,13].

Unfortunately, it soon became clear that real arrays are much less redundant for the simplest form of redundant calibration to be sufficient. For example, [14] studied redundancy in HERA and found that real-life nonredundancy produces spurious temporal structure in gain solutions. Other works such as [15–20] discuss the shortcomings and possible solutions of redundant calibration in both real telescope arrays and simulations.

In some sense, these findings indicate that redundant calibration is the victim of its own success. A possible way to look at the problem is that it is not that the arrays are too nonredundant, but that they are too sensitive. They are sensitive enough that the particularities of individual elements produce big enough effects that a good fit cannot be found for the data. One possible approach has been studied in [21]: a pair of nearly, but not perfectly, redundant baselines will have very strong, but not perfect, correlation between measured visibilities. These slight decorrelations can be propagated self-consistently using a quadratic estimator formalism to allow stable solutions and almost by construction, a good fit. This process has the advantage of gracefully dealing with outlier antennas:

there is no need to cut them out if we can instead model them as such. A more recent paper [22] proposed a unified approach to sky-based and redundant calibration, in which both can be used concurrently in a self-consistent Bayesian model, assuming we have a model of the telescope and its calibration uncertainties.

In this paper we take a different approach: instead of describing every single antenna with one complex number, a single gain, we are looking for a description in terms of multiple numbers per element that can be adjusted in order to achieve a good fit as well as give some physical insight into the type of imperfections. Note that this does not break the basic premise of redundant calibration: the number of unknowns still scales linearly with the number of antennas, while the number of measurements scales quadratically. So, for arbitrarily complex description of per-antenna nonidealities, the system will be overconstrained for large enough array (as long as the number of free parameters per antenna is finite).

We will focus on two of the canonical errors in a redundant radio array, pointing errors and geometric errors. Some of the other canonical errors include beam shape and beam size. The pointing errors refer to the fact that the beams of individual element are not aligned (i.e., the telescope points away from the zenith in a transit array configuration) and geometric errors, which means that the entire dish is displaced from its nominal lattice positions affecting the effective baseline lengths. These are the kinds of errors that we study, but we do not explicitly fit for them. Instead, we build a general pixelized description of the response of each element that is arbitrarily fine resolved given by a tunable parameter M . Ideally, our scheme would also be able to deal with beam shape and beam size issues; however, we focus on geometric and pointing errors.

II. PROBLEM SETUP

Let us consider a general interferometric array observing, for simplicity, in a single polarization mode. We consider a pair of antennas pointing at the zenith and observing in a narrow frequency range. The noiseless observed visibility for pairs of antennas i and j in the flat-sky approximation is given by

$$\begin{aligned} V_{ij} &= \int I(\theta) e^{-2\pi i(\mathbf{x}_i - \mathbf{x}_j) \cdot \theta} B_i(\theta) B_j^*(\theta) d^2\theta \\ &= \mathcal{FT}[IB_i B_j^*](\mathbf{x}_i - \mathbf{x}_j), \end{aligned} \quad (1)$$

where \mathbf{x}_i is the geometric position of antenna i measured in wavelengths, $I(\theta)$ is the intensity signal coming from the sky as a function of angle on the sky, θ , and B_i is the beam of antenna i . The operator $\mathcal{FT}[X]$ denotes the Fourier transform of X . The signal $I(\theta)$ is real and in principle spans the entire sky. The beams B_i are in general complex and compact in Fourier space (i.e., on the uv-plane). They correspond to the complex amplitude response to an incident uniform plane wave just above the aperture.

Multiplication in real space corresponds to convolution in Fourier space. Addition of the random-noise component ϵ gives the general equation for the observed visibility for a pair of antennas:

$$V_{ij}^o = \int U(\mathbf{x}_i - \mathbf{x}_j + \mathbf{u}) (\tilde{B}_i \otimes \tilde{B}_j^\dagger)(-\mathbf{u}) d^2\mathbf{u} + \epsilon, \quad (2)$$

where the uv-plane $U = \mathcal{FT}[I]$ contains the image of the sky (in Fourier space), and $\tilde{B}_i = \mathcal{FT}[B_i]$ are beam representations in this domain. Variable ϵ is a random variable describing the noise realization. We have used notation $\tilde{B}^\dagger(\mathbf{u}) = \mathcal{FT}[B^*](\mathbf{u}) = \mathcal{FT}[B](-\mathbf{u})$, i.e., complex conjugation of a real-space quantity results in mirroring across the origin in the Fourier domain. The crucial insight is that beams are compact in Fourier space, essentially corresponding to the physical extent of the dishes, and thus $\tilde{B}_i \otimes \tilde{B}_j^\dagger$ is also compact. Therefore, the limits of integration in Eq. (2) need to extend only as far as support of $\tilde{B}_i \otimes \tilde{B}_j^\dagger$.

It is instructive to compare this equation with the one that is typically used to set up redundant calibration:

$$V_{ij}^o = U_{i-j} g_i g_j^* + \epsilon, \quad (3)$$

where $i - j$ indexes redundant baselines corresponding to the baseline vector $\mathbf{x}_i - \mathbf{x}_j$. If we set our beams to be the same, scaled by only a complex gain factor for each dish, namely $B_i(\theta) = g_i B(\theta)$, we find that Eq. (2) reduces to Eq. (3) with

$$U_{i-j} = \int U(\mathbf{x}_i - \mathbf{x}_j + \mathbf{u}) (\tilde{B} \otimes \tilde{B}^\dagger)(-\mathbf{u}) d^2\mathbf{u}. \quad (4)$$

In other words, the redundant calibration assumes that all dishes have exactly the same sky response that are only allowed to vary by an overall complex gain, which is assumed to be coming from electronics, amplifiers, etc. It reconstructs these gains, but also the uv-plane information, *after the beam convolution*.

But in general, massive redundancy in compact arrays allows us to go beyond this approximation (we will quantify this in the next section). The main idea is to start with Eq. (2) as the main ingredient and reconstruct both \tilde{B} for each antenna and the uv-plane U . It is important to note, from the beginning, that we cannot unpick the variations in complex gain coming from electronics from those arising from the imperfect dish. Therefore, action of a gain g_i is equivalent to multiplying the entire beam by the same factor.

A. Discretization

We will try to build an intuition that will prepare us for formulation of general redundant calibration in steps. In short, we will try to build a redundant calibration that will replace the parameters U_{i-j} and g_i for the standard redundant calibration in Eq. (3) with a new, larger, but still finite number of parameters that can still be solved for by

using the inherent redundancy in the system but can allow for some freedom in nonredundancy.

In order to proceed, we need to express the integral of Eq. (2) as a sum over a finite number of degrees of freedom. Our approach is to pixelize the beam maps \tilde{B}_i and choose the pixelization in the uv -plane U that matches the beam pixelization in a way that expresses Eq. (2) as a sum that is cubic in input parameters without entailing any interpolation.

For concreteness, let us consider an interferometric array on regular square grid of size $N_s \times N_s$ with a total of $N_a = N_s^2$ antennas. We will measure all distances in the units of wavelength and take the lattice spacing to be L . We start by introducing the ‘‘oversampling’’ parameter M . The beam map is pixelized in $(2M + 1)^2$ pixels, i.e., a central pixel and M additional pixels on each side. For $M = 0$ the beam is reduced to a single pixel and the method reduces to a redundant calibration. The requirement on the odd number of pixels in the beam description is there to ensure that a single-pixel description of each beam is a natural limit of the problem.

