

# Calibration artefacts in radio interferometry.

## III. Phase-only calibration and primary beam correction

T. L. Grobler<sup>1,2\*</sup>, A. J. Stewart<sup>3</sup>, S. J. Wijnholds<sup>4</sup>, J. S. Kenyon<sup>1,2</sup>  
and O. M. Smirnov<sup>1,2</sup>

<sup>1</sup>*Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown, 6140, South Africa*

<sup>2</sup>*SKA South Africa, 3rd Floor, The Park, Park Road, Pinelands, 7405, South Africa*

<sup>3</sup>*Department of Astrophysics, University of Oxford, Oxford, OX1 3RH, UK*

<sup>4</sup>*ASTRON, P.O. Box 2, 7900 AA, Dwingeloo, The Netherlands*

in original form 2015 July 9

### ABSTRACT

This is the third installment in a series of papers in which we investigate calibration artefacts. Calibration artefacts (also known as ghosts or spurious sources) are created when we calibrate with an incomplete model. In the first two papers of this series we developed a mathematical framework which enabled us to study the ghosting mechanism itself. An interesting concomitant of the second paper was that ghosts appear in symmetrical pairs. This could possibly account for spurious symmetrization. Spurious symmetrization refers to the appearance of a spurious source (the anti-ghost) symmetrically opposite an unmodelled source around a modelled source. The analysis in the first two papers indicates that the anti-ghost is usually very faint, in particular when a large number of antennas are used. This suggests that spurious symmetrization will mainly occur at an almost undetectable flux level. In this paper, we show that phase-only calibration produces an anti-ghost that is  $N$ -times (where  $N$  denotes the number of antennas in the array) as bright as the one produced by phase and amplitude calibration and that this already bright ghost can be further amplified by the primary beam correction.

**Key words:** Instrumentation: interferometers, Methods: analytical, Methods: data analysis, Techniques: interferometric

### 1 INTRODUCTION

Interferometric data can be severely degraded by instrumental and environmental errors. These errors need to be removed by *calibration*. Direction independent calibration is normally achieved by finding antenna gains that minimize the difference between observed and modelled visibilities (Rau et al. 2009; Wijnholds et al. 2010; Smirnov 2011; Yatawatta et al. 2012; van der Veen & Wijnholds 2013; Salvini & Wijnholds 2014; Smirnov & Tasse 2015). However, when we calibrate with an incomplete sky model we inadvertently cause *calibration artefacts*, which appear as spurious positive or negative flux in the image. We refer to these spurious emission features as *ghost sources* or *ghosts*.

This is the third installment in a series of papers on these ghost phenomena. In the first paper (Grobler et al. 2014), which we will refer to as Paper I, we studied ghost formation by a regular east-west array using the WSRT as an example. This analysis was extended to arbitrary array

layouts in the second paper (Wijnholds et al. 2016), which we will refer to as Paper II. An interesting result of Paper II was that ghosts appear to form in symmetrical pairs. It could therefore explain the phenomenon of *spurious symmetrization*. In the context of this paper, spurious symmetrization refers to the appearance of spurious emission symmetrically opposite the unmodelled emission around modelled emission in a radio interferometric image (Cornwell & Fomalont 1999). Various cases of spurious symmetrization have been reported (Linfield 1986; Wilkinson et al. 1988; Cornwell & Fomalont 1999; Stewart 2014; Stewart et al. 2016; Wijnholds et al. 2016). Cornwell & Fomalont (1999) were one of the first to realize that a direct link exists between calibrating with an incomplete model and spurious symmetrization. They also observed that spurious symmetrization is particularly noticeable if an observation is made with a sufficiently small number of antennas; the implication being that the more antennas we use the less likely it is for spurious symmetrization to occur at a detectable flux level.

In Papers I and II, we considered a simple two-source test case with one modelled and one unmodelled source. We

\* E-mail: trienkog@gmail.com

found that a negative *suppression ghost* forms on top of the unmodelled source with a flux that is inversely proportional to the number of antennas in the array. In addition to the suppression ghost an anti-ghost also forms. The *anti-ghost* forms symmetrically opposite the unmodelled source around the modelled source. The anti-ghost is positive and has a flux that is inversely proportional to the square of the number of antennas in the array. This instance of spurious symmetrization is consistent with the assertion by Cornwell & Fomalont (1999) that the more antennas we use, the lower the flux level becomes at which spurious symmetrization will be detectable. Based on the analysis in Paper I and II, the brightness of the anti-ghost should also be independent of the location of the unmodelled source.

However, Stewart (2014) recently reported on an intriguing instance of spurious symmetrization in which the anti-ghost was much brighter than expected. Stewart was investigating whether a transient would be detected with LOFAR's Standard Imaging Pipeline (SIP) (Heald et al. 2015). He devised this test to help him make sense of a puzzling transient that was recently detected by the Transients Pipeline (TraP) (Stewart et al. 2016; Swinbank et al. 2015). For this test, he added a 60 Jy transient to an existing NCP (North Celestial Pole) dataset. He then determined whether one could detect the synthetic transient after running it through the SIP. The 60 Jy source was not included in the calibration model. He ran the experiment multiple times, varying the location of the transient. The NCP field to which the 60-Jy source was added is dominated by 3C 61.1. The outcome of his test was quite unexpected as he was able to produce a very bright anti-ghost when the transient was placed in some locations and no anti-ghost at all when placed at others. In some instances the anti-ghost was even brighter than the transient itself, which could lead to it being incorrectly identified as a source. Two snapshots of Stewart's experiment are presented in Fig. 1.

Stewart's result is surprising for two reasons:

- (i) According to Cornwell & Fomalont (1999) and Paper II, the anti-ghost should be extremely faint when we use an array like LOFAR due to its large number of elements.
- (ii) Paper II established that the brightness of the anti-ghost was independent of the location of the unmodelled source.

Obviously, we want to be able to explain the outcome of Stewart's experiment quantitatively, since understanding Stewart's experiment will deepen our understanding of ghost phenomena and spurious symmetrization in particular. There are two major differences between Stewart's experimental setup and the analysis in Paper II:

- (i) Phase-only calibration was used in Stewart's case while full-complex calibration was used in Paper II. Phase-only calibration can produce more stable solutions since we have fewer unknowns. Moreover, it is generally accepted that at the imaging stage the amplitude error will be small and constant. The phase-error, however, can still be large and varying. Furthermore, as phase-only calibration conserves phase closures it is also accepted to be quite safe<sup>1</sup>.

- (ii) The primary beam was included in Stewart's simulations, but was not taken into account in Paper II.

In this paper we will therefore study the impact of using phase-only calibration instead of full-complex calibration on ghost formation as well as the impact of primary beam correction. We will show that these two factors can indeed quantitatively explain the surprising instance of spurious symmetrization found in Stewart's experiment.

Studying the impact of phase-only calibration and primary beam correction further deepens our understanding of ghost phenomena. As mentioned earlier in this series, this is becoming crucial in light of the improved sensitivity (and, consequently, increased susceptibility to artefacts) of current and upcoming telescopes like the Low Frequency Array (LOFAR) (van Haarlem et al. 2013), the Karoo Array Telescope (MeerKAT) (Jonas 2009) and the Square Kilometre Array (SKA) (Dewdney et al. 2009). Another important motivation, as demonstrated by Stewart's experiment, is that transient pipelines run the risk of mistakingly identifying the anti-ghost as a transient source. An improved understanding of ghost phenomena may help to avoid such errors.

In the next two sections, we formulate the full-complex as well as the phase-only calibration problem and briefly review some of the ghost analysis techniques we developed in the first two papers. In Sec. 4 we present a theoretical analysis of ghost formation with phase-only calibration. We compare the results with the results for full-complex calibration described in Paper II in Sec. 5. The effect of primary beam correction is investigated in Sec. 6 before we conclude our paper with a summary of our findings. The images in this paper use the cubehelix color scheme (Green 2011).

## 2 CALIBRATION

### 2.1 Full-Complex Calibration

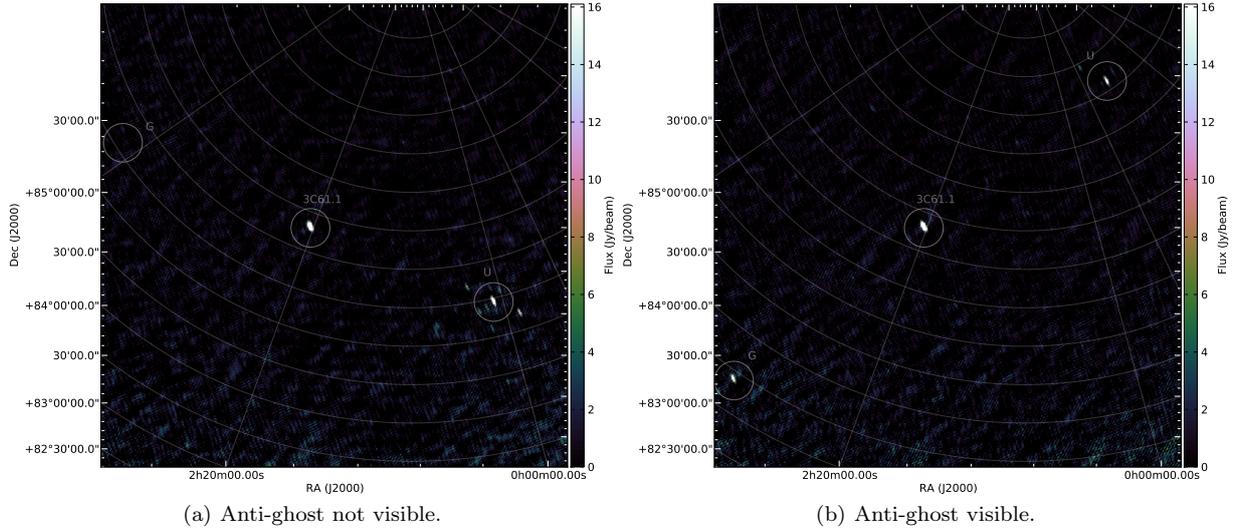
In its most basic form, calibration boils down to minimizing the difference between observed and predicted visibilities by estimating the complex instrumental gain response. Mathematically speaking, unpolarized calibration is realized by minimizing the following optimization problem

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

where:

- (i) The operator  $(\cdot)^H$  denotes the Hermitian transpose. Moreover,  $(\cdot)^T$  denotes the transpose of its operand,  $\odot$  denotes the Hadamard product and  $\|\cdot\|$  denotes the appropriate matrix or vector norm.
- (ii)  $\mathbf{R}$  is the *observed visibility matrix*. Each entry of  $\mathbf{R}$ , which we denote by  $r_{pq}$ , represents the visibility that was observed by the baseline formed by antennas  $p$  and  $q$ .
- (iii)  $\mathbf{M}$  is the *model visibility matrix* and we denote an entry of  $\mathbf{M}$  with  $m_{pq}$ . The entries of  $\mathbf{M}$  contain model visibilities, i.e. synthetic visibilities that were created from the calibration sky model.
- (iv)  $\mathbf{G}$  is the *antenna gain matrix* and is equal to  $\text{diag}(\mathbf{g})$ , with  $\mathbf{g} = [g_1, g_2, \dots, g_N]^T$  being the complex instrumental responses of the antennas. The operator  $\text{diag}(\cdot)$  creates a matrix by placing the input vector on the main diagonal of an otherwise zero matrix. We can decompose the gain of the

<sup>1</sup> <http://veraserver.mtk.nao.ac.jp/VERA/kurayama/WinterSchool/difmap6.htm>



**Figure 1.** Two snapshots created during Stewart’s experiment. Stewart added a transient to an existing LOFAR dataset which he then ran through the SIP. He found that when he placed the unmodelled transient in some positions (left image) no anti-ghost formed, but when he placed it at other locations a bright anti-ghost appeared (right image).

$p$ th element into its amplitude  $\gamma_p$  and a phasor  $\rho_p = e^{i\theta_p}$ . The amplitudes can be stacked in a vector  $\boldsymbol{\gamma} = [\gamma_1, \dots, \gamma_N]^T$  and the phasors in a vector  $\boldsymbol{\rho} = [\rho_1, \dots, \rho_N]^T$ , such that  $\boldsymbol{g} = \boldsymbol{\gamma} \odot \boldsymbol{\rho}$ . We will denote the number of antennas in our array by  $N$  and the number of baselines by  $B$ . The number of baselines can be calculated from the number of antennas with  $B = (N^2 - N) / 2$ .

(v)  $\boldsymbol{G} = \boldsymbol{g}\boldsymbol{g}^H$  is known as the *visibility gain matrix*. An entry of  $\boldsymbol{G}$  is denoted by  $g_{pq}$ . We can compute  $g_{pq}$  with

$$g_{pq} = g_p \bar{g}_q, \quad (2.2)$$

where  $\bar{(\cdot)}$  denotes complex conjugation. The Hadamard inverse of  $\boldsymbol{G}$  is denoted with  $\boldsymbol{G}^{\odot -1}$  and is known as the *visibility calibration matrix*. We use the visibility calibration matrix to correct our visibilities once we have successfully estimated  $\boldsymbol{g}$ .

## 2.2 Phase-only Calibration

When we perform full-complex calibration we assume that the phase and amplitude of our visibilities were corrupted by our antenna gains. In the case of phase-only calibration we assume that only the phases of our visibilities are affected by the antenna gains. The phase-only equivalent of Eq. (2.1) is equal to

$$\min_{\boldsymbol{P}} \left\| \boldsymbol{R} - \boldsymbol{P}\boldsymbol{M}\boldsymbol{P}^H \right\| = \min_{\boldsymbol{P}} \left\| \boldsymbol{R} - \boldsymbol{P} \odot \boldsymbol{M} \right\|, \quad (2.3)$$

where

- $\boldsymbol{P} = \text{diag}(\boldsymbol{\rho})$  is the *antenna phase matrix*.
- $\boldsymbol{P} = \boldsymbol{\rho}\boldsymbol{\rho}^H$  denotes the *visibility phase matrix*. An entry of  $\boldsymbol{P}$  is denoted by  $\rho_{pq}$ .

Once we have estimated  $\boldsymbol{\rho}$  we can correct our visibilities, i.e. remove the errors caused by the antenna gains, with

$$\boldsymbol{R}_{\boldsymbol{P}}^{(c)} = \boldsymbol{P}^{\odot -1} \odot \boldsymbol{R}, \quad (2.4)$$

where

- (i)  $\boldsymbol{R}_{\boldsymbol{P}}^{(c)}$  is the *corrected visibility matrix*.
- (ii)  $\boldsymbol{P}^{\odot -1}$  represents the *visibility phase calibration matrix*. The phase calibration matrix is computed by taking the Hadamard inverse of  $\boldsymbol{P}$ .

## 3 PRELIMINARIES

In this section, we present the remaining building blocks needed for our analysis. In Sec. 3.1, we present the experimental assumptions we adhere to in this paper. In Sec. 3.2, we define a couple of useful linear transformations. The *extrapolation* algorithm is presented in Sec. 3.3. Extrapolation employs the linear transformations in Sec. 3.2 to transform  $\boldsymbol{R}$  and  $\boldsymbol{M}$  into function-valued matrices. This enables us to use these newly formed matrices and standard calibration to create per-baseline artefact maps (a map containing only the calibration systematics). Moreover, these artefact maps contain no sidelobes as extrapolation operates on continuous functions (see Algorithm 1).

### 3.1 Two-source scenario

Unless explicitly stated otherwise, all results presented in this paper are subject to the following assumptions:

- (i) We are dealing with a monochromatic observation.
- (ii) No error was added to our visibilities by the interferometer, i.e.  $\boldsymbol{g}_t = \mathbf{1}$ . We denote an all one vector of size  $N$  with  $\mathbf{1}$ . The subscript  $t$  of  $\boldsymbol{g}_t$  is used to indicate that we are referring to the true antenna gain vector which is corrupting the observed visibilities instead of the gain vector estimate which is obtained when we calibrate.
- (iii) We assume the  $w$ -term is negligible, i.e. we are observing in a narrow field-of-view.
- (iv) We do not consider the effect of noise as O.M. Smirnov has shown that including noise does not alter the positions at which the ghosts form (Smirnov 2010).

