

Calibration artefacts in radio interferometry.

II. Ghost patterns for irregular arrays

S. J. Wijnholds^{1*}, T. L. Grobler²³, O. M. Smirnov²³

¹*ASTRON, P.O. Box 2, 7900 AA, Dwingeloo, The Netherlands*

²*Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown, 6140, South Africa*

³*SKA South Africa, 3rd Floor, The Park, Park Road, Pinelands, 7405, South Africa*

published in MNRAS, vol. 457, no. 3, pp2331-2354, April 2016, doi: 10.1093/mnras/stw118

ABSTRACT

Calibration artefacts, like the self-calibration bias, usually emerge when data is calibrated using an incomplete sky model. In the first paper of this series, in which we analyzed calibration artefacts in data from the WSRT, we showed that these artefacts take the form of spurious positive and negative sources, which we refer to as ghosts or ghost sources. We also developed a mathematical framework with which we could predict the ghost pattern of an east-west interferometer for a simple two-source test case. In this paper, we extend our analysis to more general array layouts. This provides us with a useful method for the analysis of ghosts that we refer to as extrapolation. Combining extrapolation with a perturbation analysis, we are able to: 1. analyse the ghost-pattern for a two-source test case with one modeled and one unmodeled source for an arbitrary array layout, 2. explain why some ghosts are brighter than others, 3. define a taxonomy allowing us to classify the different ghosts, 4. derive closed form expressions for the fluxes and positions of the brightest ghosts, and 5. explain the strange two-peak structure with which some ghosts manifest during imaging. We illustrate our mathematical predictions using simulations of the KAT-7 array. These results show the explanatory power of our mathematical model. The insights gained in this paper provide a solid foundation to study calibration artefacts in arbitrary, i.e., more complicated than the two-source example discussed here, incomplete sky models or full synthesis observations including direction dependent effects.

Key words: Instrumentation: interferometers, Methods: analytical, Methods: numerical, Techniques: interferometric

1 INTRODUCTION

This is the second paper in a series of papers that investigate the underlying mechanism that causes calibration artefacts. Calibration refers to the estimation and correction of instrumental and environmental errors that are induced in interferometric data. In its most basic form, calibration is accomplished by finding the antenna gains that minimize the difference between observed and predicted visibilities (Rau et al. 2009; Wijnholds et al. 2010; Smirnov 2011; van der Veen & Wijnholds 2013). Calibration artefacts are systematics that exist in interferometric data and generally emerge when you calibrate your data with incomplete sky models. Three main types of artefacts have been observed: the generation of spurious emission, the suppression of real emission and the structural deformation of extended sources (Linfield 1986; Wilkinson et al. 1988; Corn-

well & Fomalont 1999; Martí-Vidal et al. 2010; Martí-Vidal & Marcaide 2008; Kazemi & Yatawatta 2013). These artefacts can be explained by what we refer to as ghost sources or ghosts. Ghosts are spurious sources. It has been shown that ghosts are usually negative and when they form on top of real sources, they cause source suppression (Grobler et al. 2014). These ghosts manifest in specific geometric patterns, which we refer to as ghost patterns.

Understanding calibration artefacts at a deeper level is becoming crucial with the advent of modern radio telescopes like the Jansky Very Large Array (JVLA) (Chandler & Butler 2014), the Low Frequency Array (LOFAR) (van Haarlem et al. 2013), the Karoo Array Telescope (MeerKAT) (Jonas 2009) and the upcoming Square Kilometre Array (SKA) (Dewdney et al. 2009). These instruments do not only provide unprecedented sensitivity, but also have high data rates. The increase in sensitivity implies that faint artefacts will become apparent, while the high data rates necessitate automatic calibration. The latter may potentially lead to im-

* E-mail: wijnholds@astron.nl

ages that can contain even more unwanted artefacts. Moreover, modern observational techniques like intensity mapping (Furlanetto et al. 2006; Chang et al. 2010; Bernardi et al. 2009) and stacking (Delhaize et al. 2013) purport to detect signals at low signal-to-noise ratios (SNRs). Calibration artefacts are also very faint and are only noticeable at low SNR. It is therefore imperative for us to gain a deeper understanding of calibration artefacts if these new telescopes and techniques are to reach their scientific goals. The implications of ignoring calibration artefacts in future deep imaging are discussed in Grobler et al. (2014), herein referred to as “Paper I”, Smirnov & de Bruyn (2011), Zwart et al. (2014) and Asad et al. (2015).

In Paper I we gained some understanding as to why calibration artefacts form. However, the mechanism by which artefacts are formed is still poorly understood. For example, we only studied calibration artefacts produced by a regular east-west array in Paper I. Moreover, we only analyzed a simple two-source case with a bright source in the center of the field-of-view in the calibration model and a dimmer unmodeled source off-center. However, this simple setup has already led to some interesting results and additional questions.

In this paper, we extend Paper I in three ways. Firstly, we answer some of the open questions of Paper I. Secondly, we extend our analysis from a regular east-west array layout to general (irregular) array layouts. Finally, we introduce a new approach to the analysis of the ghost phenomenon based on perturbation analysis. This allows us to define a taxonomy of ghost sources, thereby laying the foundations for future work on even more general scenarios with, for example, arbitrary source models. We illustrate the results of this extended analysis using the 7-dish Karoo Array Telescope (KAT-7), the MeerKAT precursor system in South-Africa.

To comprehend the open questions from Paper I, we first present the main results from our previous paper using the following example. In Figure 1, we have a simple regular three-element east-west interferometer. The three circles represent the three antennas and the lines connecting the antennas represent the baselines. The antennas are numbered from left to right. Since the array is regular, the baselines are all multiples of some *common quotient baseline*. The common quotient baseline is depicted in Figure 1 by a thick solid black line. Let’s call the numbers (in our example the numbers are equal to 2, 3 and 5) with which the common quotient baseline needs to be multiplied to obtain the different baseline lengths the *baseline scaling factors*. We will denote the baseline scaling factor associated with baseline pq with ϕ_{pq} . Moreover, let us assume that our true sky is equal to a 1-Jy point source at the field-center and a 0.2-Jy source at $\mathbf{s}_0 = (1^\circ, 0^\circ)$.

Calibration enables us to find the antenna gains that minimize the difference between the observed and predicted visibilities. If we calibrate our example interferometer with a calibration model that only includes the 1-Jy source, our gain solutions will be biased. This causes imaging artefacts. The artefacts associated with the gain solutions are shown in the top row of Fig. 2. Each baseline has its own unique artefact map. Artefact maps are images of the spurious emission (ghost sources) caused by the incorrect antenna gains. The images in the bottom row of Fig. 2 show the artefact

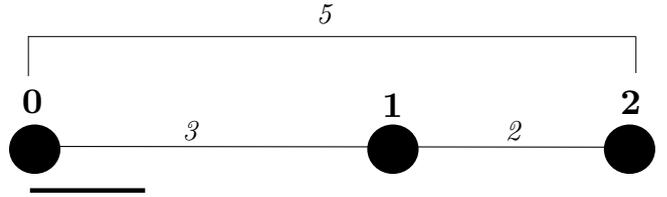


Figure 1. A simple regular east west interferometer. The baseline scaling factors of the interferometer are equal to 2,3 and 5. The thick solid line represents the common quotient baseline.

maps associated with the Hadamard inverse (the reciprocal) of the antenna gains.

We can make the following observations from these images:

- (i) Each baseline sees its own unique ghost pattern.
- (ii) The ghosts form along a straight line.
- (iii) The ghosts form at discrete positions which correspond to some rational fraction of the interval between the two sources.

In the previous paper we were able to explain why the ghost patterns have the characteristics listed above. These characteristics ultimately stem from the fact that our interferometer has a regular east-west configuration. However there were some characteristics that we could not explain:

- (i) Some ghosts are much brighter than others. We can account for this, to some degree, when the array is redundant. Due to the redundancy, some ghost positions overlap for multiple baselines, which implies that the ghosts at these positions will be brighter. However, redundancy does not explain why some ghosts are brighter in a non-redundant array. The interferometer in our example is not redundant, but some ghosts are still brighter than others.
- (ii) The brightest ghost positions are preserved under the Hadamard inverse. The images associated with the reciprocal of the antenna gains seem to be approximately equal to the “negative image” of the image associated with the antenna gains themselves.

We address these open issues in this paper.

The extension of our analysis to more general array layouts also allows us to answer the following questions:

- (i) Do ghosts form for antenna configurations other than regular east-west arrays? If they do, how do they arrange themselves?
- (ii) In the case of an east-west array, the main factors that determine the positions of the ghosts are the baseline scaling factors. Their flux is mainly determined by the flux in the unmodeled source (and, of course, the number of antennas in the array). In the case of a general layout, which parameters influence the ghost positions and fluxes?

(iii) Inspired by a recent LOFAR ghost spotting (Stewart 2014), we pose the following questions: do ghosts form in pairs? For each ghost that forms does there also form a symmetrically opposite ghost or an “anti-ghost”? This symmetrically opposite ghost is also known as the “friendly” ghost (Stewart 2014). This conjecture seems to hold for our example. In Fig. 2, the most striking ghosts with positive position vectors are marked with red crosses, while the most striking ghosts with negative position vectors are marked with

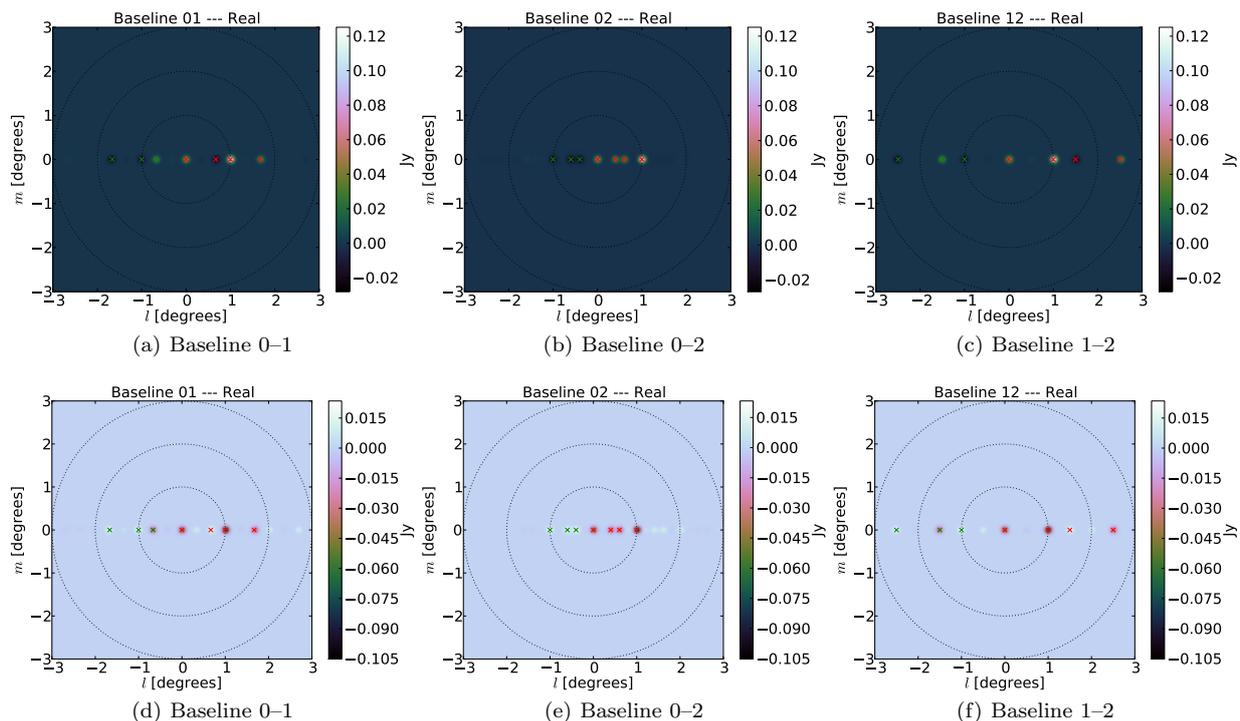


Figure 2. Artefact maps associated with the gain estimates (top row) and their reciprocals (bottom row) for the indicated baselines. Since these maps are associated with a single baseline, the patterns may be complex valued. Here, we only show the real parts, since for an east-west array the ghost pattern is completely real. Note that the positions of the ghosts do not change if we transform the gain estimates into calibration corrections by taking their reciprocal values.

green crosses. The red and green crosses form symmetrically opposite pairs.

(iv) Is there any difference between the ghosts that form (if they form) when we use a general irregular array layout and the ghosts that form when we use an east-west layout?

The first question was answered to some degree in Grobler et al. (014b). In this paper, we expand on those results and also discuss the remaining questions.

This paper is organized as follows. In the next section, we formulate the process of calibration in mathematical terms. This permits a rigorous analysis using perturbation theory, which we present in Sec. 3. In Sec. 4, we extend the linear transformation relating the actual baseline to the common quotient baseline (as presented in Paper I) to a more general set of linear transformations, that can be used to map the actual uv -tracks in a general array layout to a common reference uv -track. Once we have established the perturbation approach and the linear transformations as mathematical tools to analyze calibration artefacts, we provide an in-depth analysis of a two-source test case with one modeled and one unmodeled source. The results from this analysis will be illustrated by KAT-7-simulations and lead to a taxonomy for ghost sources.

2 CALIBRATION

Calibration can be formulated as a minimization problem:

$$\min_{\mathbf{G}} \left\| \mathcal{R} - \mathbf{G}\mathbf{M}\mathbf{G}^H \right\|, \quad (2.1)$$

where

(i) \mathcal{R} is the *observed visibility matrix*. Each entry, which we denote by r_{pq} , represents the visibility measured by the baseline formed by antennas p and q .

(ii) \mathcal{M} is the *model visibility matrix*. The entry m_{pq} of \mathcal{M} denotes a non-measured or a model visibility which was created with the calibration sky model and a uv -point on the uv -track associated with baseline pq .

(iii) $\mathbf{G} = \text{diag}(\mathbf{g})$ is the *antenna gain matrix*, where $\mathbf{g} = [g_1, g_2, \dots, g_N]^T$ denotes the *antenna gain vector*. The operator $\text{diag}(\cdot)$ forms a diagonal matrix from a vector by putting the elements of the vector on the main diagonal. The vector \mathbf{g} represents the instrumental response of the antennas, i.e. the complex antenna gains. These antenna gains are chosen in such a way that they minimize the difference between the observed and model visibilities. If your sky model is incomplete \mathbf{g} can also contain other non-physical systematics in addition to the complex antenna gains.

(iv) $\mathbf{G}\mathbf{M}\mathbf{G}^H$ is the *predicted visibility matrix*. This matrix contains the model visibilities after the antenna gains have been applied to them.

The superscript H denotes the Hermitian or conjugate transpose and $\|\cdot\|$ denotes the norm used. Most calibration algorithms use the Frobenius norm for matrices and the 2-norm or Euclidean norm for vectors, thus treating calibration as a least squares problem. In our (perturbation) analysis, we will therefore assume least squares optimization.

Eq. (2.1) can also be written as

$$\min_{\mathbf{g}} \left\| \mathcal{R} - \mathbf{g}\mathbf{g}^H \odot \mathcal{M} \right\| = \min_{\mathbf{g}} \left\| \mathcal{R} - \mathbf{g} \odot \mathcal{M} \right\|, \quad (2.2)$$

where $\mathbf{g} = \mathbf{g}\mathbf{g}^H = \check{\mathbf{G}}\mathbf{G}^H$ is the *visibility gain matrix*. When

$\mathcal{M} = \check{\mathbf{I}} = \mathbf{1}\mathbf{1}^H$ (i.e. the calibration sky model is a 1-Jy source at the field-center) then the predicted visibility matrix simplifies and becomes \mathcal{G} . The entries of \mathcal{G} are denoted by g_{pq} , which at any given time-step represents the visibility gain value associated with antennas p and q . The Hadamard or element-wise product of two matrices or vectors is denoted by \odot . We use $\mathbf{1}$ and $\check{\mathbf{I}}$ to denote an all-one column vector and an all-one matrix respectively.

Before imaging, we have to correct our observed visibilities by removing the effect that the antenna gains had on the observed visibilities. This can be accomplished by using

$$\mathcal{R}^{(c)} = \mathbf{G}^{-1} \mathcal{R} \mathbf{G}^{-H} = \mathcal{G}^{\odot -1} \odot \mathcal{R}, \quad (2.3)$$

where

- (i) $\mathcal{R}^{(c)}$ is the *corrected visibility matrix*.
- (ii) $\mathcal{G}^{\odot -1}$ denotes the *visibility calibration matrix*, which is computed by taking the Hadamard inverse of \mathcal{G} .

The superscript $^{-1}$ denotes matrix inversion, while $^{-H}$ denotes the inverse of the Hermitian transpose.

The observed visibilities can be modeled as

$$\mathcal{R} = \mathbf{G}_t (\mathcal{R}_k + \mathcal{R}_u) \mathbf{G}_t^H, \quad (2.4)$$

where \mathcal{R}_k and \mathcal{R}_u represent the visibility response to the known sources and the unknown sources respectively and \mathbf{G}_t contains the true gain values. The known sources are included in the model visibility matrix \mathcal{M} , while the unknown sources cannot be included in the model visibility matrix. We can now reformulate Eq. (2.1) as

$$\min_{\mathcal{G}} \left\| \mathbf{G}_t (\mathcal{R}_k + \mathcal{R}_u) \mathbf{G}_t^H - \mathcal{G} \mathcal{R}_k \mathcal{G}^H \right\|, \quad (2.5)$$

which is equivalent to

$$\min_{\mathcal{G}} \left\| \mathcal{G}_t \odot (\mathcal{R}_k + \mathcal{R}_u) - \mathcal{G} \odot \mathcal{R}_k \right\|. \quad (2.6)$$

If the interferometer is perfect (it induces no error) then we expect the true visibility gain matrix \mathcal{G}_t to be equal to $\check{\mathbf{I}}$. If we calibrate with an incomplete sky model then the entries of the visibility gain matrix \mathcal{G} will only be close to one. Since we assume that our sky model is reasonably accurate, we have $\mathcal{G} \approx \mathcal{G}_t$, which implies that the calibration artefacts are necessarily faint. In our analysis, we therefore have to eliminate the bright emission such that the faint artefacts become visible, as we did to obtain Fig. 2. This process is known as *distillation* (Grobler et al. 2014).

At this point, we would like to define the exact scope of this paper. Although we did not explicitly mention it, Eq. (2.1) depends on time and frequency. In our current analysis, we assume that a calibration solution is sought for each individual channel and timeslot, i.e., we do not group channels or timeslots together by assuming that the antenna gains remain constant over such solution intervals. The analysis we make in this paper is valid for snapshot and synthesis observations alike, if the per-channel and per-timeslot assumption is strictly adhered to. Interestingly, Nunhokee (2015) found that for direction dependent calibration using larger calibration solution intervals can significantly reduce the amount of suppression unmodeled sources experience. In summary, the size of your solution interval can greatly affect an array's ghost response and therefore deserves further study.