The convolution of two beams ($\tilde{B}_i \otimes \tilde{B}_j^\dagger$) is of the size $(4M + 1) \times (4M + 1)$, which sets the natural pixel size to be $L/(2M + 1)$ for both the beam grid and the U grid. We can therefore rewrite Eq. (2) in a pixelized form schematically as

$$V_{ij} = \sum_{m=-2M}^{+2M} \sum_{n=-2M}^{2M} U_{i_o+m; j_o+n} (\tilde{B}_i \otimes \tilde{B}_j^\dagger)_{mn}, \quad (5)$$

where i_o and j_o are the grid offsets corresponding to the $i - j$ redundant baseline (the baseline probed by beam \tilde{B}_i and \tilde{B}_j) on the uv plane, and m, n are the pixel indices over the convolved beam. Note that the sum is over $(4M + 1) \times (4M + 1)$ corresponding to the convolution of two $(2M + 1) \times (2M + 1)$ -sized beam maps. This is illustrated in Fig. 1. We urge the reader to spend some time trying to understand this figure as this pixelization is essential for the understanding of our approach. Consider the shortest E-W baseline of a closely packed dish array. Dishes have diameter D so the distances between pieces of reflector surface on the same dish vary between 0 and D . The possible distances between pairs of pieces of reflector surface on two dishes vary between 0 (at points where the dishes nominally touch) and $2D$. In the formalism, this is encoded by the beam being $(2M + 1)^2$ pixels in size for an individual beam and $(4M + 1)^2$ for the convolution of the two beams. This then defines the extent on the uv -plane to which this particular baseline is sensitive. In the limit of large M it comes arbitrarily close, but never reaches the origin of the uv -plane (corresponding to the monopole signal) and similarly gets arbitrarily close but never reaches two lattice spacings (corresponding to sensitivity due to pieces of reflector surface that are

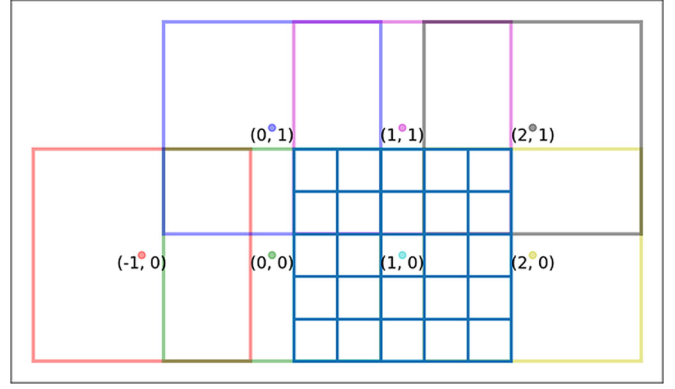


FIG. 1. Illustration of the pixelization scheme for $M = 1$. Labeled points illustrate the lattice points on the uv -plane for a square close-packed array. For example, $(1,0)$ corresponds to baselines one lattice spacing apart in the E-W direction and $(0,1)$ is the same for N-S baselines. $(0,0)$ is the origin of the uv -plane and corresponds to single-dish observations. Individual beam is described by 3×3 pixelized grid, so $(\tilde{B}_i \otimes \tilde{B}_j^\dagger)$ is the 5×5 pixelized grid shown in blue around $(1,0)$. Making prediction for a particular $(1,0)$ baseline amounts to multiplying the values of the blue beam-convolved map on the blue grid and the corresponding points in the uv -plane and summing them up. See text for further discussion.

furthest apart). In Fig. 1 this is illustrated with the blue grid. This is the correct and expected behavior for the flat-sky approximation. We discuss the nonflat-sky case further below.

In Fig. 1 we also show squares around other nominal lattice pointings. We see that most uv points are being probed by multiple baselines, except those on precise grid spacings. This severely limits the degeneracies present in the problem, which we will further comment on in a later section. We also note that we need to describe just one half of the uv -plane with the other half given by the reality of the observed field. Together with rules about Fourier transforms of conjugated fields, this ensures that prediction for the antenna pair i, j is always a complex conjugate of the prediction for the antenna pair j, i .

B. Phased-up array interpretation

To continue building intuition about the pixelization, consider the pixel m in the antenna i and pixel n in the antenna j . The visibility for the pairs of dishes i and j is given by

$$V_{ij} = \sum_{\substack{m \in \text{Beam } i \\ n \in \text{Beam } j}} U_{\Delta(i,m)-(j,n)} \tilde{B}_{i,m} \tilde{B}_{j,n}^*, \quad (6)$$

where the index to U schematically implies the uv -plane distance between pixel m on antenna i and pixel n on antenna j . This equation now has the form of Eq. (3) (with gains replaced by pixels inside the beam response), but added over all the pixel pairs formed by the two beams.

In other words, the pixelization is in effect modeling the redundant array made of $N_s \times N_s$ antennas as a grid of $N_{\text{eff}} = (2M + 1)N_s \times (2M + 1)N_s$ independent antenna elements which are phased up in blocks of $(2M + 1) \times (2M + 1)$.

We can now return to the question of the applicability of this technique to the flat-sky approximation. The redundant calibration is typically derived without considering going beyond the flat-sky approximation, but it is clear that it applies generally. An identical pair of antennas will see nominally the same signal and so the values on the uv -plane can be thought as simply keeping track of the true visibility corresponding to a certain baseline orientation without specifying how these are related to the actual sky emissions. The same argument applies to our case, but now involves an approximation. Our model implies that the individual beam can be composed from phased-up elements of equivalent antennas. Consider two sub-baselines of one such baseline pair of the same length, one stretching from central to central pixel and one stretching from two side pixels. While the model can account for different gains, it cannot account for different “primary fields of view” of two such finite-size subelements, especially far away from the pointing center. Therefore, this description will be necessarily approximate. At the same time, it is also true that in the limit of arbitrary fine pixelization, where the subelements are smaller than the wavelength, the primary fields of view of imagined subelements encompass the entire sky and can thus describe a general beam profile. However, the number of unknowns in that case will explode beyond what can be reasonably fit.

C. Redundancy factor

The total number of unknowns to be determined is thus given by (i) the total number of pixels required to describe all the beams $N_a(2M + 1)^2$ and (ii) the pixels required to describe the uv -plane, which are given by the number of redundant baselines in the effective phased-up array $2N_{\text{eff}}(N_{\text{eff}} - 1) = 2N_s(2M + 1)(2MN_s + N_s - 1)$. The total number of measurements, on the other hand, goes as the total number of baselines with measured visibilities, which is given by $N_a(N_a - 1)/2$. Thus, we can construct the redundancy factor for a square close-packed array:

$$r = \frac{\sqrt{N_a}(N_a - 1)}{2[(2M + 1)(6M\sqrt{N_a} + 3\sqrt{N_a} - 2)]}, \quad (7)$$

which describes the ratio of the number of measured quantities divided by the number of unknowns. The $M = 0$ case reduced to the standard redundant calibration scheme. Note that since we are assuming a square array, the $\sqrt{N_a}$ is always an integer. When $r \geq 1$, the system is overconstrained and we can hope to constrain it, subject to known degeneracies. Also, note that the ratio in Eq. (7) is for the particular case of square lattice. For an imperfectly filled

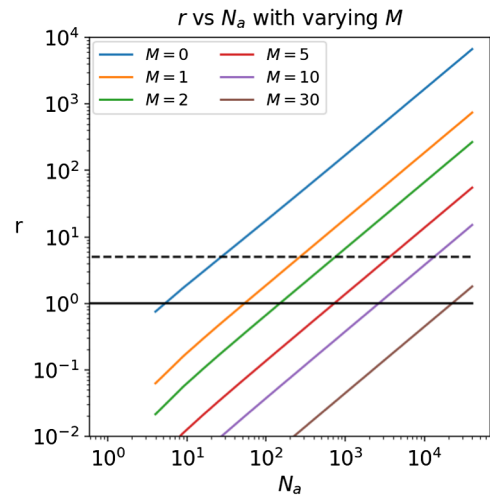


FIG. 2. Redundancy ratio as a function of N_a with varying M . Solid black line is $r = 1$ and dashed black line is $r = 5$ corresponding to significant redundancy.

array or a hexagonal close packing, the ratio would be different. But, note that this formalism does not even require antennas to be on the lattice. It can be calculated for any array and even an irregular but sufficiently packed array will have $r > 1$.

In Fig. 2 we plot r as a function of N_a for a couple of values of M . We see that the for the traditional redundant calibration, even very small arrays containing tens of antennas are sufficiently redundant. HIRAX, with a 32×32 array, can theoretically model up to $M = 5$ but realistically more likely up to $M = 2$. In what follows we will focus on the $M = 1$ case which is perhaps sufficient for the current arrays such as HERA. In this exploratory work, we have not done an explicit calculation for a cylinder array configuration like that of CHIME, although that would be a straightforward generalization.