With regards to the true sky and the calibration sky model we assume the following:

(i) Our true sky consists of two sources, one in the field center with flux  $A_1$  and another off-center with flux  $A_2$ . We denote the positional vector of the off-center source by  $\mathbf{s}_0 = (l_0, m_0)$  (a direction-cosine vector). Stated differently, if we were to observe these two sources then antennas  $p$  and  $q$  would observe

$$r_{pq} = A_1 + A_2 e^{-2\pi i \mathbf{b}_{pq}^T \cdot \mathbf{s}_0}, \quad (3.1)$$

where  $\mathbf{b}_{pq} = (u_{pq}, v_{pq})^T$  denotes the  $uv$ -coordinate at which  $r_{pq}$  was measured. We use the symbol “ $\cdot$ ” to denote the standard dot product which operates on equal length row vectors. It is important to realize that the coordinates  $u_{pq}$  and  $v_{pq}$  in Eq. (3.1) are already expressed in wavelengths and not in m (we have already divided by the observational wavelength  $\lambda$ ). In Paper II, we showed that all the  $uv$ -tracks of an array can be generated from a single circular reference track  $\mathbf{b}_0^\circ(t)$ . The circular reference track  $\mathbf{b}_0^\circ(t)$  is depicted in Fig. 2.

(ii) We only include the center source in our calibration model, i.e.

$$m_{pq} = A_1. \quad (3.2)$$

(iii) Since Eq. (2.3) is invariant to amplitude scaling and positional shifts, we may assume that our modelled source is in the center and that  $A_1 = 1$  without any loss of generality,

(iv) Unless otherwise stated  $A_2 \ll A_1$ , i.e. that our calibration model is reasonably accurate.

### 3.2 Linear transformations

All the  $uv$ -tracks of an interferometer can be mapped onto a single circular reference track  $\mathbf{b}_0^\circ(t)$  by using three linear transformations, namely

(i) a *scaling* described by

$$\phi_{pq} \mathbf{D}(\delta_0) = \phi_{pq} \begin{bmatrix} 1 & 0 \\ 0 & \sin(\delta_0) \end{bmatrix}, \quad (3.3)$$

(ii) a *rotation* described by

$$\mathbf{T}(\theta_{pq}) = \begin{bmatrix} \cos(\theta_{pq}) & -\sin(\theta_{pq}) \\ \sin(\theta_{pq}) & \cos(\theta_{pq}) \end{bmatrix}, \quad (3.4)$$

(iii) and a *translation*  $\Delta \mathbf{b}_{pq}$ .

A composition of two or more linear transformations results in another linear transformation. In this paper, we will use the following compositions:

(i) a composition of scaling, rotation and translation

$$\hat{X}_{pq}(\mathbf{b}_0^\circ) = \frac{\phi_{pq}}{\lambda} \mathbf{D}(\delta_0) \mathbf{T}(\theta_{pq}) \mathbf{b}_0^\circ + \frac{\Delta \mathbf{b}_{pq}}{\lambda} \quad (3.5)$$

(ii) a composition of scaling and translation

$$\tilde{X}_{pq}(\mathbf{b}_0^\circ) = \frac{\phi_{pq}}{\lambda} \mathbf{D}(\delta_0) \mathbf{b}_0^\circ + \frac{\Delta \mathbf{b}_{pq}}{\lambda} \quad (3.6)$$

(iii) a composition of scaling and rotation

$$X_{pq}(\mathbf{b}_0^\circ) = \frac{\phi_{pq}}{\lambda} \mathbf{D}(\delta_0) \mathbf{T}(\theta_{pq}) \mathbf{b}_0^\circ \quad (3.7)$$

In the expressions above,  $\delta_0$  is the declination of the field center. The quantities  $\phi_{pq}$ ,  $\theta_{pq}$  and  $\Delta \mathbf{b}_{pq} = (0, \Delta b_{pq})^T$  are determined by the array geometry and are formally defined in Fig. 2. The linear transformations  $\hat{X}_{pq}(\mathbf{b}_0^\circ)$  and  $\tilde{X}_{pq}^{-1}(\mathbf{b}_{pq})$  are also depicted in Fig. 2.

We denote the functional inverse of  $\hat{X}_{pq}$ ,  $\tilde{X}_{pq}$  and  $X_{pq}$  with  $\hat{X}_{pq}^{-1}$ ,  $\tilde{X}_{pq}^{-1}$  and  $X_{pq}^{-1}$  respectively.

Within this paper we use each of the aforementioned linear transformations for a specific purpose:

(i)  $\hat{X}_{pq}$ : The circular reference track  $\mathbf{b}_0^\circ(t)$  is depicted in the left most coordinate system of Fig. 2. We can generate any  $uv$ -track of an array by applying  $\hat{X}_{pq}$  to  $\mathbf{b}_0^\circ(t)$ . This fact is illustrated by the mapping of  $\mathbf{b}_0^\circ(t)$  in the left most coordinate system to  $\mathbf{b}_{pq}(t)$  in the middle coordinate system of Fig. 2. The linear transformation  $\hat{X}_{pq}(\mathbf{b}_0^\circ)$  is exploited by the extrapolation procedure (presented in Sec. 3.3) to create clean per-baseline artefact maps.

(ii)  $\tilde{X}_{pq}$ : The effect of applying  $\tilde{X}_{pq}^{-1}$  to  $\mathbf{b}_{pq}(t)$  is depicted by the mapping from  $\mathbf{b}_{pq}(t)$  in the middle coordinate system of Fig. 2 to the  $u'v'$ -track in the right most coordinate system of Fig. 2. Since there is no rotation transformation present in  $\tilde{X}_{pq}$ , the  $u'v'$ -track in the right most coordinate system of Fig. 2 does not start at coordinate (1,0). An additional derotation by  $\theta_{pq}$  is required to achieve this, i.e.  $\tilde{X}_{pq}^{-1}(\mathbf{b}_{pq})$  allows us to calculate the rotation angle  $\theta_{pq}$ ; which is a quantity that is required by  $\hat{X}_{pq}$  to function properly.

(iii)  $X_{pq}$ : It turns out that the translation  $\Delta \mathbf{b}_{pq}$  does not affect the positions of the ghosts, only their fluxes as discussed in Paper II. Since  $X_{pq}$  does not contain a translation transformation, it can be used to determine the positions at which the ghosts arise (see Eq. 4.27).

### 3.3 Extrapolation

Extrapolation is a technique, which exploits the fact that all  $uv$ -tracks of an interferometer can be derived from a single imaginary reference track (via three linear transformations) to create clean per-baseline artefact maps (an artefact map untarnished by  $uv$ -sampling).

Extrapolation requires the following three steps:

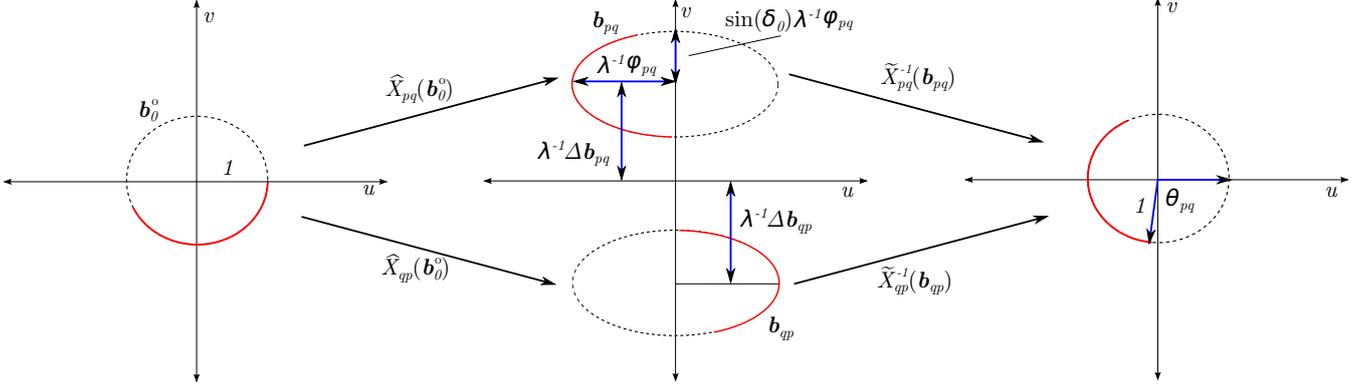
(i) Reformulate visibilities as functionals. Rewrite the visibility matrix  $\mathbf{Y}$  as a function of the imaginary reference track  $\mathbf{b}_0^\circ$ , i.e.  $\mathbf{Y}(\mathbf{b}_0^\circ)$ . The  $\mathbf{Y}$  matrix is a proxy matrix and can refer to an observed visibility, model visibility, predicted visibility, visibility gain, visibility calibration, corrected visibility matrix or any one of the distilled visibility matrices. Distillation is discussed in Sec. 4.4.

(ii) Derive the intermediate extrapolated visibility matrix. Replace  $\mathbf{b}_0^\circ$  with the free parameter  $\mathbf{b} = (u, v)^T$  to obtain  $\mathbf{Y}(\mathbf{b})$ .

(iii) Calculate the extrapolated visibility matrix. It turns out that the entries of  $\mathbf{Y}(\mathbf{b})$  have no physical meaning if left as is. To make the entries meaningful we substitute  $\mathbf{b}$  with  $\hat{X}_{pq}^{-1}(\mathbf{b})$  to obtain  $\mathbf{Y}(\hat{X}_{pq}^{-1}(\mathbf{b}))$ .

If we now image the  $pq$ th entry of  $\mathbf{Y}(\hat{X}_{pq}^{-1}(\mathbf{b}))$  then the resulting image will be completely devoid of sidelobes as the  $pq$ th entry of  $\mathbf{Y}(\hat{X}_{pq}^{-1}(\mathbf{b}))$  is defined over the entire  $uv$ -plane. Extrapolation is a per-baseline operation.

In Algorithm 1 we use the extrapolation procedure to specifically construct an empirical estimate of  $g_{pq}(\hat{X}_{pq}^{-1}(\mathbf{b}))$ .



**Figure 2.** The linear transformations  $\hat{X}_{pq}(\mathbf{b}_0^o)$  and  $\tilde{X}_{pq}^{-1}(\mathbf{b}_{pq})$ .

In this algorithm  $\mathcal{K}$  denotes the set containing all index values of sources that were “observed”, while  $\mathcal{H}$  denotes the set containing the index values of the sources that were included in the calibration model, i.e.,  $\mathcal{H} \subseteq \mathcal{K}$ . The extrapolation procedure can therefore be applied to more complicated sky model scenarios (and is not limited to the two-source test case presented in Sec. 3.1). The artefact map associated with  $g_{pq}(\hat{X}_{pq}^{-1}(\mathbf{b})) - 1$  (we subtract one to distill out the calibration artefacts) is obtained by taking the inverse Fast Fourier Transform (FFT) of the empirical estimate of  $g_{pq}(\hat{X}_{pq}^{-1}(\mathbf{b})) - 1$ , which is now trivial as the empirical estimate of  $g_{pq}(\hat{X}_{pq}^{-1}(\mathbf{b})) - 1$  is defined on a regular grid. This results in an artefact map without any sidelobes. Moreover, in Algorithm 1, “o” denotes operator composition. Also note that any calibration algorithm may be used to realise Algorithm 1.

A more detailed explanation of extrapolation can be found in Paper II.

---

**Algorithm 1** Extrapolation:  $g_{pq}(\hat{X}_{pq}^{-1}(\mathbf{b}))$

---

```

1:  $\mathbf{u} \leftarrow \text{linspace}(u_{\min}, u_{\max}, u_{\dim})$ 
2:  $\mathbf{v} \leftarrow \text{linspace}(v_{\min}, v_{\max}, v_{\dim})$ 
3:  $g_{pq} \leftarrow \text{zeros}((u_{\dim}, v_{\dim}), \text{dtype} = \text{complex})$     ▷ define
   your 2D extrapolated regular  $uv$ -grid
4:  $i \leftarrow 1$ 
5: for  $i \leq u_{\dim}$  do
6:    $j \leftarrow 1$ 
7:   for  $j \leq v_{\dim}$  do
8:      $\mathbf{b} \leftarrow (u_i, v_j)^T$ 
9:      $\mathcal{R} \leftarrow \left\{ \sum_{k \in \mathcal{K}} A_k e^{-2\pi i [\hat{X}_{rs} \circ \hat{X}_{pq}^{-1}(\mathbf{b})]^T \cdot \mathbf{s}_k} \right\}_{rs}$ 
10:     $\mathcal{M} \leftarrow \left\{ \sum_{h \in \mathcal{H}} A_h e^{-2\pi i [\hat{X}_{rs} \circ \hat{X}_{pq}^{-1}(\mathbf{b})]^T \cdot \mathbf{s}_h} \right\}_{rs}$ 
11:     $\mathcal{G} \leftarrow \text{calibrate}(\mathcal{R}, \mathcal{M})$ 
12:     $g_{pq}^{ij} \leftarrow [\mathcal{G}]_{pq}$ 
13:     $j \leftarrow j + 1$ 
14:   $i \leftarrow i + 1$ 

```

---

## 4 PHASE-ONLY GHOST PATTERN

In this section, we derive the phase-only ghost pattern that is associated with the two-source test case presented in Sec. 3.1. We do so by following the same approach we took

for full-complex calibration in Paper II. We therefore start by developing a general perturbation analysis framework for phase only calibration. We then use this general framework to analyze the two-source test case in Sec. 4.2. We then apply extrapolation in Sec. 4.3 to the result from the perturbation analysis. For the sake of completeness we also briefly touch on the topics of distillation and corrected visibilities in Sec. 4.4 and Sec. 4.5. A brief overview of imaging and how it pertains to calibration artefacts is presented in Sec. 4.6. In the last section we extend the basic per-baseline taxonomy from Paper II. This extended taxonomy makes it easier to compare the ghost patterns that are produced by full-complex and phase-only calibration respectively.

### 4.1 Perturbation Analysis

Perturbation analysis is based on the assumption that we are using a fairly accurate calibration sky model, implying that the antenna gains are only slightly perturbed due to the incomplete sky model. In the perturbation approach that we present here we estimate the amount of perturbation the gains experience, by first linearizing the estimation problem, making it easier to solve. This linearization is only possible if the perturbation we are trying to estimate is relatively small.