3 PERTURBATION ANALYSIS

It is possible to obtain approximate analytic expressions for \mathcal{G} and $\mathcal{G}^{\odot -1}$ if we employ perturbation analysis. We can model the estimated gain values with a deviation $\Delta \mathbf{g}$ from the true gain values as $\mathbf{g} = \mathbf{g}_t + \Delta \mathbf{g}$. We can then solve for $\Delta \mathbf{g}$ for the purpose of our analysis, since we are interested in the bias caused by our incomplete sky model $\mathcal{M} = \mathcal{R}_k$. The visibility gain matrix is now described by

$$\mathcal{G} = (\mathbf{g}_t + \Delta \mathbf{g}) (\mathbf{g}_t + \Delta \mathbf{g})^H. \quad (3.1)$$

This makes it possible to rewrite Eq. (2.6) to

$$\min_{\Delta \mathbf{g}} \left\| \mathbf{g}_t \mathbf{g}_t^H \odot (\mathcal{R}_k + \mathcal{R}_u) - (\mathbf{g}_t + \Delta \mathbf{g}) (\mathbf{g}_t + \Delta \mathbf{g})^H \odot \mathcal{R}_k \right\|. \quad (3.2)$$

After a little reordering of terms, we can also write this as

$$\min_{\Delta \mathbf{g}} \left\| \mathbf{g}_t \mathbf{g}_t^H \odot \mathcal{R}_u - \left(\mathbf{g}_t \Delta \mathbf{g}^H + \Delta \mathbf{g} \mathbf{g}_t^H + \Delta \mathbf{g} \Delta \mathbf{g}^H \right) \odot \mathcal{R}_k \right\|. \quad (3.3)$$

Since our model visibilities are (hopefully) based on a reasonably accurate sky model, the errors $\Delta \mathbf{g}$ will usually be small compared to \mathbf{g}_t . We will therefore assume that $\Delta \mathbf{g} \Delta \mathbf{g}^H$ is negligible compared to the other terms. This simplifies Eq. (3.3) to

$$\min_{\Delta \mathbf{g}} \left\| \mathbf{g}_t \mathbf{g}_t^H \odot \mathcal{R}_u - \left(\mathbf{g}_t \Delta \mathbf{g}^H + \Delta \mathbf{g} \mathbf{g}_t^H \right) \odot \mathcal{R}_k \right\|. \quad (3.4)$$

This equation is linear in the parameter vector $[\Delta \mathbf{g}^T, \overline{\Delta \mathbf{g}}^T]^T$ and can be solved in closed form as shown in Appendix A to obtain

$$\begin{bmatrix} \Delta \mathbf{g} \\ \overline{\Delta \mathbf{g}} \end{bmatrix} \approx \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{E} \mathbf{g}_t \\ \mathbf{F} \mathbf{g}_t \end{bmatrix}, \quad (3.5)$$

where

$$\begin{aligned} \mathbf{A} &= \overline{\mathcal{R}_k} \mathbf{G}_t \mathbf{G}_t^H \overline{\mathcal{R}_k} \odot \mathbf{I} \\ \mathbf{B} &= \overline{\mathcal{R}_k} \mathbf{G}_t \odot \mathbf{G}_t \mathcal{R}_k \\ \mathbf{C} &= \mathbf{G}_t^H \overline{\mathcal{R}_k} \odot \mathcal{R}_k \mathbf{G}_t^H = \overline{\mathbf{B}} \\ \mathbf{D} &= \mathbf{I} \odot \mathcal{R}_k \mathbf{G}_t^H \mathbf{G}_t \mathcal{R}_k = \overline{\mathbf{A}} \\ \mathbf{E} &= \overline{\mathcal{R}_k} \mathbf{G}_t \mathbf{G}_t^H \overline{\mathcal{R}_u} \odot \mathbf{I} \\ \mathbf{F} &= \mathbf{G}_t^H \overline{\mathcal{R}_u} \odot \mathcal{R}_k \mathbf{G}_t^H. \end{aligned} \quad (3.6)$$

4 ARRAY GEOMETRY

In the first paper, we exploited the fact that all baselines of a regular east-west array are scaled versions of a common quotient baseline. After publication of Paper I, we realized that by mapping each baseline on the common quotient baseline, or reference baseline, by a linear transformation (scaling), the uv -track produced by a baseline can be mapped on a reference track traced out by the reference baseline (Grobler et al. 014b). In this section, we extend the idea of mapping all baselines on a common reference baseline by linear transformations to general irregular array layouts. The goal is to map each uv -track on a circular reference track centered at the origin of the uv -plane tracing out (a fraction of) the unit circle. We will see that this requires three linear transformations, namely *scaling, translation and rotation*. The composition of three separate linear transformations is

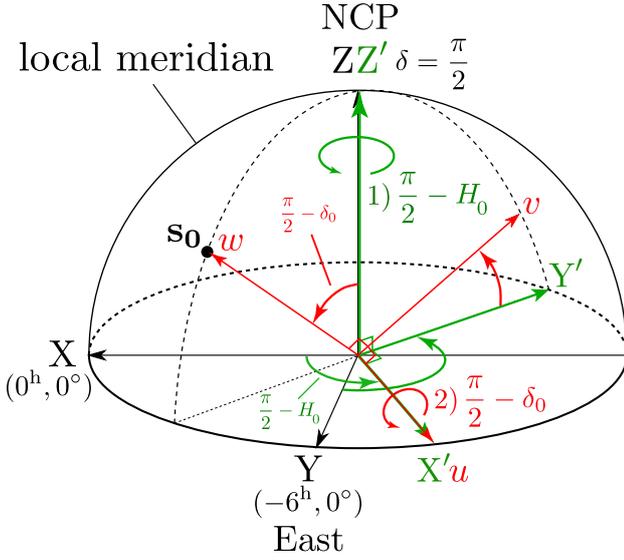


Figure 3. The relationship between the XYZ and uvw axes. The phase reference (field-center of the observation) is denoted by \mathbf{s}_c . The equatorial coordinates of \mathbf{s}_c are (H_0, δ_0) . Note that the hour angle of \mathbf{s}_c is negative, since the field center is in the eastern half of the hemisphere. Two rotations are required to convert between the XYZ and uvw axes. First, we rotate around the Z -axis by the angle $90^\circ - H_0$ and then about the X' -axis by the angle $90^\circ - \delta_0$.

in itself just a linear transformation. It is well worth mentioning that the regular east-west case is nothing more than a special case of the more generic array layout. Before we can introduce the linear transformations, we need to clearly define our coordinate systems, which we will do in the next section.

4.1 uv -Coverage

It is common practice to specify the coordinate difference vectors, i.e. baseline vectors, of the antenna positions of an array in the Cartesian terrestrial coordinate system XYZ . We can define the orientation of the XYZ coordinate system by using equatorial coordinates (H, δ) , with H being the hour angle and δ the declination. In the coordinate system XYZ , the X axis points towards $(0^{\text{h}}, 0^\circ)$ (the point where the vernal equinox crosses the local meridian), the Y axis towards $(-6^{\text{h}}, 0^\circ)$ (due east) and the Z axis towards the North Celestial Pole (NCP). We generate uv -coverage by converting the baseline vectors in the XYZ coordinate system to the uvw coordinate system as the hour angle of the center of the observational field changes. The uvw coordinate system is fixed to the center of the observation field. The relationship between these two coordinate systems is depicted in Fig. 3.

For the remainder of the paper assume that λ and ν respectively denote the wavelength and frequency at which we are observing. Let

$$\mathbf{b}_{pq}^{xyz} = \left(\frac{\Delta x_{pq}}{\lambda}, \frac{\Delta y_{pq}}{\lambda}, \frac{\Delta z_{pq}}{\lambda} \right)^T, \quad (4.1)$$

where Δx_{pq} , Δy_{pq} and Δz_{pq} are the coordinate differences between antennas p and q in the XYZ coordinate system.

The vector \mathbf{b}_{pq}^{xyz} in the XYZ coordinate system can be converted to a vector $\mathbf{b}_{pq}^{uvw} = (u, v, w)^T$ in the uvw coordinate system by using (Thompson et al. 2001; Cornwell & Fomalont 1999)

$$\mathbf{b}_{pq}^{uvw} = \mathbf{A} \mathbf{b}_{pq}^{xyz}, \quad (4.2)$$

where

$$\mathbf{A} = \begin{bmatrix} \sin H_0 & \cos H_0 & 0 \\ -\sin \delta_0 \cos H_0 & \sin \delta_0 \sin H_0 & \cos \delta_0 \\ \cos \delta_0 \cos H_0 & -\cos \delta_0 \sin H_0 & \sin \delta_0 \end{bmatrix}. \quad (4.3)$$

In Eq. (4.3), H_0 denotes the hour angle and δ_0 denotes the declination of the field-center (\mathbf{s}_c in Fig. 3). This is the product of two rotation matrices; the first matrix representing a rotation around the Z -axis by the angle $90^\circ - H_0$ and the second matrix describing a rotation around the X' -axis by the angle $90^\circ - \delta_0$. The vector \mathbf{b}_{pq}^{xyz} does not change as the Earth rotates, since it resides in a terrestrial reference frame. The vector \mathbf{b}_{pq}^{uvw} however changes continuously, since the hour angle of the field-center changes due to the rotation of the Earth, i.e., we may write $\mathbf{b}_{pq}^{uvw}(t)$, since $H_0(t)$. If we disregard w we obtain the baseline vector $\mathbf{b}_{pq} = (u, v)^T$, i.e.

$$\mathbf{b}_{pq} = \mathbf{A}[\alpha, \beta] \mathbf{b}_{pq}^{xyz}, \quad (4.4)$$

where $\alpha = \{1, 2\}$ and $\beta = \{1, 2, 3\}$. The notation $\mathbf{A}[\alpha, \beta]$ denotes the submatrix of \mathbf{A} that is obtained by only keeping the rows and columns respectively listed in α and β . As the Earth rotates the baseline vector \mathbf{b}_{pq} traces out a uv -track $\mathbf{b}_{pq}(t)$ in the uv -plane. The traced out uv -track has an elliptical locus whose equation is given by (Thompson et al. 2001; Cornwell & Fomalont 1999)

$$u^2 + \left(\frac{v - \frac{\Delta z_{pq} \cos \delta_0}{\lambda}}{\sin \delta_0} \right)^2 = \frac{\Delta x_{pq}^2 + \Delta y_{pq}^2}{\lambda^2}. \quad (4.5)$$

The baseline vector \mathbf{b}_{qp} is equal to $-\mathbf{b}_{pq}$.

4.2 Linear transformations

In this section it will become apparent that all uv -tracks of an interferometer can be derived from a single *imaginary reference track* $\mathbf{b}_0(t)$ via linear transformations. The most generic imaginary reference track $\mathbf{b}_0(t)$ has an elliptical locus with equation

$$u^2 + \left(\frac{v}{\sin \hat{\delta}_0} \right)^2 = \phi_0^2, \quad (4.6)$$

where ϕ_0 is equal to the semi-major axis of the elliptical locus and $\hat{\delta}_0$ denotes the declination of an imaginary field-center.

In this section we will discuss the following linear transformations:

(i) $\bar{X}_{pq}(\mathbf{b}_0)$: This linear transformation only consists of a scaling. It maps the elliptical imaginary reference track associated with a regular east-west array to \mathbf{b}_{pq} .

(ii) $X_{pq}(\mathbf{b}_0)$: This linear transformation consists of a scaling and a rotation. It will be shown that the translation transformation can influence the flux of the ghosts, but not their positions. We therefore use this transformation when we analyze the positions of the ghosts.

(iii) $\hat{X}_{pq}(\mathbf{b}_0)$: This linear transformation consists of a scaling, a translation and a rotation. This transformation maps the circular imaginary reference track associated with a general array layout to \mathbf{b}_{pq} .

(iv) $\tilde{X}_{pq}(\mathbf{b}_0)$: This linear transformation consists of a scaling and a translation. The rotation angle needed to define $\tilde{X}_{pq}(\mathbf{b}_0)$ is derived using this transformation.

We explicitly define the above mentioned linear transformations in Appendix B, which is a very cumbersome procedure. This derivation is not reused in the rest of the paper. However, the reader should inspect Fig. 4 before continuing with Sec. 5. Fig. 4 visually defines the symbols from Sec. B1 and Sec. B2 that we often use in the rest of the paper.

5 PRELIMINARIES

In this section, we lay the groundwork for predicting where the brightest ghosts will form for a general array layout. We start by introducing the simple two-source model we will use for our analysis. In Sec. 5.2, we introduce the concept of *extrapolation* (Grobler et al. 014b). This procedure enables us to produce clean artefact maps that contain all ghosts, i.e., bright and faint ghosts. It gives us an exact solution, albeit an empirical one.

We then discuss *perturbation* (Sec. 5.3) with which we can, in conjunction with extrapolation, derive analytic expressions that can predict both the position and a reasonable estimate of the amplitude of the brightest ghosts. In Sec. 6, we combine the two approaches to analyze the ghost pattern resulting from the two-source model for an arbitrary array layout. The limitation of this conglomerated analytic technique is that it can only make predictions about the brightest ghosts. However, these brightest ghosts are the most important ones, since they produce the dominant artefacts in the image.

5.1 Two-source scenario

In our analysis, we assume that we are observing with a narrow bandwidth (a monochromatic observation) and that our interferometer adds no error to our visibilities, i.e., $\mathcal{G}_t = \check{\mathbf{I}}$. We further assume that we are only working with a narrow field-of-view which implies that we may assume that the w -term is negligible. We also do not consider the effect of noise in our analysis (as we mentioned in Grobler et al. (2014), earlier simulations by Smirnov (2010) showed that including noise did not alter the positions at which the ghosts arise).

Let us assume that our true sky consists of two point sources, one in the field center and another off-center, with fluxes A_1 and A_2 respectively. The positional vector of the off-center source is denoted by $\mathbf{s}_0 = (l_0, m_0)$ (a direction-cosine vector). Furthermore, we assume that only the center source is included in our calibration model. We may assume without any loss of generality that $A_1 = 1$ and that the source in the calibration model is located at the field center, since Eq. (2.1) is invariant to amplitude scaling and positional shifts. We also assume that our calibration model is reasonably accurate, i.e., $A_2 \ll A_1$. If we had a perfect interferometer that was observing this two-source sky and that could sample the entire uv -plane, our observed visibilities

would be described by

$$r(\mathbf{u}) = A_1 + A_2 e^{-2\pi i \mathbf{u}^T \cdot \mathbf{s}_0}, \quad (5.1)$$

where $\mathbf{u} = (u, v)^T$. The real part of $r(\mathbf{u})$ is a periodic sinusoidal fringe pattern, while the imaginary part of $r(\mathbf{u})$ is a periodic sinusoidal fringe pattern. Both the real and imaginary parts of $r(\mathbf{u})$ are periodic in the u and v direction with periods respectively equal to $\frac{1}{l_0}$ and $\frac{1}{m_0}$.

5.2 Extrapolated visibilities

In this section we discuss a technique that enables us to create “clean” *artefact maps*. An artefact map is an image of the ghost sources caused by the calibration systematics (we create artefact maps by imaging distilled visibilities). This can be either a clean or dirty map depending on whether the sampling function of the interferometer is taken into account. We image the distilled visibilities, since we want to eliminate everything else from the image except for the calibration bias. Distillation is discussed in more detail in Sec. 6.2.

The actual visibilities are sampled along uv -tracks which implies that the entries of the best fitting visibilities \mathcal{G} also lie on the uv -tracks of an interferometer. As the ghosts are quite faint, directly imaging the distilled best-fitted visibilities (i.e. $\mathcal{G} - \check{\mathbf{I}}$) to produce an artefact map generally turns out to be futile, because the faint ghosts are masked by the sidelobes of the brighter ghosts. Deconvolving this dirty artefact map turns out to be very hard, since the ghost spread function of the ghosts is not equal to the point spread function of the interferometer (see Sec. 6.4). This problem can be circumvented by reformulating Eq. (2.1) as a function of an imaginary reference track. This allows us to extrapolate each entry of the matrices involved to arbitrary points in the uv -plane. This results in a continuous uv -plane coverage that allows us to make clean artefact maps by using a continuous Fourier transform.

This procedure, however, can only be used to create a clean artefact map for each baseline, i.e., it provides us with a virtual uv -plane per baseline that is consistent with the measurements on the specific uv -track associated with that baseline. By resampling these virtual best-fitted uv -planes we can create physically meaningful visibilities (step 3 below). Imaging the per-baseline resampled best-fitted uv -planes (after performing distillation) allows us to create per-baseline clean artefact maps. To obtain the conglomerated artefact map of the interferometer we can just average all per-baseline artefact maps (assuming natural weighting). A more detailed explanation of extrapolation can be found in paper I.

More formally, the extrapolation process works as follows:

Step 1: Reformulate visibilities as functionals. Every visibility matrix $\mathcal{V}(\mathbf{b}_0)$ can be written as a function of the imaginary reference track \mathbf{b}_0° . We use \mathcal{V} as a proxy matrix, which can refer to an observed visibility, model visibility, predicted visibility, visibility gain, visibility calibration or a corrected visibility matrix (\mathcal{V} can also double for any of the distilled visibility matrices). Its entries are denoted by y_{pq} . Let us take the observed visibility matrix \mathcal{R} as an example. Its

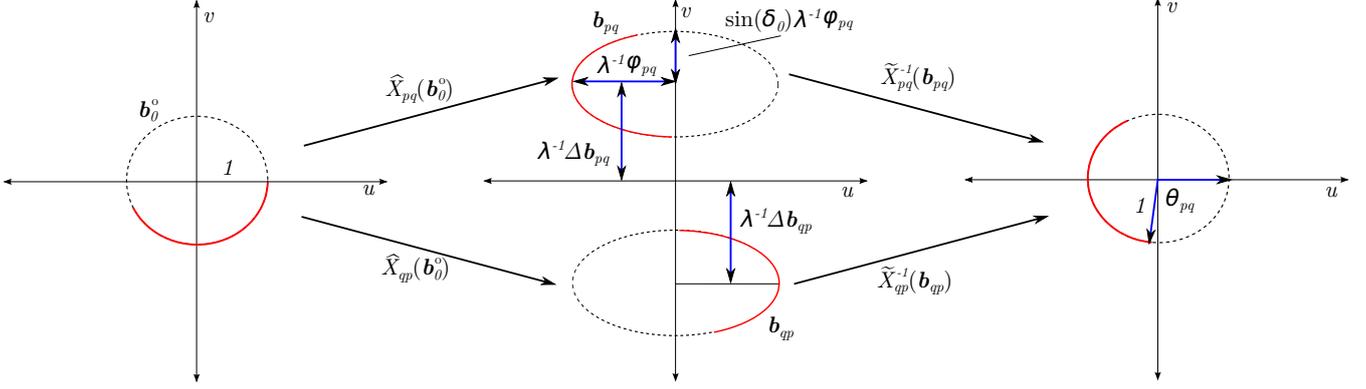


Figure 4. The linear transformations $\hat{X}_{pq}(\mathbf{b}_0^\circ)$ and $\hat{X}_{pq}^{-1}(\mathbf{b}_0^\circ)$. With the linear transformation $\hat{X}_{pq}(\mathbf{b}_0^\circ)$ we can create the uv -track $\mathbf{b}_{pq}(t)$ from the imaginary reference track $\mathbf{b}_0^\circ(t)$. We use $\hat{X}_{pq}^{-1}(\mathbf{b}_0^\circ)$ to define θ_{pq} . In this case θ_{pq} will be negative since it is the anti-clockwise rotation angle between $(0, 1)$ and \mathbf{y}_{pq} .

entries are equal to

$$r_{pq}(\mathbf{b}_0^\circ) = r(\hat{X}_{pq}(\mathbf{b}_0^\circ)) = r(\mathbf{b}_{pq}). \quad (5.2)$$

Step 2: Derive the intermediate extrapolated visibility matrix $\mathcal{Y}(\mathbf{b}) : \mathbf{R}^2 \rightarrow \mathbf{C}^{N \times N}$. By replacing the imaginary reference baseline \mathbf{b}_0° with a free parameter $\mathbf{b} = (u, v)^T$, the elements of \mathcal{Y} become functions that are defined over the entire visibility plane, i.e., the extrapolated visibility plane. Whenever \mathbf{b} is equal to \mathbf{b}_0° then $y_{pq}(\mathbf{b})$ represents a true visibility value. For all other values of \mathbf{b} , $y_{pq}(\mathbf{b})$ represents an extrapolated or imaginary visibility, i.e., a nonphysical visibility with a purely mathematical definition which is used to fill in the uv -plane.

The visibility gain matrix plays a crucial role in the remainder of this paper. Let us therefore take a closer look at $g_{pq}(\mathbf{b})$. Since $g_{pq}(\mathbf{b})m_{pq}(\mathbf{b})$ is the best possible fit of the intermediate extrapolated visibility plane $r_{pq}(\mathbf{b})$, given Eq. (2.1), $g_{pq}(\mathbf{b})$ represents an intermediate extrapolated visibility gain plane.

Step 3: Find the extrapolated visibility matrix $\mathcal{Y}(\hat{X}_{pq}^{-1}(\mathbf{b}))$. Once we have derived the intermediate extrapolated visibility matrix, we need to reverse the effect of \hat{X}_{pq} on $y_{pq}(\mathbf{b})$ to make our results physically meaningful. We therefore apply the inverse transformation \hat{X}_{pq}^{-1} to each entry of $\mathcal{Y}(\mathbf{b})$. To see what this does, let us investigate what happens if we apply $\hat{X}_{pq}^{-1}(\mathbf{b})$ to the entries of $\mathcal{R}(\mathbf{b})$:

$$\left[\mathcal{R}(\hat{X}_{pq}^{-1}(\mathbf{b})) \right]_{pq} = r_{pq}(\hat{X}_{pq}^{-1}(\mathbf{b})) = r(\mathbf{b}) \quad (5.3)$$

$$\left[\mathcal{R}(\hat{X}_{pq}^{-1}(\mathbf{b})) \right]_{rs} = r_{rs}(\hat{X}_{pq}^{-1}(\mathbf{b})) = r(\hat{X}_{rs} \circ \hat{X}_{pq}^{-1}(\mathbf{b})), \quad (5.4)$$

where “ \circ ” represents functional composition. For the entry $r_{pq}(\mathbf{b})$ the effect of \hat{X}_{pq} is completely eliminated. However, not all of the unwanted transformations are removed from the matrix $\mathcal{R}(\hat{X}_{pq}^{-1}(\mathbf{b}))$, since the transformations from the other baselines are still present in the other entries of $\mathcal{R}(\hat{X}_{pq}^{-1}(\mathbf{b}))$. Note that the baseline is chosen arbitrarily. We have successfully recast $r_{pq}(\mathbf{b})$ so that it becomes physically meaningful (it is equal to the true sky).

Since the visibility gain matrix plays an important role, we need to understand how the above result affects $\mathcal{G}(\mathbf{b})$. Since, \hat{X}_{pq} is eliminated from $r_{pq}(\hat{X}_{pq}^{-1}(\mathbf{b}))$, it implies that $g_{pq}(\hat{X}_{pq}^{-1}(\mathbf{b}))$ is the extrapolated visibility gain plane that is

associated with baseline pq , i.e., $g_{pq}(\hat{X}_{pq}^{-1}(\mathbf{b}))m_{pq}(\hat{X}_{pq}^{-1}(\mathbf{b}))$ is the best possible visibility plane one can construct for baseline pq that approximates $r(\mathbf{u})$, given the restrictions imposed by Eq. (2.1). If we therefore take the Fourier inverse of $g_{pq}(\hat{X}_{pq}^{-1}(\mathbf{b})) - 1$ (we subtract one to eliminate everything except the calibration systematics) we will obtain a “clean” artefact map, that is it will be completely untarnished by uv -sampling. From the sky model associated with this artefact map (which is the sky brightness distribution of the calibration systematics) we will be able to reproduce $g_{pq}(\mathbf{b}_{pq})$.