D. Comparison with other approaches

This is not the first paper to extend the minimal redundant calibration. In fact, in the seminal paper [12] they considered an option to Taylor expand the uv -plane to first order around the lattice positions. In other words, the model for visibilities is expanded from Eq. (3) to

$$V_{ij}^o = (U_{i-j} + \nabla U_{i-j} \cdot \Delta \mathbf{x}_{i,j}) g_i g_j^* + \epsilon, \quad (8)$$

where $\Delta \mathbf{x}_{i,j} = \mathbf{x}_i - \mathbf{x}_j$ is the vector difference between the nominal and actual baseline length, so that the term in brackets corrects for the slightly nonredundant baseline length. One can then self-consistently fit for all the derivatives ΔU_{i-j} and individual antenna positions.

Given the discussion above [see Eq. (3)], the standard redundant calibration assumes the primary beams to be perfectly aligned and so the quantity U_{i-j} is the actual uv -plane after the convolution with primary beams. Therefore,

the Taylor expansion method can deal with geometrical errors (see below) exactly (within the limits of Taylor expansion), but not with pointing errors. We see that an almost natural extension of this approach is to extend the Taylor expansion in the beam mispointing direction as follows:

$$V_{ij}^o = (U_{i-j} + \nabla U_{i-j} \cdot \Delta \mathbf{b}_{i,j} + \Phi_{i-j} \cdot \Delta \mathbf{p}_{i,j}) g_i g_j^* + \epsilon, \quad (9)$$

where Φ is a 2D field defined on the grid points that represents a response of the system to mispointing and $\Delta \mathbf{p}_{i,j} = (\mathbf{p}_i + \mathbf{p}_j)/2$ is the mean mispointing.¹ The system will still be overconstrained for a sufficiently large array. We have not pursued this approach in the name of generality. Since sufficiently big arrays can be very redundant, the idea is to let the system be capable of absorbing arbitrary imperfections, even those that fall in the standard expectation of geometry, pointing, and beam-size errors.

An alternative approach has also been described in [21]. The idea here is to deal with imperfect redundancy by sweeping it under the rug, statistically. The main idea is that in a perfectly redundant array, the redundant baselines would measure the same number; as we dial up the array imperfections, the nominally redundant arrays will not measure the same values, but only highly correlated ones. One can self-consistently write a statistical quadratic model that tracks those correlations and is able to gracefully describe the array nonredundancy. The main issue with this approach is that the differences are not exploited for their information content; they are simply treated as random fields rather than information that can be used to understand in precisely what way the array is imperfect.

III. NUMERICAL IMPLEMENTATION

In previous sections we have presented a general method for modeling the signals from an imperfect interferometer. The method contains the “resolution” parameter M which controls how fine the pixels which describe the response of each individual beam are. $M = 0$ corresponds to standard redundant calibration and $M \rightarrow \infty$ is a completely general description. The hope is that a relatively small M will be sufficient to describe typical level of nonredundancy. The purpose of this section is to find to what extent this is true and our logic is to attempt to use $M = 1$ on relatively small arrays to focus on pointing and geometry errors. The path we follow is to assume a concrete perfectly redundant array observing a given sky and then see how the general redundant calibration performs compared to

¹While seemingly counterintuitive, the mispointing goes with the mean displacement, not the difference. If one dish points off center to the north, while the other off-center to the south, then to first order they will see the same signal, which is suppressed by factor that is quadratic in mispointing. If they both mispoint north, the signal will change in proportion to the mean mispointing.

a standard redundant calibration as we introduce beam nonredundancies.

A. Model beam and its imperfections

Our model beam is a circularly symmetric tapered filled circle in the Fourier space:

$$\tilde{B}_c(r) = 1 - \frac{1}{1 + e^{-2\frac{r-r_0}{\Delta r}}}. \quad (10)$$

This is a more realistic than the usual approximations which usually assume a purely Gaussian or purely Airy disk beams and suffices for this exploratory work. We plot the function in Fig. 3. Note that this is the description of the Fourier transform of the beam, rather than the beam itself, i.e., the real-space beam becomes an Airy disk in the limit of Δr going to zero. The reason why we specify the beam in Fourier space is partly because this is the input quantity we need for making predictions, but also more importantly it corresponds to space with a compact beam representation. Since the antenna is physically compact, the complex electric field in a plane just above the antenna is compact. This approximation breaks down in two ways in realistic instruments. First, when we Fourier transform this to obtain the shape of the beam on the sky we are implicitly making a flat-sky approximation. Second, coupling to neighboring elements will make the antenna physical size bigger than what we mechanically consider to be the antenna element.

We use units of lattice spacing, so radius $r_0 = 0.5L$ would correspond to dishes that just touch. In our simulation we use $r_0 = 0.4L$ and $\Delta r = 0.05L$. We show the beam in real and Fourier space in Fig. 4. The beam shows the expected sidelobes in real space.

In this work we focus on the most common types of errors encountered in redundant arrays: the pointing errors (the center of the beam of the dish is misplaced) and geometry errors (the dish is not at its nominal location on the lattice). Pointing errors are displacement of the beam in the Fourier space and thus correspond to applying a phase

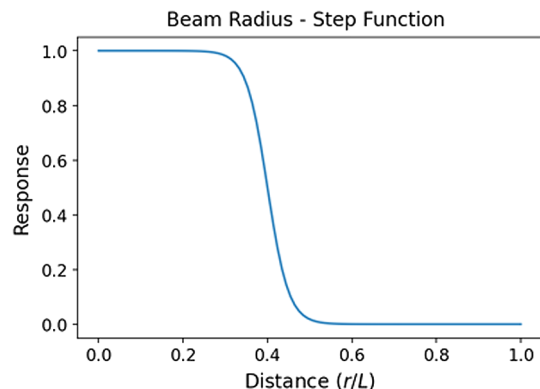


FIG. 3. Radius response function with $R = 0.4L$ and taper = $0.05L$.

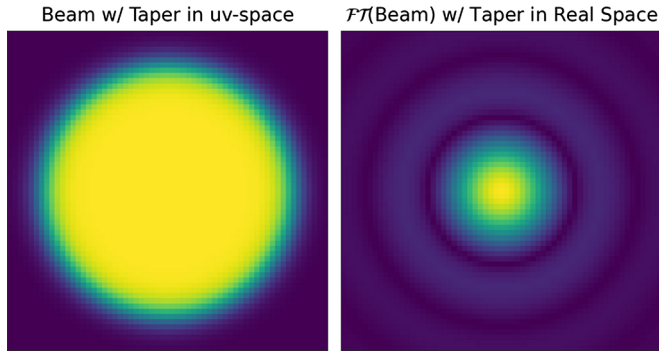


FIG. 4. 2D visualization of the beam \tilde{B} used on the left and its Fourier transform, the actual primary beam B , on the right. The left image spans the uv -plane while the right is in the x - y plane. In the limit of no taper, the beam on the right would be an Airy disk pattern.

gradient across the complex beam response of a single dish. The geometry errors on the other hand are just simple displacement of the dish in the Fourier domain. A full expression for the beam is thus given by

$$\tilde{B}(x, y) = A \tilde{B}_c \left(\sqrt{(x-x_o)^2 + (y-y_o)^2} \right) e^{-i(p_x x + p_y y)}, \quad (11)$$

where A corresponds to the overall complex gain factor, (x_o, y_o) is the geometric error, and (p_x, p_y) is the pointing error. We draw both the overall complex gains from a Gaussian distribution centered around $1 + 0i$ with a variance of σ_g . Geometric and pointing errors are drawn from 2D Gaussians with 1D variances of σ_g and σ_p , respectively.

In Fig. 5 we illustrate a perfect and an imperfect array with the types of errors discussed above.

B. Simulating signal

We simulate the signal relying on a flat-sky approximation. We do so using two approaches. In the first approach, we simply generate data using a large value of M . We use $M = 14$ (corresponding to beam images of 29×29 pixels). In this case we take the true values of the uv -plane as random Gaussian variates, corresponding to a white-noise signal on the sky. We have checked that increasing M beyond the chosen value used does not affect our results.