The observed visibility matrix is equal to

$$\mathcal{R} = \mathbf{G}_t(\mathcal{R}_k + \mathcal{R}_u)\mathbf{G}_t^H = \mathbf{g}_t\mathbf{g}_t^H \odot (\mathcal{R}_k + \mathcal{R}_u), \quad (4.1)$$

where  $\mathcal{R}_k$  contains the visibilities of the sources that we know exist, i.e.  $\mathcal{R}_k = \mathcal{M}$ ,  $\mathcal{R}_u$  contains the visibilities of the sources that we are unaware of,  $\mathbf{g}_t$  is the true antenna gain vector and  $\mathbf{G}_t = \text{diag}(\mathbf{g}_t)$ . We can also decompose  $\mathbf{g}_t$  into an amplitude vector  $\boldsymbol{\gamma}_t = [\gamma_1^t, \dots, \gamma_N^t]^T$  and phasor vector  $\boldsymbol{\rho}_t = [e^{i\varphi_1^t}, \dots, e^{i\varphi_N^t}]^T$ , such that  $\mathbf{g}_t = \boldsymbol{\gamma}_t \odot \boldsymbol{\rho}_t$ . Eq. (4.1) allows us to reformulate Eq. (2.3) as

$$\underset{\boldsymbol{\rho}}{\text{argmin}} \left\| \mathbf{g}_t\mathbf{g}_t^H \odot (\mathcal{R}_k + \mathcal{R}_u) - \boldsymbol{\rho}\boldsymbol{\rho}^H \odot \mathcal{R}_k \right\|. \quad (4.2)$$

If we ignore the true amplitude variations  $\boldsymbol{\gamma}_t$  or we neglect the mismatch between the model visibilities (described by  $\mathcal{R}_k$ ) and the observed visibilities (described by  $\mathcal{R}_k + \mathcal{R}_u$ ), then we will induce a phase bias  $\Delta\boldsymbol{\varrho}$  into  $\boldsymbol{\rho}$ , i.e.

$$\boldsymbol{\rho} = e^{i(\boldsymbol{\varrho}_t + \Delta\boldsymbol{\varrho})} \quad (4.3)$$

Without loss of generality, we may assume that  $\boldsymbol{\rho}_t = \mathbf{0}$ . Assuming that we are justified in using phase-only calibration and that we have a reasonably accurate sky model, the biases will be small. We may therefore use the approximation  $e^{i\Delta\boldsymbol{\rho}} \approx 1 + i\Delta\boldsymbol{\rho}$ . It then follows that

$$\mathcal{P} \approx (\mathbf{1} + i\Delta\boldsymbol{\rho})(\mathbf{1} + i\Delta\boldsymbol{\rho})^H. \quad (4.4)$$

Eq. (4.4) allows us to rewrite Eq. (4.2) to

$$\underset{\Delta\boldsymbol{\rho}}{\operatorname{argmin}} \left\| \mathbf{g}_t \mathbf{g}_t^H \odot (\mathbf{R}_k + \mathbf{R}_u) - (\mathbf{1} + i\Delta\boldsymbol{\rho})(\mathbf{1} + i\Delta\boldsymbol{\rho})^H \odot \mathbf{R}_k \right\| \quad (4.5)$$

Since  $\Delta\boldsymbol{\rho}$  is small, we may ignore the product  $\Delta\boldsymbol{\rho}\Delta\boldsymbol{\rho}^H$  and write our minimization problem as

$$\begin{aligned} \underset{\Delta\boldsymbol{\rho}}{\operatorname{argmin}} & \left\| \mathbf{g}_t \mathbf{g}_t^H \odot \mathbf{R}_k + \mathbf{g}_t \mathbf{g}_t^H \odot \mathbf{R}_u - \mathbf{1}\mathbf{1}^H \odot \mathbf{R}_k + i\mathbf{1}\Delta\boldsymbol{\rho}^H \odot \mathbf{R}_k - i\Delta\boldsymbol{\rho}\mathbf{1}^H \odot \mathbf{R}_k \right\| \\ = & \underset{\Delta\boldsymbol{\rho}}{\operatorname{argmin}} \left\| \operatorname{vec}(\mathbf{I}\mathbf{G}_t\mathbf{R}_k\mathbf{G}_t^H) - \operatorname{vec}(\mathbf{I}\mathbf{I}\mathbf{R}_k\mathbf{I}^H) + \operatorname{vec}(\mathbf{I}\mathbf{G}_t\mathbf{R}_u\mathbf{G}_t^H) + i\operatorname{vec}(\mathbf{R}_k\operatorname{diag}(\Delta\boldsymbol{\rho})^H\mathbf{I}) - i\operatorname{vec}(\mathbf{I}\operatorname{diag}(\Delta\boldsymbol{\rho})\mathbf{R}_k) \right\| \end{aligned} \quad (4.6)$$

In Eq. (4.6), we made use of the fact that  $\mathbf{x}\mathbf{x}^H \odot \mathbf{X} = \operatorname{diag}(\mathbf{x})\mathbf{X}\operatorname{diag}(\mathbf{x})^H$ , where  $\mathbf{x}$  is a vector and  $\mathbf{X}$  is matrix of compatible size. In the next step, we will use

$$\operatorname{vec}(\mathcal{A}\operatorname{diag}(\mathbf{b})\mathcal{C}) = (\mathcal{C}^T \bullet \mathcal{A})\mathbf{b}, \quad (4.7)$$

where  $\mathcal{A}$  and  $\mathcal{C}$  are matrices of compatible size and  $\bullet$  denotes the Khatri-Rao or column-wise Kronecker product of two matrices. We swapped around and added identity matrices to Eq. 4.6, denoted by  $\mathbf{I}$ , as visual cues to enable the reader to follow how one can apply Eq. (4.7) to Eq. (4.6) to obtain Eq. (4.8). We thus find

$$\begin{aligned} \underset{\Delta\boldsymbol{\rho}}{\operatorname{argmin}} & \left\| (\mathbf{G}_t^H\mathbf{R}_k^T \bullet \mathbf{I})\mathbf{g}_t - (\mathbf{I}\mathbf{R}_k^T \bullet \mathbf{I})\mathbf{1} + (\mathbf{G}_t^H\mathbf{R}_u^T \bullet \mathbf{I})\mathbf{g}_t + i(\mathbf{I} \bullet \mathbf{R}_k)\Delta\boldsymbol{\rho} - i(\mathbf{R}_k^T \bullet \mathbf{I})\Delta\boldsymbol{\rho} \right\| \\ = & \underset{\Delta\boldsymbol{\rho}}{\operatorname{argmin}} \left\| (\overline{\mathbf{G}_t\mathbf{R}_k} \bullet \mathbf{I})\mathbf{g}_t - (\overline{\mathbf{I}\mathbf{R}_k} \bullet \mathbf{I})\mathbf{1} + (\overline{\mathbf{G}_t\mathbf{R}_u} \bullet \mathbf{I})\mathbf{g}_t - i((\overline{\mathbf{R}_k} \bullet \mathbf{I}) - (\mathbf{I} \bullet \mathbf{R}_k))\Delta\boldsymbol{\rho} \right\| \end{aligned} \quad (4.8)$$

Note that the first mentioned  $\Delta\boldsymbol{\rho}$  in the first equation of Eq. 4.8 is not conjugated, although strictly speaking it should be. We do not conjugate  $\Delta\boldsymbol{\rho}$  in Eq. 4.8, since  $\Delta\boldsymbol{\rho}$  is per definition a real valued vector.

If we apply the normal equation method to Eq. (4.8) we readily obtain

$$\Delta\boldsymbol{\rho} \approx -i \left[ \mathbf{A}^H \mathbf{A} \right]^{-1} \mathbf{A}^H \times \{\mathbf{B}\mathbf{g}_t - \mathbf{C}\mathbf{1} + \mathbf{D}\mathbf{g}_t\}, \quad (4.9)$$

where

$$\begin{aligned} \mathbf{A} &= (\overline{\mathbf{R}_k} \bullet \mathbf{I}) - (\mathbf{I} \bullet \mathbf{R}_k), \\ \mathbf{B} &= \overline{\mathbf{G}_t\mathbf{R}_k} \bullet \mathbf{I}, \\ \mathbf{C} &= \overline{\mathbf{R}_k} \bullet \mathbf{I}, \\ \mathbf{D} &= \overline{\mathbf{G}_t\mathbf{R}_u} \bullet \mathbf{I}. \end{aligned} \quad (4.10)$$

There is an inherent ambiguity present in Eq. (4.9), as we

did not explicitly select a reference antenna. However, this ambiguity has no effect on the outcome of our analysis as it drops out when we calculate  $\mathcal{P}$ .

## 4.2 Two-source Perturbation Results

In this section we apply the general result derived in the previous section to our two-source test case. It follows from Eq. (3.1) and Eq. (3.2) (if  $A_1 = 1$ ) that

$$\mathcal{R}_k = \mathbf{1}\mathbf{1}^H, \quad (4.11)$$

$$\mathcal{R}_u = A_2 \mathbf{a}\mathbf{a}^H, \quad (4.12)$$

where  $\mathbf{a} = [e^{-2\pi i \mathbf{b}_1^T \cdot \mathbf{s}_0}, \dots, e^{-2\pi i \mathbf{b}_N^T \cdot \mathbf{s}_0}]^T$  and  $\mathbf{b}_n$  is the  $(x, y, z)$ -position of the  $n$ -th antenna in units of  $\lambda$ . It also follows from Sec. 3.1 that  $\mathbf{g}_t = \mathbf{1}$ . Substitution of these results in Eq. (4.10) gives

$$\begin{aligned} \mathbf{A} &= (\mathbf{1}\mathbf{1}^H \bullet \mathbf{I}) - (\mathbf{I} \bullet \mathbf{1}\mathbf{1}^H) \\ \mathbf{B} &= \mathbf{1}\mathbf{1}^H \bullet \mathbf{I} \\ \mathbf{C} &= \mathbf{1}\mathbf{1}^H \bullet \mathbf{I} = \mathbf{B} \\ \mathbf{D} &= A_2 \overline{(\mathbf{a}\mathbf{a}^H)} \bullet \mathbf{I} \end{aligned} \quad (4.13)$$

Note that the dimensions of the matrices in Eq. (4.13) are  $N^2 \times N$ .

To find an expression for our two source model using Eq. (4.9), we need to find the inverse of

$$\mathbf{A}^H \mathbf{A} = 2(N\mathbf{I} - \mathbf{1}\mathbf{1}^H). \quad (4.14)$$

Since this matrix is rank deficient, we will resort to the pseudo-inverse, denoted by  $(\cdot)^\dagger$  to emulate the behaviour of most least squares fitting routines. We find that

$$[\mathbf{A}^H \mathbf{A}]^\dagger = \frac{1}{2N} \mathbf{I} - \frac{1}{2N^2} \mathbf{1}\mathbf{1}^H, \quad (4.15)$$

for which a proof can be found in Appendix A.

If we substitute Eqs. (4.13) and (4.15) in Eq. (4.9), we obtain

$$\Delta\boldsymbol{\rho} \approx \frac{A_2}{N} \mathcal{I} \left\{ (\mathbf{a}^H \mathbf{1}) \mathbf{a} \right\}, \quad (4.16)$$

where  $\mathcal{I}\{\cdot\}$  denotes the imaginary operator. Substitution of this result into Eq. (4.4) gives

$$\begin{aligned} \mathcal{P} &\approx \mathbf{1}\mathbf{1}^H - i\mathbf{1}\Delta\boldsymbol{\rho}^H + i\Delta\boldsymbol{\rho}\mathbf{1}^H \\ &= \mathbf{1}\mathbf{1}^H - c \mathcal{I} \left\{ (\mathbf{a}^H \mathbf{1}) \mathbf{1}\mathbf{a}^T \right\} + c \mathcal{I} \left\{ (\mathbf{a}^H \mathbf{1}) \mathbf{a}\mathbf{1}^H \right\}, \end{aligned} \quad (4.17)$$

where  $c = A_2 i / N$ . Concentrating on a single element of the visibility phase matrix, we find

$$[\mathcal{P}]_{pq} = \rho_{pq} \approx 1 + i(\Delta\varrho_p - \Delta\varrho_q). \quad (4.18)$$

Simplifying Eq. (4.18) results in

$$\rho_{pq} \approx 1 - \sum_{s \neq p} c \sin(2\pi \mathbf{b}_{ps}^T \cdot \mathbf{s}_0) - \sum_{r \neq q} c \sin(2\pi \mathbf{b}_{rq}^T \cdot \mathbf{s}_0), \quad (4.19)$$

where  $\mathbf{b}_{ps} = \mathbf{b}_p - \mathbf{b}_s$  and  $\mathbf{b}_{rq} = \mathbf{b}_r - \mathbf{b}_q$  (also see Sec. 3.1).

The fact that a sine can be written in terms of two complex exponentials allows us to rewrite Eq. (4.19) to

$$\rho_{pq} \approx \sigma_{pq,0} + \sum_{rs \in \mathcal{I}} \sigma_{pq,rs} e^{-2\pi i \mathbf{b}_{rs}^T \cdot \mathbf{s}_0}, \quad (4.20)$$

where

$$\sigma_{pq,0} = 1, \quad (4.21)$$

and

$$\sigma_{pq,rs} = \begin{cases} \frac{A_2}{N} & \text{if } rs \in \mathcal{I}_1 \\ -\frac{A_2}{N} & \text{if } rs \in \mathcal{I}'_1 \\ \frac{A_2}{2N} & \text{if } rs \in \mathcal{I}_2 \\ -\frac{A_2}{2N} & \text{if } rs \in \mathcal{I}'_2 \\ \frac{A_2}{2N} & \text{if } rs \in \mathcal{I}_3 \\ -\frac{A_2}{2N} & \text{if } rs \in \mathcal{I}'_3 \\ 0 & \text{otherwise} \end{cases}, \quad (4.22)$$

with

$$\begin{aligned} \mathcal{I}_1 &= \{rs | (r=p) \wedge (s=q) \wedge (r \neq s)\} \\ \mathcal{I}'_1 &= \{rs | (r=q) \wedge (s=p) \wedge (r \neq s)\} \\ \mathcal{I}_2 &= \{rs | (r \neq p) \wedge (s=q) \wedge (r \neq s)\} \\ \mathcal{I}'_2 &= \{rs | (r \neq q) \wedge (s=p) \wedge (r \neq s)\} \\ \mathcal{I}_3 &= \{rs | (r=p) \wedge (s \neq q) \wedge (r \neq s)\} \\ \mathcal{I}'_3 &= \{rs | (r=q) \wedge (s \neq p) \wedge (r \neq s)\}, \end{aligned} \quad (4.23)$$

where “ $\wedge$ ” denotes the logical and. Moreover, we define

$$\mathcal{I} = \bigcup_{k=1}^3 \mathcal{I}_k \cup \mathcal{I}'_k. \quad (4.24)$$

### 4.3 Extrapolation

We will use extrapolation to show that each term in Eq. (4.20) produces a ghost artefact in the image. We can express  $\rho_{pq}$  as a function of the reference baseline  $\mathbf{b}_0^\circ$ , i.e.

$$\rho_{pq}(\mathbf{b}_0^\circ) = \sigma_{pq,0} + \sum_{rs \in \mathcal{I}} \sigma_{pq,rs} e^{-2\pi i [\hat{X}_{rs}(\mathbf{b}_0^\circ)]^T \cdot \mathbf{s}_0}. \quad (4.25)$$

To extrapolate, we replace  $\mathbf{b}_0^\circ$  with  $\mathbf{b} = (u, v)^T$  to obtain

$$\rho_{pq}(\mathbf{b}) \approx \sigma_{pq,0} + \sum_{rs \in \mathcal{I}} \sigma_{pq,rs} e^{-2\pi i [\hat{X}_{rs}(\mathbf{b})]^T \cdot \mathbf{s}_0}. \quad (4.26)$$

As discussed in Section 3.3, we can create physically meaningful extrapolated visibilities by replacing  $\mathbf{b}$  in Eq. (4.20) with  $\hat{X}_{pq}^{-1}(\mathbf{b})$  to obtain

$$\begin{aligned} \rho_{pq}(\hat{X}_{pq}^{-1}(\mathbf{b})) &\approx \sigma_{pq,0} + \sum_{rs \in \mathcal{I}} \sigma_{pq,rs} e^{-2\pi i [\hat{X}_{rs} \circ \hat{X}_{pq}^{-1}(\mathbf{b})]^T \cdot \mathbf{s}_0}, \\ &= \tau_{pq,0} + \sum_{rs \in \mathcal{I}} \tau_{pq,rs} e^{-2\pi i [X_{rs} \circ X_{pq}^{-1}(\mathbf{b})]^T \cdot \mathbf{s}_0}, \\ &= \tau_{pq,0} + \sum_{rs \in \mathcal{I}} \tau_{pq,rs} e^{-2\pi i \mathbf{b}^T \cdot \mathbf{s}_{pq}^{rs}} \end{aligned} \quad (4.27)$$

where  $\tau_{pq,0} = \sigma_{pq,0}$ ,  $\tau_{pq,rs} = \sigma_{pq,rs} \beta_{pq,rs}$  and

$$\beta_{pq,rs} = e^{2\pi i \left[ \frac{\phi_{rs}}{\lambda \phi_{pq}} \mathcal{D}(\delta_0) \mathbf{T}(\theta_{rs} - \theta_{pq}) \mathcal{D}^{-1}(\delta_0) \Delta \mathbf{b}_{pq} - \frac{\Delta \mathbf{b}_{rs}}{\lambda} \right]^T \cdot \mathbf{s}_0} \quad (4.28)$$

and

$$\mathbf{s}_{pq}^{rs} = \frac{\phi_{rs}}{\phi_{pq}} \mathbf{s}_0 \mathbf{Z}, \quad (4.29)$$

with

$$\mathbf{Z} = \begin{bmatrix} \cos(\theta_{rs} - \theta_{pq}) & -\frac{\sin(\theta_{rs} - \theta_{pq})}{\sin(\delta_0)} \\ \sin(\delta_0) \sin(\theta_{rs} - \theta_{pq}) & \cos(\theta_{rs} - \theta_{pq}) \end{bmatrix}. \quad (4.30)$$

Since the Fourier inverse of a complex exponential is a delta function, each term in the sum of Eq. (4.27) can be interpreted as a spurious point source, i.e. as a calibration artefact or ghost. The amplitude of the spurious point source is equal to  $\tau_{pq,rs} \in \mathbf{C}$  and its position vector is equal to  $\mathbf{s}_{pq}^{rs}$ . Furthermore, the amplitudes of the ghosts exhibit the following symmetry

$$\tau_{pq,rs} = -\overline{\tau_{pq,rs}}. \quad (4.31)$$

In Eq. (4.24), we defined a set of baseline indices to uniquely identify all baselines that produce a ghost in the ghost pattern for baseline  $pq$ . Since each term in Eq. (4.27) is associated with a specific ghost, we can map our set of baseline indices to a set of ghost locations for baseline  $pq$ , i.e., we can define the mapping  $M_{pq} : \mathbf{P}(\mathcal{I}) \rightarrow \mathbf{P}(\mathcal{Q}_{pq})$  as

$$M_{pq}(\mathcal{I}) = \bigcup_{rs \in \mathcal{I}} \{\mathbf{s}_{pq}^{rs}\}, \quad (4.32)$$

where  $\mathcal{I} = \{rs\}_{r \neq s}$ ,  $\mathcal{Q}_{pq} = \{\mathbf{s}_{pq}^{rs}\}_{r \neq s}$  and  $\mathbf{P}(\cdot)$  denotes the power set of its operand (which has to be a set). We can use Eq. (4.32) to group all per-baseline phase-only ghost positions that are predicted by our perturbation analysis into one set

$$\mathcal{T}_{pq} = M_{pq}(\mathcal{I}). \quad (4.33)$$

We are actually interested in studying the ghost pattern that is generated by the instrument as a whole. It is therefore more appropriate to merge the per-baseline perturbation results into one set  $\mathcal{T}$ . We can construct  $\mathcal{T}$  as follows

$$\mathcal{T} = \bigcup_{pq} \mathcal{T}_{pq}. \quad (4.34)$$

Note that  $\mathcal{T}_{pq}$  and  $\mathcal{T}$  contain no recurring elements<sup>2</sup>.