5.3 Perturbation

We derived an approximation of \mathcal{G} in Section 3 by using perturbation theory. In that derivation we did not assume anything about our modeled and unmodeled sky. In this section we will take a closer look at the two-source scenario discussed in Section 5.1, so we take

$$\mathcal{R}_k = \mathbf{1}\mathbf{1}^H \quad (5.5)$$

$$\mathcal{R}_u = A_2 \mathbf{a} \mathbf{a}^H, \quad (5.6)$$

where $\mathbf{a} = [e^{-2\pi i \mathbf{x}_1^T \cdot \mathbf{s}_0}, \dots, e^{-2\pi i \mathbf{x}_N^T \cdot \mathbf{s}_0}]^T$ and \mathbf{x}_n is the (x, y, z) -position of the n th antenna in units of λ .

Substituting Eq. (5.5) and Eq. (5.6) into Eq. (3.6) allows us to calculate

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{N}\mathbf{I} & \mathbf{1}\mathbf{1}^H \\ \mathbf{1}\mathbf{1}^H & \mathbf{N}\mathbf{I} \end{bmatrix}, \quad (5.7)$$

$$\mathbf{E} \mathbf{g}_t = A_2 (\mathbf{a}^H \mathbf{1}) \mathbf{a} = A_2 \begin{bmatrix} \sum_s e^{-2\pi i \mathbf{b}_{1s}^T \cdot \mathbf{s}_0} \\ \sum_s e^{-2\pi i \mathbf{b}_{2s}^T \cdot \mathbf{s}_0} \\ \vdots \\ \sum_s e^{-2\pi i \mathbf{b}_{Ns}^T \cdot \mathbf{s}_0} \end{bmatrix}, \quad (5.8)$$

and

$$\mathbf{F} \mathbf{g}_t = A_2 (\bar{\mathbf{a}}^H \mathbf{1}) \bar{\mathbf{a}} = A_2 \begin{bmatrix} \sum_r e^{-2\pi i \mathbf{b}_{r1}^T \cdot \mathbf{s}_0} \\ \sum_r e^{-2\pi i \mathbf{b}_{r2}^T \cdot \mathbf{s}_0} \\ \vdots \\ \sum_r e^{-2\pi i \mathbf{b}_{rN}^T \cdot \mathbf{s}_0} \end{bmatrix}. \quad (5.9)$$

Moreover,

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^\dagger = \frac{1}{N} \mathbf{I} - \frac{1}{4N^2} \begin{bmatrix} \mathbf{311}^H & -\mathbf{11}^H \\ -\mathbf{11}^H & \mathbf{311}^H \end{bmatrix}, \quad (5.10)$$

where $()^\dagger$ denotes the pseudo-inverse. We have replaced the inverse in Eq. (3.5) by the pseudo-inverse, since the matrix to be inverted may have a rank deficiency in the two-source case. Replacement of the inverse by the pseudo-inverse mimics the behavior of a least squares solver in such scenarios. A derivation of this result can be found in Appendix C. Substituting Eqs. (5.8) through (5.10) into Eq. (3.5) we find

$$\Delta \mathbf{g} \approx \frac{A_2}{N} (\mathbf{a}^H \mathbf{1}) \mathbf{a} - \frac{A_2}{2N^2} (\mathbf{1}^H \mathbf{a}) (\mathbf{a}^H \mathbf{1}) \mathbf{1}. \quad (5.11)$$

Substitution of this result in Eq. (3.1) and assuming small deviations from the true gain values, we find

$$\begin{aligned} \mathcal{G} &= (\mathbf{1} + \Delta \mathbf{g}) (\mathbf{1} + \Delta \mathbf{g})^H \\ &\approx \mathbf{11}^H + \Delta \mathbf{g} \mathbf{1}^H + \mathbf{1} \Delta \mathbf{g}^H \\ &= \mathbf{11}^H + \frac{A_2}{N} (\mathbf{a}^H \mathbf{1}) \mathbf{a} \mathbf{1}^H + \frac{A_2}{N} (\mathbf{1}^H \mathbf{a}) \mathbf{1} \mathbf{a}^H \\ &\quad - \frac{A_2}{N^2} |\mathbf{1}^H \mathbf{a}|^2 \mathbf{11}^H \end{aligned} \quad (5.12)$$

In Eq. (5.12) (and in our analysis in Sec. 6), we recognize a number of products of the form $\mathbf{a}_1^H \mathbf{a}_2$. These products can be interpreted as follows: if a tied array beam is formed using weights given by $\bar{\mathbf{a}}_1$ while a signal is received by the array with array response described by \mathbf{a}_2 , the inner product $\mathbf{a}_1^H \mathbf{a}_2$ describes the voltage domain output signal of the beamformer. The inner product $\mathbf{1}^H \mathbf{a}$ therefore describes the complex valued sidelobe level of the tied array voltage beam formed towards the modeled calibration source in the direction of the unmodeled source. Hence, Eq. (5.12) shows that if the unmodeled source happens to be in the null of the tied array beam of the modeled source, i.e., if $\mathbf{1}^H \mathbf{a} = 0$, the gain solutions are not affected by the presence of the unmodeled source.

Since the visibility calibration matrix $\mathcal{G}^{\odot -1}$ is simply the Hadamard inverse of the visibility gain matrix, the assumption that the gain deviations are small, has the interesting consequence that

$$\begin{aligned} \mathcal{G}^{\odot -1} &\approx \left(\mathbf{11}^H + \Delta \mathbf{g} \mathbf{1}^H + \mathbf{1} \Delta \mathbf{g}^H \right)^{\odot -1} \\ &\approx \mathbf{11}^H - \Delta \mathbf{g} \mathbf{1}^H - \mathbf{1} \Delta \mathbf{g}^H, \end{aligned} \quad (5.13)$$

since we can use the approximation $(1+x)^{-1} \approx (1-x)$, which holds for $x \ll 1$. This explains why, to first order, the ghost structure of the visibility gain matrix (top row in Fig. 2) is identical to the ghost structure of its reciprocal, the visibility calibration matrix (bottom row in Fig. 2), except for the fact that positive ghosts have become negative ones and vice versa.

In our analysis of the ghost pattern of an irregular array in Sec. 6, we will use extrapolation to identify the positions of the ghost sources. It is therefore convenient to express the pq -th entry of the visibility gain matrix in the following form

$$[\mathcal{G}]_{pq} = g_{pq} = c_{pq,0} + \sum_{r \neq s} c_{pq,rs} e^{-2\pi i \mathbf{b}_{rs}^T \cdot \mathbf{s}_0}. \quad (5.14)$$

Substitution of the definition of \mathbf{a} below Eq. (5.6) in

Eq. (5.12) and selecting the pq -th entry, we find that

$$c_{pq,0} = 1 + \frac{A_2}{N}, \quad (5.15)$$

and

$$c_{pq,rs} = \begin{cases} \frac{2A_2}{N} - \frac{A_2}{N^2} & \text{if } rs \in \mathcal{I}_1 \\ \frac{A_2}{N} - \frac{A_2}{N^2} & \text{if } rs \in \mathcal{I}_2 \\ \frac{A_2}{N} - \frac{A_2}{N^2} & \text{if } rs \in \mathcal{I}_3, \\ -\frac{A_2}{N^2} & \text{if } rs \in \mathcal{I}_4 \\ 0 & \text{if } rs \in \mathcal{I}_5 \end{cases}, \quad (5.16)$$

where

$$\begin{aligned} \mathcal{I}_1 &= \{rs | (r=p) \wedge (s=q) \wedge (r \neq s)\} \\ \mathcal{I}_2 &= \{rs | (r=p) \wedge (s \neq q) \wedge (r \neq s)\} \\ \mathcal{I}_3 &= \{rs | (r \neq p) \wedge (s=q) \wedge (r \neq s)\} \\ \mathcal{I}_4 &= \{rs | (r \neq p) \wedge (s \neq q) \wedge (r \neq s)\} \\ \mathcal{I}_5 &= \{rs | (r=s)\}, \end{aligned} \quad (5.17)$$

where “ \wedge ” denotes logical and.

We will see in the next section that each term in Eq. (5.14) can be associated with a ghost and that the value of the corresponding coefficient $c_{pq,rs}$ determines the amplitude of that ghost. A look at the values of the coefficients $c_{pq,rs}$ quickly reveals that some coefficients are inversely proportional to the number of elements in the array, N , while others are inversely proportional to N^2 . The ghosts associated with the first set of coefficients, i.e., those with $\{r,s\} \in \mathcal{I}_P$ with $\mathcal{I}_P = \mathcal{I}_1 \cup \mathcal{I}_2 \cup \mathcal{I}_3$ where “ \cup ” denotes set union, will be referred to as *proto-ghosts*. The ghosts associated with the second set of coefficients, i.e., those with $\{r,s\} \in \mathcal{I}_D$ with $\mathcal{I}_D = \mathcal{I}_4$, will be referred to as *deutero-ghosts*.

6 THEORETICAL ANALYSIS OF GHOST PATTERNS

In this section we combine extrapolation (Sec. 5.2) and perturbation (Sec. 5.3) to create a framework with which we can derive estimates of the fluxes and the position vectors of the brightest ghosts. The perturbation approach gave us an analytic expression with which we could approximate g_{pq} (see Eq. (5.14)). Extrapolation will make Eq. (5.14) intelligible, translating the coefficients and the exponents in Eq. (5.14) into fluxes and position vectors. Extrapolation can provide us with an exact empirical solution but is used here, together with perturbation, to provide us with an approximate analytic solution. We will also investigate how the ghosts manifest during imaging, how to perform distillation and what exactly happens to the ghosts in the antenna gains when we create corrected visibilities.

6.1 Visibility gain matrix

Following our perturbation analysis, g_{pq} is approximately described by Eq. (5.14). This equation is hard to interpret in this form. However, using extrapolation, we can give a meaningful interpretation to the individual terms in Eq. (5.14). We can express g_{pq} as function of the reference baseline \mathbf{b}_0^0 ,

i.e.,

$$g_{pq}(\mathbf{b}_0^\circ) = c_{pq,0} + \sum_{r \neq s} c_{pq,rs} e^{-2\pi i [\hat{X}_{rs}(\mathbf{b}_0^\circ)]^T \cdot \mathbf{s}_0}. \quad (6.1)$$

We are now able to create intermediate extrapolated visibilities by replacing \mathbf{b}_0° with $\mathbf{b} = (u, v)^T$, which is a free parameter, to obtain

$$g_{pq}(\mathbf{b}) = c_{pq,0} + \sum_{r \neq s} c_{pq,rs} e^{-2\pi i [\hat{X}_{rs}(\mathbf{b})]^T \cdot \mathbf{s}_0}. \quad (6.2)$$

As explained in Section 5.2, Eq. (6.2) is physically meaningless if left in this form. We need to create meaningful extrapolated visibilities. This is accomplished by replacing \mathbf{b} in Eq. (6.2) with $\hat{X}_{pq}^{-1}(\mathbf{b})$. This gives

$$\begin{aligned} g_{pq}(\hat{X}_{pq}^{-1}(\mathbf{b})) &\approx c_{pq,0} + \sum_{r \neq s} c_{pq,rs} e^{-2\pi i [\hat{X}_{rs} \circ \hat{X}_{pq}^{-1}(\mathbf{b})]^T \cdot \mathbf{s}_0}, \\ &= d_{pq,0} + \sum_{r \neq s} d_{pq,rs} e^{-2\pi i [X_{rs} \circ X_{pq}^{-1}(\mathbf{b})]^T \cdot \mathbf{s}_0}, \\ &= d_{pq,0} + \sum_{r \neq s} d_{pq,rs} e^{-2\pi i \mathbf{b}^T \cdot \mathbf{s}_{pq}^{rs}} \end{aligned} \quad (6.3)$$

where $d_{pq,0} = c_{pq,0}$, $d_{pq,rs} = c_{pq,rs} \beta_{pq,rs}$ and

$$\beta_{pq,rs} = e^{2\pi i \left[\frac{\phi_{rs}}{\chi_{\phi pq}} \mathbf{D}(\delta_0) \mathbf{T}(\theta_{rs} - \theta_{pq}) \mathbf{D}^{-1}(\delta_0) \Delta \mathbf{b}_{pq} - \frac{\Delta \mathbf{b}_{rs}}{\chi} \right]^T \cdot \mathbf{s}_0} \quad (6.4)$$

and

$$\mathbf{s}_{pq}^{rs} = \frac{\phi_{rs}}{\phi_{pq}} \mathbf{s}_0 \mathbf{Z}, \quad (6.5)$$

with

$$\mathbf{Z} = \begin{bmatrix} \cos(\theta_{rs} - \theta_{pq}) & -\frac{\sin(\theta_{rs} - \theta_{pq})}{\sin(\delta_0)} \\ \sin(\delta_0) \sin(\theta_{rs} - \theta_{pq}) & \cos(\theta_{rs} - \theta_{pq}) \end{bmatrix}. \quad (6.6)$$

Since the inverse Fourier transform of a complex exponential is a delta function, each term of the sum in Eq. (6.3) can be interpreted as a calibration artefact that will manifest itself as a spurious point source or ghost when imaged. The amplitude of each ghost is approximately equal to $d_{pq,rs}$ and its position vector is equal to \mathbf{s}_{pq}^{rs} . Moreover, the amplitude of the ghosts are complex, i.e. $d_{pq,rs} \in \mathbf{C}$. These ghost sources are generated when we calibrate with an incomplete sky model. During the calibration procedure ‘‘leakage’’ occurs, the uncanceled transformations present in $\mathcal{R}(\hat{X}_{pq}^{-1}(\mathbf{b}))$ (which stem from the linear dependence that exists between the uv -tracks of an interferometer) leak into $g_{pq}(\hat{X}_{pq}^{-1}(\mathbf{b}))$. These uncanceled transformations can be roughly interpreted as spurious spatial frequencies and ultimately cause the ghosts during imaging.

Let us now formally define the set that contains all ghost positions for baseline pq as predicted by our perturbation analysis as

$$\mathcal{S}_{pq} = \bigcup_{rs} \{\mathbf{s}_{pq}^{rs}\} \cup \mathbf{0}. \quad (6.7)$$

In Eq. (6.7), $\mathbf{0}$ denotes $(0, 0)$. With the aid of Eq. (6.7) we can construct a set that contains all predicted ghost positions of an array

$$\mathcal{S} = \bigcup_{pq} \mathcal{S}_{pq}. \quad (6.8)$$

Note that \mathcal{S}_{pq} and \mathcal{S} contain no reoccurring elements¹. In the previous section, we have made a distinction between proto-ghosts and deuterio-ghosts. For baseline pq , we can therefore define the set

$$\mathcal{S}_{pq}^P = \bigcup_{rs \in \mathcal{I}_P} \{\mathbf{s}_{pq}^{rs}\} \cup \mathbf{0}. \quad (6.9)$$

of proto-ghost positions and the set

$$\mathcal{S}_{pq}^D = \bigcup_{rs \in \mathcal{I}_D} \{\mathbf{s}_{pq}^{rs}\}. \quad (6.10)$$

of deuterio-ghost positions. Based on Eq. (5.14), we obviously have $\mathcal{S}_{pq} = \mathcal{S}_{pq}^P \cup \mathcal{S}_{pq}^D$ and \mathcal{S}_{pq}^P and \mathcal{S}_{pq}^D are non-overlapping. For the array, we can define $\mathcal{S}^P = \bigcup_{pq} \mathcal{S}_{pq}^P$ and $\mathcal{S}^D = \bigcup_{pq} \mathcal{S}_{pq}^D$.

Eq. (6.8) allows us to reformulate Eq. (6.3) as

$$g_{pq}(\hat{X}_{pq}^{-1}(\mathbf{b})) \approx \sum_{\tilde{\mathbf{s}} \in \mathcal{S}} d_{pq,\tilde{\mathbf{s}}} e^{-2\pi i \mathbf{b}^T \cdot \tilde{\mathbf{s}}}, \quad (6.11)$$

where

$$d_{pq,\tilde{\mathbf{s}}} = \begin{cases} \sum_{rs \in \mathcal{I}_{pq,\tilde{\mathbf{s}}}} d_{pq,rs} & \text{if } \mathcal{I}_{pq,\tilde{\mathbf{s}}} \neq \emptyset \\ d_{pq,0} & \text{if } \tilde{\mathbf{s}} = \mathbf{0} \\ 0 & \text{otherwise} \end{cases}. \quad (6.12)$$

In Eq. (6.12), $\mathcal{I}_{pq,\tilde{\mathbf{s}}} = \{rs | (\mathbf{s}_{pq}^{rs} \in \mathcal{S}_{pq}) \wedge (\tilde{\mathbf{s}} = \mathbf{s}_{pq}^{rs}) \wedge (\tilde{\mathbf{s}} \neq \mathbf{0})\}$.

6.1.1 Regular east-west layout

In this section we will reformulate the results of Paper I using the general transformations introduced in Sec. 4 and, in doing so, we will be able relate the perturbation approach to the main result of Paper I. Recall that in Paper I we used the properties of $\mathcal{G}(\bar{X}_{pq}^{-1}(\mathbf{b}))$ to show that

$$g_{pq}(\bar{X}_{pq}^{-1}(\mathbf{b})) = \sum_{j=-\infty}^{\infty} h_{pq,j} e^{-2\pi i \frac{j}{\phi_{pq}} \mathbf{b}^T \cdot \mathbf{s}_0}, \quad (6.13)$$

where $h_{pq,j} \in \mathbf{R}$.

The coefficients are real, since the elements of $\mathcal{R}(\bar{X}_{pq}^{-1}(\mathbf{b}))$ are Hermitian functions, which would imply that the elements of $\mathcal{G}(\bar{X}_{pq}^{-1}(\mathbf{b}))$ are also Hermitian functions. If we take the Fourier inverse of Eq. (6.13), we end up with a string of spurious point sources that lie on a discrete line with lm -coordinates $\{\frac{j\mathbf{s}_0}{\phi_{pq}}\}$ and fluxes $\{h_{pq,j}\}$. The validity of Eq. (6.13) is supported by Fig. 2, which are the clean artefact maps of the three element example introduced in Sec. 1. The images in the top row of Fig. 2 respectively present $\mathcal{F}^{-1}\{g_{01}(\bar{X}_{01}^{-1}(\mathbf{b})) - 1\}$, $\mathcal{F}^{-1}\{g_{02}(\bar{X}_{02}^{-1}(\mathbf{b})) - 1\}$ and $\mathcal{F}^{-1}\{g_{12}(\bar{X}_{12}^{-1}(\mathbf{b})) - 1\}$.

If we use the perturbation approach in conjunction with extrapolation we can derive the regular east-west equivalent of Eq. (6.3) which is equal to

$$\begin{aligned} g_{pq}(\bar{X}_{pq}^{-1}(\mathbf{b})) &\approx c_{pq,0} + \sum_{r \neq s} c_{pq,rs} e^{-2\pi i [\bar{X}_{rs} \circ \bar{X}_{pq}^{-1}(\mathbf{b})]^T \cdot \mathbf{s}_0} \\ &= c_{pq,0} + \sum_{r \neq s} c_{pq,rs} e^{-2\pi i \frac{\phi_{rs}}{\phi_{pq}} \mathbf{b}^T \cdot \mathbf{s}_0} \\ &= c_{pq,0} + \sum_{r \neq s} c_{pq,rs} e^{-2\pi i \mathbf{b}^T \cdot \mathbf{s}_{pq}^{rs}}, \end{aligned} \quad (6.14)$$

¹ Recall that $\{1, 2, 3\} \cup \{1, 2, 4\} = \{1, 2, 3, 4\}$

where $\mathbf{s}_{pq}^{rs} = \frac{\phi_{rs}}{\phi_{pq}} \mathbf{s}_0$. In Eq. (6.14), $c_{pq,0} \approx h_{pq,0}$ and $c_{pq,rs} \approx h_{pq,\phi_{rs}}$. The difference between Eq. (6.3) and Eq. (6.14) is that they were derived by using different baseline reference tracks. We used the circular track \mathbf{b}_0^c to derive Eq. (6.3) and the elliptical track \mathbf{b}_0^e to derive Eq. (6.14). Eq. (6.14) is nothing more than a truncated version of Eq. (6.13), implying that the series in Eq. (6.13) is dominated by a few terms. Moreover, Eq. (6.14) indicates that the brightest ghosts are real ($c_{pq,rs} \in \mathbf{R}$) and located at $\{\frac{\phi_{rs}}{\phi_{pq}} \mathbf{s}_0\}$ which corresponds nicely with the results shown in Fig. 2. The regular east-west layout can therefore be treated as a special case in which we may use the more elaborate elliptical imaginary reference track \mathbf{b}_0^e when we perform extrapolation which leads to a much simpler ghost pattern.

Comparing Eq. (6.3), Eq. (6.13) and Eq. (6.14) enables us to highlight the following differences between the ghost patterns of an east-west and a general irregular array layout:

(i) For a general layout the ghosts no longer manifest in straight lines.

(ii) The ghosts associated with a baseline in a general layout may be complex (they have complex flux), in contrast to baselines in an east-west array, whose ghosts are real valued. Since the sky itself is completely real, the idea that the ghosts in the antenna gains can have complex flux therefore warrants further explanation. For an east-west array, the entries of $\mathcal{G}(\mathbf{b})$ are Hermitian functions, thereby ensuring that all ghosts will have real flux. For a general array layout, we find that the entries of $\mathcal{G}(\mathbf{b})$ are not Hermitian functions, which potentially leads to ghosts with complex flux. This difference stems from the fact that the uv -tracks are not homocentric for a general irregular array. In Sec. 6.3 we will see that, even though the function-valued entries of $\mathcal{G}(X_{pq}^{-1}(\mathbf{b}))$ are not Hermitian functions, the fact that $\mathcal{G}(\mathbf{b})$ is a Hermitian matrix (if evaluated at any $\mathbf{b} \in \mathbf{R}^2$) enforces a specific type of symmetry to exist between ghosts formed by baselines pq and qp . In Sec. 6.4 we show that this symmetry eventually leads to the formation of a real image.

