Calibration of stations of the low band antenna system of the LOw Frequency ARray (LOFAR-LBA) by Holography

by

Srivardhan Sivadevuni

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Thursday November 12, 2020 at 09:00 AM.

Student number:	4812158	
Project duration:	November 4, 2019 – September 1, 2020	
Thesis committee:	Prof. A. Yarovoy,	TU Delft, supervisor
	Prof. Stefan Wijnholds,	ASTRON, supervisor
	Prof. Richard Hendriks,	TU Delft
	Dr. Michiel Brentjens,	ASTRON

TU Delft

Acknowledgement

This Masters project has been a very memorable and valuable learning experience for me. I would like to acknowledge the following people, without whose help and support I could not have finished this project.

I express my sincere gratitude to my supervisors Prof. Alexander Yarovoy and Prof. Stefan Wijnholds for their guidance, invaluable support and motivation during this project. I am grateful for your inspiring vision, constructive feedback, extensive knowledge and great patience throughout this thesis, and especially during these exceptional times of COVID-19.

I thank my daily supervisor Prof. Jianping Wang for his feedback and assistance during this project. I appreciate your calmness and constructive feedback. I extend my sincere thanks to Prof. Sander ter Veen for his assistance, valuable inputs and for performing the LOFAR measurements needed for this project. I thank the people of the Department of Microwave Sensing, Signals and Systems at TU Delft for creating a welcoming environment and providing a pleasant workplace during this project. I extend my gratitude to my colleagues at ASTRON for their support, providing a pleasant environment and a friendly workplace during this project.

I thank TU Delft and ASTRON for providing me with this exemplary opportunity and setting up a platform for an outstanding learning experience. I thank my friends for their support throughout this journey. I thank all my family members for always being there for me. Finally and most importantly, I thank my parents and my brother for making this journey possible. I would not be here without your sacrifices and support.

Abstract

The LOw Frequency ARray (LOFAR) radio telescope in the Netherlands covers an area of 300,000 sq.metre. LOFAR is considered as a pathfinder to the new and the largest radio telescope under construction, the Square Kilometer Array (SKA) covering an area of 1000,000 sq. metre. The similarities between the systems of the LOFAR and the SKA allow extension of calibration strategies from LOFAR to SKA. The calibration strategy employed for the LOFAR radio telescope assumes identical embedded element patterns (EEPs) among the antenna elements, which leads to systematic calibration errors.

The aim of this thesis is to investigate the impact this assumption on calibration errors for the low band antenna system of the LOFAR radio telescope (LOFAR-LBA). To this end, a simulation of calibration of the stations of the LOFAR-LBA system by holography is developed and implemented during this thesis. Results from the simulation are compared to investigate the impact of calibration errors with respect to the systemic errors calculated from a LOFAR-LBA observation.

This work starts with analysing the similarities between the LOFAR and the low frequency system of the SKA (SKA-low) that allow the extension of calibration strategies from the former to the later. After identifying these similarities, the reason and effect of the errors caused by the assumption of identical EEPs is discussed in detail with regards to the LOFAR system. Certain novel calibration methodologies in the form of calibration by holography, superstation calibration and stand alone station calibration are analysed, with the possibility of inclusion of individual EEPs as a priority, to facilitate comparison of computational costs involved in their respective implementations based on systemic structure of LOFAR.

Holography is identified and chosen as the computationally optimal methodology for development of an instrument model with the inclusion of individual EEPs and implemented both as a simulation and on LOFAR-LBA measurement data. Holography on LOFAR-LBA is first implemented through simulation within the observed range of the signal-to-noise ratio from the measurement data. Although results clearly indicate errors in calibration after including individual EEPs in the instrument model, it was identified that these errors are swamped by the noise in the system owing to the low signal-to-noise ratio in the observed data. An attempt was made to implement holography on LOFAR-LBA measurement data and several issues with the data were identified. The identified issues with LOFAR-LBA data were resolved and the code from the implementation of holography on LOFAR-HBA is made compatible to LOFAR-LBA data.