Let  $\mathbf{0} = (0, 0)$ . Eq. (4.34) allows us to reformulate Eq. (4.27) as

$$\rho_{pq}(\hat{X}_{pq}^{-1}(\mathbf{b})) \approx \sum_{\tilde{\mathbf{s}} \in \hat{\mathcal{T}}} \tau_{pq,\tilde{\mathbf{s}}} e^{-2\pi i \mathbf{b}^T \cdot \tilde{\mathbf{s}}}, \quad (4.35)$$

where  $\hat{\mathcal{T}} = \mathcal{T} \cup \{\mathbf{0}\}$  and

$$\tau_{pq,\tilde{\mathbf{s}}} = \begin{cases} \sum_{rs \in \mathcal{I}_{pq,\tilde{\mathbf{s}}}} \tau_{pq,rs} & \text{if } \mathcal{I}_{pq,\tilde{\mathbf{s}}} \neq \emptyset \\ \tau_{pq,0} & \text{if } \tilde{\mathbf{s}} = \mathbf{0} \\ 0 & \text{otherwise} \end{cases}. \quad (4.36)$$

In Eq. (4.36),  $\mathcal{I}_{pq,\tilde{\mathbf{s}}} = \{rs | (\mathbf{s}_{pq}^{rs} \in \mathcal{T}_{pq}) \wedge (\tilde{\mathbf{s}} = \mathbf{s}_{pq}^{rs}) \wedge (\tilde{\mathbf{s}} \neq \mathbf{0})\}$ .

### 4.4 Distillation

In our analysis, we assume that our model is reasonably accurate, such that the deviations from the true gain values will be small. As a result, the ghosts will be very faint. We therefore have to distill  $\mathcal{P}$  and  $\mathcal{P}^{\circ-1}$  to make the ghosts visible. Distillation effectively removes all real emission from the image, such that we can image only the deviations from the true gain values (see Papers I and II). We can distill  $\mathcal{P}$  by subtracting  $\check{\mathbf{I}} = \mathbf{1}^H$  from it, i.e.

$$[\mathcal{P}(\hat{X}_{pq}^{-1}(\mathbf{b})) - \check{\mathbf{I}}]_{pq} \approx \sum_{\tilde{\mathbf{s}} \in \mathcal{T}} \tau_{pq,\tilde{\mathbf{s}}} e^{-2\pi i \mathbf{b}^T \cdot \tilde{\mathbf{s}}}. \quad (4.37)$$

<sup>2</sup> Recall that  $\{1, 2, 3\} \cup \{1, 2, 4\} = \{1, 2, 3, 4\}$

The visibility phase calibration matrix can be described by

$$\mathcal{P}^{\circ-1} \approx \mathbf{1}\mathbf{1}^H + i\mathbf{1}\Delta\boldsymbol{\rho}^H - i\Delta\boldsymbol{\rho}\mathbf{1}^H, \quad (4.38)$$

since  $(\rho_{pq})^{-1} = \bar{\rho}_{pq}$ . As was the case in Paper II, Eq. (4.38) implies that the brightest ghost positions are preserved under the Hadamard inverse. Although the positions are preserved, the fluxes of the ghosts change sign. Comparing Eqs. (4.17) and (4.38), we can easily see that

$$[\mathcal{P}^{\circ-1}(\hat{X}_{pq}^{-1}(\mathbf{b})) - \check{\mathbf{1}}]_{pq} \approx \sum_{\tilde{\mathbf{s}} \in \mathcal{T}} \tau_{pq, \tilde{\mathbf{s}}}^{\circ-1} e^{-2\pi i \mathbf{b}^T \cdot \tilde{\mathbf{s}}}. \quad (4.39)$$

where  $\tau_{pq, \tilde{\mathbf{s}}}^{\circ-1} = -\tau_{pq, \tilde{\mathbf{s}}}$ . Note that the index set  $\mathcal{T}$  is used in Eq. (4.37) and Eq. (4.39), while the index set  $\hat{\mathcal{T}}$  was used in Eq. (4.35). The reason being, according to the perturbation approach presented in Sec. 4.2, there is no ghost at  $\mathbf{0}$ .

#### 4.5 Corrected Visibilities

The corrected visibility matrix  $\mathcal{R}_{\mathcal{P}}^{(c)}$  was defined in Eq. (2.4). As in the case of  $\mathcal{P}$ , distilling  $\mathcal{R}_{\mathcal{P}}^{(c)}$  allows us to study the fainter spurious emission in  $\mathcal{R}_{\mathcal{P}}^{(c)}$ . The distilled corrected visibility matrix  $\mathcal{R}_{\mathcal{P}}^{\Delta}$  (also known as the residual visibility matrix) can be calculated with

$$\mathcal{R}_{\mathcal{P}}^{\Delta} = \mathcal{R}_{\mathcal{P}}^{(c)} - \mathcal{R} = (\mathcal{P}^{\circ-1} - \check{\mathbf{1}}) \odot \mathcal{R}, \quad (4.40)$$

implying that

$$[\mathcal{R}_{\mathcal{P}}^{\Delta}(\hat{X}_{pq}^{-1}(\mathbf{b}))]_{pq} \approx \sum_{\tilde{\mathbf{s}} \in \mathcal{T}} \tau_{pq, \tilde{\mathbf{s}}}^{\circ-1} e^{-2\pi i \mathbf{b}^T \cdot \tilde{\mathbf{s}}} (1 + A_2 e^{-2\pi i \mathbf{b}^T \cdot \mathbf{s}_0}). \quad (4.41)$$

Inspecting Eq. (4.41) reveals that when we correct we actually convolve the atomic ghost pattern in  $\mathcal{P}^{\circ-1} - \check{\mathbf{1}}$  with the modelled and unmodelled source. Our corrected residuals are therefore made up of two ghost patterns, one around the modelled source and one around the unmodelled source, that were added together coherently (Papers I and II).

#### 4.6 Imaging

Imaging and how it relates to calibration systematics was discussed at length in Paper II. In this section we will only summarize the most important quantities we obtained in Paper II. We denote the dirty images associated with  $\mathcal{P}^{\circ-1} - \check{\mathbf{1}}$  and  $\mathcal{R}_{\mathcal{P}}^{\Delta}$  with  $I_D^{\mathcal{P}^{\circ-1} - \check{\mathbf{1}}}$  and  $I_D^{\mathcal{R}_{\mathcal{P}}^{\Delta}}$  respectively. As we showed in Sec. 4.3 each baseline has its own unique ghost pattern. When we image, we effectively average all the individual per-baseline ghost patterns together. If a ghost happens to form consistently in each baseline its per-baseline flux will add up coherently. On the other hand, if a ghost is only present in one baseline averaging the different baselines together will result in a dimmer ghost. In mathematical terms, the aforementioned result translates into the following: if we assume natural weighting then we can calculate the flux of the ghost at  $\tilde{\mathbf{s}}$  in  $I_D^{\mathcal{P}^{\circ-1} - \check{\mathbf{1}}}$  with

$$\xi_{\tilde{\mathbf{s}}}^{\circ-1} = \langle \Re \{ \tau_{pq, \tilde{\mathbf{s}}}^{\circ-1} \} \rangle_{p < q}, \quad (4.42)$$

where  $\langle \cdot \rangle_{p < q}$  denotes averaging over all the baselines and  $\Re\{\cdot\}$  denotes the real operator. In Paper II we showed that

only the real flux of a ghost manifests at  $\tilde{\mathbf{s}}$ , which explains why we only average the real flux together in Eq. (4.42). It is actually more sensible to calculate the maximum possible flux that the ghost at position  $\tilde{\mathbf{s}}$  can contribute to the image, since some of the imaginary flux of a ghost manifests in the vicinity of  $\tilde{\mathbf{s}}$  due to the Ghost Spread Function (GSF) (Papers I and II). We therefore also define

$$\hat{\xi}_{\tilde{\mathbf{s}}}^{\circ-1} = \langle |\tau_{pq, \tilde{\mathbf{s}}}^{\circ-1}| \rangle_{p < q}. \quad (4.43)$$

Eq. (4.43) represents the maximum flux the ghost at  $\tilde{\mathbf{s}}$  could have contributed to the image. It contributes this maximum amount if it is completely real, and can therefore be interpreted as a type of upper bound. Eq. (4.41) also allows us to define

$$\xi_{\tilde{\mathbf{s}}}^{\Delta} = \xi_{\tilde{\mathbf{s}}}^{\circ-1} + A_2 \xi_{\tilde{\mathbf{s}} - \mathbf{s}_0}^{\circ-1} \quad (4.44)$$

and

$$\hat{\xi}_{\tilde{\mathbf{s}}}^{\Delta} = \hat{\xi}_{\tilde{\mathbf{s}}}^{\circ-1} + A_2 \hat{\xi}_{\tilde{\mathbf{s}} - \mathbf{s}_0}^{\circ-1}, \quad (4.45)$$

which are the  $I_D^{\mathcal{R}_{\mathcal{P}}^{\Delta}}$  equivalent of Eq. (4.42) and Eq. (4.43).

#### 4.7 Augmented Per-baseline Ghost Taxonomy

Based on Paper II, the result equivalent of Eq. (4.20) for full-complex calibration is

$$[\mathcal{G}]_{pq} = g_{pq} = c_{pq,0} + \sum_{rs \in \mathcal{J}} c_{pq,rs} e^{-2\pi i \mathbf{b}_{rs}^T \cdot \mathbf{s}_0}, \quad (4.46)$$

where

$$c_{pq,0} = 1 + \frac{A_2}{N}, \quad (4.47)$$

and

$$c_{pq,rs} = \begin{cases} \frac{2A_2}{N} - \frac{A_2}{N^2} & \text{if } rs \in \mathcal{I}_1 \\ \frac{A_2}{N} - \frac{A_2}{N^2} & \text{if } rs \in \mathcal{I}_2 \\ \frac{A_2}{N} - \frac{A_2}{N^2} & \text{if } rs \in \mathcal{I}_3, \\ -\frac{A_2}{N^2} & \text{if } rs \in \mathcal{I}_4 \\ 0 & \text{otherwise} \end{cases}, \quad (4.48)$$

The index sets  $\mathcal{I}_1$ ,  $\mathcal{I}_2$  and  $\mathcal{I}_3$  are already defined in Eq. (4.23). The fourth index set is defined as

$$\mathcal{I}_4 = \{rs | (r \neq p) \wedge (s \neq q) \wedge (r \neq s)\}. \quad (4.49)$$

The index set  $\mathcal{J}$  was defined in Sec. 4.3.

We can make the following observations if we inspect Eq. (4.20) and Eq. (4.46):

(i) Some of the coefficients ( $\sigma_{pq,rs}$  and  $c_{pq,rs}$ ) in Eq. (4.20) and Eq. (4.46) are inversely proportional to  $N$  while others are inversely proportional to  $N^2$ . The ghosts that are associated with the coefficients that are inversely proportional to  $N$  are called proto-ghosts, and those associated with the coefficients that are inversely proportional to  $N^2$  are called deutero-ghosts (Paper II).

(ii) Some coefficients in Eq. (4.20) and Eq. (4.46) have a positive sign, while others have a negative sign. If the sign is positive and the ghost flux is completely real, it will manifest itself as a negative ghost, i.e. as a suppressing ghost. If the sign is negative and the ghost flux is completely real, it will manifest itself as a positive ghost, i.e. as an amplifying ghost. The inverse relation between the sign of the coefficients and

whether a ghost manifests as a positive or negative ghost is due to the fact that  $(\rho_{pq})^{-1} = \bar{\rho}_{pq}$  (also see Eq. (4.38)).

We can augment the per-baseline taxonomy of Paper II by using the second observation. The details of this proposed augmentation follows below.

Assigning the false color blue to the coefficients with a positive sign and the false color red to those with a negative sign, enables us to divide  $\mathcal{I}$  into the following mutually exclusive subsets:  $\mathcal{I}_{PB} = \bigcup_{i=1}^3 \mathcal{I}_i$ ,  $\mathcal{I}_{PR} = \bigcup_{i=1}^3 \mathcal{I}'_i$ ,  $\mathcal{I}_{DB} = \emptyset$  and  $\mathcal{I}_{DR} = \emptyset$ . Similarly, we can divide  $\mathcal{J}$  into  $\mathcal{J}_{PB} = \bigcup_{i=1}^3 \mathcal{I}_i$ ,  $\mathcal{J}_{PR} = \emptyset$ ,  $\mathcal{J}_{DB} = \emptyset$  and  $\mathcal{J}_{DR} = \mathcal{I}_4$ . The aforementioned index sets are also presented in Table 1.

Using  $M_{pq}$  we can map the sets of baseline indices to sets of ghost positions. Interestingly enough, when the array is not redundant then  $M_{pq}$  is an isomorphism, i.e., it is structure preserving. If we inspect Table 1 whilst keeping the aforementioned fact in mind we realize that  $\mathcal{T}_{pq} = M_{pq}(\mathcal{I})$  can be divided into the following mutually exclusive subsets:  $\mathcal{T}_{pq}^{PR}$ ,  $\mathcal{T}_{pq}^{PB}$ ,  $\mathcal{T}_{pq}^{DR}$  and  $\mathcal{T}_{pq}^{DB}$ . We can use a similar argument to divide the set  $\mathcal{S}_{pq} = M_{pq}(\mathcal{J}) \cup \{\mathbf{0}\}$  into four mutually exclusive subsets, namely  $\mathcal{S}_{pq}^{PR}$ ,  $\mathcal{S}_{pq}^{PB}$ ,  $\mathcal{S}_{pq}^{DR}$  and  $\mathcal{S}_{pq}^{DB}$ . The per-baseline ghosts can therefore, by extension, also be divided into four categories, namely *red proto-ghosts* (associated with  $\mathcal{T}_{pq}^{PR}$  and  $\mathcal{S}_{pq}^{PR}$ ), *blue proto-ghosts* (associated with  $\mathcal{T}_{pq}^{PB}$  and  $\mathcal{S}_{pq}^{PB}$ ), *red deutero-ghosts* (associated with  $\mathcal{T}_{pq}^{DR}$  and  $\mathcal{S}_{pq}^{DR}$ ) and *blue deutero-ghosts* (associated with  $\mathcal{T}_{pq}^{DB}$  and  $\mathcal{S}_{pq}^{DB}$ ). The augmented per-baseline ghost taxonomy is summarized in Table 2.