(iii) In contrast with a regular east-west array layout we do not have an analytic expression to calculate the exact ghost pattern that is associated with a general irregular array layout. Using extrapolation in conjunction with perturbation allows us to derive a finite sum, with which we can approximate Eq. (6.13).

6.2 Distillation

As mentioned in Sec. 2 we expect the entries of \mathcal{G} to be close to unity, which implies that our ghosts will be relatively faint. We therefore need to distill \mathcal{G} and $\mathcal{G}^{\odot-1}$ to spot the ghosts, i.e., we need to subtract the real emission. Since we have $\mathcal{G}_t = \check{\mathbf{1}}$, distillation merely boils down to subtracting one from Eq. (6.11) to obtain the pq -th entry of the distilled visibility gain matrix, i.e.

$$[\mathcal{G}(\hat{X}_{pq}^{-1}(\mathbf{b})) - \check{\mathbf{1}}]_{pq} \approx \sum_{\tilde{\mathbf{s}} \in \mathcal{S}} \hat{d}_{pq,\tilde{\mathbf{s}}} e^{-2\pi i \mathbf{b}^T \cdot \tilde{\mathbf{s}}}, \quad (6.15)$$

where $\hat{d}_{pq,\tilde{\mathbf{s}}} = d_{pq,\tilde{\mathbf{s}}} - 1_{\{\tilde{\mathbf{s}}=\mathbf{0}\}}$ and

$$1_{\{\tilde{\mathbf{s}}=\mathbf{0}\}} = \begin{cases} 1 & \text{if } \tilde{\mathbf{s}} = \mathbf{0} \\ 0 & \text{otherwise} \end{cases}. \quad (6.16)$$

In analogy with Eqs. (5.12) and (5.13), it obviously follows from Eq. (6.3) that

$$\begin{aligned} g_{pq}^{-1}(\hat{X}_{pq}^{-1}(\mathbf{b})) &\approx d_{pq,0}^{\odot-1} + \sum_{r \neq s} d_{pq,rs}^{\odot-1} e^{-2\pi i \mathbf{b}^T \cdot \mathbf{s}_{pq}^{rs}} \\ &= \sum_{\tilde{\mathbf{s}} \in \mathcal{S}} d_{pq,\tilde{\mathbf{s}}}^{\odot-1} e^{-2\pi i \mathbf{b}^T \cdot \tilde{\mathbf{s}}}, \end{aligned} \quad (6.17)$$

where $d_{pq,0}^{\odot-1} = 1 - A_2/N$, $d_{pq,rs}^{\odot-1} = -d_{pq,rs}$ and $d_{pq,\tilde{\mathbf{s}}}^{\odot-1}$ has a piece-wise definition similar to Eq. (6.12). Similarly,

$$[\mathcal{G}^{\odot-1}(\hat{X}_{pq}^{-1}(\mathbf{b})) - \check{\mathbf{1}}]_{pq} \approx \sum_{\tilde{\mathbf{s}} \in \mathcal{S}} \hat{d}_{pq,\tilde{\mathbf{s}}}^{\odot-1} e^{-2\pi i \mathbf{b}^T \cdot \tilde{\mathbf{s}}}, \quad (6.18)$$

where $\hat{d}_{pq,\tilde{\mathbf{s}}}^{\odot-1} = d_{pq,\tilde{\mathbf{s}}}^{\odot-1} - 1_{\{\tilde{\mathbf{s}}=\mathbf{0}\}}$.

6.3 Corrected visibilities

The corrected visibility matrix $\mathcal{R}^{(c)}$ can be calculated with Eq. (2.3). As in the case of \mathcal{G} , distilling $\mathcal{R}^{(c)}$ makes it possible to observe and study the fainter spurious emission in $\mathcal{R}^{(c)}$. The distilled corrected visibility matrix \mathcal{R}^Δ (also known as the residual visibility matrix) can be calculated with

$$\mathcal{R}^\Delta = \mathcal{R}^{(c)} - \mathcal{R} = (\mathcal{G}^{\odot-1} - \check{\mathbf{1}}) \odot \mathcal{R}. \quad (6.19)$$

Substitution of our two-source model as given by Eqs. (5.5) and (5.6) as well as Eq. (5.13) gives

$$\begin{aligned} \mathcal{R}^\Delta &= \left(-\Delta \mathbf{g} \mathbf{1}^H - \mathbf{1} \Delta \mathbf{g}^H \right) \odot \left(\mathbf{1} \mathbf{1}^H + A_2 \mathbf{a} \mathbf{a}^H \right) \\ &= -\left(\Delta \mathbf{g} \mathbf{1}^H + \mathbf{1} \Delta \mathbf{g}^H \right) + \\ &\quad -A_2 \left((\Delta \mathbf{g} \odot \mathbf{a}) \mathbf{a}^H + \mathbf{a} (\Delta \mathbf{g} \odot \mathbf{a})^H \right) \end{aligned} \quad (6.20)$$

This equation shows two important characteristics of the ghost phenomenon:

(i) The first term describes the ghost pattern caused by the biases on the gain solutions associated with the calibration source, while the second term describes the ghost pattern associated with the unmodeled source. Since the latter term is just a scaled copy of the first term after phase rotation to the unmodeled source, the ghost pattern centered around each source will be the same.

(ii) The terms associated with the two sources each consist of two terms, that are each other's Hermitian transpose. This implies that the ghost pattern around each source is caused by a perturbation that can be described by a Hermitian matrix. As a result, the ghost pattern on a specific baseline pq may be complex valued, but the ghost patterns that we see in an image are real valued owing to the fact that the contributions from baselines pq and qp are each other's complex conjugate. This result can be confirmed by extrapolation as demonstrated in Appendix D.

In the next section, we will see, that extrapolation is a very convenient tool to locate the individual ghost sources. We therefore use Eqs. (5.3) and (6.18) to express the distilled corrected visibility matrix in the following form:

$$[\mathcal{R}^\Delta(\hat{X}_{pq}^{-1}(\mathbf{b}))]_{pq} \approx \sum_{\tilde{\mathbf{s}} \in \mathcal{S}} \hat{d}_{pq,\tilde{\mathbf{s}}}^{\odot-1} e^{-2\pi i \mathbf{b}^T \cdot \tilde{\mathbf{s}}} (1 + A_2 e^{-2\pi i \mathbf{b}^T \cdot \mathbf{s}_0}) \quad (6.21)$$

6.4 Imaging

In this section we will use two different imaging frameworks to investigate how ghosts manifest during imaging. In Sec 6.4.1 we will be using the idea of beamforming to determine the pixel values in the dirty images associated with the distilled visibility calibration matrix at the position of the modeled and unmodeled source. We will essentially use this framework to verify the validity of the coefficients we derived in Eq. (6.18). We will then create a conventional dirty map of our residuals by taking the inverse Fourier transform of the sampled extrapolated residual visibility plane in Sec. 6.4.2. These conventional dirty maps will demonstrate that we can produce real images even though our ghosts can be complex valued.

6.4.1 Beamforming

If a direct Fourier transform is used with equal weights applied to all visibilities, the image obtained by the direct Fourier transform of the visibilities is the same as the image obtained by consecutively beamforming towards every individual pixel in the image (van der Veen & Wijnholds 2013). We can therefore analyze the pixel values in the distilled image by applying appropriate weights \mathbf{w} to the distilled corrected visibility matrix, i.e.,

$$\begin{aligned} \mathbf{w}^H \mathcal{R}^\Delta \mathbf{w} &= -\mathbf{w}^H \left(\Delta \mathbf{g} \mathbf{1}^H + \mathbf{1} \Delta \mathbf{g}^H \right) \mathbf{w} \\ &\quad - A_2 \mathbf{w}^H \left((\Delta \mathbf{g} \odot \mathbf{a}) \mathbf{a}^H + \mathbf{a} (\Delta \mathbf{g} \odot \mathbf{a})^H \right) \mathbf{w}. \end{aligned} \quad (6.22)$$

Concentrating on the ‘‘atomic’’ ghost pattern, i.e. on $\mathcal{G}^{\odot-1} - \check{\mathbf{1}}$, which is created by the first term of Eq. (6.22) and substituting Eq. (5.11) into it, we find the pixel value

$$\begin{aligned} I_{\mathcal{G}^{\odot-1} - \check{\mathbf{1}}}^D(\mathbf{w}) &= -2\mathcal{R} \left\{ \frac{A_2}{N} (\mathbf{a}^H \mathbf{1}) (\mathbf{w}^H \mathbf{a}) (\mathbf{1}^H \mathbf{w}) \right\} \\ &\quad + \frac{A_2}{N^2} |\mathbf{a}^H \mathbf{1}|^2 (\mathbf{w}^H \mathbf{1}) (\mathbf{1}^H \mathbf{w}). \end{aligned} \quad (6.23)$$

Based on this equation, there are only two locations in $I_{\mathcal{G}^{\odot-1} - \check{\mathbf{1}}}^D$ (the dirty image associated with the atomic ghost pattern) that are obviously special: the location of the modeled source ($\mathbf{w} = \mathbf{1}/N$, normalized such that $|\mathbf{w}^H \mathbf{1}|^2 = 1$ to preserve power) and the location of the unmodeled source ($\mathbf{w} = \mathbf{a}/N$).

At the location of the modeled source, we have

$$\zeta_{s_0}^{\odot-1}(\mathbf{a}) = I_{\mathcal{G}^{\odot-1} - \check{\mathbf{1}}}^D(\mathbf{1}/N) = -\frac{A_2}{N^2} |\mathbf{a}^H \mathbf{1}|^2. \quad (6.24)$$

The minus sign indicates that a negative ghost forms on top of the calibration source, which causes suppression of the flux of the calibration source. We will therefore refer to this ghost as the *primary suppression ghost*.

Let us now shift our focus to the factor $|\mathbf{a}^H \mathbf{1}|^2$. Since we can express this factor as $\sum_{pq} 1 \times e^{2\pi i b_{pq}^T \cdot s_0}$ we may interpret it as the unnormalized tied array beam of our interferometer evaluated at the position of the unmodeled source. The normalized tied array beam is therefore equal to $|\mathbf{a}^H \mathbf{1}|^2 / N^2$. The magnitude of the primary suppressor therefore scales proportionally to the amount of unmodelled flux and the tied array beam sidelobe level measured at the position of the unmodeled source.

At the location of the unmodeled source, we find

$$\zeta_{s_0}^{\odot-1}(\mathbf{a}) = I_{\mathcal{G}^{\odot-1} - \check{\mathbf{1}}}^D(\mathbf{a}/N) = -\frac{A_2}{N^2} |\mathbf{a}^H \mathbf{1}|^2 \left(2 - \frac{|\mathbf{a}^H \mathbf{1}|^2}{N^2} \right). \quad (6.25)$$

Since $|\mathbf{a}^H \mathbf{1}| \leq N$, this negative ghost, which forms on top of the unmodeled source, is at least as bright as the primary suppression ghost. Usually, $|\mathbf{a}^H \mathbf{1}| \ll N$ and this ghost is about twice as bright as the primary suppression ghost. We will refer to this ghost as the *secondary suppression ghost*.

At other points in $I_{\mathcal{G}^{\odot-1} - \check{\mathbf{1}}}^D$, Eq. (6.23) indicates that the ghost pattern is determined by the tied array beam sidelobe response around the modeled source towards the unmodeled source, the tied array beam sidelobe response around the direction of interest towards the modeled source and the tied array beam sidelobe response around the direction of interest towards the unmodeled source. However, Eq. (6.23) does not provide an obvious procedure to predict where the interplay between these sidelobes will result in a significant ghost response. Fortunately, extrapolation allows us to find the ghost positions and fluxes quite easily. We will therefore return to Eq. (6.21) to see what we can learn from extrapolation.

6.4.2 Conventional imaging

In reality we only have access to the visibilities that lie on the uv -tracks of an interferometer. To emulate the real world scenario we need to sample the extrapolated residual visibilities of each baseline and then add them together to obtain our sampled distilled uv -plane $V_{\mathcal{R}^\Delta}$ (assuming natural weighting). This sampled uv -plane is exactly the same plane we would have obtained in the absence of extrapolation. Our dirty distilled image $I_{\mathcal{R}^\Delta}^D$ corresponds to the Fourier inverse of $V_{\mathcal{R}^\Delta}$. The sampled distilled uv -plane is equal to (assuming natural weighting)

$$V_{\mathcal{R}^\Delta} = \sum_{p \neq q} \left[\mathcal{R}^\Delta \left(\hat{X}_{pq}^{-1}(\mathbf{b}) \right) \right]_{pq} S_{pq}(\mathbf{b}), \quad (6.26)$$

where $S_{pq}(\mathbf{b})$ is the sampling function of baseline pq .

Substituting Eq. (6.21) in Eq. (6.26) results in

$$V_{\mathcal{R}^\Delta} = \sum_{p \neq q} \left(\sum_{\bar{s} \in \mathcal{S}} \hat{d}_{pq, \bar{s}}^{\odot-1} e_{\bar{s}} (1 + A_2 e_{s_0}) \right) S_{pq}(\mathbf{b}), \quad (6.27)$$

where $e_{\bar{s}} = e^{-2\pi i \mathbf{b}^T \cdot \bar{s}}$. Changing the order of summation, we obtain

$$V_{\mathcal{R}^\Delta} = \sum_{\bar{s} \in \mathcal{S}} \left(\sum_{p \neq q} \hat{d}_{pq, \bar{s}}^{\odot-1} S_{pq}(\mathbf{b}) \right) (e_{\bar{s}} + A_2 e_{\bar{s} + s_0}) \quad (6.28)$$

This is equivalent to

$$\begin{aligned} V_{\mathcal{R}^\Delta} &= \sum_{\bar{s} \in \mathcal{S}} \left(\sum_{p < q} \hat{d}_{pq, \bar{s}}^{\odot-1} S_{pq}(\mathbf{b}) + \hat{d}_{qp, \bar{s}}^{\odot-1} S_{qp}(\mathbf{b}) \right) \\ &\quad \times (e_{\bar{s}} + A_2 e_{\bar{s} + s_0}). \end{aligned} \quad (6.29)$$

Since $\hat{d}_{qp, \bar{s}}^{\odot-1} = \overline{\hat{d}_{pq, \bar{s}}^{\odot-1}}$ (due to the Hermiticity of the visibility calibration matrix), we can reformulate Eq. (6.29) as

$$\begin{aligned} V_{\mathcal{R}^\Delta} &= \sum_{\bar{s} \in \mathcal{S}} \left(\sum_{p < q} \mathcal{R} \left\{ \hat{d}_{pq, \bar{s}}^{\odot-1} \right\} S_{pq}^+(\mathbf{b}) + \mathcal{I} \left\{ \hat{d}_{pq, \bar{s}}^{\odot-1} \right\} S_{pq}^-(\mathbf{b}) \right) \\ &\quad \times (e_{\bar{s}} + A_2 e_{\bar{s} + s_0}), \end{aligned} \quad (6.30)$$

where

$$S_{pq}^+(\mathbf{b}) = S_{pq}(\mathbf{b}) + S_{qp}(\mathbf{b}), \quad (6.31)$$

and

$$S_{pq}^-(\mathbf{b}) = iS_{pq}(\mathbf{b}) - iS_{qp}(\mathbf{b}). \quad (6.32)$$

The Fourier inverse of Eq. (6.30) is equal to

$$I_{\mathcal{R}\Delta}^D = \sum_{\tilde{\mathbf{s}} \in \mathcal{S}} \left(\sum_{p < q} \mathcal{R} \left\{ \hat{d}_{pq, \tilde{\mathbf{s}}}^{\circ-1} \right\} P_{pq}^+(\mathbf{s}) + \mathcal{I} \left\{ \hat{d}_{pq, \tilde{\mathbf{s}}}^{\circ-1} \right\} P_{pq}^-(\mathbf{s}) \right) \star (\delta_{\tilde{\mathbf{s}}} + \delta_{\tilde{\mathbf{s}} + \mathbf{s}_0}), \quad (6.33)$$

where $P_{pq}^+(\mathbf{s}) = \mathcal{F}^{-1}\{S_{pq}^+(\mathbf{b})\}$, $P_{pq}^-(\mathbf{s}) = \mathcal{F}^{-1}\{S_{pq}^-(\mathbf{b})\}$, $\delta_{\tilde{\mathbf{s}}} = \delta(\mathbf{s} - \tilde{\mathbf{s}})$ and “ \star ” denotes convolution.

Note that $P_{pq}^+(\mathbf{s})$ is the (unnormalized) point spread function (PSF) associated with baseline pq . Since $S_{pq}^-(\mathbf{b})$ is imaginary and anti-symmetric, $P_{pq}^-(\mathbf{s})$ is real valued and anti-symmetric. We can therefore also see it as a type of PSF that we can associate with baseline pq . To avoid confusion let us respectively refer to $P_{pq}^+(\mathbf{s})$ and $P_{pq}^-(\mathbf{s})$ as the *real point spread function* (RPSF) and *imaginary point spread function* (IPSF) of baseline pq . Due to the properties of $P_{pq}^+(\mathbf{s})$ and $P_{pq}^-(\mathbf{s})$ the RPSF has one peak at the origin (which is its maximum) whereas the IPSF has two peaks close to the origin, one negative and one positive (its minimum and maximum). In Paper I we defined the *ghost spread function* (GSF) $P_{\tilde{\mathbf{s}}}(\mathbf{s})$ in terms of $P_{pq}^+(\mathbf{s})$. This definition is only correct if the ghosts are not complex (which is the case for an east-west array). A more general definition however is:

$$P_{\tilde{\mathbf{s}}}(\mathbf{s}) = \left(\sum_{p < q} \mathcal{R} \left\{ \hat{d}_{pq, \tilde{\mathbf{s}}}^{\circ-1} \right\} P_{pq}^+(\mathbf{s}) + \mathcal{I} \left\{ \hat{d}_{pq, \tilde{\mathbf{s}}}^{\circ-1} \right\} P_{pq}^-(\mathbf{s}) \right). \quad (6.34)$$

The GSF of the ghost at $\tilde{\mathbf{s}}$ is therefore equal to the sum of two terms, where the first term is equal to the sum of its real flux in each baseline multiplied by the (unnormalized) RPSF of each baseline and the second term is equal to the sum of its imaginary flux in each baseline multiplied by the (unnormalized) IPSF of each baseline. Eq. (6.34) enables us to rewrite Eq. (6.33) to

$$I_{\mathcal{R}\Delta}^D = \sum_{\tilde{\mathbf{s}} \in \mathcal{S}} P_{\tilde{\mathbf{s}}}(\mathbf{s}) \star (\delta(\mathbf{s} - \tilde{\mathbf{s}}) + A_2 \delta(\mathbf{s} - \tilde{\mathbf{s}} - \mathbf{s}_0)) \quad (6.35)$$

$$= \sum_{\tilde{\mathbf{s}} \in \mathcal{S}} P_{\tilde{\mathbf{s}}}(\mathbf{s} - \tilde{\mathbf{s}}) + A_2 P_{\tilde{\mathbf{s}}}(\mathbf{s} - \tilde{\mathbf{s}} - \mathbf{s}_0) \quad (6.36)$$

This equation confirms our earlier observation, that the distilled dirty image for our two-source model is a superposition of two identical ghost patterns, one centered around the modeled source and one centered around the unmodeled source.

We now add the position $2\mathbf{s}_0$ to \mathcal{S} to obtain the set $\hat{\mathcal{S}}$. We can now approximate Eq. (6.36) with

$$\sum_{\tilde{\mathbf{s}} \in \hat{\mathcal{S}}} P_{\tilde{\mathbf{s}}}(\mathbf{s} - \tilde{\mathbf{s}}) + A_2 P_{\tilde{\mathbf{s}} - \mathbf{s}_0}(\mathbf{s} - \tilde{\mathbf{s}}). \quad (6.37)$$

The reason for adding $2\mathbf{s}_0$ is that this position becomes significant in the conglomerated ghost pattern. For each individual baseline, a very faint ghost is located here in the distilled visibility calibration matrix, but it is so faint that it is not predicted by the perturbation approach. We restrict the sum in Eq. (6.37) to the set $\hat{\mathcal{S}}$ (without any loss of generality) as it contains the positions of the brightest ghosts in

$I_{\mathcal{R}\Delta}^D$. Eq. (6.37) merely adds up the ghost responses of the ghost pattern around the modeled and unmodeled source if they both contain a ghost at the same position $\tilde{\mathbf{s}}$.

We are now able to determine the amplitude of the ghosts contained in $\mathcal{G}^{\circ-1} - \hat{\mathbf{I}}$ and \mathcal{R}^Δ . For the matrix $\mathcal{G}^{\circ-1} - \hat{\mathbf{I}}$ the amplitude of the ghost at $\tilde{\mathbf{s}}$ is equal to $P_{\tilde{\mathbf{s}}}(\mathbf{0})$. Based on Eq. (6.34), this amplitude can be determined by the weighted sum

$$P_{\tilde{\mathbf{s}}}(\mathbf{0}) = \sum_{p < q} \mathcal{R} \left\{ \hat{d}_{pq, \tilde{\mathbf{s}}}^{\circ-1} \right\} P_{pq}^+(\mathbf{0}), \quad (6.38)$$

since $P_{pq}^-(\mathbf{0}) = 0$. In the case of natural weighting, Eq. (6.38) allows us to define

$$\zeta_{\tilde{\mathbf{s}}}^{\circ-1} = \left\langle \mathcal{R} \left\{ \hat{d}_{pq, \tilde{\mathbf{s}}}^{\circ-1} \right\} \right\rangle_{p < q}, \quad (6.39)$$

where $\langle \cdot \rangle_{p < q}$ denotes averaging over all the baselines. Eq. (6.39) denotes the total real flux of the ghost at position $\tilde{\mathbf{s}}$ in $I_{\mathcal{G}^{\circ-1} - \hat{\mathbf{I}}}^D$.