Improvements in calibration results were expected with increasing SNR during the simulation. However, it was observed that the incorporation of individual EEPs into the instrument model to perform holography resulted in gain estimates very close to the gain estimates with the assumption of identical EEPs. This illustrates that while incorporation of individual EEPs does have an impact on the calibration results, the computational effort is not justified compared to the implementation with identical EEPs.

List of Figures

List of Tables

Contents

1
Introduction

With increased access to new windows in the electromagnetic spectrum, researchers in the last few decades have discovered quite a few astonishing facts about the unravelling universe. There have been significant discoveries about the most extravagant happenings even in the farthest reaches of our universe in the last few years which can be attributed to decades of research and extra-terrestrial observations using radio waves, infrared signatures, ultraviolet rays, x-rays, optical waves and γ-rays. However, one part of the electromagnetic spectrum that has not been given much attention pertains to the lower frequencies in the spectral window, below a few hundred MHz, until the emergence of low-frequency aperture arrays such as the LOFAR.

During the early stages of growth of radio astronomy, frequencies of about a few hundred MHz were used in detecting some very significant extra-terrestrial phenomena such as cyclotron radio emission from Jupiter [3], discovery of the cosmic microwave background [4], discovery of pulsars and study of radio galaxies and quasars. The detection of the 21-cm line transition in hydrogen at 1420 MHz [5] (shown in figure 1.1) propelled radio observations at higher frequencies.

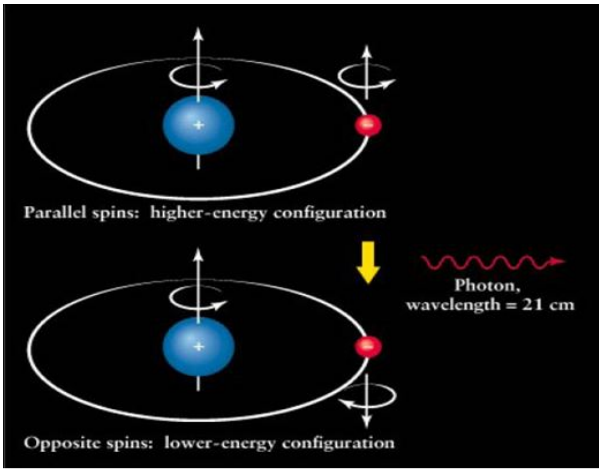


Figure 1.1: Energy released when an electron in hydrogen atom flips it's rotation from parallel spin (to the proton spin) to anti-parallel observed at 1420MHz. Figure taken from [1].

The increasing need for higher angular resolution and developments in receiver technology, interferometry, aperture synthesis, continental and intercontinental very long baseline interferometry (VLBI) have further inclined the researchers towards high-frequency end of the radio spectrum.

However, the astronomical realisations during 1980s and 1990s resurrected low-frequency radio astronomy. Observations such as the inverted radio spectra in radio sources, steep spectra in pulsars and high red-shift radio galaxies need low-frequency observations. Another leap in low-frequency radio astronomy occurred with the development of sky surveys [6] using low-frequency receiver systems during the same period.

Radio astronomers around the globe began to look towards a bigger collaborative goal of detecting hydrogen at cosmological distances which requires a large collecting area of about a square kilometer, later came to be known as the Square Kilometre Array (SKA) [7] [8].

Although several other large antenna array radio telescopes existed before the proposed SKA such as the Long Wavelength Array (LWA) in New Mexico [9], the Murchison Widefield Array (MWA) in Australia [10], the LOw Frequency ARray (LOFAR) in the Netherlands [11] and the Giant Meterwave Radio Telescope (GMRT) in India [12], SKA was envisioned to be the largest radio telescope ever built that can amplify the power of radio telescopes by many a fold.

Construction and operation of the SKA was also deemed possible with the developments in digital electronics, high performance computing, storage capacity and fibre-optics data networks. The SKA is expected to be operational by 2025 which makes the calibration and design ideas for its operation timely.