We can make a number of interesting observations from Table 2. First, phase-only calibration does not produce deutero-ghosts, which implies that the flux of all ghosts only decreases in proportion to the number of antennas in the array. Moreover, phase-only calibration produces an equal amount of red (positive/amplification) and blue (negative/suppressing)<sup>3</sup> ghosts. In the case of full-complex calibration, all red ghosts are deutero-ghosts while all blue ghosts are proto-ghosts. As a result, the red ghosts will typically be less bright in the case of full-complex calibration than they will be in the case of phase-only calibration. Moreover, the red ghosts will be less bright than the blue ghosts in the case of full-complex calibration, while they are of similar brightness in the case of phase-only calibration.

Note that our classification of ghosts assumes that the array is not redundant. If the array is redundant, some per-baseline-ghosts would have the same position vector and would therefore be added together; changing their characteristics. This would require a more elaborate classification scheme, which we do not pursue here, since it not necessary to understand the symmetry relationships that exist between pairs of ghosts.

## 5 COMPARISON

In this section we compare the full-complex ghost response of Paper II with the phase-only results we obtained in Sec. 4. We first consider the ghost-pattern of an individual baseline in Sec. 5.1 before studying the conglomerated patterns in

<sup>3</sup> The sign refers to the sign of the ghost itself and not to the sign of the coefficient associated with it.

Sec. 5.2. All of the results in this section were obtained using the seven dish Karoo Array Telescope (KAT-7) array layout. KAT-7 is a precursor to the MeerKAT array and is located in the Karoo (South Africa). The layout of this array was depicted in Paper II. Moreover, we only consider the two-source test case from Sec. 3.1 with the unmodelled source being 0.2 Jy at  $1^\circ$ ,  $\delta_0 = -74.66^\circ$  and  $\nu = 1.445$  GHz. We denote the observational frequency with  $\nu$ .

### 5.1 Per-baseline Results

In Fig. 3 we present extrapolated artefact maps which were created by using the KAT-7 array layout. More specifically, we have plotted  $\mathcal{F}^{-1}\{g_{45}^{-1}(\hat{X}_{45}^{-1}(\mathbf{b})) - 1\}$  and  $\mathcal{F}^{-1}\{\rho_{45}^{-1}(\hat{X}_{45}^{-1}(\mathbf{b})) - 1\}$  in Figs. 3(a)–3(b) and Figs. 3(d)–3(e) respectively. The inverse Fourier transform is denoted by  $\mathcal{F}^{-1}\{\cdot\}$ . Fig. 3(c) and Fig. 3(f) show the theoretical ghost patterns based on the index sets summarized in Table 2. The theoretical ghost positions shown in Figs. 3(c) and 3(f) help us visualize the similarities and differences that exist between full-complex calibration and phase-only calibration. The brightest ghosts in the artefact maps match perfectly with the theoretical ghost positions; indicating that the results from raw extrapolation and perturbation align well.

We can draw the following conclusions from Table 2 and Fig. 3:

(i) *Proto-ghosts*: Full-complex calibration only produces blue proto-ghosts. In contrast, phase-only calibration produces an equal number of red and blue proto-ghosts. The blue and red proto-ghosts form symmetrically opposite each other around the modelled source. Moreover, the set of blue proto-ghosts that are created by phase-only calibration and full-complex calibration are almost equal.

(ii) *Deutero-ghosts*: All the deutero-ghosts that are produced by full-complex calibration are red. No deutero-ghosts are generated by phase-only calibration. The position vector set associated with the red proto-ghosts is a subset of the position vector set that is associated with the full-complex deutero-ghosts, i.e.  $\mathcal{T}_{pq}^{PR} \subset \mathcal{S}_{pq}^{DR}$ . We therefore produce fewer first order ghosts, i.e. the ghosts associated with the first order perturbation of the visibility gains, with phase-only calibration when compared with full-complex calibration. We do however produce more proto-ghosts.

In Paper II we realized that the following three ghosts were exceptionally bright:

- The *primary suppressor*. The primary suppressor forms on top of the modelled source, and it can therefore suppress the modelled source. It is important to mention here that, in the full-complex case, the primary suppressor only forms when we incorporate the autocorrelations during calibration, which is rarely done in practice (Paper II). The primary suppressor forms at  $\mathbf{0}$ .
- The *secondary suppressor*. The secondary suppressor lies on top of the unmodelled source, i.e. at  $\mathbf{s}_0$ . This ghost can suppress the unmodelled source.
- The *anti-ghost*. The anti-ghost appears symmetrically opposite the secondary suppressor around the primary suppressor, i.e. at  $-\mathbf{s}_0$ .

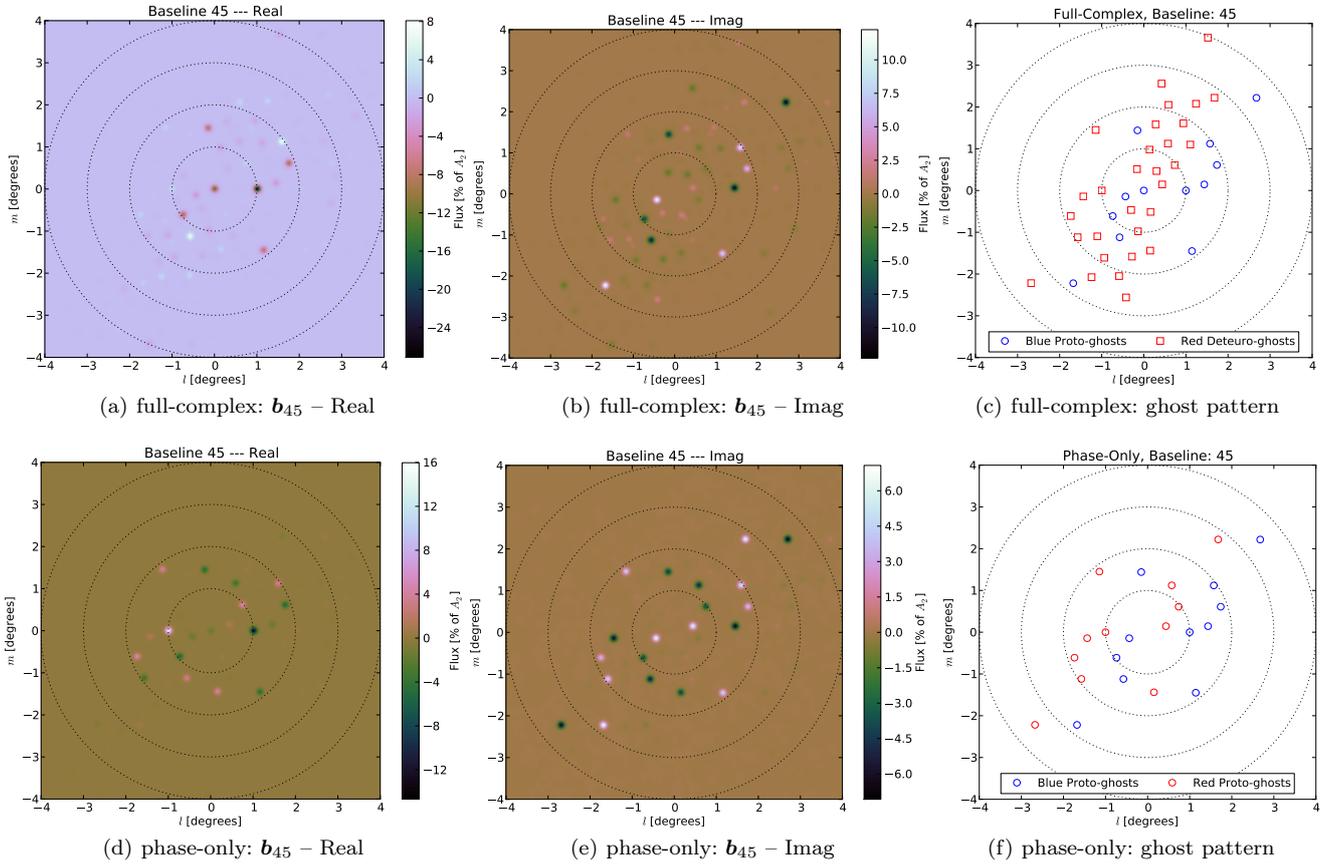
Concentrating on these specific sources in Table 2 and Fig. 3, we find:

**Table 1.** Comparison between the composite baseline index sets produced by phase-only calibration and full-complex calibration. The entries in the last row of each subtable were obtained by taking the union of all the other entries in the column in which they reside. The same applies to the entries in the last column of each subtable.

sign	phase-only			full-complex		
	$\frac{1}{N}$	$\frac{1}{N^2}$	all	$\frac{1}{N}$	$\frac{1}{N^2}$	all
+	$\mathcal{I}_{PB} = \bigcup_{i=1}^3 \mathcal{I}_i$	$\mathcal{I}_{DB} = \emptyset$	$\mathcal{I}_B = \mathcal{I}_{PB} \cup \mathcal{I}_{DB}$	$\mathcal{J}_{PB} = \bigcup_{i=1}^3 \mathcal{I}_i$	$\mathcal{J}_{DB} = \emptyset$	$\mathcal{J}_B = \mathcal{J}_{PB} \cup \mathcal{J}_{DB}$
-	$\mathcal{I}_{PR} = \bigcup_{i=1}^3 \mathcal{I}'_i$	$\mathcal{I}_{DR} = \emptyset$	$\mathcal{I}_R = \mathcal{I}_{PR} \cup \mathcal{I}_{DR}$	$\mathcal{J}_{PR} = \emptyset$	$\mathcal{J}_{DR} = \mathcal{I}_A$	$\mathcal{J}_R = \mathcal{J}_{PR} \cup \mathcal{J}_{DR}$
	$\mathcal{I}_P = \mathcal{I}_{PB} \cup \mathcal{I}_{PR}$	$\mathcal{I}_D = \emptyset$	$\mathcal{I}$	$\mathcal{J}_P = \mathcal{J}_{PB} \cup \mathcal{J}_{PR}$	$\mathcal{J}_D = \mathcal{J}_{DB} \cup \mathcal{J}_{DR}$	$\mathcal{J}$

**Table 2.** Comparison between the index sets of ghost locations produced by phase-only calibration and full-complex calibration using the augmented taxonomy with blue and red proto-ghosts and blue and red deutero-ghosts. Since  $M_{pq}$  is structure preserving, the entries in the last row of each subtable can also be obtained by taking the union of all the other entries in the column in which they reside. The same applies to the entries in the last column of each subtable. Note that in the table we have used the shorthand  $\tilde{\mathbf{0}}$  to denote  $\{\mathbf{0}\}$ .

color	phase-only			full-complex		
	proto	deutero	all	proto	deutero	all
B	$\mathcal{T}_{pq}^{PB} = M_{pq}(\mathcal{I}_{PB})$	$\mathcal{T}_{pq}^{DB} = \emptyset$	$\mathcal{T}_{pq}^B = M_{pq}(\mathcal{I}_B)$	$\mathcal{S}_{pq}^{PB} = M_{pq}(\mathcal{J}_{PB}) \cup \tilde{\mathbf{0}}$	$\mathcal{S}_{pq}^{DB} = \emptyset$	$\mathcal{S}_{pq}^B = M_{pq}(\mathcal{J}_B) \cup \tilde{\mathbf{0}}$
R	$\mathcal{T}_{pq}^{PR} = M_{pq}(\mathcal{I}_{PR})$	$\mathcal{T}_{pq}^{DR} = \emptyset$	$\mathcal{T}_{pq}^R = M_{pq}(\mathcal{I}_R)$	$\mathcal{S}_{pq}^{PR} = \emptyset$	$\mathcal{S}_{pq}^{DR} = M_{pq}(\mathcal{J}^{DR})$	$\mathcal{S}_{pq}^R = M_{pq}(\mathcal{J}_R)$
	$\mathcal{T}_{pq}^P = M_{pq}(\mathcal{I}_P)$	$\mathcal{T}_{pq}^D = \emptyset$	$\mathcal{T}_{pq} = M_{pq}(\mathcal{I})$	$\mathcal{S}_{pq}^P = M_{pq}(\mathcal{J}_P) \cup \tilde{\mathbf{0}}$	$\mathcal{S}_{pq}^D = M_{pq}(\mathcal{J}_D)$	$\mathcal{S}_{pq} = M_{pq}(\mathcal{J}) \cup \tilde{\mathbf{0}}$



**Figure 3.** The left and middle column show the real (left column) and imaginary (middle column) part of the extrapolated artefact map for baseline  $b_{45}$  obtained for full-complex calibration (top row) and phase-only calibration (bottom row). These maps were created by using KAT-7. The modelled source is in the center and the unmodelled source (0.2 Jy) was at  $1^\circ$ . Moreover,  $\delta_0 = -74.66^\circ$  and  $\nu = 1.445$  Ghz. The two panels in the right column show the theoretical ghost patterns for full-complex (top) and phase-only (bottom) calibration for this scenario.

(i) *Primary suppressor*: The primary suppressor is a blue proto-ghost, but only forms if full-complex calibration is employed. Mathematically speaking  $\mathcal{S}_{pq}^{PB} \setminus \mathcal{T}_{pq}^{PB} = \{\mathbf{0}\}$ . We denote set-theoretic difference by “ $\setminus$ ”. Intuitively, this is in agreement with the fact that the primary suppressor disappears when the autocorrelations are ignored in full-complex

calibration, since the autocorrelations do not contain phase information. As we showed in Paper II, the primary suppressor is completely real and since it is blue it will appear as a negative source. This can be confirmed by inspecting Figs. 3(a)–3(b).

(ii) *Secondary suppressor and its anti-ghost*: In the case

of full-complex calibration, the secondary suppressor is a blue proto-ghost, and its anti-ghost is a red deuterio-ghost. As is the case with full-complex calibration the phase-only secondary suppressor is also a blue proto-ghost, but its anti-ghost is a red proto-ghost. This explains why the phase-only anti-ghost is so much brighter than the full-complex anti-ghost in Fig. 3. In Paper II we showed that the secondary suppressor and its anti-ghost are real valued. The fact that the secondary suppressor is real valued and blue explains why the secondary suppressor ends up being a negative ghost. The anti-ghost of the secondary suppressor manifests as a positive ghost, because it is real valued and red.

## 5.2 Conglomerated Results

In the previous section we only looked at the per-baseline ghost response but, as we mentioned in Sec. 4.6, the conglomerated ghost pattern and the ghost pattern of each individual baseline are completely different. It turns out that some ghost add up coherently for all baselines while others do not. In Paper II we developed the following taxonomy for ghosts in the conglomerated pattern to address this issue:

(i) *Line ghosts*: All ghosts that form at  $\tilde{\mathbf{s}} \in \mathbf{s}_0 \mathbf{Z}$  are line ghosts. Line ghosts are present in each baseline and therefore add up coherently when we image.

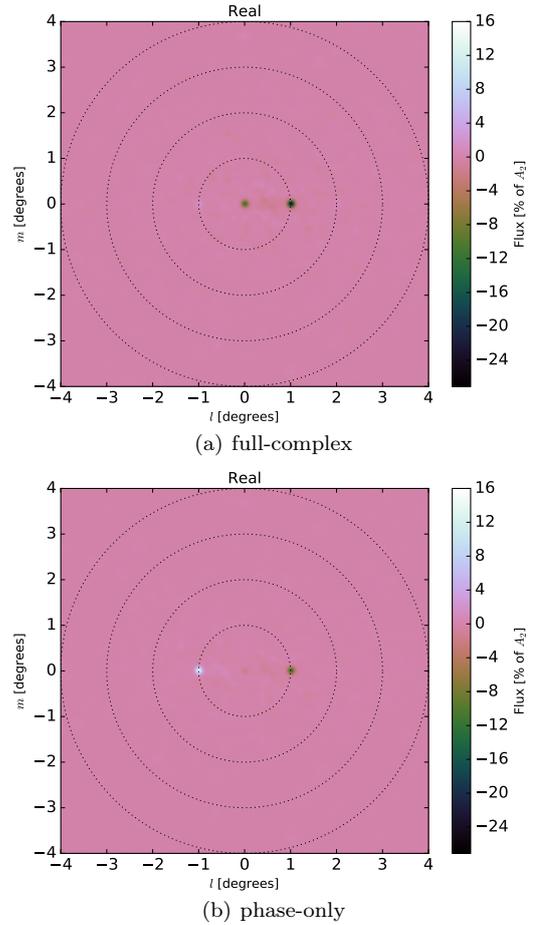
(ii) *Scattered ghosts*: If a ghost is not a line ghost, it is a scattered ghost. In the non-redundant case, scattered ghosts are only present in one baseline, which is why they appear dimmer in the conglomerated ghost pattern.