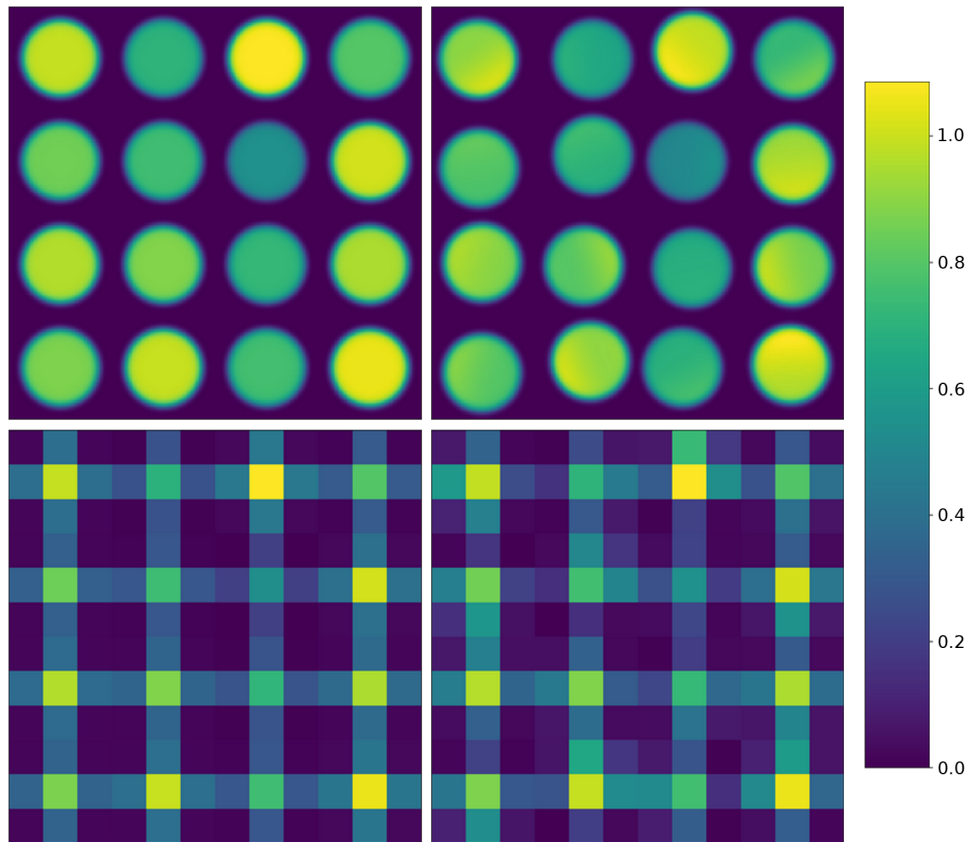


FIG. 5. Example image of a 4×4 array in uv space with a perfectly redundant system on the left and a perturbed version of the same systems on the right. The top images are generated using $M = 30$ and bottom are the same system but downsampled to $M = 1$ resolution. The complex gain of each pixel in the telescope dish is turned into an RGB value for this image with the color bar referring to the overall amplitude ranging from 0 to slightly greater than 1. The perturbations for the right-side graphs were generated with geometric errors at $\sigma_g = 0.1L$ and pointing errors at $\sigma_p = 0.5$. Pointing errors which lead to phase changes are modeled as a gradient on each beam and geometry errors are modeled as circle offsets.

In this approach, the code using for generating the signal and fitting is essentially the same with only the value of M being different. On one hand this is elegant, but it is also prone to a potential bug inadvertently canceling out between data generation and fitting. Moreover, the realistic skies are often dominated by a few bright sources.

In the second approach we generate the signal as a sum over discrete sources. This is to take into account the fact that some sources are considerably brighter than the others. For each source and baseline we do the following:

- (i) Calculate the primary beam response at the location of the source for both dishes. Given pointing errors, this is different for each baseline.
- (ii) Directly calculate the response of the interferometer for the baseline length spanned by a given pair of dishes (taking into account geometry errors).

We use sources whose fluxes are randomly drawn uniformly in log from 1 to 1000. A comparison between the two methods can be found in Fig 6. Aside from an overall scale factor the two are quite similar.

In both approaches, once the noiseless visibilities are calculated, we multiply visibility by the total complex gain contribution $g_i g_j^*$, where g_i is the gain for dish i . Note that the method is completely general with respect to individual gains, i.e., they can be perfectly absorbed into the beam description. Therefore, we keep the variance small to avoid dealing with solver falling into local minima.

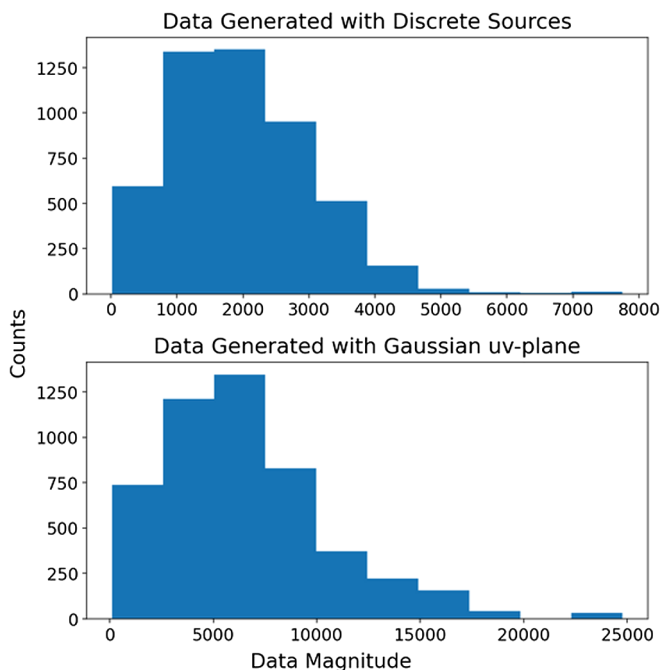


FIG. 6. Histogram comparing the magnitude of observed visibilities generated between the two methods. There are 100 discrete sources used to model the sky. The overall statistics of the two are comparable, alleviating concern about the sky simulation with M .

Finally, we add complex noise ϵ drawn from a Gaussian with variance σ_ϵ^2 . Since the only quantity that matters is the level of noise compared to the level of signal, we will report our results in terms of signal-to-noise ratio (SNR). We define the SNR per visibility as

$$\text{SNR}^2 = \frac{1}{N_{\text{visibilities}} \text{baselines } i-j} \sum \left(\frac{|V_{i-j}^2|}{\sigma_\epsilon^2} \right)^2. \quad (12)$$

There are thus six additional degrees of freedom per each dish: two for pointing error, two for geometry error, and two for overall gain. Even with $M = 1$, each beam is described by 9 complex numbers (i.e., 18 degrees of freedom) so there is reasonable hope that the description is sufficient. However, on purpose, we decided to consider the types of errors that are not perfectly reproducible by our model to assess its flexibility. Had we instead decided to model dish imperfections as random deviations the size of a pixel on our $(2M + 1)^2$ grid, say as dirt 1/9th the size of our dish for $M = 1$, our model would be guaranteed to perform better compared to redundant calibration.

IV. SOLVING FOR \tilde{B}_i AND U

In this work we use simple iterative solvers for beams \tilde{B}_i and the true uv -plane U . Since we are focusing on the modeling side, i.e., how well the solutions perform, these solvers are not designed to be either particularly stable or fast. That being said, some notes and improvements are listed in the relevant sections below.

In general, we are trying to maximize the log likelihood of the model, which is equivalent to minimizing the χ^2 , given by

$$\chi^2 = \sum_{\text{all pairs } i,j} \frac{(V_{ij}^o - V_{ij}^p(\tilde{B}, U))^2}{\sigma_\epsilon^2}, \quad (13)$$

where V_{ij}^p denotes the predicted visibilities which are a function of all beam parameters and uv -plane values. At each step we optimize for either visibilities or beam parameters.

A. Visibility

To solve for visibilities, U , we rely on the fact that observed visibilities are linear in the input visibilities:

$$V_i^o = M_{ik} U_k + \epsilon. \quad (14)$$

Here, i iterates over observed visibilities and k over the input, or true, visibilities. While the specific values will depend on the geometry of the array, $i \sim N_a^2$ and $k \sim N_a^2 M^2$. The matrix M is quite sparse, but because beams overlap, it is not a block matrix. In other words, neighboring baselines do probe some of the same sky signals as illustrated

in Fig. 1. Matrix M depends on all beams and so we assume the current best guess for the beams (which improve with every iteration). We rely on SciPy [23] for actual calculations. Specifically, to calculate $(\tilde{B}_i \otimes \tilde{B}_j^*)$, we use `scipy.signal.convolve1d` and to solve the sparse system we rely on `scipy.sparse.linalg.lsqr`. The rest is rather painful but otherwise straightforward housekeeping. For interested readers we point at some of the more subtle technical issues in the Appendix.