It is actually more prudent to calculate the maximum possible flux that the ghost at position $\tilde{\mathbf{s}}$ can contribute to the image. We therefore also define

$$\hat{\zeta}_{\tilde{\mathbf{s}}}^{\circ-1} = \left\langle \left| \hat{d}_{pq, \tilde{\mathbf{s}}}^{\circ-1} \right| \right\rangle_{p < q}. \quad (6.40)$$

Eq. (6.40) is important as Eq. (6.39) can be deceiving. Although only the real flux manifest at $\tilde{\mathbf{s}}$, some of the imaginary flux does manifest in the proximity of $\tilde{\mathbf{s}}$ due to the IPSF. Eq. (6.40) should be interpreted as a type of upper bound. It gives the maximum flux that the source could have contributed to the image. The source contributes this maximum amount of flux if it is completely real. With the aid of Eq. (6.37) we may also define

$$\zeta_{\tilde{\mathbf{s}}}^\Delta = \zeta_{\tilde{\mathbf{s}}}^{\circ-1} + A_2 \zeta_{\tilde{\mathbf{s}} - \mathbf{s}_0}^{\circ-1} \quad (6.41)$$

and

$$\hat{\zeta}_{\tilde{\mathbf{s}}}^\Delta = \hat{\zeta}_{\tilde{\mathbf{s}}}^{\circ-1} + A_2 \hat{\zeta}_{\tilde{\mathbf{s}} - \mathbf{s}_0}^{\circ-1}, \quad (6.42)$$

which are the $I_{\mathcal{R}\Delta}^D$ equivalent of Eq. (6.39) and Eq. (6.40).

6.4.3 Comparison

We have derived two ways of estimating the flux of the primary and secondary suppressor (associated with $\mathcal{G}^{\circ-1} - \hat{\mathbf{I}}$). We can estimate the primary suppressor with either $\zeta_0^{\circ-1}$ using Eq. (6.24) or $\zeta_0^{\circ-1}$ using Eq. (6.39), while we can approximate the secondary suppressor with either $\zeta_{\mathbf{s}_0}^{\circ-1}$ using Eq. (6.25) or $\zeta_{\mathbf{s}_0}^{\circ-1}$ using Eq. (6.39). The first approach ($\zeta_*^{\circ-1}$) uses the location of the unmodeled source as an input parameter (it depends on the tied array beam sidelobe level at the location of the unmodeled source) while the latter ($\zeta_*^{\circ-1}$) does not. The two approaches therefore seem to be incompatible with one another, since they will yield different results as you move the unmodeled source around in the lm -plane. The argument below shows why this is not the case.

During a synthesis observation, the unmodeled source will move through a (large) number of sidelobes of the tied array beam pointed at the modeled source. Moreover, we usually do not know where an unmodeled source will appear.

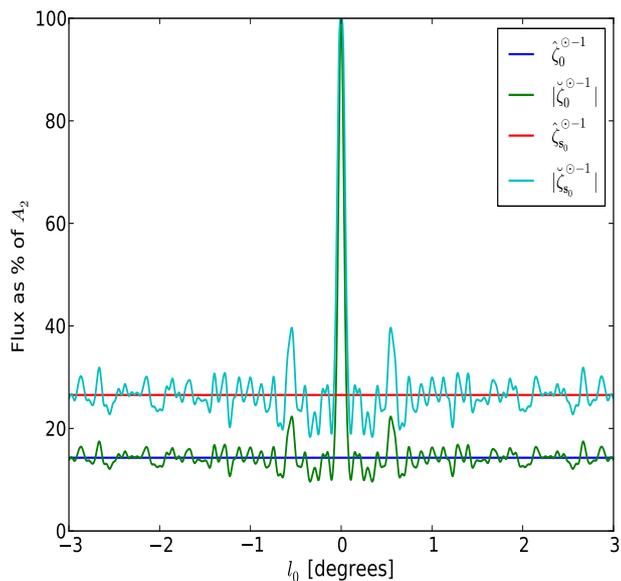


Figure 5. The flux in the primary and secondary suppressor as a percentage of A_2 ($A_2 = 0.2$ Jy) as predicted by Eq. (6.24), Eq. (6.25) and Eq. (6.40) for $l_0 \in [-3^\circ, 3^\circ]$ and $m_0 = 0^\circ$. This result was generated with the KAT-7 antenna layout with $\delta_0 = -74.66^\circ$ and $\nu = 1.445$ GHz. The primary difference between the constant and varying curves is that the constant curves represent the flux of the ghosts as measured in a perfectly “clean” image, while the varying curves depict the pixel values at the ghost positions in a dirty image.

It is therefore reasonable to see whether the two approaches produce the same average value. We therefore calculate

$$\mathcal{E}_{s_0} \left\{ \zeta_0^{\circ-1}(\mathbf{a}) \right\} = -\frac{A_2}{N^2} \mathcal{E}_{s_0} \left\{ \left| \mathbf{a}^H \mathbf{1} \right|^2 \right\} = -\frac{A_2}{N} = \zeta_0^{\circ-1}, \quad (6.43)$$

where $\mathcal{E}_{s_0} \{ \cdot \}$ denotes expectation over all the possible positions that the unmodeled source can occupy in the lm -plane. The second and third expression in Eq. (6.43) are equal, since

$$\begin{aligned} \mathcal{E}_{s_0} \left\{ \left| \mathbf{a}^H \mathbf{1} \right|^2 \right\} &= \mathcal{E}_{s_0} \left\{ \sum_{pq} e^{-2\pi i \mathbf{b}_{pq}^T \cdot \mathbf{s}_0} \right\} \\ &= N + \sum_{p < q} \mathcal{E}_{s_0} \left\{ \cos(2\pi \mathbf{b}_{pq}^T \cdot \mathbf{s}_0) \right\} \\ &= N. \end{aligned} \quad (6.44)$$

This also agrees with the well known result that the average sidelobe level of the normalized tied array beam of an irregular array is $1/N$ (if the autocorrelations are included during imaging). Similarly,

$$\mathcal{E}_{s_0} \left\{ \zeta_{s_0}^{\circ-1}(\mathbf{a}) \right\} = \zeta_{s_0}^{\circ-1}. \quad (6.45)$$

We have now established that on average the two approaches with which we can estimate the primary and secondary suppressor produce similar results. We plot $\zeta_0^{\circ-1}$, $\hat{\zeta}_0^{\circ-1}$, $\zeta_{s_0}^{\circ-1}$ and $\hat{\zeta}_{s_0}^{\circ-1}$ in Fig. 5 for a KAT-7 full synthesis simulation which also verifies this result.

The curves associated with the approach defined in terms of the tied array beam vary as a function of \mathbf{s}_0 , because they represent actual pixel values at the position of the modeled and unmodeled source in a dirty image. The

approach that gives us constant values, on the other hand, gives the ghost values that one would measure in an image that contained no psf sidelobes. The two peaks that are observable in Fig. 5 occur when the modeled and unmodeled source lie on top of each other. When the modeled and unmodeled source lie on top of each other then the gains can completely compensate for the missing flux, which explains why the peak values are 100%.

7 TAXONOMY

Based on our analysis, we can now identify different classes of ghost sources. There are two main ways of classifying the ghosts based on their amplitude. The first approach is to concentrate on the ghosts produced by each individual baseline, while the second approach entails looking at the ghosts that form when we combine the ghost patterns of all baselines. The reason for creating a separate taxonomy for the conglomerated ghost pattern is that some ghost positions are common to a number of baselines, while others are not, which implies that a ghost that is dim in the per-baseline patterns may turn out to be bright in the conglomerated pattern if it occurs in all baselines when compared to ghosts that may be bright in one baseline but do not re-occur in any other baseline.

If we focus on each individual baseline we can construct the following taxonomy of ghosts:

First order ghosts: These ghosts are predicted by the perturbation approach, i.e., they are associated with the first order deviation from the true gain values. The first order ghosts, which appear at locations $\tilde{\mathbf{s}} \in \mathcal{S}$, include the brightest ghost sources and can be subdivided in two categories:

(i) In Sec. 5.3 we saw that the amplitude of some ghosts are inversely proportional to the number of antennas in the array, N , while the amplitude of other ghosts are inversely proportional to N^2 . We will refer to the first as *proto-ghosts*. These can be found at positions $\tilde{\mathbf{s}} \in \mathcal{S}^P$.

(ii) The *deutero-ghosts* are ghosts whose amplitude is inversely proportional to N^2 . These ghosts can be found at locations $\tilde{\mathbf{s}} \in \mathcal{S}^D$ and are usually significantly fainter than the proto-ghosts.

Higher order ghosts: These are the faintest ghosts associated by the second and higher order deviations from the true gain values. As a result, the higher order ghosts cannot be predicted using the perturbation approach, but they can be found by extrapolation.

If we concentrate on the conglomerated ghost pattern we can create the following taxonomy of ghosts:

Line ghosts: All ghosts with position vector $\tilde{\mathbf{s}} \in \mathbf{s}_0 \mathbf{Z}$ belong to this group. The brightest ghosts in our conglomerated pattern belong to this group. The following four important ghosts belong to this category:

(i) The *primary suppression ghost* is a negative ghost that forms on top of the modeled source and is responsible for the observed suppression of the modeled source. It is also a *proto-ghost*. Its flux is denoted by ζ_0^Δ (associated with the distilled corrected visibility matrix, the distilled visibility calibration matrix equivalent of ζ_0^Δ is $\zeta_0^{\circ-1}$).

(ii) The *secondary suppression ghost* is the negative ghost that forms on top of the unmodeled source and therefore suppresses the unmodeled source. This source is responsible for the well-known self-cal bias. It is also a *proto-ghost*. Its flux is denoted by $\zeta_{s_0}^\Delta$. The flux of the secondary suppression ghost associated with the visibility calibration matrix is denoted by $\zeta_{s_0}^{\circ-1}$.

(iii) The *anti-ghost of the secondary suppressor* forms symmetrically opposite to the position of the secondary suppression ghost, i.e., at $-s_0$. It is also a *deutero-ghost*. The flux of the ghosts at this position for the distilled corrected visibility matrix and the distilled calibrated visibility matrix are respectively denoted by $\zeta_{-s_0}^\Delta$ and $\zeta_{-s_0}^{\circ-1}$.

(iv) The *anti-ghost of the primary suppressor* is a negative ghost that forms at $2s_0$, i.e., at the symmetrically opposite position of the modeled source around the unmodeled source. However, $2s_0 \notin \mathcal{S}$, so this is a *higher order ghost*.

Scattered ghosts: All ghosts that are not line ghosts belong to this group. We can divide this group into three main groups:

(i) The *scattered proto-ghosts* include all scattered ghosts that are proto-ghosts. The primary and secondary suppressor are excluded from this category. These ghosts are generally brighter than the remaining categories.

(ii) The *scattered deutero-ghosts* include all the scattered ghosts that are deutero-ghosts. The anti-ghost of the secondary suppressor is excluded from this category. These ghosts are generally dimmer than the scattered proto-ghosts.

(iii) The *scattered higher order ghosts* include all the higher order ghosts that are not line ghosts. Note that this restriction excludes the anti-ghost of the primary suppressor from this category. This group of ghosts usually contains the faintest ghosts.

8 GHOST PATTERN OF INDIVIDUAL BASELINES

The aim of this section is to present baseline specific results. In Sec. 8.1 we image the systematics in the antenna gains created via simulations. We then compare the simulation results with the theory in Sec. 6. In Sec. 8.2 we visually explore the main results we derived in Sec. 6. The reader should keep in mind that Sec. 6 is actually self-contained, every final outcome of that section logically follows from previous results. We end this section by exploring the parameter space that influences the positions and amplitudes of the brightest ghosts. The conglomerated ghost pattern, which is created by combining the ghost patterns of all baselines into one global ghost pattern, is analyzed in Sec. 9.

The simulated results presented here and in Sec. 9 are created using the KAT-7 antenna layout. KAT-7 is a seven-dish interferometer located in the Karoo in South Africa and is a precursor of the MeerKAT telescope². The KAT-7 layout is depicted in Fig. 6. Choosing an antenna layout with which to verify the theory introduced in this paper is

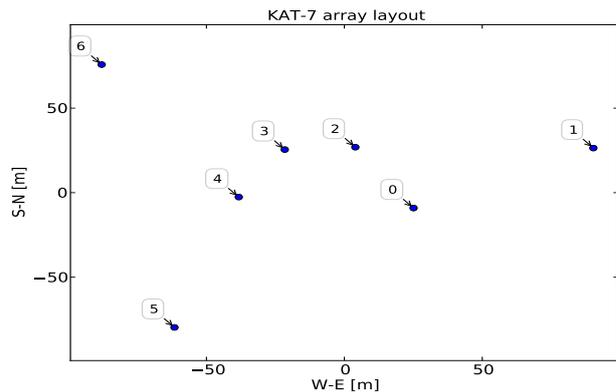


Figure 6. The KAT-7 antenna layout. The latitude of the KAT-7 array is $-30^\circ 43' 17.34''$.

actually quite arbitrary. However, using the KAT-7 layout does offer two benefits. Firstly, an array configuration of only seven dishes makes the analysis tractable, since the number of ghosts predicted by the perturbation approach increases with the number of antennas. Secondly, the KAT-7 layout produces a spatially compact ghost pattern since the array itself is compact (see Fig. 9 to Fig. 11).

8.1 Simulated results

The results in this section were created via simulation, i.e. extrapolation and perturbation were not used here. The results will therefore be able to serve as an independent check of the theoretical derivations we made in Sec. 6. To make these images, we simulated a KAT-7 dataset which contained a central source of 1 Jy and one off-center source of 0.2 Jy at 1° . We then calibrated the data by only keeping the 1-Jy source in the calibration model. No thermal noise or gain errors were added in the simulation. We then made a dirty image of the distilled calibrated visibilities of baseline 45 (i.e. $g_{45} - 1$), which is shown in the left panel of Fig. 7. We then ran CLEAN on the obtained dirty image with a cleaning mask that contained the theoretical ghost positions predicted by Eq. (6.3) (which are indicated with white markers in the left panel of Fig. 7). The right image in Fig. 7 is the result of this procedure.

We can make the following observations from Fig. 7:

(i) The theoretical ghost positions that Eq. (6.3) predicts (the white markers), align nicely with the visible ghost sources in the right image of Fig. 7. This result supports the validity of Eq. (6.3).

(ii) Some of the ghost sources in Fig. 7 contain two peaks, a bright negative peak and an almost equal symmetrically opposite positive peak. This squares with the theoretical results derived in Section 6.4. Each ghost is complex, but since the ghosts of baseline pq are the complex conjugates of the ghosts of baseline qp the ghosts manifest as a weighted sum of two PSFs, the RPSF and the IPSF of baseline 45. The RPSF and IPSF of baseline 45 are shown in Fig. 8. The ghosts that contain the two peak structure have the same orientation as the IPSF in Fig. 8. Not all ghosts have this two peak structure. This is easily explained by realizing that if a ghost contains more real flux than imaginary flux it will appear to manifest with a RPSF, and if it contains more

² <http://public.ska.ac.za/kat-7>

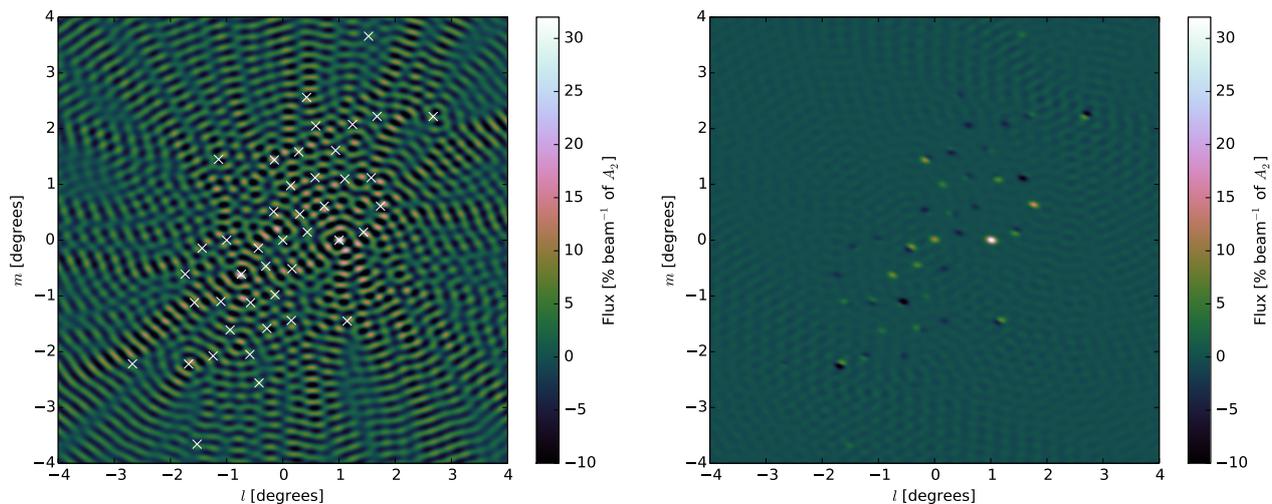


Figure 7. A dirty (left) and clean (right) image of $g_{45}-1$, with the unmodeled source being 0.2 Jy at 1° , $\delta_0 = -74.66^\circ$ and $\nu = 1.445 \text{ GHz}$.

imaginary flux than real flux it will appear to manifest with an IPSF.

(iii) The clean image is not that clean. This should not be surprising, since the image was created with traditional CLEAN, which does not take into account that the GSF of each ghost is not equal to the traditional PSF of baseline 45.

8.2 Theoretical results

In this section we visually explore the theory we derived in Sec. 6. Here we will only concentrate on the distilled visibility calibration matrix $\mathcal{G}^{\circ-1} - \mathbf{1}$, as we have shown in Sec. 5.3 that there exists a simple relation between \mathcal{G} and $\mathcal{G}^{\circ-1}$. We carry out our theoretical analysis with KAT-7.

We plot the clean artefact maps of baselines 23, 45 and 15 respectively in the top and middle panels of Fig. 9, Fig. 10 and Fig. 11. These figures were created via extrapolation (see Sec. 5.2) and can therefore be interpreted as an exact empirical representation of the ghost patterns of baselines 23, 45 and 15. These images are complex, since the ghosts have complex flux (also note the geometric resemblance between Fig. 7 and Fig. 10).

The difference between the images in the top and middle panels is that the top panel was created by taking into account the autocorrelations during calibration, while the autocorrelations were excluded when creating the images in the middle panel. The theory we derived in Sec. 6 is for the case when the autocorrelations are included. This nicely extends the regular east-west derivations we made in Paper I, which also included the autocorrelations. We show the extrapolated results for the case where the autocorrelations are excluded as well to check whether leaving them out has a significant impact on the results we obtain from our theoretical analysis in Sec. 6.

The images in the top and middle panels are virtually identical (the bright ghosts are located at the same positions). There are, however, some differences that should be highlighted:

(i) The primary suppressor is not present in the middle

panels. When we do not include the autocorrelations during calibration then we do not suppress the modeled source. The reason why the primary suppressor disappears is explained in Paper I.

(ii) Close inspection of the top and middle panels reveals that the fluxes of the remaining ghosts are not equal. However, the observable difference in the fluxes of the top and middle panels does not have a significant impact on the major results in this paper.

These minor differences imply that the major conclusions (see Sec. 10) that we make based on the theory we derived in Sec. 6 generally hold for the case when we exclude the autocorrelations during calibration as well. We came to a similar conclusion when we studied the regular east-west case in Paper I. We can therefore assume that excluding the autocorrelations during calibration will not significantly alter the results obtained in this paper, with the primary suppressor as the major exception.

Although the extrapolated images in the top panel are useful in their own right, having an analytic equation with which we can predict the positions and fluxes of the ghosts would be even more advantageous. We can accomplish this if we combine perturbation and extrapolation. This culminates in the derivation of Eq. (6.17), which gives us an analytic expression with which we can calculate estimates of the fluxes and the position vectors of the brightest ghosts. As we discussed in Sec. 5.3, the coefficients of this sum allow us to subdivide the ghosts into two main groups, namely proto-ghosts and deuterio-ghosts. Proto-ghosts are much brighter than deuterio-ghosts. The bottom left images of Fig. 9, Fig. 10 and Fig. 11 graphically presents the positions of the proto-ghosts (blue circles) and deuterio-ghosts (green squares) of baselines 23, 45 and 15, which are commensurate with the images on the top rows of Fig. 9, Fig. 10 and Fig. 11.

Another interesting concomitant of Eq. (6.17) is that we can more or less interpret each ghost as an unwanted frequency (which emanates from the linear dependence between the uv -tracks of an interferometer) that “leaked” into baseline pq from an entry in $\mathcal{R}(\hat{X}_{pq}^{-1}(\mathbf{b}))$. The baselines (entries) that we can associate with each ghost are depicted in the bottom right images of Fig. 9, Fig. 10 and Fig. 11.