The extrapolated conglomerated ghost patterns for full-complex and phase-only calibration for the full KAT-7 array are shown in Fig. 4(a) and Fig. 4(b) respectively. These maps are clearly different from the per-baseline results shown in Fig. 3. One of the major differences is that the secondary suppressor and its anti-ghost are much brighter relative to the remaining background ghosts.

The fluxes of the three most important line ghosts in the conglomerated ghost pattern are shown in Fig. 5(a) and Fig. 6(a). The scattered equivalents of Fig. 5(a) and Fig. 6(a) are depicted in Fig. 5(b) and Fig. 6(b). Fig. 5 and Fig. 6 were produced by employing full-complex and phase-only calibration respectively. To be more descriptive, Figs. 5(a)–5(b) and Figs. 6(a)–6(b) illustrate the full-complex equivalent of Eq. (4.43) (which can be found in Paper II) and Eq. (4.43) respectively. Similarly, Figs. 5(c)–5(d) and Figs. 6(c)–6(d) were created by employing the full-complex equivalent of Eq. (4.45) and Eq. (4.45) respectively.

In Fig. 5(a) and Fig. 6(a) the blue curve represents the primary suppression ghost, the red curve the secondary suppression ghost and the green curve the anti-ghost of the secondary suppressor. In Fig. 5(b) the blue curve represents a random blue full-complex proto-ghost, while its anti-ghost is depicted in red. Its anti-ghost is a red deuterio-ghost. In Fig. 6(b) the blue curve is associated with a random blue phase-only proto-ghost, the red curve is associated with its anti-ghost, a red proto-ghost.

The dashed curves in Fig. 5 and Fig. 6 are the perturbation theoretical estimates we derived in Sec. 4.2 and in Paper II. In effect, we plot Eq. (4.22) and Eq. (4.48). The fact that the results we obtained from perturbation and raw extrapolation agree supports the validity of the first order

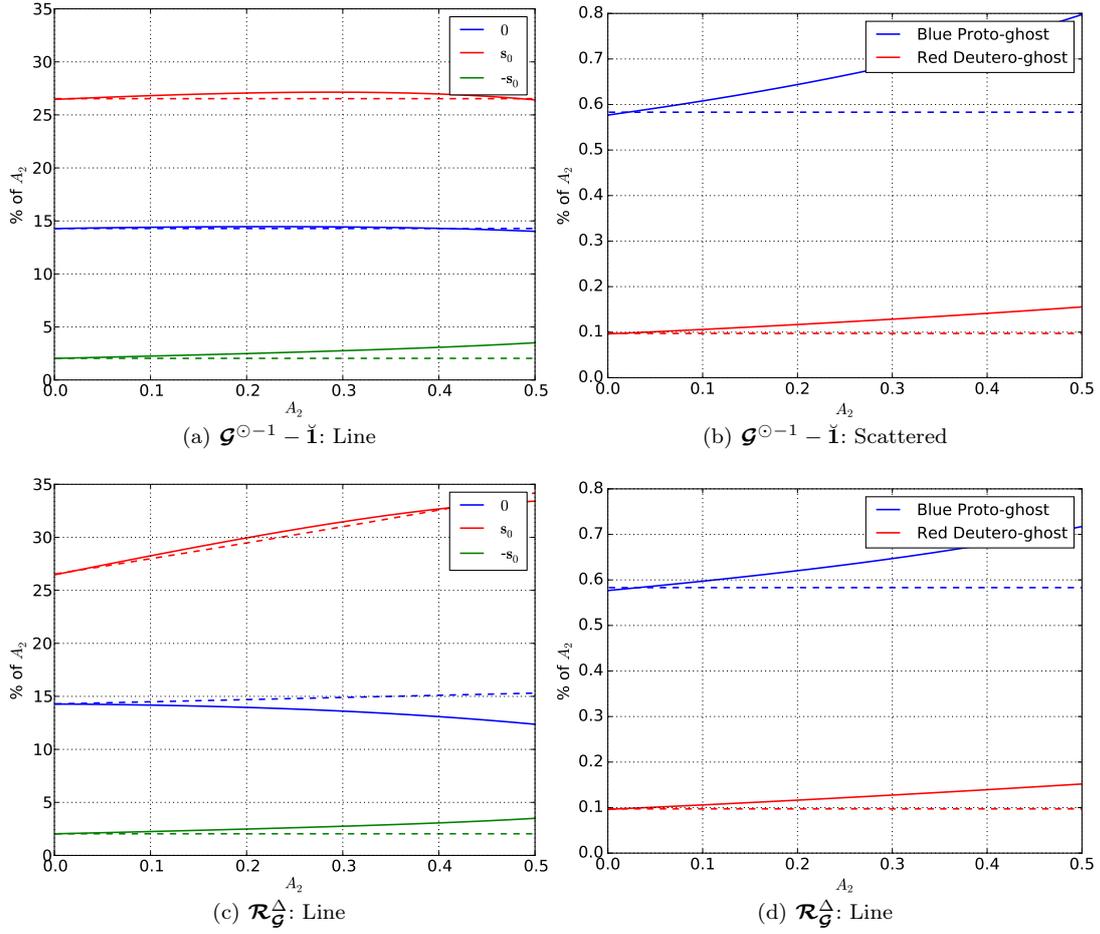


**Figure 4.** Conglomerated ghost pattern for full-complex calibration (top) and phase-only calibration (bottom). The suppression ghost, the ghost forming on top of the unmodelled source, is brighter in the top panel than in the bottom panel. In the case of the anti-ghost, the ghost forming symmetrically opposite the suppression ghost around the modelled source, we observe the exact opposite. Moreover, fewer bright ghosts are visible in the bottom panel than in the top panel.

perturbation derivation in Sec. 4. The colors of the dashed lines indicate with which source we can associate each theoretical curve. In Fig. 6 we use black dash lines instead, because in the case of phase-only calibration perturbation predicts the same absolute flux value for every ghost and anti-ghost pair. This can be verified by inspecting Eq. (4.22) and Eq. (4.31). Perturbation also predicts that the primary suppressor will always have zero flux. This explains why the dashed blue line is absent in Fig. 6.

The primary suppressor is on average 14% of  $A_2$  and 2.5% of  $A_2$  when we employ full-complex and phase-only calibration respectively. Moreover, in the case of full-complex calibration the secondary suppressor and its anti-ghost are on average 30% of  $A_2$  and 2% of  $A_2$  respectively. For the phase-only setup, we find that on average both the secondary suppressor and its anti-ghost are 15% of  $A_2$ .

A random blue proto-ghost is on average 0.7% of  $A_2$  and 0.35% of  $A_2$  when we perform full-complex and phase-only calibration respectively. The blue proto-ghost's anti-ghost is



**Figure 5.** The relative amplitude of the primary suppressor, secondary suppressor and its anti-ghost on the left and the relative amplitude of a random proto-ghost and its anti-ghost on the right. These curves were created by employing full-complex calibration. 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.

on average 0.1% of  $A_2$  and 0.35% of  $A_2$  when we use the full-complex and phase-only calibration algorithms respectively.

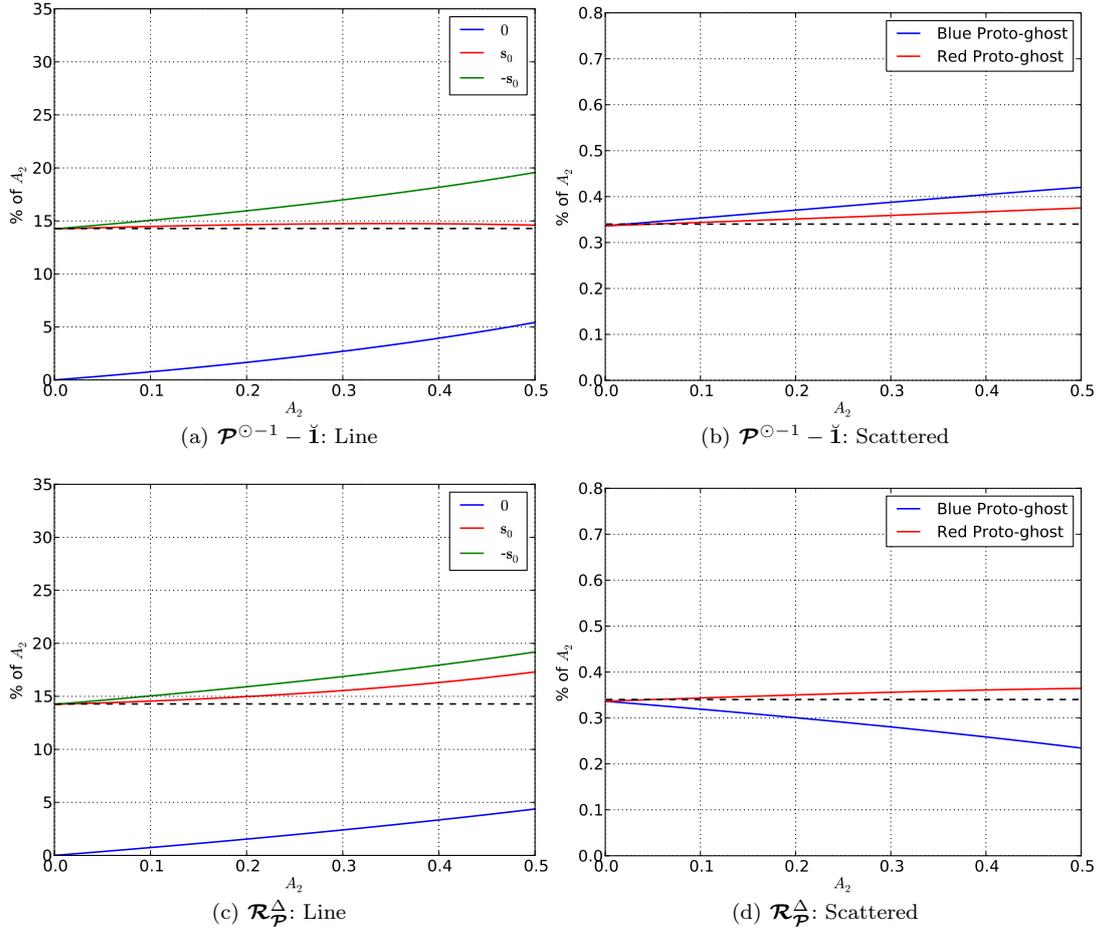
Also note that correcting our visibilities does not significantly impact the brightness of our ghosts (compare the top panel of Fig. 5 with the bottom panel in Fig. 5). The slight differences between the top and bottom panel Fig. 5 are discussed in Paper I and II. The same reasoning can be applied to Fig. 6.

We can reach the following conclusions by comparing Fig. 4(a), Fig. 4(b), Fig. 5 and Fig. 6 with each other:

(i) *Line ghosts*: The primary suppressor, the secondary suppressor and its anti-ghost are all line ghosts, which is why they are so much brighter than the other ghosts in Fig. 4. According to perturbation theory, the primary suppressor only forms when performing full-complex calibration. However, Fig. 6 clearly indicates that a phase-only primary suppressor does form when  $A_2$  becomes large enough. The phase-only primary suppressor, however, is much dimmer than its full-complex counterpart. Moreover, the phase-only anti-ghost in Fig. 4(b) and Fig. 6 is much brighter than the anti-ghost in Fig. 4(a) and Fig. 5 since the anti-ghost which is produced by phase-only calibration is a proto-ghost, while the anti-ghost produced by full-complex calibration is a deutero-ghost. The

full-complex secondary suppressor is also brighter than its phase-only counterpart. This can be explained by comparing Eq. (4.22) and Eq. (4.48). These two equations tell us that the secondary suppressor is proportional to  $2/N$  and  $1/N$  when we employ full-complex and phase-only calibration respectively.

(ii) *Scattered ghosts*: We see that we produce less scattered ghosts with phase-only calibration than with full-complex calibration, since we do not produce any deutero-ghosts while performing phase-only calibration. Furthermore, in the case of phase-only calibration, the blue and red symmetric ghost pairs in Fig. 6 are of the same order of magnitude (both are proto-ghosts). This is not true when performing full-complex calibration (Fig. 3(c) clearly shows that at least one of the ghosts in each symmetric pair is a deutero-ghost). This follows trivially from Eq. (4.31), which is only true for phase-only calibration. Moreover, the full-complex proto-ghosts in Fig. 5 are brighter than the phase-only proto-ghosts in Fig. 6, which is corroborated by Eq. (4.22) and Eq. (4.48).



**Figure 6.** The relative amplitude of the primary suppressor, secondary suppressor and its anti-ghost on the left and the relative amplitude of a random proto-ghost and its anti-ghost on the right. These curves were created by employing phase-only calibration. 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.

## 6 PRIMARY BEAM

In this section we replicate and then try to explain the synthetic anti-ghost experiment in Stewart (2014), which we discussed in detail in Sec. 1 (see also Fig. 1). In Sec. 6.1, we will modify Stewart’s experiment slightly to be able to explain his results using just a single graph. Our findings are discussed in detail in Sec. 6.2. Although the strange behaviour of the anti-ghost appears to contradict some of the results in Paper II and Sec. 4, we will see that the primary beam correction causes the apparent dependence of the anti-ghost’s brightness on the position of the unmodelled source. The results in this section were obtained using standard reduction software packages and techniques.

### 6.1 Spiral Experiment

In this section we present an experiment that can shed some light on Stewart’s puzzling result (see Sec. 1). We actually discuss a simple, modified version of Stewart’s experiment below. Instead of sampling the locations of the unmodelled source from a rectangular grid, we propose that the possible locations of the unmodelled source all lie on a spiral. As will become apparent later, altering Stewart’s experiment in this

manner allows us to create a one dimensional plot which can help us shed some light on Stewart’s findings.

Just as with Stewart’s experiment, we decided to conduct our experiment with the LOFAR antenna layout. To mirror Stewart’s original experiment as much as possible we used 11min observations. Moreover, the field center and the NCP coincided. The measurement set we used was created assuming 33 LOFAR stations<sup>4</sup>. We also assumed a monochromatic observation at a frequency of  $\nu = 58.66$  MHz.

Furthermore, we only considered a two-source scenario. We place an 88.7 Jy source at (02h22m52.1496s, 86°18′59.184″). This bright source is included in our calibration model. We also introduce a second source of 60 Jy to our simulation. This second source is not included in our calibration model. Moreover, the modelled source is placed at the same location as 3C 61.1.

We start our experiment by placing the unmodelled source at the field center. We then generate the visibili-

<sup>4</sup> LBA CS 1–7, 11, 13, 17, 21, 24, 26, 28, 30–32, 101, 103, 201, 301, 302, 401, 501; LBA RS 106, 205, 208, 306–307, 406, 503, 508–509.

ties associated with the modelled and unmodelled source, which we then store in the measurement set. The modelled and unmodelled source were attenuated by the array’s primary beam before the visibilities were generated. The effect of beam rotation was ignored. We then perform phase-only calibration. We then image and clean the resulting corrected visibilities and measure the *apparent and intrinsic flux* of the modelled source, the unmodelled source and the anti-ghost.