At this point we can use any number of methods to solve a linear equation, with `scipy.sparse.linalg.lsqr` being the most direct in application. We can also employ any number of tools to better solve our linear system such as the Wiener filter, which we motivate in the Regularization section. Using Eq. (19) we can use `scipy.sparse.linalg.spsolve` to solve the system, avoid any tough inversions, and although there is more overhead, it is ultimately faster than directly solving Eq. (14). Notably, employing a Wiener filter is also considerably better at reducing χ^2 per iteration over `scipy.sparse.linalg.lsqr`, which makes the filter a worthwhile implementation.

B. Beams

While Eq. (2) is nominally quadratic in the beams, this is not an issue in our actual problem, because we do not consider autocorrelation signal. To solve for the beam B_i , we fix all the remaining beams and the solved uv -plane U , so that

$$V_k^o = M_{kl} \tilde{B}_{i,l} + \epsilon. \quad (15)$$

Here, index k runs over all visibilities that depend on the beam i (i.e., all baselines that contain antenna i) and l over all pixels of beam \tilde{B}_i . This is a dense system that we solve using `scipy.optimize.lsqr_linear` separately for each beam i . While this can be embarrassingly parallel, we found improvements to χ^2 per iteration when done sequentially. After solving for \tilde{B}_i , the updated solution is used when writing M_{kl} for \tilde{B}_{i+1} and so on. Of course this can be circumvented with some clever distributed computing which we leave for the future.

V. PERFECT DEGENERACIES

The standard redundant calibration has perfect degeneracies spanning four degrees of freedom:

- (i) Multiply gains by complex factor α and divide sky signal by $\alpha\alpha^*$ (this is often split into the overall amplitude degeneracy and the overall phase degeneracy);
- (ii) Translate the sky by a vector \mathbf{p} and apply a commensurate phase gradient across the gain solution.

The same degeneracies continue to exist in our case. One would naively expect that we also have a similar set of per-

element degeneracies; however, these are not present, in general, because neighboring antennas actually measure many of the same points in the uv -plane, thus introducing “interlocking” of the uv -plane solutions.

However, we have a different kind of degeneracy present. We know that if an array is truly redundant, then one needs only $\sim 3N_s^2$ numbers to describe the data. So, if the array is actually fully redundant, we are free to pick any “shape” of the beam \tilde{B} for $M > 0$ and still have sufficient freedom in the U array to form a model that gives precisely the same predictions. In the other limit, if the array is really nonredundant then this degeneracy disappears. Therefore, this is not really a model degeneracy, but a degeneracy associated with perfectly redundant array solutions which are, from a mathematical perspective, pathological. This is analogous to solving a matrix equation $Mx = y$, which for a general matrix M is solvable by $x = M^{-1}y$, unless matrix M is singular. Removing these degeneracies using polarization data, as suggested by [24], should be plausible but tenuous with nonredundant arrays. With our array formalism it is possible to tackle this issue but left as a future endeavour.

In practice, the presence of noise will instead use the extra model freedom to “fit the noise” and find a nominally better solution as we will describe in the Results section. A formally correct way would be to perform a strict Bayesian model comparison, where we weight the solutions by the Bayesian evidence in favor of a certain model: if the model with $M > 0$ can fit the data equally well as the standard redundant calibration, then the standard redundant calibration would be strongly favored due to having many fewer priors.

VI. REGULARIZATION

To prevent overfitting, we implement two regularization schemes as follows. We first introduce a prior on the beam parameters that attempts to pull the beam solution back to the fiducial, redundant beams, and also include an option to minimize the variance on solved visibilities. The total likelihood with the beam prior is thus given by

$$\log L = -\frac{1}{2}\chi^2 + \sum_{\text{beam } i} \sum_{\text{pixel } k} \left(-2 \log \sigma_B - \frac{|\tilde{B}_{ik} - \tilde{B}_k^0|^2}{2\sigma_B^2} \right), \quad (16)$$

where χ^2 is given by Eq. (13) and \tilde{B}^0 corresponds to the beam prior. As stated above we choose this to be the perfect, unperturbed beam.

The σ_B describes typical deviation from perfect beams. Taking derivative of the $\log L$ with respect to σ_B (with other parameters fixed), one gets

$$\sigma_B = \sqrt{\frac{1}{2N} \sum_{\text{beam } i} \sum_{\text{pixel } k} (|\tilde{B}_{ik} - \tilde{B}_k^0|^2)}, \quad (17)$$

with N being the total number of pixels over all beams, $(2M + 1)^2 N_a$.

If the array is close to truly redundant, the system can achieve a good fit by floating beams towards nominal beam values and lowering σ_B (and thus achieve a large likelihood improvement through the normalization term $-\log \sigma_B$). However, for a nonredundant array, it is better to raise σ_B and instead improve the χ^2 . This prior will bias the beams towards having a magnitude of 1, or whatever the normalization of the perfect beam is set to. To counteract this bias, we introduce an overall gain for each dish which applies to each pixel. Our solver implements the *omnical* [14] method to iteratively solve for overall gains.

Finally, we put a prior on visibilities. In a more complete treatment, we would solve for the visibility power spectrum at the same time as we solve for the visibility map, similar to the Gibbs' sampling approach to CMB (see [25]). In our case we just add a penalty term to the solved visibilities in the form of

$$f = \|\Sigma_b^{-1} \cdot (M_{ik} U_k - V_i^o) + \Sigma_v^{-1} \cdot U_k\| \quad (18)$$

with respect to U_k , where $\Sigma_b^{-2} = N^{-1}$ is the noise covariance and $\Sigma_v^{-2} = S^{-1}$ is the data covariance. In our simulations we assume both are diagonal to simplify calculations. This leads to a Wiener filter. Following the form written in [26] we specifically solve for visibilities using the following linear equation:

$$[S^{-1} + (M_{ik})^\dagger N^{-1} M_{ik}] U_k = ((M_{ik})^\dagger N^{-1}) V_i^o. \quad (19)$$

VII. RECAP OF THE ITERATIVE SOLVER

To recap, we maximize the likelihood with iterative steps. In order, each step involves fitting for:

- (1) The visibilities by solving either Eq. (14) or (19).
- (2) The beam shape by solving (15) and incorporating the prior parameter, σ_B , by adding rows to the bottom of our linear system.
- (3) The overall per-antenna gain parameters A_i via *omnical* [14].
- (4) Finally, the beam parameter σ_B which amounts to calculating the variance of beam solutions with respect to their nominal values as written in (17).

VIII. RESULTS

The purpose of telescopes is to image the sky. If we have an imperfect description of the physical effects that affect the instrument's response, and calibration vectors that are only approximately correct and whose values are "effective," it should not matter as long as the map of the sky is

faithfully reproduced. Therefore, the method that will best recover the input uv -plane is the most successful method.

We have condensed our results into two figures, Figs. 7 and 8, that succinctly summarize our results regarding the method. For both figures, the upper row of plots corresponds to rather modest geometry and pointing errors, while the bottom row contains considerably worse errors. All figures show various quantities plotted as a function of per-visibility SNR [see Eq. (12)]. The left column shows the values of χ^2 : high values of χ^2 indicate a poor fit, while lower values a better fit; values below the shaded region indicating the expected χ^2 given degrees of freedom indicate overfitting. The central left column shows the goodness of the uv -plane recovery as measured by the variance between the true and recovered uv -plane, labeled σ_V . This is only done for the uv -plane pixels that lie on the nominal beam-array lattice positions where the SNR is concentrated. The lower the value, the more faithful recovery of the uv -plane a certain method is giving. The central right column plots show the variance of the recovered beams relative to the true beams which we call Σ_B , while the variance between recovered beams and prior beams, σ_B , is plotted in the rightmost column. The lower the Σ_B , the closer we are in recovering the true values of the beam. The value of σ_B on the other hand tells us not only how far the beams are from their nominal (unperturbed) values but also how far our method thinks they are in average, since σ_B is at the same time a regularization parameter (see Sec. V). Remember that the overall phase factor is set in a separate parameter and that the central beam pixels were set to unity—central pixels are not used in the variance calculation.