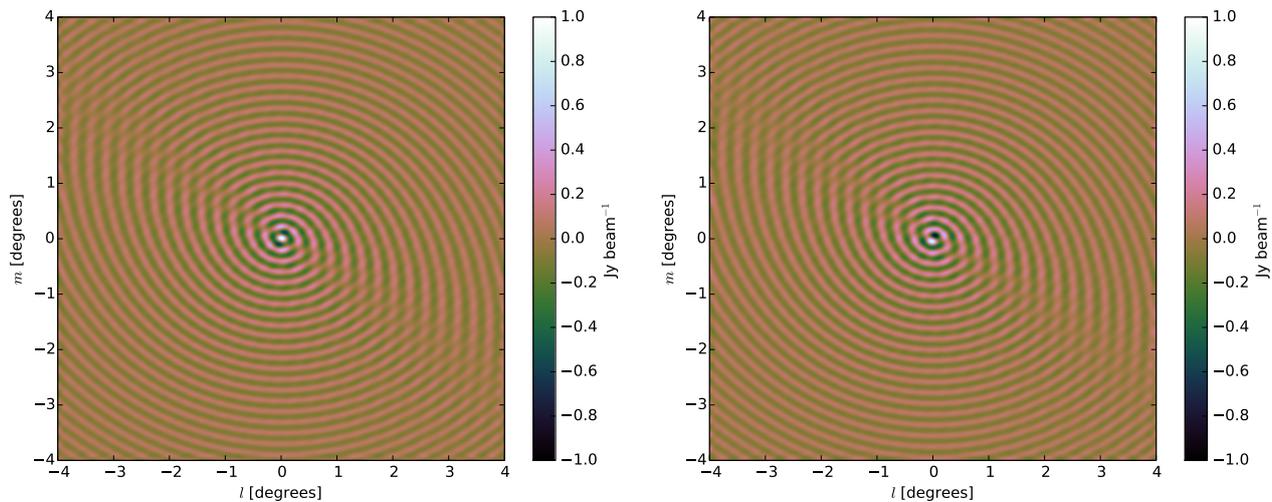


Figure 8. The RPSF (left) and IPSF (right) of baseline 45, with $\delta_0 = -74.66^\circ$ and $\nu = 1.445$ GHz.

Fig. 9, Fig. 10 and Fig. 11 also reveal that ghosts do form in symmetrically opposite pairs. A close examination of these plots reveals two types of symmetrically opposite ghost pairs. Intuitively, one would already expect the ghost associated with baseline rs to form a pair with the ghost associated with baseline sr . These ghost pairs, which are symmetrically opposite around the phase center, can indeed be found in Figs. 9, 10 and 11. However, it turns out that for the ghost pattern of baseline pq the baselines ps and rq also form ghost pairs, which are symmetrically opposite around the mid-point of the line connecting the modeled and the unmodeled source. These ghosts can be considered to be associated with phase closure errors introduced by the incomplete sky model.

8.3 Parameter dependencies

The aim of this section is to explore the parameter space that determines the positions and fluxes of the ghosts. Let us first explore the parameter space of the fluxes of the ghosts. The coefficients in Eq. (6.17) imply that the flux of the ghosts are complex. The absolute fluxes of the ghosts are only influenced by A_2 and N (see Eq. (5.15) and Eq. (5.16)), while the amount of real and imaginary flux contained in the amplitude of a ghost has a more complicated parameter space, which also depends on the observing wavelength λ , the declination of the source δ_0 and the orientation and length of the baseline as described by ϕ_{pq} , θ_{pq} and $\Delta\mathbf{b}_{pq}$ (see Eq. (6.4) for further details).

It is also interesting to note that some of the ghosts are always real. The anti-ghost and the secondary suppressor are always real, since they are caused by leaked signal from every individual baseline itself, i.e., the relevant terms in Eq. (6.3) are the $(r, s) = (p, q)$ and $(r, s) = (q, p)$, for which Eq. (6.4) is equal to unity. It is also trivially true that the primary suppressor is real. We explore this parameter space further in Sec. 9.2.

According to Eq. (6.5) the position vectors of the ghosts are primarily influenced by the geometry of the array, the declination of the observation and the position of the unmodeled source. Fig. 12 graphically illustrates the depen-

dence that exists between the declination of the observation and the position vectors of the ghost sources. Based on the results shown in Figs. 9, 10 and 11 one may conclude that a shorter baseline will have a ghost pattern with a larger spatial extent and one may therefore expect that an observation at a large zenith angle will be subject to a ghost pattern with a larger spatial extent than an observation at a small zenith angle due to projection effects. Fig. 12 seems to contradict this intuition. This is caused by the larger w -component associated with observations at larger zenith angles.

9 CONGLOMERATED GHOST PATTERN

In this section we will be investigating how the ghosts behave in the final conglomerated image. We will discuss the theoretical results of Sec. 6 and explore the parameter space that influences the fluxes of the ghosts.

9.1 Theoretical results

Up to this point we have been looking at individual baseline results. In reality we are interested in creating a conglomerated image that contains the ghost contributions of all baselines. We again focus here on $\mathcal{G}^{\ominus-1}$. The left image of Fig. 13 displays the clean artefact map of the full KAT-7 array and is obtained by averaging the per-baseline extrapolated results. Note that the images in Fig. 13 are real, since the imaginary fluxes of baselines pq and qp simply cancel, which is actually a bit deceptive. When we take into account the sampling function, the imaginary flux does not simply cancel and will manifest through the GSF of each ghost when we image. The right image of Fig. 13 is just the left image after we have removed the ghosts at \mathbf{s}_0 and $\mathbf{0}$. When we inspect Fig. 13 the first thing we notice, is that the secondary suppressor and the primary suppressor are now much brighter than the other proto-ghosts. Even more interesting, the anti-ghost, which is actually a deutero-ghost, is now the third brightest ghost. What has happened is that only these three ghosts are consistently present on all baselines (i.e. $\hat{X}_{pq} \circ \hat{X}_{pq}^{-1}(\mathbf{b}) = \mathbf{b}$ and $\hat{X}_{qp} \circ \hat{X}_{qp}^{-1}(\mathbf{b}) = -\mathbf{b}$), which means that they add up

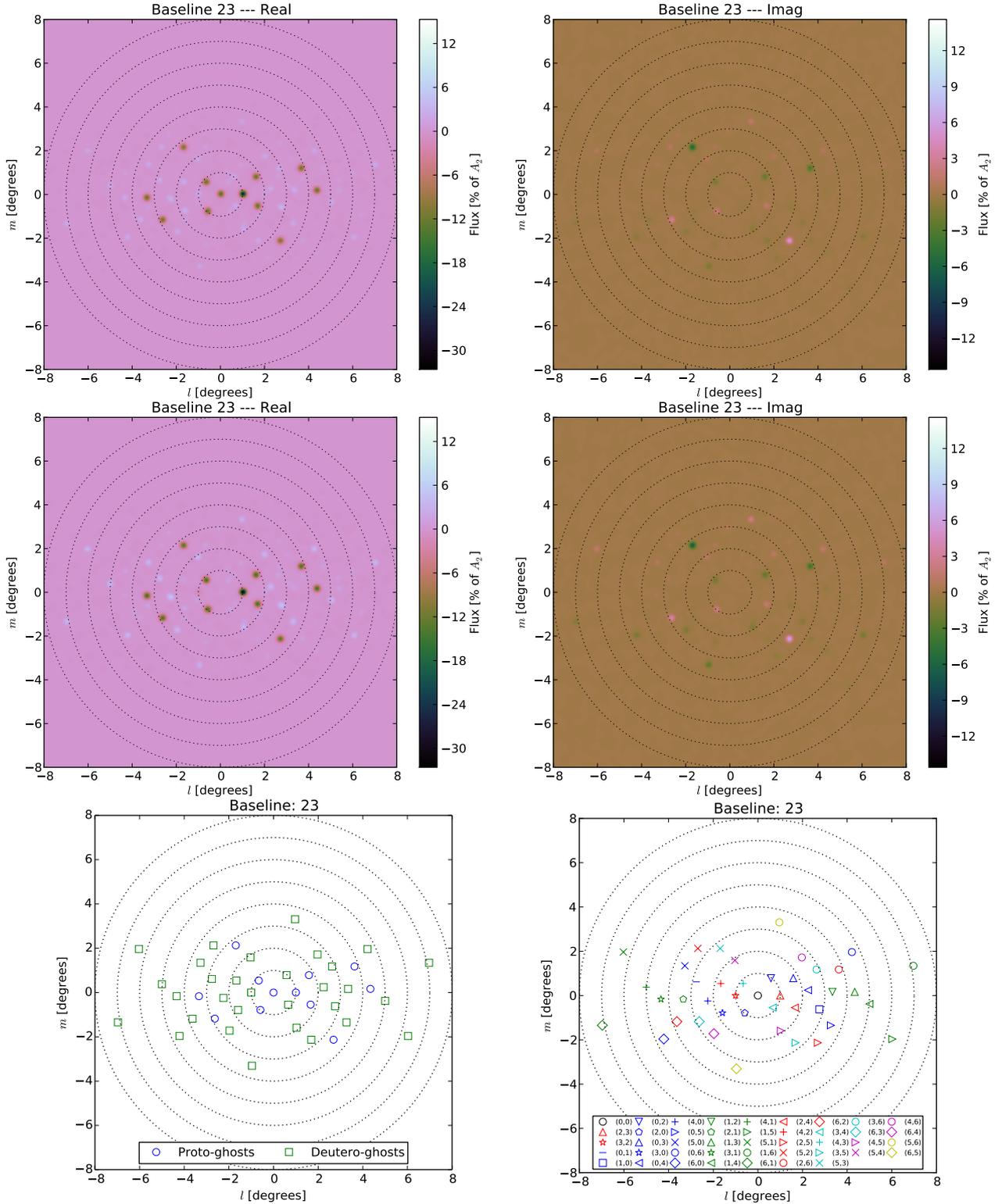


Figure 9. The images on the top and middle row show the real (left) and imaginary (right) part of the theoretical distilled ghost pattern for the shortest baseline 23, i.e. $\mathcal{F}^{-1}\{g_{23}^{-1}(X_{23}^{-1}(\mathbf{b})) - 1\}$, with the unmodeled source being 0.2 Jy at 1° , $\delta_0 = -74.66^\circ$ and $\nu = 1.445$ Ghz. The images on the top row were created by including the autocorrelations during calibration, while the middle row was created by excluding them. The images on the bottom row show the theoretical distilled ghost pattern (positions only) for this baseline. The left image displays where the proto-ghosts and the deutero-ghosts of this baseline are located. The right image indicates from which baseline each ghost was “leaked”.

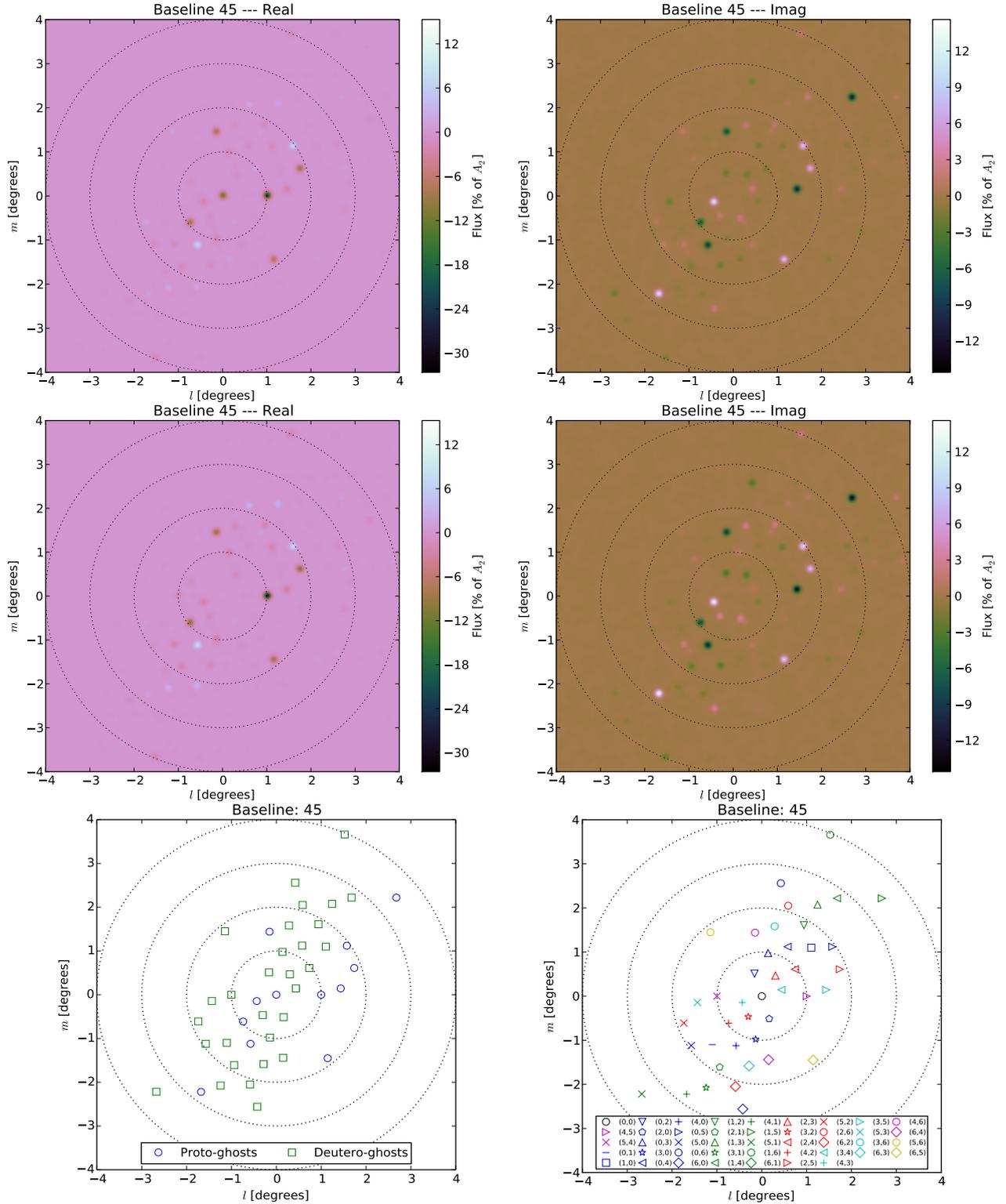


Figure 10. The images on the top and middle row show the real (left) and imaginary (right) part of the theoretical distilled ghost pattern for baseline 45, i.e. $\mathcal{F}^{-1}\{g_{45}^{-1}(X_{45}^{-1}(\mathbf{b}) - 1)\}$, with the unmodeled source being 0.2 Jy at 1° , $\delta_0 = -74.66^\circ$ and $\nu = 1.445$ GHz. The images on the top row were created by including the autocorrelations during calibration, while the middle row was created by excluding them. The images on the bottom row show the theoretical distilled ghost pattern (positions only) for this baseline. The left image displays where the proto-ghosts and the deutero-ghosts of this baseline are located. The right image indicates from which baseline each ghost was “leaked”.

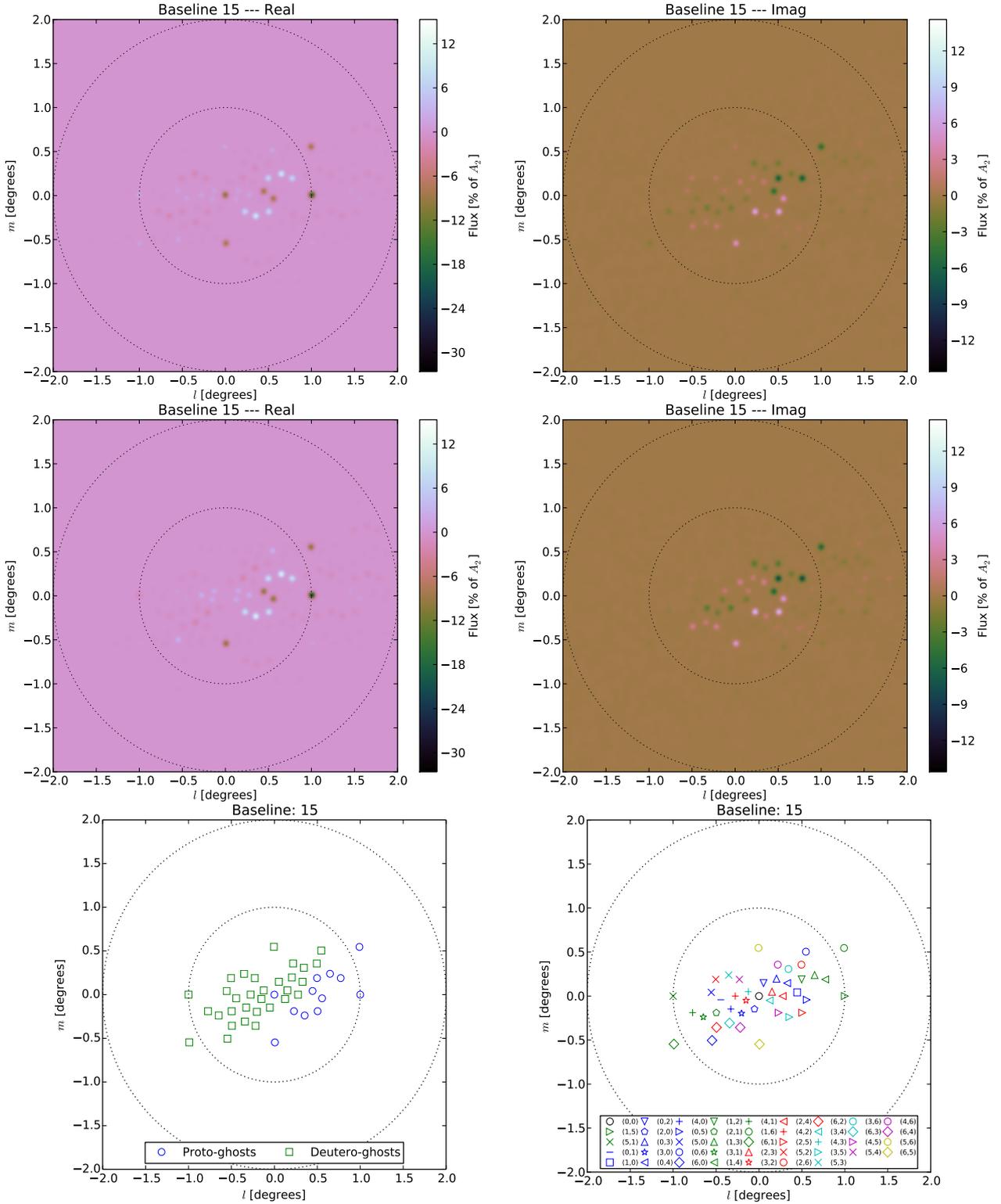


Figure 11. The images on the top and middle row show the real (left) and imaginary (right) part of the theoretical distilled ghost pattern for the longest baseline 15, i.e. $\mathcal{F}^{-1}\{g_{15}^{-1}(X_{15}^{-1}(\mathbf{b}) - 1)\}$, with the unmodeled source being 0.2 Jy at 1° , $\delta_0 = -74.66^\circ$ and $\nu = 1.445$ GHz. The images on the top row were created by including the autocorrelations during calibration, while the middle row was created by excluding them. The images on the bottom row show the theoretical distilled ghost pattern (positions only) for this baseline. The left image displays where the proto-ghosts and the deuterio-ghosts of this baseline are located. The right image indicates from which baseline each ghost was “leaked”.

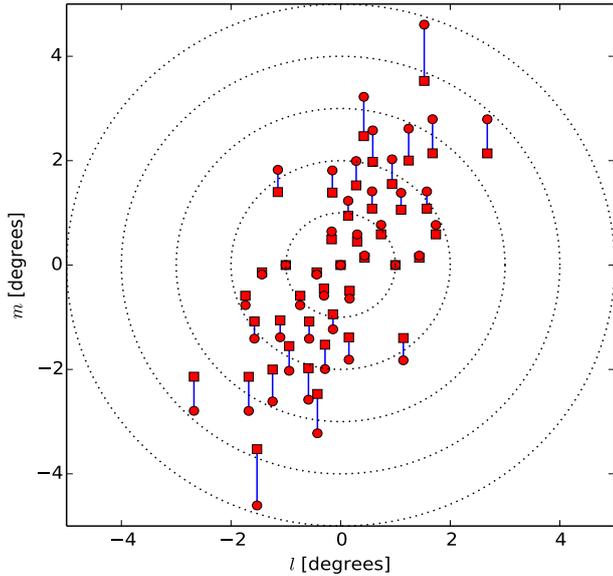


Figure 12. Theoretical ghost pattern (positions only) of baseline 45, with the unmodeled source at 1° and $\delta_0 \in [-90^\circ, -50^\circ]$. The squares represent the ghost positions at -90° , while the circles represent the ghost positions at -50° .

coherently during imaging. All the remaining proto-ghosts and deuterio-ghosts are only present on one baseline, since KAT-7 is non-redundant. This means that their amplitudes decrease with a factor $N(N-1)/2$. It should now also be clear to the reader why and how we were able to separate the ghosts into line and scattered ghosts in Sec. 7. Fig. 14 contains the theoretically calculated ghost positions of the full KAT-7 array (i.e. \hat{S}). The ghost positions have been divided into the taxonomy that we provided in Sec. 7. Fig. 14 corresponds with Fig. 13. Note that the brightest ghosts in Fig. 13 are all line ghosts, that the second brightest group in Fig. 13 are scattered proto-ghosts and that the dimmest (visible) ghosts in Fig. 13 are scattered deuterio-ghosts.

The line ghosts, in particular the primary suppressor, the secondary suppressor and the anti-ghost of the secondary suppressor, are the dominant ghosts for an irregular array like KAT-7, since these are the only ghosts that are associated with all baselines. This is a very important observation for further studies. In these studies, we will consider more complicated source models, for which it is neither manageable nor meaningful to try and find the origin of every individual ghost. What we want to understand is the level at which ghosts form and where we should expect the brightest ghosts to appear. For that purpose, it is sufficient to concentrate on the dominant ghost sources, which we can now clearly identify for irregular, i.e., non-redundant, array configurations based on the analysis in this paper.