The intrinsic flux density of a source is the true flux value of a specific source and its apparent flux is the flux it appears to have after it has been attenuated by the observing array’s primary beam response. In our spiral experiment, there are two different ways to use the terms intrinsic and apparent flux. The first way is to use them in the context of the true modelled flux of the source and the true modelled flux after primary beam attenuation. These are the input / a priori intrinsic and apparent flux respectively. The second way is to use them in the context of the flux of the source and the flux of the source after primary beam attenuation obtained from the images produced by the simulation. These are the output / measured / a posteriori intrinsic and apparent flux respectively.

We can now change the position of the unmodelled source and then repeat the aforementioned simulation, calibration and imaging steps to generate another simulated observation. If we continue in this way, varying the position of the unmodelled source followed by the generation of a simulated observation, we will be able to produce multiple snapshot images (the unmodelled source will be in a different location in every snapshot).

The unmodelled source can be placed at completely random positions or it can follow a deterministic trajectory. We opted for a deterministic trajectory. During the course of the spiral experiment, the unmodelled source follows an Archimedean spiral trajectory, which can be described by the equation

$$r(\theta) = \frac{-\theta}{2\pi}, \quad (6.1)$$

where  $r$  is the radial distance (in degrees) of the unmodelled source from the field center and  $\theta$  is the polar angle of the spiral. In Eq. (6.1),  $\theta$  is a positive real number and is measured in radians. The results we obtained by performing this experiment are presented in Sec. 6.2.

The red spiral in Fig. 7(a) depicts the trajectory that the unmodelled source follows during the experiment. The anti-ghost in turn follows the blue trajectory in Fig. 7(a). Moreover, the LOFAR primary beam which we used during the spiral experiment is presented in Fig. 7(a).

Fig. 7 also contains three other images: Fig. 7(b) – Fig. 7(d). Each of these figures represents a spiral experiment snapshot. Fig. 7(b), Fig. 7(c) and Fig. 7(d) were obtained by placing the unmodelled source  $0.32^\circ$ ,  $3.08^\circ$  and  $3.48^\circ$  away from the field center respectively. The anti-ghost is clearly visible in Fig. 7(b) and Fig. 7(d), but is not visible in Fig. 7(c). This confirms Stewart’s observation (see Sec. 1) that the brightness of the anti-ghost is highly dependent on the position of the unmodelled source. The contours of the primary beam that was used in this experiment are depicted in Fig. 7 with yellow dashed lines.

## 6.2 Spiral Experiment Results

### 6.2.1 Preliminary observations

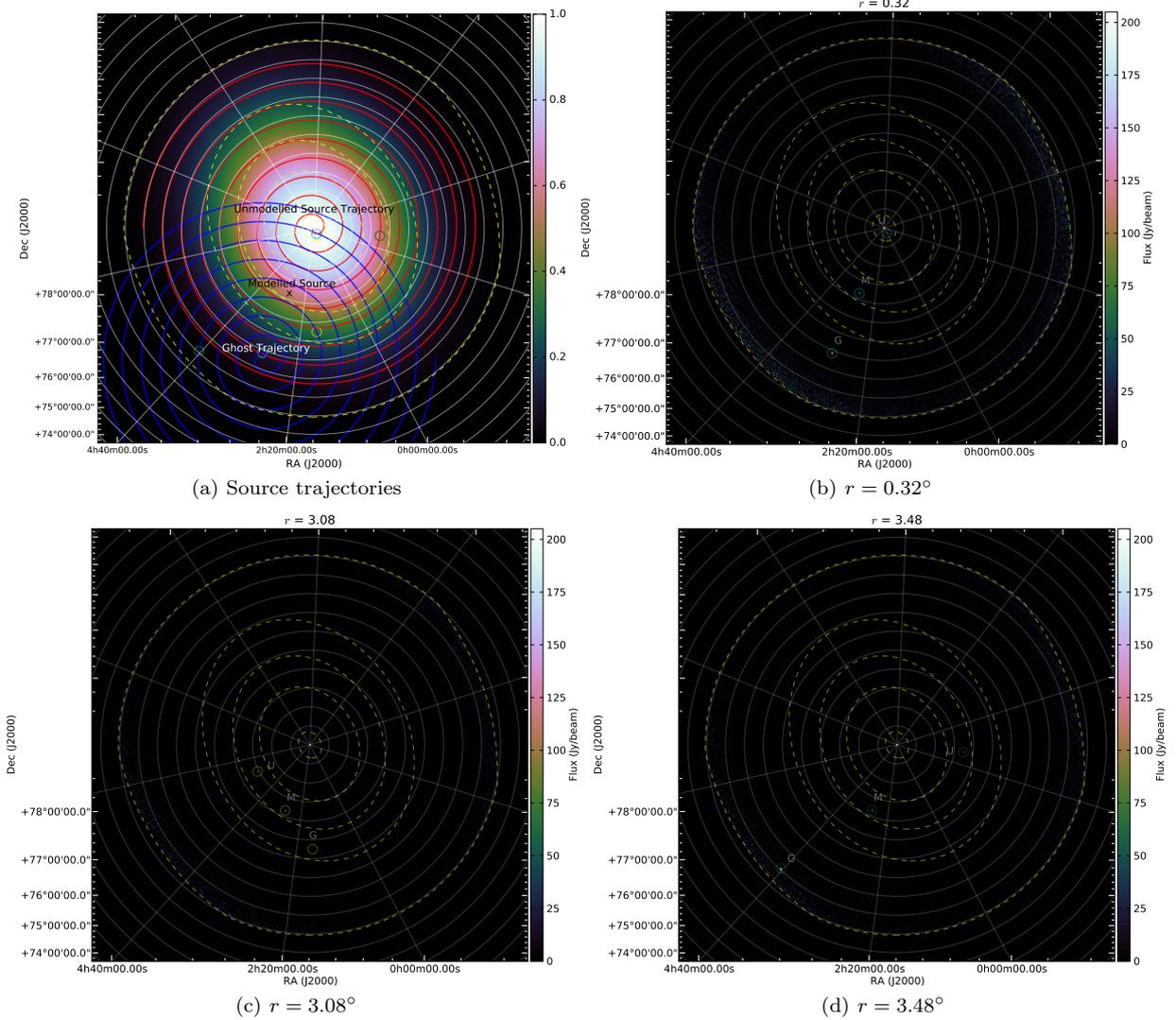
The results of the spiral experiment are shown in Fig. 8. We find the different flux values of the modelled source, the unmodelled source and the anti-ghost for the spiral experiment as a function of the radial distance  $r$  (see Eq. (6.1)). The line-types indicate which curves can be associated with which source: the dotted curves, the dashed curves and the solid curves in Fig. 8 can be associated with the modelled source, the unmodelled source and the anti-ghost respectively. The only exception is the solid magenta line, which denotes the beam gain the anti-ghost experiences when the unmodelled source is  $r$  degrees from the field center (and is not measured in Jy, but is a unitless gain value).

In Fig. 8, there are four different-colored curves for each of the three aforementioned sources/line-types (i.e. there are twelve curves in total). The blue and red curves denote the true intrinsic and the true apparent flux values of the sources with which they are associated, i.e., true simulated model and true primary beam attenuated simulated model flux values. The values of the red and blue curves are known a priori. The solid red and blue line, however, warrant further explanation. The solid red line represents the theoretical flux density of the anti-ghost, as derived in Sec. 4, and the solid blue curve represents the product of the primary beam gain (magenta curve) and the theoretical flux density of the anti-ghost, i.e. the effect that the primary beam had on the theoretical flux density of the anti-ghost once we correct for it. The green and black curves denote the measured intrinsic and the apparent flux values of the sources with which they are associated, i.e., raw measured flux values and primary beam corrected measured flux values. The green and black curve values were obtained from the images produced by the spiral experiment. To summarize, the red and blue curves can be interpreted as theoretical predictions, i.e. as the values that are fed into the spiral experiment, while the green and black curves can be interpreted as observed values, i.e. the quantities that are produced by the spiral experiment.

The yellow bars in Fig. 8 indicate regions where the anti-ghost lies outside the boundaries of the images that were created by the spiral experiment. This explains why the yellow regions do not contain any solid curves.

Note that, since the modelled source remains at the same location during the experiment, its intrinsic flux curves are scalar multiples of its apparent flux curves. This assertion can be verified by comparing the dotted blue curve with the dotted red curve. A similar conclusion can be drawn by comparing the dotted black curve with the dotted green curve. In contrast, the apparent flux of the unmodelled source slowly dims as it spirals outward due to the array’s primary beam response. This can be verified by inspecting the dashed red curve.

We have indicated two regions in Fig. 8. In the red region with  $r \leq 2.5^\circ$ , the apparent flux of the unmodelled source is larger than the apparent flux of the modelled source, i.e. the red dashed line is above the red dotted line. In this region, the assumption  $A_2 \ll A_1$  of our perturbation analysis is clearly violated. The opposite is true in the blue region with  $r > 2.5^\circ$ . Not unexpectedly, we find that when the apparent flux of the unmodelled source is larger than the modelled source we get a very different ghost response,



**Figure 7.** The source trajectories that the unmodelled source and its anti-ghost follow during the spiral experiment are depicted in Fig 7(a). The LOFAR primary beam model, which we used in the spiral experiment, is also shown in Fig 7(a). Figs. 7(b)–7(d) depict three distinct spiral experiment snapshots. The radial distance the unmodelled source is from the field center is given in the individual sub-captions of the three figures. The contours of the LOFAR primary beam model are depicted in Figs. 7(a)–7(d) with dashed yellow lines. The positions of the modelled source, the unmodelled source and the anti-ghost in Figs. 7(b)–7(d) are indicated with labeled circular markers. The labels M, U, and G were used to tag the modelled source, the unmodelled source and the anti-ghost respectively. The positions of the modelled source, the unmodelled source and the anti-ghost in each of the three snapshot can also be found in Fig 7(a) and are indicated with circular markers. The color of the marker should be used as a key to identify to which snapshot each position belongs.

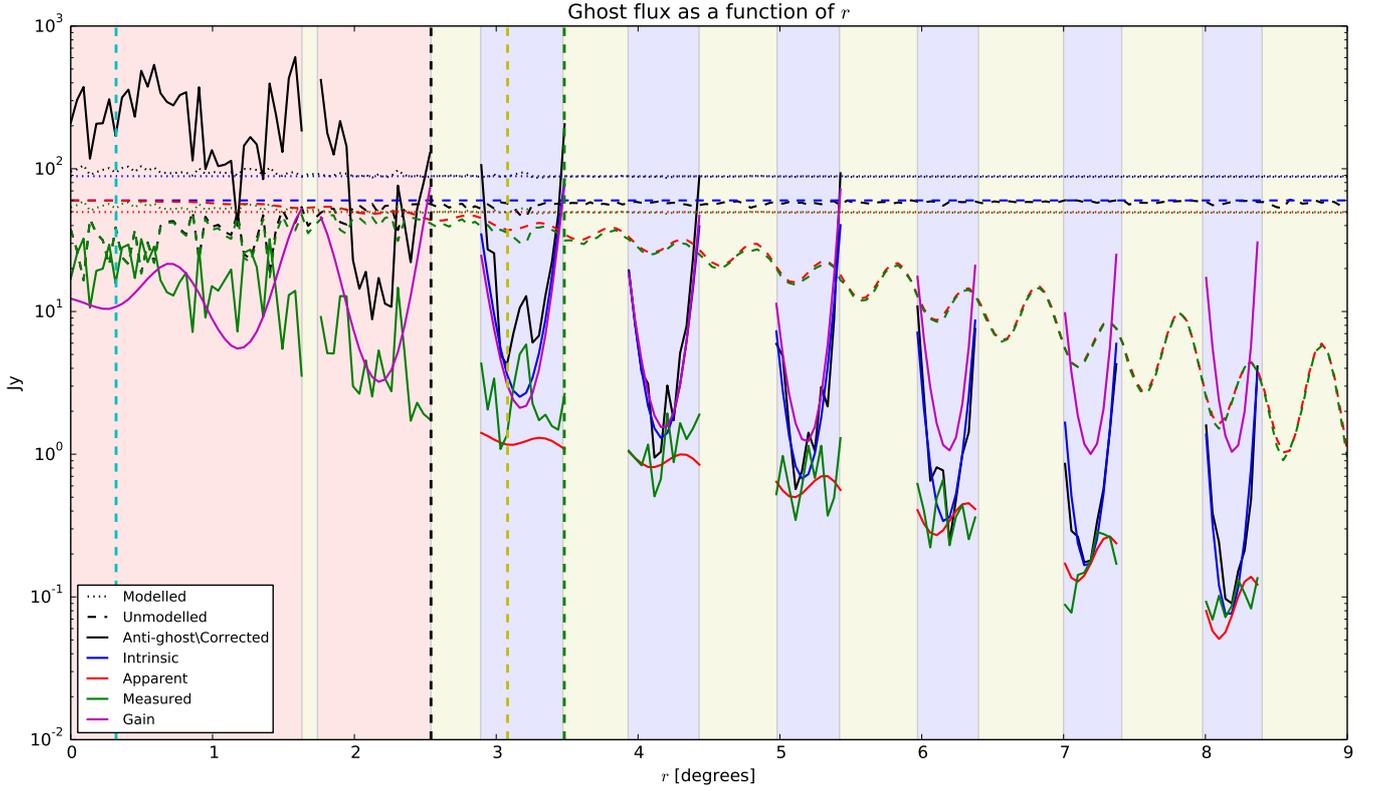
than the well studied case in which the apparent flux of the unmodelled source is smaller than the apparent flux of the modelled source.

### 6.2.2 Model Accuracy

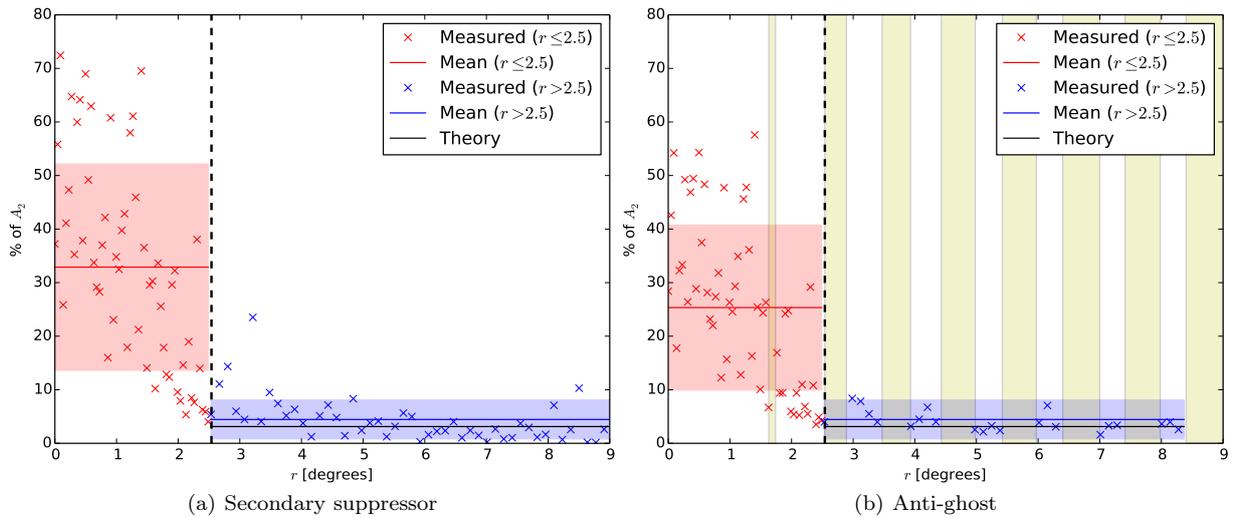
The fact that the ghost response is different in the red and blue regions of Fig. 8 is easier to spot if Fig. 9 is inspected instead. In Fig. 9(a), we have the measured value of the secondary suppressor (as a % of  $A_2$ ) as a function of the radial distance the unmodelled source is from the field center. In Fig. 9(b), we have the measured value of the anti-ghost (as a % of  $A_2$ ) as a function of the radial distance the unmodelled source is from the field center. The crosses in each region of

Fig. 9 represent actual measured flux values (as a % of  $A_2$ ). These values were obtained from the images that were produced during the spiral experiment. The coloured lines in each region represent the mean flux value measured within each region and the filled in block in each region depicts the standard deviation about the mean flux in each region. The black line in both of these figures is the theoretical value of the secondary suppressor and the anti-ghost, as derived in Sec. 4.