Three different colors correspond to three different methods. In blue we plot the standard redundant calibration which we implement by simply forcing beams to be at their nominal, unperturbed values and recover just the gains and the uv values. Note that it is still not exactly the same as standard redundant calibration, because the recovery is on the uv -plane quantities before convolution with the beam. In orange we plot the same, but this time we instead fix the beams to their true value by assuming that somehow these were independently calibrated with infinite precision. This sets the upper limit of how well a generalized redundant calibration should perform. Finally, in green we show the results of our method in its full self-calibrating glory.

We start with a simplified case, Fig. 7, in which we generate the data using exactly the same theory as we fit (the data are generated with $M = 1$ and fit with $M = 1$). In this case, the fitted theory is by construction a perfect description of the data. The purpose of this exercise is to isolate the effects of overfitting from effects of using an approximate theory.

At low signal to noise, the method cannot really detect with certainty that the beams are different from their unperturbed values. The regularization thus follows the beam prior, resulting in a small σ_B and a relatively large Σ_B .

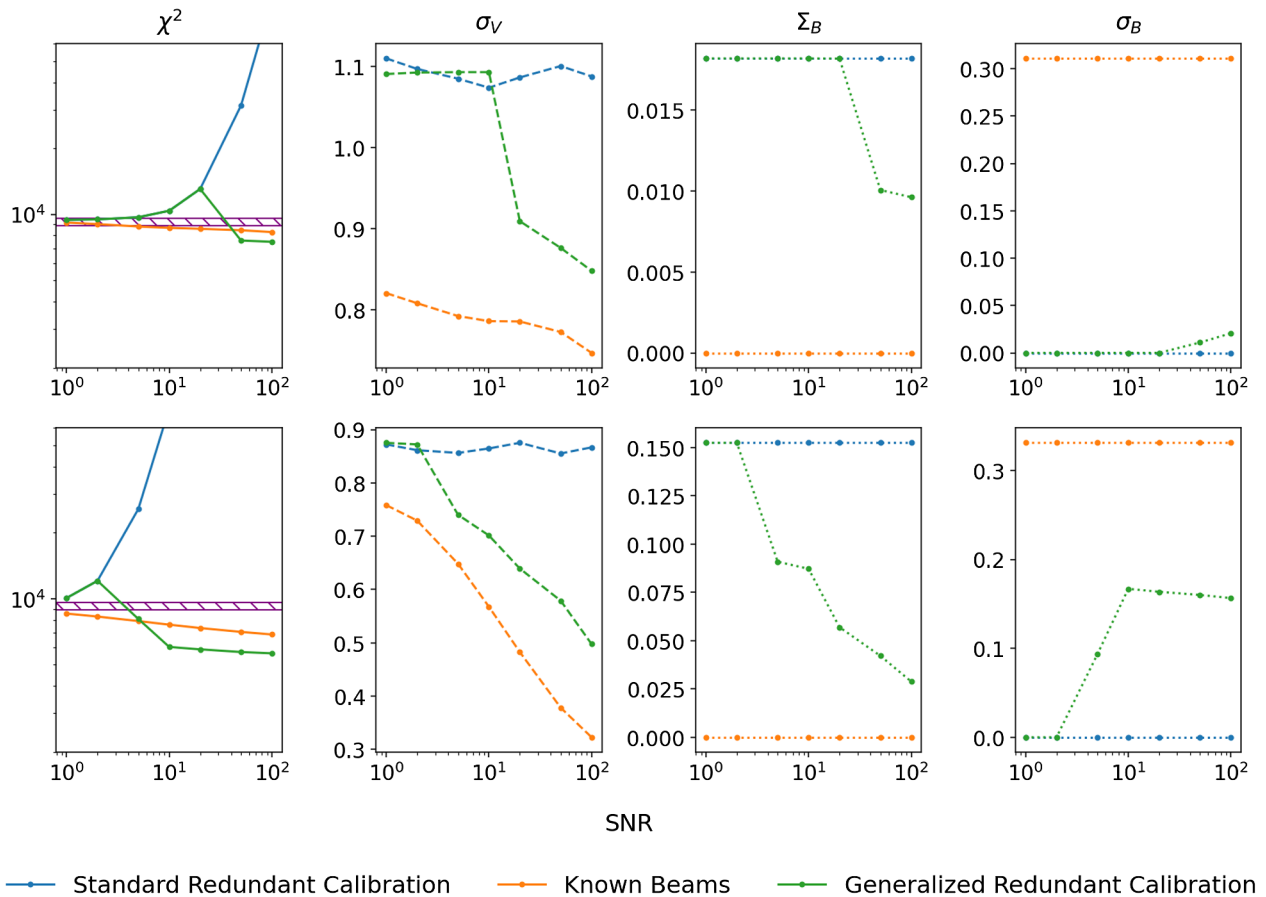


FIG. 7. Comparison of our algorithm (green) against perfect knowledge of beam shape (orange) and also against an assumption of a perfectly redundant array (blue). Blue and orange have fixed beams, at the prior and truth, respectively, and can only vary the visibilities and overall gain factors. As such, these two methods are forms of standard redundant calibration. Our algorithm solves using $M = 1$ and the data are generated with $M = 1$, meaning that the model is a perfect description of the data. The top row corresponds to a weak departure from perfect redundancy with the geometric and pointing errors set to 0.01, while the bottom row is for relatively strong departure from redundancy with geometric and pointing errors set to 0.1. The plots in the first column show χ^2 for each model to see if we can recreate the data. The purple dashed box is the 5% percentile to 95% percentile bound of the expected χ^2 for generalized redundant calibration with no regularization. χ^2 below this box is indicative of overfitting. The second column plots the variance between the solved and true visibilities, σ_V , to see if we are accurately recovering the true sky—since we are generating and solving using $M = 1$ we can compare the solved to true values directly. The third column plots the variance between the solved and true beam shapes, labeled Σ_B , to see if we recover the correct array. As the blue and orange lines use a fixed beam shape (the beam prior and true beams, respectively), their values for Σ_B are constant. The fourth column plots the variance between the solved and prior beam shape, σ_B from Eq. (17).

As the signal to noise increases, the tension becomes significant, and the method relaxes σ_B , allowing beams to unstick from their unperturbed prior values. Σ_B therefore undergoes a transition to smaller values with beam values moving towards their true values. In the bottom panel, as SNR increases, our algorithm asymptotes towards $\Sigma_B = 0$, matching the orange line. The process works better for the more nonredundant array, because while nonredundancy is a source of noise generically, in our case it is also a source of signal. In χ^2 this is manifested as χ^2 increasing with signal to noise and then “snapping” down towards lower values when σ_B is relaxed. We find that even in this case the statistical system is prone to overfitting. This indicates that

further regularization is likely warranted in the uv -plane sector. At the same time, we see the redundant calibration simply unable to explain the data with the χ^2 monotonically increasing away from good fit with either increasing SNR or the nonredundancy of the array.

In the second column we see consistent behavior. The orange line with known beams performs best and its variance on the visibilities monotonically decreases as the SNR increases—the fidelity of the solution increases. The standard redundant calibration is systematically limited and as the χ^2 increases it is simply unable to improve its solution. Overall, the generalized redundant calibration is between the blue and orange lines with issues of falling into