9.2 Parameter dependencies

As we explained in Sec. 8.3 the main factor that influences the absolute flux of the ghosts is the flux of the unmodeled source A_2 . In this section we will be investigating the relationship between the relative absolute amplitude of the ghosts and A_2 . We first focus our attention on $\mathcal{G}^{\ominus-1} - \check{\mathbf{I}}$ and then move onto \mathcal{R}^Δ . We also look at \mathcal{R}^Δ as we are pri-

marily interested in the flux that the ghosts have after we have corrected our visibilities. Roughly speaking \mathcal{R}^Δ can be interpreted as the sky we obtain after we convolve the ghost pattern contained in $\mathcal{G}^{\ominus-1} - \check{\mathbf{I}}$ with the two source model in Eq. (5.1), which explains why we have been paying so much attention to the atomic ghost pattern contained in $\mathcal{G}^{\ominus-1} - \check{\mathbf{I}}$ (Grobler et al. 2014). The relative absolute amplitude of the most important ghosts contained in $\mathcal{G}^{\ominus-1} - \check{\mathbf{I}}$ are presented in Fig. 15, while Fig. 16 illustrates the relative absolute amplitude of the most important ghosts contained in \mathcal{R}^Δ . Both figures contain solid lines and dashed lines. The dashed lines represent theoretically derived values obtained with perturbation (essentially they are created with Eq. (5.15) and Eq. (5.16)). The solid lines are measured values and were obtained from extrapolated images (see Sec. 5.2). Monte Carlo analysis revealed that the flux values from which the solid lines were derived exhibited a percentage error (measurement error) of $\sim 6\%$ ³. Recall that extrapolation provides us with an exact solution, albeit empirical. The results obtained from extrapolation and perturbation match very well. This result shows us that perturbation does provide us with a good approximate analytic expression with which we can model the behavior of the brightest ghosts. In future analysis of ghost patterns of more complex source models, this will allow us to derive useful upper bounds for ghost fluxes in observations with future interferometers, whose increased sensitivity makes this problem more important.

Only one arbitrary proto-ghost and deuterio-ghost is plotted as Eq. (5.16) indicates that all proto-ghosts and deuterio-ghosts will have an absolute amplitude with the same order of magnitude. Fig. 16 reveals that the the source suppression factor (secondary suppressor) of KAT-7 is approximately 30%. The primary suppressor has an amplitude of approximately 14% that of A_2 . The anti-ghost is about 2% that of A_2 . The amplitude of all remaining ghosts are below 1% that of A_2 . More generally, Eqs. (5.15) and (5.16) enable us to estimate the relative flux of the most important ghosts (associated with $\mathcal{G}^{\ominus-1} - \check{\mathbf{I}}$) of a non-redundant irregular array of N antennas. The approximate fluxes (as percentage of A_2) of the most important ghosts are summarized in Table 1 for N equal to 7 (e.g., KAT-7), 14 (e.g., WSRT), 27 (e.g., VLA) and 100.

10 CONCLUSION AND FUTURE WORK

In this section, we summarise the main conclusions of this paper and discuss a number of issues that will be addressed in future endeavours.

10.1 Conclusions

In this paper, we significantly extended our analysis in Paper I, which allowed us to analyse the two-source test case for arbitrary array layouts. We created a new analytic tool by combining extrapolation (Sec. 5.2) with perturbation analysis (Sec. 5.3). This enabled us to answer the questions we posed in the introduction:

³ <http://www.mathsisfun.com/numbers/percentage-error.html>

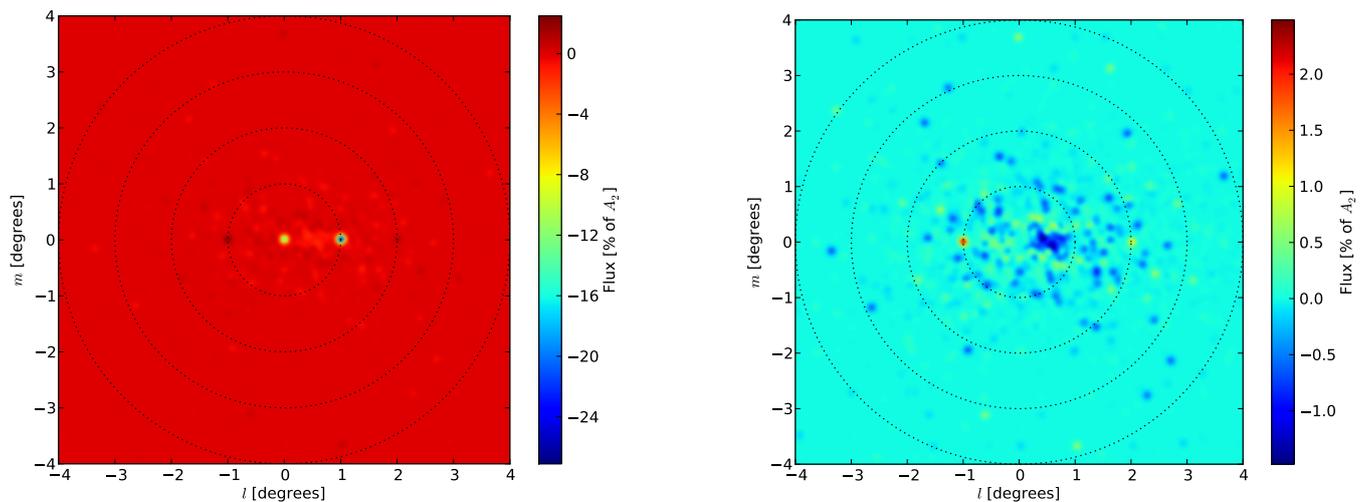


Figure 13. Theoretical distilled ghost pattern (left image) for the full KAT-7 array, with the unmodeled source being 0.2 Jy at 1° and $\delta_0 = -74.66^\circ$. The right image was obtained by eliminating the primary and secondary suppressor. The right image clearly shows that the anti-ghost of the secondary suppressor is the third brightest ghost.

Table 1. The flux of the primary suppressor, the secondary suppressor, the anti-ghost of the secondary suppressor, a proto-ghost and a deuterio-ghost for a non-redundant interferometer with 7, 14, 27 and 100 antennas as a percentage of A_2 . In this table N and N_b respectively denote the number of antennas and baselines in the array.

Ghost	Formula fraction of A_2	N			
		7	14	27	100
Primary	$\frac{1}{N}$	14%	7%	4%	1%
Secondary	$\frac{2}{N} - \frac{1}{N^2}$	27%	14%	7%	2%
Anti-ghost	$\frac{1}{N^2}$	2%	0.5%	0.1%	0.01%
Proto	$\frac{1}{N_b} (\frac{1}{N} - \frac{1}{N^2})$	0.6%	0.07%	0.01%	0.0002%
Deutero	$\frac{1}{N_b N^2}$	0.1%	0.006%	0.0004%	0.000002%

Taxonomy: The approximate analytic expressions we derived in Sec. 5 and Sec. 6 enabled us to develop a taxonomy classifying the ghosts into classes based on their fluxes. This taxonomy also explains why some ghosts are brighter than others. We refer the reader to Sec. 7 for more details.

Reciprocal independence: We also found a simple inverse relation between the flux of the ghosts of \mathcal{G} and $\mathcal{G}^{\circ-1}$, which does not affect their position, and implies that the brightest ghost positions are preserved under the Hadamard inverse. This is discussed in more detail in Sec. 5.3.

Parameter dependence: The expressions from Sec. 5 and Sec. 6 show the parameter space of the ghost flux. The absolute ghost flux is influenced by the number of antennas in the array (as illustrated by Table 1) and the amount of unmodeled flux A_2 (see Eq. (5.15) and Eq. (5.16)). The amount of real and imaginary flux of a ghost has a more complicated parameter space (see Eq. (6.4)). For an irregular array, ghosts do not form in straight lines, but in complicated geometric patterns that depend on the geometry of the array and the declination of the field center (see Eq. (6.5)).

Two peaks: The fluxes of the ghosts for a general irregular layout are complex, and not completely real as was the case for an east-west array. The only difference between the ghosts of baseline pq and qp are that the amplitudes of the

ghosts are each others complex conjugates. This implies that the ghosts manifest as a weighted sum of two PSFs during imaging. The real part manifests with an RPSF and the imaginary part manifests with an IPSF (see Sec. 6.4). Interestingly enough, we also found that the suppression ghost and its anti-ghost are always real. The weighted addition of the RPSF and IPSF during imaging explains the two peak structure observable in Fig. 7.

Symmetry: Ghosts form in symmetrically opposite pairs (see Fig. 9, Fig. 10 and Fig. 11).

The analysis in this paper includes the autocorrelations, which are ignored by most astronomical data reduction packages. In our simulations, we verified that this does not have a significant impact on the results with the primary suppressor as the major exception. In our analysis, we assumed that all time slots and frequency channels are calibrated independently. Sec. 6.4.3 briefly illustrates how to interpret the results when combining snapshots into synthesis observations. This analysis indicates that the impact of the unmodeled source does not average out over time, which is consistent with the empirical results presented in paper I. Eq. (6.4) shows that the ratio of real and imaginary flux of a ghost depends on the observing wavelength. This will produce a fringe-like pattern, which indicates that averag-

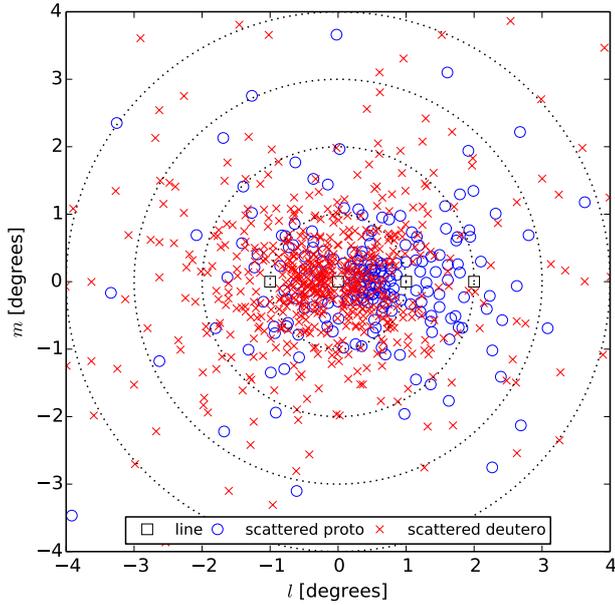


Figure 14. Theoretical ghost pattern (positions only) for the full KAT-7 array, with the unmodeled source at 1° and $\delta_0 = -74.66^\circ$. The line ghost positions are marked with black squares, the scattered proto-ghost positions are marked with blue circles and the scattered deuterio-ghost positions are marked with red crosses.

ing over frequency may provide a powerful tool to reduce the ghost flux. Unfortunately, the phasor described by (6.4) equals unity for the secondary suppressor and its anti-ghost, so a wider frequency span may not reduce the flux of these very bright ghosts.

An important conclusion from this work, is that a non-redundant array will create weaker scattered ghosts at more positions, while a redundant array produces brighter scattered ghosts at fewer positions. Since the latter will form on a grid, they are easy to identify, as illustrated in paper I, while the flux of the many weak scattered ghosts produced by a non-redundant array may make them negligible. What constitutes an optimal array lay-out thus depends on the application. In either case, as illustrated by Table 1, interferometers with a larger number of receiving elements provide better ghost suppression.

10.2 Future work

The mathematical tools we developed and the above results lay a good foundation from which we can tackle the more relevant problem of studying the formation of ghosts and source suppression in the case of more complicated sky models (which should also contain extended emission) and full synthesis observations. The fact that this study shows that the line ghosts are the dominant ghosts for interferometers with an irregular layout will greatly simplify the analysis of more complicated real-world scenarios. The following aspects will be addressed in future work:

General sky models: We will investigate whether the conclusions drawn for the two-source model also hold for more general sky models. Some progress in this regard

has already been made by Nunhokee (2015), who analysed source suppression in a 3C147 observation with the WSRT.

Solution intervals: Our analysis assumed that all snapshots in time and frequency are calibrated independently. In reality, the gain parameters are assumed to be constant over a solution interval that may span both time and frequency. Nunhokee (2015) found that for direction dependent calibration using larger calibration solution intervals can significantly reduce the amount of suppression of unmodeled sources. This indicates that the size of the solution intervals can have a significant effect and therefore deserves further study.

Time and frequency smearing: In many data reduction pipelines, the time and frequency resolution of the data is reduced after initial processing (RFI flagging, subtraction of bright sources from the sidelobes, etc.). This reduces the data volume at the expense of some decorrelation on the longest baselines. The latter results in a small reduction of the peak flux of point sources at some distance from the fringe stopping center. In the visibility domain, the decorrelation level is baseline dependent, while the analysis in this paper assumes the same amount of flux for the unmodeled sources on all baselines. Time and frequency smearing effects thus warrant further study to see how this dependence on baseline affects the ghost patterns.

Direction dependent calibration: In this paper (and in paper I), we only studied artefacts produced by direction independent calibration using incomplete sky models. We want to extend our analysis to direction dependent calibration. This topic was already touched upon by Nunhokee (2015) and deserves further attention in view of the fact that direction dependent calibration schemes are now commonly used in high dynamic range observations.

Phase-only calibration: Studying the ghost response of phase-only calibration is important as it is often used as the last step in transient calibration pipelines, in which incomplete sky models are a given (Stewart 2014).

Once we have quantified the flux level of the image artefacts and have a proper understanding of the mechanisms by which they form, we should be able to indicate whether such faint systematics will hamper the performance of increasingly sensitive future interferometers, in particular the SKA. If they do hamper performance, our knowledge of the underlying formation mechanism will help to develop strategies to mitigate their effect. One obvious question, for example, is whether self-calibration can reliably recover from an initial sky model that is biased by ghost artefacts and, if so, to what level the ghosts can be suppressed. This is particularly important for proposed cosmological observations that will attempt to infer cosmological parameters from the statistics of the background noise in an image.

ACKNOWLEDGEMENTS

This work is funded by IBM, ASTRON, the Dutch Ministry of Economic Affairs and the Province of Drenthe. It is also part of the SKA-TSM project and supported by The Northern Netherlands Provinces Alliance (SNN), Koers Noord and the Province of Drenthe, and the European Community FP7 programme MIDPREP, Grant Agreement PIRSES-GA-2013-612599.

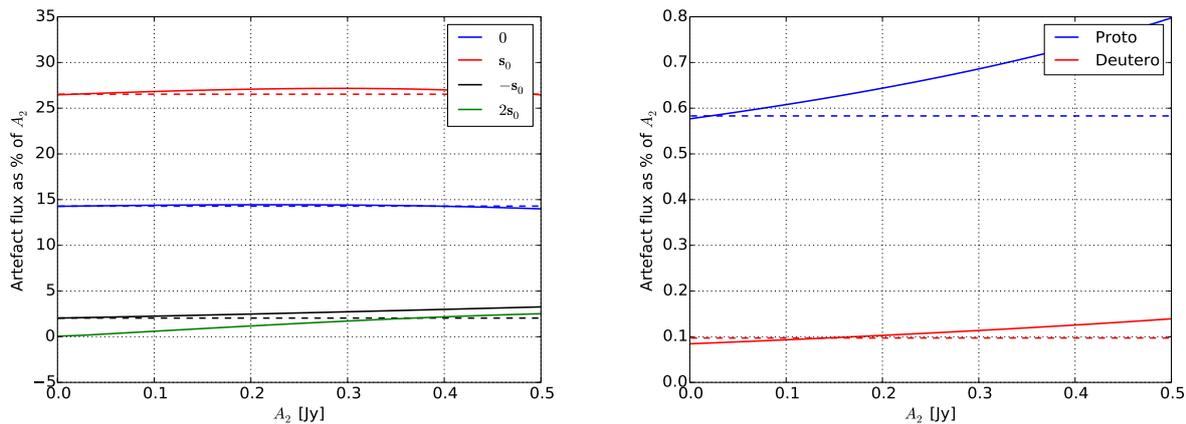


Figure 15. The relative amplitude $\hat{\zeta}_s^{\odot-1}/A_2$ of the primary suppressor, secondary suppressor and their anti-ghosts on the left and the relative amplitude of a random proto-ghost and deutero-ghost on the right. The dashed line represents the theoretically derived values that we obtained with perturbation, while the solid line represents a measured value which was obtained from a clean artefact map that was created via extrapolation.

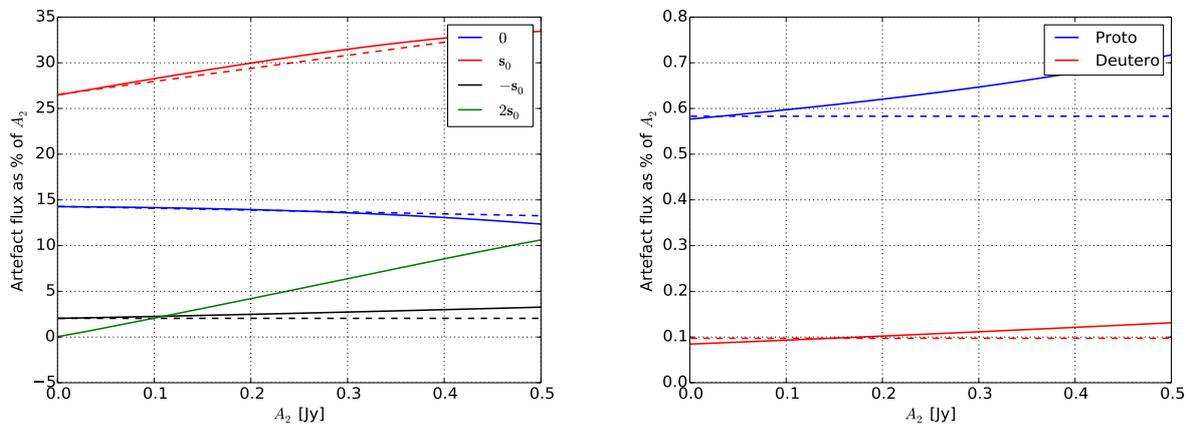


Figure 16. The relative amplitude $\hat{\zeta}_s^{\Delta}/A_2$ of the primary suppressor, secondary suppressor and their anti-ghosts on the left and the relative amplitude of a random proto-ghost and deutero-ghost on the right. The dashed line represents the theoretically derived values that we obtained with perturbation, while the solid line represents a measured value which was obtained from a clean artefact map that was created via extrapolation.

This work is based upon research supported by the South African Research Chairs Initiative of the Department of Science and Technology and National Research Foundation. The authors would like to thank Jonathan Kenyon, Julien Girard and Gianni Bernardi for their valuable suggestions. The authors would also like to thank the anonymous reviewer for his constructive feedback.