Inspection of Fig. 8 and Fig. 9 reveals the following: when the calibration model is accurate enough, i.e.  $A_2 < A_1$  or  $r > 2.5^\circ$ , the measured flux values are close to the theoretical value, which we derived in Sec. 4. However, when the model becomes too inaccurate, i.e.  $A_2 \geq A_1$  or  $r \leq 2.5^\circ$ ,



**Figure 8.** The flux values of the spiral experiment for the modelled source, the unmodelled source and the anti-ghost as a function of the radial distance  $r$  (the distance the unmodelled source is located from the field center). The curves that are associated with the modelled source, the unmodelled source and the anti-ghost are depicted by dotted, dashed and solid lines respectively. Four curves are associated with each source, each one depicted in a different color. The blue curves represent the true intrinsic flux of a source, the red curves depict the true apparent flux of a source, the green curves represent the measured apparent flux values of a source and the black curves denote the measured intrinsic flux of a source. The solid magenta line is an exception and denotes the beam gain the anti-ghost experiences when the unmodelled source is  $r$  degrees from the field center. It is not measured in Jy, but is actually a unitless quantity. The yellow blocks represent regions where the radial distance the ghost is from the field center is larger than  $9^\circ$ . The graph is also divided into two regions, a red region ( $r \leq 2.5^\circ$ ) and a blue region ( $r > 2.5^\circ$ ). The anti-ghost appears much brighter in the red region than in the blue region on average. The colored vertical dashed lines indicate the values of  $r$  at which the snapshots in Fig. 7 were made.



**Figure 9.** The crosses represent the measured flux values of the secondary suppressor and the anti-ghost (as a % of  $A_2$ ) as a function of the radial distance  $r$ . The plots are divided into two regions: a red region and a blue region. The colored horizontal lines represent the mean flux values within each region, while the horizontal filled in blocks represent the standard deviation about the mean within each region. The black line represents the theoretical value of the secondary suppressor and the anti-ghost as predicted in Sec. 4. In Fig. 9(b), the yellow blocks represent regions where the radial distance the ghost is from the field center is larger than  $9^\circ$ .

then it deviates completely from the two black lines in Fig. 9. Fig. 9, therefore, shows us that when the unmodelled source is much brighter than the modelled source we produce a much brighter secondary suppressor and anti-ghost than that predicted in Sec. 4. This result does not contradict anything in Sec. 4, because one of the main assumptions that we make in Sec. 3 is that  $A_2 \ll A_1$ . A similar result was obtained in Grobler & Smirnov (2015) for the WSRT array. Fig. 9 also shows us that the theory we derived in Sec. 4 holds up if  $A_2 < A_1$ . We can now use the insight we gained from inspecting Fig. 9 to make sense of the remaining curves in Fig. 8.

### 6.2.3 Modelled source

The only interesting thing one can conclude by inspecting the curves of the modelled source is that the modelled source seems to be amplified a bit when we calibrate with a very inaccurate model ( $A_2 \geq A_1$ ), but that this quickly improves when the calibration model becomes more accurate. This conclusion can be reached by comparing either the dotted blue and black curves or the dotted red and green curves with each other.

### 6.2.4 Unmodelled source

The first thing one notices when inspecting the curves associated with the unmodelled source is that the measured apparent flux curve (dashed green) of the unmodelled source is far below the true apparent flux curve (dashed red) when  $r \leq 2.5^\circ$ . We are now able to explain this as, according to Sec. 6.2.2, the suppression ghost becomes exceptionally bright if the model is too inaccurate ( $A_2 \geq A_1$ ). In contrast, when  $r > 2.5^\circ$  we see that the dashed green and dashed red curves are virtually indistinguishable, which can also be explained with the aid of Sec. 6.2.2. Sec. 6.2.2 also tells us that if an accurate enough model ( $A_2 < A_1$ ) is used during calibration the unmodelled source is only suppressed slightly if  $N$  is large. Moreover, when  $r \leq 2.5^\circ$  the black dashed line and the green dashed line are fairly close to each other, as the primary beam gain is close to unity when  $r \leq 2.5^\circ$ . As we mentioned previously, when  $r > 2.5^\circ$ , the degree of suppression is reduced significantly, which also explains why the black dashed curve ends up lying just below the intrinsic flux value of the unmodelled source (the dashed blue curve).

### 6.2.5 Anti-ghost

When we inspect the anti-ghost curves we see that the anti-ghost behaves differently when  $r \leq 2.5^\circ$  than when  $r > 2.5^\circ$ . When  $r \leq 2.5^\circ$ , we are actually using a very inaccurate calibration model ( $A_2 \geq A_1$ ), implying that the anti-ghost will be exceptionally bright in this region of the graph (see Sec. 6.2.2). We can confirm this by inspecting the solid green curve in the red region of Fig. 8. The anti-ghost can now be made even brighter if we also correct for the primary beam. This can be verified by inspecting the solid black curve in the red region of Fig. 8.

The combination of these two factors explain why the anti-ghost is clearly visible during the course of the spiral experiment when  $r \leq 2.5^\circ$ . Fig. 7(b) is a snapshot of the

spiral experiment. In this snapshot the unmodelled source is  $0.34^\circ$  from the field center. Fig. 7(b) should be used as a visual aid that shows us just how bright the anti-ghost can be when  $r \leq 2.5^\circ$ .

When  $r > 2.5^\circ$  we see that the anti-ghost becomes very dim and becomes close to the theoretically predicted value (the solid red and green lines are close together). This is expected as, when  $r > 2.5^\circ$ , we satisfy the criterion  $A_2 < A_1$  (see Sec. 6.2.2). However, if we now apply the beam gain we see that the anti-ghost can still become exceptionally bright even if the ‘‘apparent’’ anti-ghost is fairly weak. This is due to the relative positions of the modelled and unmodelled source just happens to be of such an orientation that the anti-ghost that forms lies in a region of the primary beam where the primary beam attenuation is high, then the anti-ghost can be significantly amplified when we do the primary beam correction. If the anti-ghost forms in a region of the primary beam where there is little attenuation then the anti-ghost will remain fairly faint even after we do the primary beam correction. Note that the theoretical ‘‘intrinsic’’ flux curve (the solid blue curve which is associated with the anti-ghost) accurately describes the behaviour of the measured intrinsic anti-ghost flux values (solid black curve).

Fig. 7(c) is a snapshot from the spiral experiment which shows us how hard it can be to spot the anti-ghost when it forms in a region of the primary beam where there is little attenuation. In stark contrast, Fig. 7(d) is a snapshot from the spiral experiment that shows that the anti-ghost can be very bright even if the model is quite accurate, provided the anti-ghost happens to form in a region of the primary beam where the attenuation is high. The radial distances at which the three snapshots were made are depicted in Fig. 8 with three vertical dashed lines.

The most important observation that we can make while inspecting Fig. 8 is that the solid black line is highly correlated with the solid magenta line. This clearly shows the large impact of the primary beam on the visibility of the anti-ghost. This observation helps explain Stewart’s original result. Stewart observed that the brightness of the anti-ghost was highly dependent on the location of the unmodelled source, since the location of the unmodelled source determines the location of its anti-ghost, which may thus end up in a spot where the primary beam has hardly any gain.

## 7 CONCLUSION

In this paper we showed that spurious symmetrization can be quite bright even if we employ a large number of antennas. We found that two factors working in unison are needed for this to occur: phase-only calibration and the array’s primary beam. We showed that phase-only calibration produces a brighter anti-ghost than full-complex calibration. This already bright anti-ghost can in turn be further amplified by the primary beam correction. We studied both of these factors in detail. It is crucial for us to be cognisant of these results while reducing transient data, in order to prevent the misclassification of artefacts as sources.

We first present the major results we obtained while investigating phase-only calibration. We used extrapolation

(see Sec. 4.3) in conjunction with perturbation analysis (see Sec. 4.2) to derive the two-source phase-only ghost-pattern, which we then compared to the full-complex pattern in Sec. 5. We summarize the main conclusions we made in Sec. 5 below.

**Bright anti-ghost.** Phase-only calibration produces a brighter anti-ghost if we compare it to the anti-ghost which forms when we perform full-complex calibration. The reason for this is that, in the case of phase-only calibration, the anti-ghost is a proto-ghost, while for full-complex calibration it is a deutero-ghost. This can be confirmed by inspecting Table 2.

**Fewer ghosts.** Phase only calibration produces fewer scattered ghosts than full-complex calibration as no deutero-ghosts form when we employ phase-only calibration.

**Color.** We extended the basic per-baseline taxonomy in Paper II in Sec. 4.7. We realized that we can use the sign of a ghost's real amplitude factor to assign one of the two following false colors to it: red and blue. If a ghost is completely real, its color determines whether it manifests as a positive or negative ghost and, consequently, whether it suppresses or amplifies. Blue ghosts are suppressors and red ghosts are amplifiers. The primary suppressor, secondary suppressor and its anti-ghost are all usually real valued which in turn means that only their colour is generally important. The primary suppressor and secondary suppressor are blue which explains why they suppress flux. The anti-ghost of the secondary suppressor is red which explains why it manifests as a positive source during imaging. The fact that the anti-ghost is red is important. If this were not the case, spurious symmetrization would not be possible.

**Symmetric pairs.** The blue proto-ghosts of phase-only calibration and full-complex calibration in general form a symmetrical pair with a red anti-ghost. However, the red anti-ghosts are proto-ghosts in the case of phase-only calibration and deutero-ghosts when full-complex calibration is employed. This explains why the anti-ghosts of the phase-only blue proto-ghosts are brighter than their full-complex counterparts.

To conclude our phase-only calibration summary, we present ghost flux values as predicted by Eq. (4.22) and Eq. (4.48) for different values of  $N$  in Table 3. The reader can use this table to directly compare the ghost responses of full-complex and phase-only calibration and by implication enable them to validate most of the aforementioned conclusions.

We investigated the effect that primary beam correction can have on the anti-ghost in Sec. 6. We found that the primary beam can amplify the brightness of the anti-ghost in two main ways.

**Model accuracy.** If the apparent flux of the unmodelled source happens to be larger than the apparent flux of the brightest source in the field then our calibration model becomes very inaccurate. If we calibrate with a very inaccurate model ( $A_2 \geq A_1$ ) then a very bright anti-ghost is created. We discussed this in detail in Sec. 6.2.2 and Sec. 6.2.5.

**Source positions.** If the relative positions of the modelled and unmodelled source is such that the anti-ghost forms in a region of the primary beam where the attenu-

ation is high, then the anti-ghost is significantly amplified when we perform our beam correction (see Section 6.2.5).

## ACKNOWLEDGEMENTS

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. 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), Kokers Noord and the Province of Drenthe, and the European Community FP7 programme MIDPREP, Grant Agreement PIRSES-GA-2013-612599.

The authors would like to thank Gianni Bernardi and Sandeep Sirothia for the valuable contributions that they made. We would also like to thank the reviewer, Dr Maxim Voronkov, for the useful feedback he provided; his feedback has definitely improved the quality of the paper.

## REFERENCES

- 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
- Dewdney P. E., Hall P. J., Schilizzi R. T., Lazio T. J. L. W., 2009, Proc. of the IEEE, 97, 1482
- Green D., 2011, Bull. Astr. Soc. India, 39, 289
- 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., 2015, in Radio Science Conference (URSI AT-RASC), 2015 1st URSI Atlantic Calibration artefacts: Phase only calibration. pp 1–1
- Heald G., Pizzo R., Orrú E., Breton R., Carbone D., Ferrari C., Hardcastle M., Jurusik W., Macario G., Mulcahy D., et al., 2015, A&A, 582, A123
- Jonas J. L., 2009, Proc. of the IEEE, 97, 1522
- Linfield R. P., 1986, AJ, 92, 213
- Rau U., Bhatnagar S., Voronkov M. A., Cornwell T. J., 2009, Proc. of the IEEE, 97, 1472
- Salvini S., Wijnholds S. J., 2014, A&A, 571, A97
- Smirnov O., Tasse C., 2015, MNRAS, 449, 2668
- 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](http://www.astron.nl/calim2010/presentations/14_Ghostbusters_Smirnov.pdf)
- Smirnov O. M., 2011, A&A, 527, A106
- Stewart A., Fender R., Broderick J., Hassall T., Muñoz-Darias T., Rowlinson A., Swinbank J., Staley T., Molenaar G., Scheers B., et al., 2016, MNRAS, 456, 2321
- 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>

**Table 3.** 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  $B$  denote the number of antennas and baselines in the array respectively. We generated the values in this table by employing two different calibration strategies, namely full-complex and phase-only calibration. We use the abbreviations FC and PO to denote full-complex and phase-only calibration respectively.

Ghost	Approach	Formula fraction of $A_2$	$N$			
			7	14	27	100
Primary	FC	$\frac{1}{N}$	14%	7%	4%	1%
	PO	N/A	N/A	N/A	N/A	N/A
Secondary	FC	$\frac{2}{N} - \frac{1}{N^2}$	27%	14%	7%	2%
	PO	$\frac{1}{N}$	14%	7%	4%	1%
Anti-ghost	FC	$\frac{1}{N^2}$	2%	0.5%	0.1%	0.01%
	PO	$\frac{1}{N}$	14%	7%	4%	1%
Proto	FC	$\frac{1}{B}(\frac{1}{N} - \frac{1}{N^2})$	0.6%	0.07%	0.01%	0.0002%
	PO	$\frac{1}{2BN}$	0.3%	0.04%	0.005%	0.0001%
Deutero	FC	$\frac{1}{BN^2}$	0.1%	0.006%	0.0004%	0.000002%
	PO	N/A	N/A	N/A	N/A	N/A

Swinbank J. D., Staley T. D., Molenaar G. J., Rol E., Rowlinson A., Scheers B., Spreeuw H., Bell M. E., Broderick J. W., Carbone D., et al., 2015, *Astr. and Comp.*, 11, 25  
 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., Grobler T. L., Smirnov O. M., 2016, *MNRAS*, 457, 2331  
 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  
 Yatawatta S., Kazemi S., Zaroubi S., 2012, in *Innovative Parallel Computing (InPar)*, 2012 GPU accelerated non-linear optimization in radio interferometric calibration. pp 1–6

We therefore have

$$\begin{aligned}\Lambda^\dagger &= \text{diag} \left( \left[ 0, \frac{1}{2N}, \dots, \frac{1}{2N} \right] \right) \\ &= \frac{1}{2N} \mathbf{I} - \frac{1}{2N} \text{diag}([1, 0, \dots, 0]).\end{aligned}\quad (\text{A5})$$

Hence

$$\begin{aligned}\mathbf{V} \Lambda^\dagger \mathbf{V}^H &= \frac{1}{2N} \mathbf{V} (\mathbf{I} - \text{diag}([1, 0, \dots, 0])) \mathbf{V}^H \\ &= \frac{1}{2N} \mathbf{I} - \frac{1}{2N^2} \mathbf{1} \mathbf{1}^H,\end{aligned}\quad (\text{A6})$$

which completes our derivation of Eq. (4.15).

## APPENDIX A: DERIVATION OF EQ. (4.15)

In this appendix we derive a closed form expression for the pseudo-inverse of Eq. (4.14). We do this by employing eigenvalue decomposition, i.e. we first find

$$2(N\mathbf{I} - \mathbf{1}\mathbf{1}^H) = \mathbf{V} \Lambda \mathbf{V}^H, \quad (\text{A1})$$

which enables us to compute

$$\left[ 2(N\mathbf{I} - \mathbf{1}\mathbf{1}^H) \right]^\dagger = \left( \mathbf{V} \Lambda \mathbf{V}^H \right)^\dagger = \mathbf{V} \Lambda^\dagger \mathbf{V}^H, \quad (\text{A2})$$

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

It is rather easy to show that

$$\mathbf{V} = \left[ \frac{1}{\sqrt{N}} \mathbf{1}, \mathbf{V}_0 \right], \quad (\text{A3})$$

where the vector of ones has length  $N$  and  $\mathbf{V}_0$  provides an orthonormal basis for the null space of the explicitly stated eigenvector. Substituting Eq. (A3) into (A1) shows that

$$\Lambda = \text{diag}([0, 2N, \dots, 2N]). \quad (\text{A4})$$