A low-frequency radio telescope with extreme sensitivity and angular resolution is operational in the form of the LOw Frequency ARray (LOFAR) [11], operating between 10-240 MHz frequencies, having the closest attributes to that of the low-frequency receiving system of SKA (SKA-Low). Because of the similarities in LOFAR and SKA-low systems, improvements in LOFAR calibration can be used to improve the calibration strategy of the SKA-low system. Calibration of LOFAR can hence be considered as a pathfinder to the more ambitious SKA since it has been operational in the Netherlands and across other parts of Europe from 2010 [11] [13], providing a receiving area of 300,000 square meter as opposed to 1,000,000 square meter area of the SKA.

The motivation behind this thesis is to improve the accuracy of calibration of LOFAR by reducing the systematic errors caused by the underlying assumptions made in the calibration strategy.

This chapter introduces phased array antennas starting with linear arrays in section 1.1 followed by 2-D arrays in section 1.2. Section 1.3 describes the current LOFAR calibration strategy and its accuracy issues based on the underlying assumptions. The motivation of the thesis is introduced in the form of problem statement and followed by the outline of the thesis in section 1.4.

1.1. Linear arrays

Linear arrays form the basic building block for most communication applications. They are used to meet requirements such as improvement of signal to noise ratio, improvement in the direction dependent response of the signal, null steering and beamforming.

A linear array can be formed as a collection of antenna elements equally spaced along a straight line as shown in figure 1.2. The antenna elements are assumed to be omnidirectional.

Consider a single source illuminating the linear array. Assuming that the source is celestial and hence located at a distance much farther than the size of the array, the assumption that the signal reaching the array is a plane wave can be made.

The received signal vector of the linear array is the product of the attenuated source signal, gain of the antennas and the geometric delay towards the source with respect to antenna positions. The attenuation of the source signal is ignored here for the sake of simplicity.

Assuming that the source signal is narrow band, centered at frequency f_c, geometric delays can be described by phases that are a function of frequency f_c, corresponding to the wavelength λ, and distance of each antenna from a reference point, referred to as the phase reference centre of the station.

The mathematical expression for geometric delay τ_j and received signal for any antenna j from the source signal $s(t)$ positioned at an elevation of θ with respect to the uniform linear array are given by equation 1.1a and equation 1.1b respectively. Similarly the received signal vector for the linear array can be obtained by stacking the received signals at all the antennas as shown in equation 1.1c [2].

$$\tau_j(\theta) = \frac{\Delta_j sin(\theta)}{c} \tag{1.1a}$$

$$x_j(t) = s(t) g_j(\theta) e^{-2\pi\tau_j f_c} + n_j(t) \tag{1.1b}$$

$$\mathbf{x}(t) = \mathbf{G}(\boldsymbol{\theta}) s(t) \begin{bmatrix} e^{-2\pi f_c \tau_1(\theta)} \\ \vdots \\ e^{-2\pi f_c \tau_J(\theta)} \end{bmatrix} + \mathbf{n}(t) = \mathbf{G}(\boldsymbol{\theta})\; s(t)\mathbf{a}(\theta) + \mathbf{n}(t) \tag{1.1c}$$

In above equations, c is the speed of light, $g_j(\theta)$ is the gain of antenna j towards direction θ, $\mathbf{G}(\boldsymbol{\theta}) = diag[g_1(\theta), g_2(\theta), ..., g_J(\theta)]$ is the antenna gain matrix for the linear array, $n_j(t)$ is the noise in the system for antenna j, $\mathbf{n}(t)$ is the noise vector for the linear array, $(\boldsymbol{\tau}) = [\tau_1, \tau_2, ..., \tau_J]^T$, $\mathbf{a}(\boldsymbol{\theta}) = [e^{-2\pi\tau_1(\theta) f_c}, ..., e^{-2\pi\tau_J(\theta) f_c}]^T$ and Δ_j is the distance of antenna j from antenna J (chosen as reference here) from figure 1.2.

The source signal $s(t-\tau_j)$ at antenna element j is approximated by $s(t)$ as in equation 1.1b because of the narrow-band assumption. In practice, the received radio frequency signal $\tilde{x}_j(t)$ is not narrow band and is first converted into base-band signal $x_j(t)$, sampled and split into narrow subbands to satisfy the narrowband assumption.