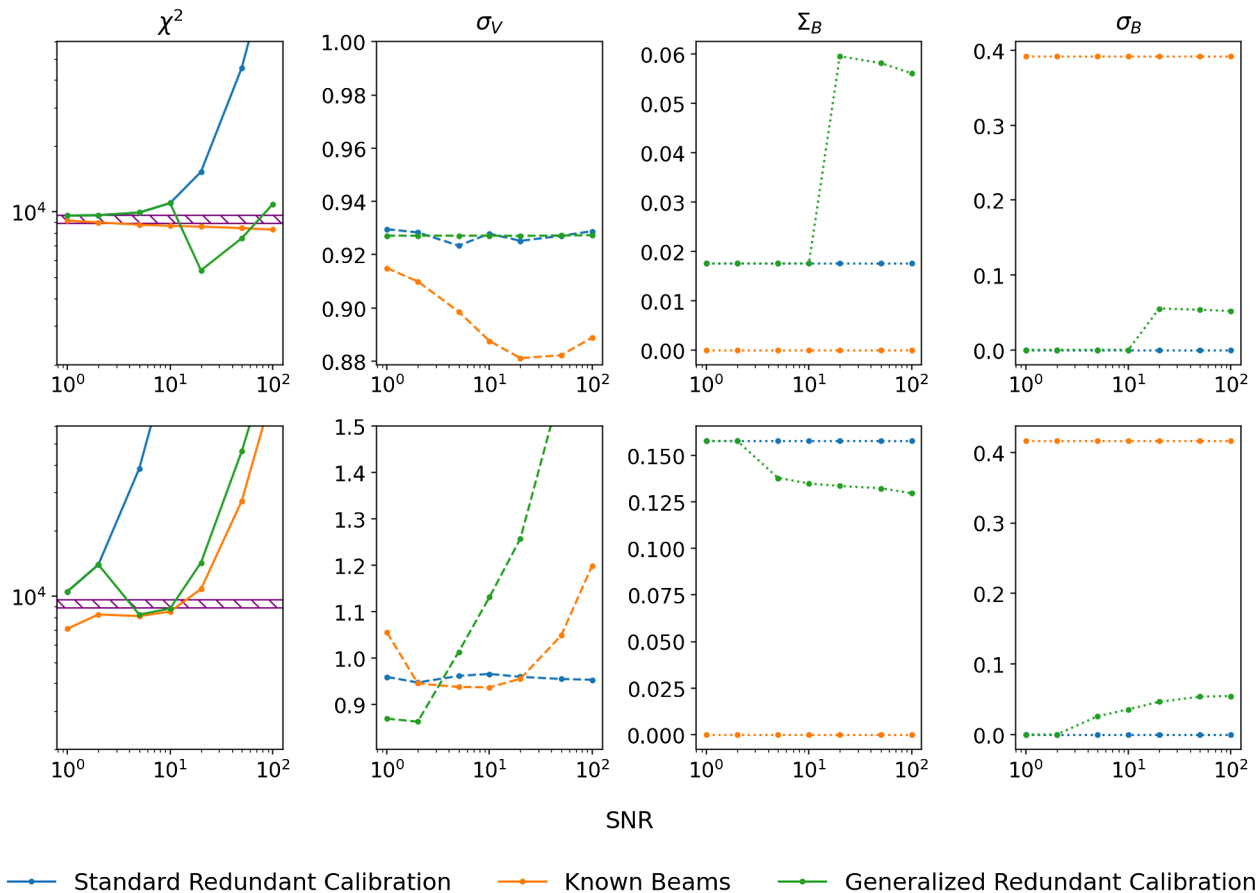


FIG. 8. Same as Fig. 8 but with data generated using $M = 14$ which is essentially indistinguishable from the continuous model. The data were still fitted using $M = 1$ model, which is now an approximate model. For σ_v , the central pixels from the fitted uv data have been compared with the averaged uv -plane pixels from the corresponding area from the model representing the truth. See text and caption to Fig. 8 for complete description.

local minima at low SNR. But, for sufficiently high SNR it outperforms the standard redundant calibration as expected. It never performs quite as well as the case with known beams, because the latter utilizes significant extra amounts of information.

In Fig. 8 we show the same set of plots for more realistic cases where we simulated data with $M = 14$ and recover them with $M = 1$ theory as before. Here, the beams are described in a true theory by a 29×29 matrix which we average to a 3×3 matrix in cells in order to get “true” beams. Similarly, we average the true uv -plane into a reduced resolution uv -plane that we compare with recovered values. The details of how this is performed matter, but largely do not affect our results. The right-hand side plots showing the σ_B behavior are largely unchanged. However, we see that the χ^2 plot keeps increasing as the SNR increases. This indicates that we are fitting a model that is actually unable to fit the data. The main assumption of this paper, namely that a low number of extra degrees of freedom in individual beams parametrized as pixels will be sufficiently flexible to describe a general geometrical and pointing errors has proven to be insufficient at SNR per

measurement of over ~ 10 . As we increase the signal to noise, the recovery of the map (second panel) first improves, but then starts to get worse. Interestingly, we find that for large level of array nonredundancy all methods perform worse at higher SNR; this is possible when methods fall into the wrong local minima or where the imperfections of the model are such that better formal fits actually perform worse in the quantities of interest. But, it is also possible that our recipe to convert the $M = 14$ truth into $M = 1$ pixelized uv -plane is just too simplistic. In this case, the implication is that we have a solution that is formally good but with a poorly understood relation to the underlying truth.

IX. CONCLUSIONS

We have presented a new method for calibrating imperfect redundant arrays. The method is a derivative of redundant calibration and models each independent beam element as a phased-up array of $(2M + 1) \times (2M + 1)$ subelements, each with its own complex gain factors. In the limit of large M , the method is capable of modeling any array by having a complete freedom to represent the

response of each dish as pixelized $(2M + 1) \times (2M + 1)$ complex beam response. In the limit of $M = 0$, the method reduces to the standard redundant calibration.

To avoid fitting for the noise we have attempted a regularization scheme that models the departures from the perfect beams using a Gaussian with diagonal scatter. The magnitude of scatter σ_B is a fitted parameter. As expected, we found that when signal to noise is low, and the data are sufficiently well described by the standard redundant calibration, the beam solutions relax to their priors and σ_B becomes essentially zero. In this limit, the system has less tendency to fit for the noise, although we find that the χ^2 remains too low and noise fitting remains an issue. When the signal to noise is sufficient to detect non-redundancy, the value of σ_B rises and for sufficiently nonredundant array, the solutions approach those without regularization.

In this paper we have focused on methodological aspects of this method, namely, is the method capable of producing good fits to the data? In practice, while this might be true, the very high dimensionality of this problem makes finding these solutions difficult. We found that, unless we start with a good approximate guess, the method is likely to fall into local minima. Therefore, in order to make this method practically usable, it is necessary to first find efficient minimizers. Moreover, the method currently works with a single sky snapshot and should be generalized to time-stream data.

Unfortunately, we have found that low- M configurations are not good at describing generic array errors. The high- M configurations will likely perform considerably better, however, given the increased model complexity in that case is even more likely to suffer from overfitting and preponderance of local minima. The correct solution to these issues is to employ a much more sophisticated marginalization scheme than the maximum likelihood scheme employed in this work.

While we have focused on nearly redundant arrays, our method is trivially generalizable to only partially redundant arrays in arbitrary configuration. A fixed value of M defines a grid with $D/(\lambda M)$ in the uv -plane. Any array containing dishes (even heterogeneous ones!), whose response with respect to some arbitrary original can be satisfactorily described on the uv -plane in this grid can be in general fit with generalized redundant calibration. Of course, this model is interesting only if the array has a sufficient redundancy that the number of unknowns does not exceed the number of observed visibilities, since otherwise it is capable of trivially explaining any measurement.

We found that the main downside of this method is that it is nontrivial to connect the measured values to the underlying quantities of interest. We have seen that the central uv -plane pixels (i.e., those on the nominal lattice positions) are well recovered, but this has not been investigated in any

detail. In comparison, in the standard redundant calibration, the interpretation of the fitted uv values is clear: they are exactly the values of the true uv -plane convolved with the appropriate beam responses. Unfortunately, in generalized redundant calibration, however, the actual beams are not made up from identical sub-beams. Therefore, it is non-trivial to precisely connect the recovered $(2M + 1) \times (2M + 1)$ pixelized beam back to actual pointing and geometric offsets.

The beauty of the proposed scheme is that it is very general, especially for a large array that could afford to go beyond $M = 1$. On the other hand, if we did know that the dominant errors are pointing and geometry errors, one could design a model that would fit explicitly for those. In that case, the contents of the uv -plane could be modeled by the values and its spatial derivatives at the lattice points. We leave extended comparisons of these methods for the future.

ACKNOWLEDGMENTS

The authors acknowledge useful feedback obtained during the PUMA collaboration seminar series. This work was supported in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internships Program (SULI).

APPENDIX: TECHNICAL CONSIDERATIONS IN FITTING

Since we are probing nearby baselines, it is possible that we probe the conjugate of a known visibility, i.e., while looking at the baseline $(0,1)$ we will pick up some amount of signal from the baseline $(-\frac{1}{3}, 1)$. It is important that we are careful to enforce that

$$U_{(-\frac{1}{3}, 1)} = (U_{(\frac{1}{3}, -1)})^*. \quad (\text{A1})$$

This means that even if we construct a large sparse matrix M [from Eq. (14)] with each row corresponding to the probed baseline $\alpha = (i, j)$ with coefficients from P_{ij} (the convolution of two beams), we cannot simply solve for \vec{U} using this system.