REFERENCES

- Asad K. M. B., et al., 2015, *MNRAS*, 451, 3709
 Bernardi G., et al., 2009, *A&A*, 500, 965
 Chandler C. J., Butler B. J., 2014, in Peck A. B., Benn C. R., Seaman R. L., eds, *Proc. of the SPIE Vol. 9149, Observatory Operations: Strategies, Processes, and Systems V*. SPIE, Washington
 Chang T. C., Pen U. L., Bandura K., Peterson J. B., 2010, *Nature*, 466, 463
 Cornwell T. J., Fomalont E. B., 1999, in Taylor G. B., Carilli C. L., Perley R. A., eds, *ASP Conf. Ser. Vol. 180, Synthesis Imaging in Radio Astronomy II*. Astron. Soc. Pac., San Francisco, p. 197
 Delhaize J., Meyer M. J., Staveley-Smith L., Boyle B. J., 2013, *MNRAS*, 433, 1398
 Dewdney P. E., Hall P. J., Schilizzi R. T., Lazio T. J. L. W., 2009, *Proc. of the IEEE*, 97, 1482
 Furlanetto S. R., Oh S. P., Briggs F. H., 2006, *Phys. Rep.*, 433, 181
 Grobler T. L., Nunhokee C. D., Smirnov O. M., van Zyl A. J., de Bruyn A. G., 2014, *MNRAS*, 439, 4030
 Grobler T. L., Smirnov O. M., Nunhokee C. D., 2014b, in *Proc. of the XXXIth URSI GASS, Calibration artefacts in KAT-7 data*. URSI, Beijing
 Jonas J. L., 2009, *Proc. of the IEEE*, 97, 1522
 Kazemi S., Yatawatta S., 2013, *MNRAS*, 435, 597
 Linfield R. P., 1986, *AJ*, 92, 213
 Martí-Vidal I., Guirado J. C., Jiménez-Monferrer S., Marcaide J. M., 2010, *A&A*, 517, A70
 Martí-Vidal I., Marcaide J. M., 2008, *A&A*, 480, 289
 Nunhokee C. D., 2015, Master's thesis, Rhodes Univ.
 Rau U., Bhatnagar S., Voronkov M. A., Cornwell T. J.,

2009, Proc. of the IEEE, 97, 1472

Smirnov O. M., 2010, Ghostbusters: The Unknown Unknowns Of Selfcal, presentation at CALIM2010 conference, Dwingeloo,

http://www.astron.nl/calim2010/presentations/14_Ghostbusters_Smirnov.pdf

Smirnov O. M., 2011, A&A, 527, A106

Smirnov O. M., de Bruyn A. G., 2011, in Proc. of the XXXth URSI GASS, MeqTrees and direction-dependent effects. URSI, Istanbul

Stewart A. J., 2014, LOFAR Image Plane Transient Candidate #3, presentation at LOFAR TKP meeting, Amsterdam,

<https://speakerdeck.com/transientskp/lofar-image-plane-transient-candidate-number-3>

Thompson A. R., Moran J. M., Swenson, Jr. G. W., 2001, Interferometry and Synthesis in Radio Astronomy, 2 edn. Wiley, New York

van der Veen A.-J., Wijnholds S. J., 2013, in Bhattacharyya S. S., Deprettere E. F., Leupers R., Takala J., eds, Signal Processing Tools for Radio Astronomy, Handbook of Signal Processing Systems. Springer, New York, p. 421

van Haarlem M. P., et al., 2013, A&A, 556, 1

Wijnholds S. J., van der Tol S., Nijboer R., van der Veen A.-J., 2010, IEEE Sig. Process. Mag., 27, 30

Wilkinson P. N., Conway J., Biretta J., 1988, in Reid M. J., Moran J. M., eds, Proc. IAU Symp. 129 Vol. 129, The impact of VLBI on Astrophysics and Geophysics. Springer, Netherlands, p. 509

Zwart J. T. L., et al., 2014, in Bourke T. L., et al. eds, Advancing Astrophysics with the Square Kilometre Array, Astronomy below the Survey Threshold. Proc. of Sci., Gardini Naxos

APPENDIX A: DERIVATION OF EQ. (3.5)

To solve Eq. (3.4) in closed form, it is convenient to write this equation as

$$\min_{\Delta \mathbf{g}} \left\| \begin{array}{l} \text{vec} \left(\mathbf{I} \mathbf{G}_t \left[\mathcal{R}_u \mathbf{G}_t^H \right] \right) - \text{vec} \left(\left[\mathbf{G}_t \mathcal{R}_k \right] \Delta \mathbf{G}^H \mathbf{I} \right) - \\ \text{vec} \left(\mathbf{I} \Delta \mathbf{G} \left[\mathcal{R}_k \mathbf{G}_t^H \right] \right) \end{array} \right\|, \quad (\text{A1})$$

where $\text{vec}(\cdot)$ denotes the vectorization operator. The Khatri-Rao product is the column-wise Kronecker product of two matrices and will be denoted by \bullet in this paper. It has a number of useful properties, including

$$\text{vec}(\mathcal{A} \text{diag}(\mathbf{b}) \mathcal{C}) = \left(\mathcal{C}^T \bullet \mathcal{A} \right) \mathbf{b} \quad (\text{A2})$$

where \mathcal{A} and \mathcal{C} are matrices of matching size and \mathbf{b} is a vector. Using this property of the Khatri-Rao product, we can rewrite Eq. (A1) to

$$\min_{\Delta \mathbf{g}} \left\| \begin{array}{l} \left(\left[\mathbf{G}_t^H \mathcal{R}_u^T \right] \bullet \mathbf{I} \right) \mathbf{g}_t - \left(\mathbf{I} \bullet \left[\mathbf{G}_t \mathcal{R}_k \right] \right) \overline{\Delta \mathbf{g}} - \\ \left(\left[\mathbf{G}_t^H \mathcal{R}_k^T \right] \bullet \mathbf{I} \right) \Delta \mathbf{g} \end{array} \right\|, \quad (\text{A3})$$

where $\overline{(\cdot)}$ denotes conjugation. We can also express Eq. (A3) as

$$\min_{\Delta \mathbf{g}} \left\| \begin{array}{l} \left(\left[\mathbf{G}_t^H \mathcal{R}_u^T \right] \bullet \mathbf{I} \right) \mathbf{g}_t - \\ \left(\left(\left[\mathbf{G}_t^H \mathcal{R}_k^T \right] \bullet \mathbf{I} \right), \left(\mathbf{I} \bullet \left[\mathbf{G}_t \mathcal{R}_k \right] \right) \right) \left[\Delta \mathbf{g}^T, \overline{\Delta \mathbf{g}^T} \right]^T \end{array} \right\|. \quad (\text{A4})$$

By using the Moore-Penrose inverse and Eq. (A4) we can solve for $\left[\Delta \mathbf{g}^T, \overline{\Delta \mathbf{g}^T} \right]^T$ to obtain

$$\begin{aligned} \left[\begin{array}{l} \Delta \mathbf{g} \\ \overline{\Delta \mathbf{g}} \end{array} \right] &\approx \left\{ \begin{array}{l} \left(\left(\left[\mathbf{G}_t^H \mathcal{R}_k \right] \bullet \mathbf{I} \right)^H \right) \left(\left(\left[\mathbf{G}_t^H \mathcal{R}_k \right] \bullet \mathbf{I} \right)^T \right)^{-1} \\ \left(\mathbf{I} \bullet \left[\mathbf{G}_t \mathcal{R}_k \right] \right)^H \left(\mathbf{I} \bullet \left[\mathbf{G}_t \mathcal{R}_k \right] \right)^T \end{array} \right\}^{-1} \\ &\times \left[\begin{array}{l} \left(\left[\mathbf{G}_t^H \mathcal{R}_k \right] \bullet \mathbf{I} \right)^H \left(\left[\mathbf{G}_t^H \mathcal{R}_u \right] \bullet \mathbf{I} \right) \mathbf{g}_t \\ \left(\mathbf{I} \bullet \left[\mathbf{G}_t \mathcal{R}_k \right] \right)^H \end{array} \right] \end{aligned} \quad (\text{A5})$$

This result can be simplified using the following property of the Khatri-Rao product:

$$\left(\mathcal{A} \bullet \mathcal{B} \right)^H \left(\mathcal{C} \bullet \mathcal{D} \right) = \mathcal{A}^H \mathcal{C} \odot \mathcal{B}^H \mathcal{D}. \quad (\text{A6})$$

Applying this property to Eq. (A5) gives Eq. (3.5).

APPENDIX B: LINEAR TRANSFORMATIONS TO REFERENCE BASELINE

B1 Regular east-west layout

When the array has a regular east-west configuration we use an elliptical locus instead of a circular locus to represent our imaginary reference track, since it simplifies the derivation of the ghost pattern. When we are dealing with a regular east-west array, $\phi_0 = \sqrt{\Delta x_0^2 + \Delta y_0^2}$, where Δx_0^e and Δy_0^e are antenna differences of the *common quotient baseline* \mathbf{b}_0^e in the X and Y direction, while $\hat{\delta}_0 = \delta_0$. We will denote the imaginary reference track of a regular east-west array with $\mathbf{b}_0^e(t)$. Moreover, $\forall q \geq p$ let

$$\phi_{pq} = \frac{\sqrt{\Delta x_{pq}^2 + \Delta y_{pq}^2}}{\phi_0}, \quad (\text{B1})$$

and

$$\phi_{qp} = -\phi_{pq}. \quad (\text{B2})$$

We are now able to define our first transformation

$$\bar{X}_{pq}(\mathbf{b}_0^e) = \frac{\phi_{pq}}{\lambda} \mathbf{b}_0^e. \quad (\text{B3})$$

With the elliptical imaginary reference track $\mathbf{b}_0^e(t)$ and Eq. (B3) we can derive all the other uv -tracks of a regular east-west array.

B2 General irregular layout

When we are dealing with a general irregular array layout, we take $\phi_0 = 1$ and $\hat{\delta}_0 = 90^\circ$. This turns the elliptical locus in Eq. (4.6) into a circular locus, centered at the origin, with a radius equal to one. The imaginary reference track associated with an irregular layout is denoted by $\mathbf{b}_0^o(t)$. Furthermore let, $\Delta \mathbf{z}_{pq} = (0, \Delta z_{pq})^T$ and $\Delta \mathbf{b}_{pq} = (0, \Delta b_{pq})^T$, where $\Delta b_{pq} = \Delta z_{pq} \cos \delta_0$.

For our analysis of a general irregular array layout, we need a linear transformation that can map a circular locus

centered at the origin to an elliptical locus centered at an arbitrary point. If we want to map $\mathbf{b}_0^\circ(t)$ onto an elliptical locus $\mathbf{b}_{pq}(t)$, we first need to scale it so that its semi-major axis and semi-minor axis become equal to $\frac{\phi_{pq}}{\lambda}$ and $\frac{\phi_{pq} \sin \delta_0}{\lambda}$ respectively and then translate it such that its center coincides with the point $(0, \frac{\Delta b_{pq}}{\lambda})$. We can achieve this with the following linear transformation:

$$\tilde{X}_{pq}(\mathbf{b}_0^\circ) = \frac{\phi_{pq}}{\lambda} \mathbf{D}(\delta_0) \mathbf{b}_0^\circ + \frac{\Delta \mathbf{b}_{pq}}{\lambda}, \quad (\text{B4})$$

where

$$\mathbf{D}(\delta_0) = \begin{bmatrix} 1 & 0 \\ 0 & \sin \delta_0 \end{bmatrix}. \quad (\text{B5})$$

To reverse this mapping we can use the inverse transformation $\tilde{X}_{pq}^{-1}(\mathbf{b}_{pq})$.

However, Eq. (B4) is not sufficient to reconstruct the uv -coverage of a general array. To understand why this is impossible, let us try to use $\tilde{X}_{pq}^{-1}(\mathbf{b}_{pq})$ and $\tilde{X}_{rs}^{-1}(\mathbf{b}_{rs})$ to map $\mathbf{b}_{pq}(t)$ and $\mathbf{b}_{rs}(t)$ to $\mathbf{b}_0^\circ(t)$. Although these transformations successfully map $\mathbf{b}_{pq}(t)$ and $\mathbf{b}_{rs}(t)$ onto circular loci, the start and end points of the circular loci may not coincide. The transformations $\tilde{X}_{pq}^{-1}(\mathbf{b}_{pq})$ and $\tilde{X}_{rs}^{-1}(\mathbf{b}_{rs})$ therefore fail to map $\mathbf{b}_{pq}(t)$ and $\mathbf{b}_{rs}(t)$ to $\mathbf{b}_0^\circ(t)$.

Let us denote the starting point of the circular locus associated with baseline pq with \mathbf{y}_{pq} , i.e. $\mathbf{y}_{pq} = \tilde{X}_{pq}^{-1}(\mathbf{b}_{pq}(t_0))$. We now need a rotation transformation that will enable us to derive the starting points of the circular loci associated with all baselines from a single starting point $\mathbf{b}_0^\circ(t_0)$. If we do not include this rotation, then we will not be able to construct all uv -tracks from a single imaginary reference track $\mathbf{b}_0^\circ(t)$ with a single locus starting point $\mathbf{b}_0^\circ(t_0)$. Without loss of generality, we use $(1, 0)$ as starting point in this paper.

Let $\mathbf{T}(\theta)$ denote an anti-clockwise rotation matrix, which by definition has the following form

$$\mathbf{T}(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}. \quad (\text{B6})$$

Moreover, let $\mathbf{b}_{pq}(t_0)$ denote the first uv -point on the uv -track $\mathbf{b}_{pq}(t)$ and $\mathbf{x}_{pq} = \phi_{pq}^{-1}(\Delta x_{pq}, \Delta y_{pq})^T$. We can now write

$$\mathbf{y}_{pq} = \frac{\lambda}{\phi_{pq}} \mathbf{D}^{-1}(\delta_0) \left(\mathbf{b}_{pq}(t_0) - \frac{\Delta \mathbf{b}_{pq}}{\lambda} \right) \quad (\text{B7})$$

$$= \mathbf{T}(H_0(t_0) - 90^\circ) \mathbf{x}_{pq}. \quad (\text{B8})$$

If θ_{pq} is defined as the anti-clockwise rotation angle between the vectors $\mathbf{b}_0^\circ(t_0) = (1, 0)^T$ and \mathbf{y}_{pq} then Eq. (B8) implies that

$$\theta_{pq} = \alpha_{pq} + (H_0(t_0) - 90^\circ), \quad (\text{B9})$$

where

$$\alpha_{pq} = \cos^{-1} \left(\frac{\Delta x_{pq}}{\phi_{pq}} \right) \text{sgn} \left(\frac{\Delta y_{pq}}{\phi_{pq}} \right), \quad (\text{B10})$$

and $\text{sgn}(\cdot)$ denotes the sign function.

With the aid of θ_{pq} we can now define the following transformation

$$\hat{X}_{pq}(\mathbf{b}_0^\circ) = \frac{\phi_{pq}}{\lambda} \mathbf{D}(\delta_0) \mathbf{T}(\theta_{pq}) \mathbf{b}_0^\circ + \frac{\Delta \mathbf{b}_{pq}}{\lambda} \quad (\text{B11})$$

$$= \frac{\phi_{pq}}{\lambda} \mathbf{D}(\delta_0) \mathbf{T}(\theta_{pq}) \mathbf{b}_0^\circ + \frac{\Delta \mathbf{z}_{pq} \cos \delta_0}{\lambda}, \quad (\text{B12})$$

which allows us to derive all the uv -tracks of a general layout

from one imaginary reference track $\mathbf{b}_0^\circ(t)$. The transformations $\hat{X}_{pq}(\mathbf{b}_0^\circ)$ and $\tilde{X}_{pq}^{-1}(\mathbf{b}_0^\circ)$ are depicted in Fig. 4.

The last linear transformation, consisting of a scaling and a rotation, is equal to

$$X_{pq}(\mathbf{b}_0^\circ) = \frac{\phi_{pq}}{\lambda} \mathbf{D}(\delta_0) \mathbf{T}(\theta_{pq}) \mathbf{b}_0^\circ. \quad (\text{B13})$$

As mentioned earlier, this transformation is very convenient as it will become apparent later, that the positions of the ghosts are not influenced by the translation transformation.

To conclude this section, we summarize the three basic linear operations required to map the imaginary reference uv -track onto an uv -track of an arbitrary array layout:

(i) Scaling: a scaling matrix can transform a circle into an ellipse. The semi-major axis of $\mathbf{b}_{pq}(t)$ depends mainly on the XY projected baseline length ϕ_{pq} and the observational wavelength λ . The eccentricity of $\mathbf{b}_{pq}(t)$ depends on the declination of the field-center δ_0 . Using $\frac{\phi_{pq}}{\lambda}$ in conjunction with $\mathbf{D}(\delta_0)$ ensures that the ellipse obtained from \hat{X}_{pq} and $\mathbf{b}_0^\circ(t)$ has the same dimensions as $\mathbf{b}_{pq}(t)$.

(ii) Translation: a translation is needed to change the center of the ellipse. The center point of $\mathbf{b}_{pq}(t)$ is determined by Δz_{pq} , δ_0 and λ . We use $\frac{\Delta \mathbf{b}_{pq}}{\lambda}$ to ensure that the center of the ellipse formed by \hat{X}_{pq} and $\mathbf{b}_0^\circ(t)$ coincides with $\mathbf{b}_{pq}(t)$.

(iii) Rotation: a rotation by $\mathbf{T}(\theta_{pq})$ ensures that we can use one circular locus to derive all uv -tracks of an array.

APPENDIX C: DERIVATION OF EQ. (5.10)

We want to find a closed form expression for the pseudo-inverse of Eq. (5.7). We will do this by finding the eigenvalue decomposition

$$\begin{bmatrix} \mathbf{N}\mathbf{I} & \mathbf{1}\mathbf{1}^T \\ \mathbf{1}\mathbf{1}^T & \mathbf{N}\mathbf{I} \end{bmatrix} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H \quad (\text{C1})$$

and using the fact that

$$(\mathbf{V}\mathbf{\Lambda}\mathbf{V}^H)^\dagger = \mathbf{V}\mathbf{\Lambda}^\dagger\mathbf{V}^H, \quad (\text{C2})$$

where $\mathbf{\Lambda}^\dagger$ is defined such that all non-zero eigenvalues are replaced by their reciprocals while the zero eigenvalues remain zero.

Since (5.7) can be written as

$$\begin{bmatrix} \mathbf{N}\mathbf{I} & \mathbf{1}\mathbf{1}^T \\ \mathbf{1}\mathbf{1}^T & \mathbf{N}\mathbf{I} \end{bmatrix} = \mathbf{N}\mathbf{I} + \begin{bmatrix} \mathbf{0}\mathbf{0}^T & \mathbf{1}\mathbf{1}^T \\ \mathbf{1}\mathbf{1}^T & \mathbf{0}\mathbf{0}^T \end{bmatrix} \quad (\text{C3})$$

we only need to find the eigenvectors of the second matrix on the right hand side, since a (scaled) identity matrix does only affect the eigenvalues, not the eigenvectors. It is easily found that

$$\mathbf{V} = \left[\frac{1}{\sqrt{2N}} \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix}, \frac{1}{\sqrt{2N}} \begin{bmatrix} -\mathbf{1} \\ \mathbf{1} \end{bmatrix}, \mathbf{V}_0 \right], \quad (\text{C4})$$

where the vectors with ones have length N and \mathbf{V}_0 provides an orthonormal basis for the null space of the two explicitly stated eigenvectors. Substitution of this result in (C3) shows that

$$\mathbf{\Lambda} = \text{diag}([2N, 0, N, \dots, N]). \quad (\text{C5})$$

We therefore have

$$\begin{aligned}\mathbf{\Lambda}^\dagger &= \text{diag} \left(\left[\frac{1}{2N}, 0, \frac{1}{N}, \dots, \frac{1}{N} \right] \right) \\ &= \frac{1}{N} \mathbf{I} - \frac{1}{N} \text{diag} \left(\left[\frac{1}{2}, 1, 0, \dots, 0 \right] \right).\end{aligned}\quad (\text{C6})$$

Hence

$$\begin{aligned}\mathbf{V} \mathbf{\Lambda}^\dagger \mathbf{V}^H &= \frac{1}{N} \mathbf{V} \left(\mathbf{I} - \text{diag} \left(\left[\frac{1}{2}, 1, 0, \dots, 0 \right] \right) \right) \mathbf{V}^H \\ &= \frac{1}{N} \mathbf{I} - \frac{1}{2N^2} \begin{bmatrix} \mathbf{1} & -\mathbf{1} \\ \mathbf{1} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{1} & -\mathbf{1} \\ \mathbf{1} & \mathbf{1} \end{bmatrix}^H \\ &= \frac{1}{N} \mathbf{I} - \frac{1}{4N^2} \begin{bmatrix} \mathbf{3}\mathbf{1}\mathbf{1}^H & -\mathbf{1}\mathbf{1}^H \\ -\mathbf{1}\mathbf{1}^H & \mathbf{3}\mathbf{1}\mathbf{1}^H \end{bmatrix},\end{aligned}\quad (\text{C7})$$

which completes our derivation of Eq. (5.10).

APPENDIX D: BASELINE PQ AND QP

In this section we study the relationship that exists between the ghosts found in baseline pq and qp which we will accomplish by taking a closer look at Eq. (6.3). If we undo the resampling function (replace \mathbf{b} with $\hat{X}_{pq}(\mathbf{b})$) in Eq. (6.3) we obtain

$$g_{pq}(\mathbf{b}) \approx \omega_{pq,0} + \sum_{r \neq s} \omega_{pq,rs} e^{-2\pi i [X_{rs}(\mathbf{b})]^T \cdot \mathbf{s}_0}, \quad (\text{D1})$$

where $\omega_{pq,0} = c_{pq,0}$, $\omega_{pq,rs} = c_{pq,rs} e^{-2\pi i \lambda^{-1} [\Delta \mathbf{b}_{rs}]^T \cdot \mathbf{s}_0}$. Since the visibility gain matrix is Hermitian, we also have

$$g_{qp}(\mathbf{b}) \approx \omega_{pq,0} + \sum_{r \neq s} \overline{\omega_{pq,rs}} e^{2\pi i [X_{rs}(\mathbf{b})]^T \cdot \mathbf{s}_0}. \quad (\text{D2})$$

The extrapolation technique only provides physically meaningful results for baseline pq if we replace \mathbf{b} with $\hat{X}_{pq}^{-1}(\mathbf{b})$. If we do this in Eq. (D1), we find

$$g_{pq}(\hat{X}_{pq}^{-1}(\mathbf{b})) \approx \omega_{pq,0} + \sum_{r \neq s} \omega_{pq,rs} y_{pq,rs} e^{-2\pi i [X_{rs} \circ X_{pq}^{-1}(\mathbf{b})]^T \cdot \mathbf{s}_0} \quad (\text{D3})$$

where

$$y_{pq,rs} = e^{2\pi i \left[\frac{\phi_{rs}}{\lambda \phi_{pq}} \mathbf{D}(\delta_0) \mathbf{T}(\theta_{rs} - \theta_{pq}) \mathbf{D}^{-1}(\delta_0) \Delta \mathbf{b}_{pq} \right]^T \cdot \mathbf{s}_0}. \quad (\text{D4})$$

When we want to apply extrapolation to Eq. (D2), we need to replace \mathbf{b} with $\hat{X}_{qp}^{-1}(\mathbf{b})$, which results in

$$g_{qp}(\hat{X}_{qp}^{-1}(\mathbf{b})) \approx \omega_{pq,0} + \sum_{r \neq s} \overline{\omega_{pq,rs}} y_{pq,rs} e^{-2\pi i [X_{rs} \circ X_{pq}^{-1}(\mathbf{b})]^T \cdot \mathbf{s}_0} \quad (\text{D5})$$

Since $\omega_{pq,0}$ is real valued, the physical significance of Eqs. (D3) and (D5) is, that if there is a spurious point source present in $g_{pq}(\hat{X}_{pq}^{-1}(\mathbf{b}))$ with position vector \mathbf{s} and flux $A_{\mathbf{s}}$, there will also be a spurious point source present in $g_{qp}(\hat{X}_{qp}^{-1}(\mathbf{b}))$ with the same position vector \mathbf{s} , but amplitude $A_{\mathbf{s}}$. In other words, the complex valued ghosts of baselines pq and qp are located at the same positions and are each others complex conjugates.