Uniform linear arrays:

A uniform linear array (ULA) is a linear array with the antennas in the array equidistant from each other. This changes the received signal vector in equation 1.1c as the spacing Δ between every two consecutive antennas is equal and subsequently simplifies beamforming.

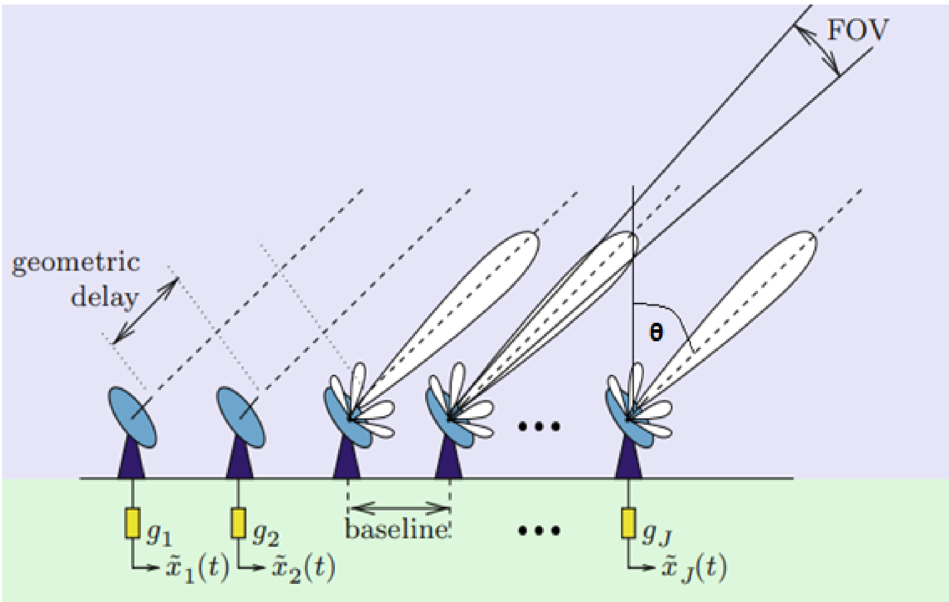


Figure 1.2: Linear array of J antenna elements. Figure taken from [2].

In practice, the influence of antennas in the array on each other in the form of mutual coupling is neglected for the sake of computational simplicity, leading to the assumption that antenna elements in the array are identical. This assumption causes serious errors in phased array systems if the arrays are too compact.

1.2. Two dimensional arrays

Mathematical expressions for the received signals of a linear array can be extended to a 2-dimensional array. The assumption of a single source is held along with another reasonable assumption that the antennas lie on the x-y plane, restraining the array to two dimensions. The origin in (x,y,z) coordinate system is assumed to be at the centre of the station. The time delay $\tau_j(t)$ for a signal arriving at antenna j from a single source located at direction $\mathbf{l}(t)$ is given by equation 3.1,

$$\tau_j(t) = \frac{\mathbf{p}_j \cdot \mathbf{l}(t)}{c}, \tag{1.2}$$

where $\{\cdot\}$ represents inner product, c is the speed of light, $\mathbf{p}_j$ is the position of antenna j with respect to origin and $\mathbf{l}(t)$ is the unit vector pointing from the phase reference centre of the station (at origin) to the source $s(t)$. Since the source in radio astronomy applications is usually celestial, the position of the source is recorded with time to counter Earth rotation.

The position $\mathbf{l}(t)$ of the source is determined in the (l, m, n) coordinate system as a function of it's positions in zenith θ and azimuth α as:

$$\mathbf{l}(t) = [l = sin(\theta)cos(\alpha),\ m = sin(\theta)sin(\alpha),\ n = cos(\theta)],$$

such that $l^2 + m^2 + n^2 = 1$.

The time delay at antenna j for multiple sources can be obtained by stacking the time delays for each source into a vector. For the sake of simplicity, received signal vector model is estimated for a single narrowband frequency channel using a single source signal similar to the assumption made for linear arrays.