Thus, the correct way to implement Eq. (14) would be to split our \vec{U} into real and complex components and doing the same with our data. Explicitly, given a generic $a, b \in \mathbb{C}$, with $a = a_1 + a_2i$ and similarly for b , we can say

$$a \cdot b = (a_1b_1 - a_2b_2) + (a_1b_2 + a_2b_1)i, \quad (\text{A2})$$

$$a^* \cdot b = (a_1b_1 + a_2b_2) + (a_1b_2 - a_2b_1)i. \quad (\text{A3})$$

So, it simply comes down to flipping the sign on a few of the coefficients of our convolution, P_{ij} , for some of the baselines.

While the housekeeping for visibility solving boils down to changing signs and splitting our complex data into two parts, writing out Eq. (15) to solve for the beams is a bit more involved.

To write our 2D convolution as a matrix product we can use a blocked Toeplitz matrix. For each \tilde{B}_i we can write its effect in convolution as a matrix β_i by writing it as a blocked Toeplitz matrix, and for \tilde{B}_j we simply have β_j^\dagger . Then, our data can be written as

$$V_{ij} = (\tilde{B}_j^\dagger \otimes \tilde{B}_i) \cdot \vec{U}_\alpha \quad (\text{A4})$$

$$= \vec{U}_\alpha \cdot (\beta_j^\dagger \tilde{B}_i) \quad (\text{A5})$$

$$= (\vec{U}_\alpha^T \beta_j^\dagger) \cdot \tilde{B}_i \quad (\text{A6})$$

$$= M_{ij} \tilde{B}_i \quad (\text{A7})$$

Here, we have a fixed i and iterate over the other beams j . Unlike the visibility solver, there is no relation between any of the beams and thus we are able to solve each one as an independent linear system.

-
- [1] K. Bandura *et al.*, Canadian hydrogen intensity mapping experiment (CHIME) pathfinder, *Proc. SPIE Int. Soc. Opt. Eng.* **9145**, 22 (2014).
- [2] F. Wu, J. Zhang, Y. Wang, X. Chen, and H. Shi, Tianlai: A 21 cm radio telescope array for BAO and dark energy, status and progress, in *Proceedings, 51st Rencontres de Moriond, Cosmology session: La Thuile, Italy, 2016* (ARISF, Paris, 2016), p. 315.
- [3] D. Jacobs and HERA Collaboration, Hydrogen epoch of reionization array—report on last phases of construction and first light, in *American Astronomical Society Meeting Abstracts #236*, American Astronomical Society Meeting Abstracts Vol. 236 (AAS, Washington DC, 2020), p. 308.04.
- [4] C. M. Trott *et al.*, Deep multiredshift limits on Epoch of Reionization 21 cm power spectra from four seasons of murchison widefield array observations, *Mon. Not. R. Astron. Soc.* **493**, 4711 (2020).
- [5] A. R. Parsons, D. C. Backer, G. S. Foster, M. C. H. Wright, R. F. Bradley, N. E. Gugliucci, C. R. Parashare, E. E. Benoit, J. E. Aguirre, D. C. Jacobs *et al.*, The precision array for probing the epoch of re-ionization: Eight station results, *Astron. J.* **139**, 1468 (2010).
- [6] G. Swarup, Giant metrewave radio telescope (GMRT)-Scientific objectives and design aspects, *Indian J. Radio Space Phys.* **19**, 493 (1990).
- [7] L. B. Newburgh *et al.*, HIRAX: A probe of dark energy and radio transients, *Proc. SPIE Int. Soc. Opt. Eng.* **9906**, 99065X (2016).
- [8] C. A. Wuensche and the BINGO Collaboration, The BINGO telescope: A new instrument exploring the new 21 cm cosmology window, *J. Phys. Conf. Ser.* **1269**, 012002 (2019).
- [9] P. Dewdney, W. Turner, R. Millenaar, R. McCool, J. Lazio, and T. Cornwell, Skat system baseline design, Report No. SKA-TEL-SKO-DD-001 Revision, 2013.
- [10] K. Vanderlinde *et al.*, The Canadian hydrogen observatory and radio-transient detector (CHORD), in *Canadian Long Range Plan for Astronomy and Astrophysics White Papers* (Zenodo, Geneva, 2019), Vol. 2020, p. 28.
- [11] E. Castorina, S. Foreman, D. Karagiannis, A. Liu, K. W. Masui, P. D. Meerburg, L. B. Newburgh, P. O’Connor, A. Obuljen, H. Padmanabhan, J. R. Shaw, A. Slosar, P. Stankus, P. T. Timbie, B. Wallisch, and M. White, Packed ultra-wideband mapping array (PUMA): Astro2020 RFI response, [arXiv:2002.05072](https://arxiv.org/abs/2002.05072).
- [12] A. Liu, M. Tegmark, S. Morrison, A. Lutomirski, and M. Zaldarriaga, Precision calibration of radio interferometers using redundant baselines, *Mon. Not. R. Astron. Soc.* **408**, 1029 (2010).
- [13] J. S. Dillon and A. R. Parsons, Redundant array configurations for 21 cm cosmology, *Astrophys. J.* **826**, 181 (2016).
- [14] J. S. Dillon *et al.*, Redundant-baseline calibration of the hydrogen epoch of reionization array, *Mon. Not. R. Astron. Soc.* **499**, 5840 (2020).
- [15] S. Choudhuri, P. Bull, and H. Garsden, Patterns of primary beam non-redundancy in close-packed 21 cm array observations, *Mon. Not. R. Astron. Soc.* **506**, 2066 (2021).
- [16] R. C. Joseph, C. M. Trott, and R. B. Wayth, The bias and uncertainty of redundant and sky-based calibration under realistic sky and telescope conditions, *Astron. J.* **156**, 285 (2018).
- [17] M. H. Wieringa, An investigation of the telescope based calibration methods ‘redundancy’ and ‘self-cal’, *Exp. Astron.* **2**, 203 (1992).
- [18] H. Zheng, M. Tegmark, V. Buza, J. S. Dillon, H. Gharibyan, J. Hickish, E. Kunz, A. Liu, J. Losh, A. Lutomirski *et al.*, MITEOR: A scalable interferometer for precision 21 cm cosmology, *Mon. Not. R. Astron. Soc.* **445**, 1084 (2014).
- [19] H. Zheng, M. Tegmark, J. S. Dillon, A. Liu, A. R. Neben, S. M. Tribiano, R. F. Bradley, V. Buza, A. Ewall-Wice, H. Gharibyan *et al.*, Brute-force mapmaking with compact interferometers: A miteor northern sky map from 128 to 175 mhz, *Mon. Not. R. Astron. Soc.* **465**, 2901 (2016).
- [20] N. Orosz, J. S. Dillon, A. Ewall-Wice, A. R. Parsons, and N. Thyagarajan, Mitigating the effects of antenna-to-antenna variation on redundant-baseline calibration for 21 cm cosmology, *Mon. Not. R. Astron. Soc.* **487**, 537 (2019).
- [21] J. L. Sievers, Calibration of quasi-redundant interferometers, [arXiv:1701.01860](https://arxiv.org/abs/1701.01860).

- [22] R. Byrne, M. F. Morales, B. J. Hazelton, and M. Wilensky, A unified calibration framework for 21 cm cosmology, *Mon. Not. R. Astron. Soc.* **503**, 2457 (2021).
- [23] P. Virtanen *et al.*, and *sciPy 1.0 Contributors*, *SciPy 1.0: Fundamental algorithms for scientific computing in python*, *Nat. Methods* **17**, 261 (2020).
- [24] J. S. Dillon, S. A. Kohn, A. R. Parsons, J. E. Aguirre, Z. S. Ali, G. Bernardi, N. S. Kern, W. Li, A. Liu, C. D. Nunhokee *et al.*, Polarized redundant-baseline calibration for 21 cm cosmology without adding spectral structure, *Mon. Not. R. Astron. Soc.* **477**, 5670 (2018).
- [25] B. D. Wandelt, D. L. Larson, and A. Lakshminarayanan, Global, exact cosmic microwave background data analysis using Gibbs sampling, *Phys. Rev. D* **70**, 083511 (2004).
- [26] M. Tegmark, How to make maps from cosmic microwave background data without losing information, *Astrophys. J. Lett.* **480**, L87 (1997).