This gives us equation 1.3a for signal received at antenna j in the randomised 2-D array, similar to that of the expression in equation 1.1b.

$$x_j(t) = s(t)g_j(\theta)e^{-2\pi\tau_j f_c} + n_j(t) \tag{1.3a}$$

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_J(t) \end{bmatrix} = \mathbf{G}(\boldsymbol{\theta})\ s(t)\mathbf{a}(\boldsymbol{\theta}) + \mathbf{n}(t) \tag{1.3b}$$

Stacking the received signals for all the elements into a single vector gives us the received signal vector for the array with random antenna layout as shown in equation 1.3b. Figure 1.3 and figure 1.4 show randomised array layout of the inner and outer 48 antennas in a LOFAR-LBA station respectively. Neglecting the mutual coupling between antennas in such layouts lead to systematic errors. A more elaborate introduction to the problems caused by this assumption in station calibration for LOFAR and the need to arrive at an improved calibration strategy is given in section 1.3.

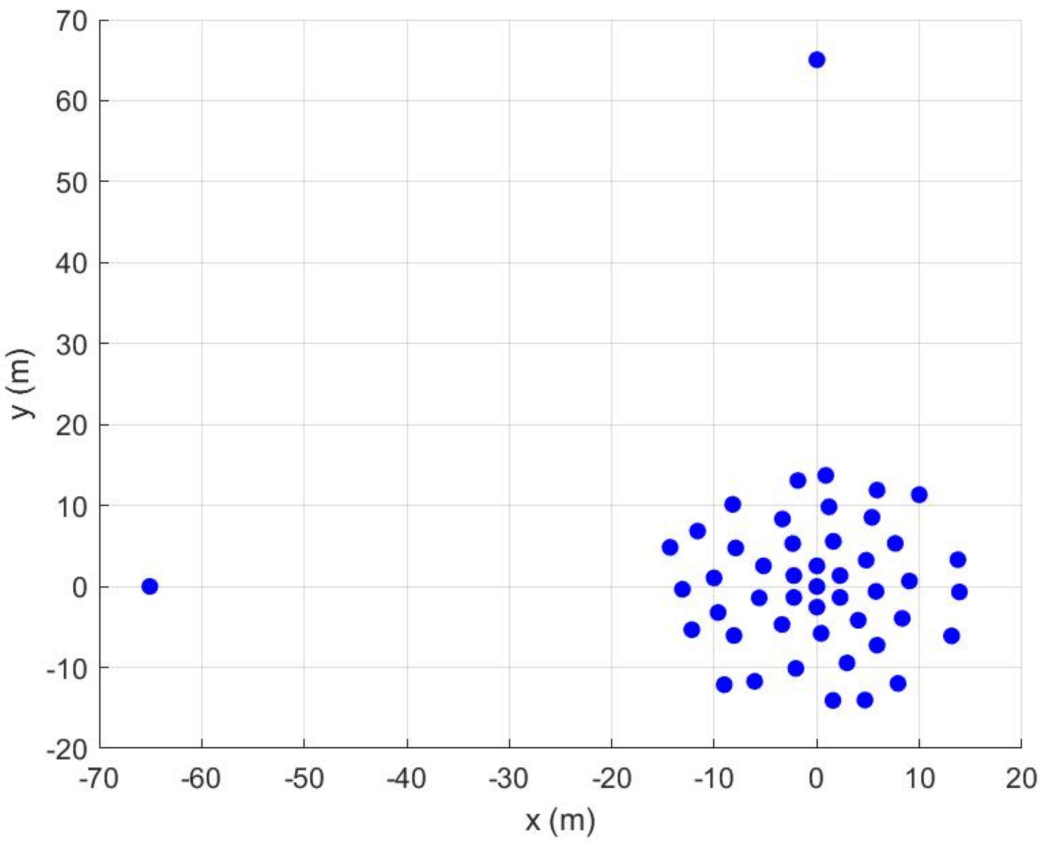


Figure 1.3: LOFAR-LBA inner station layout.

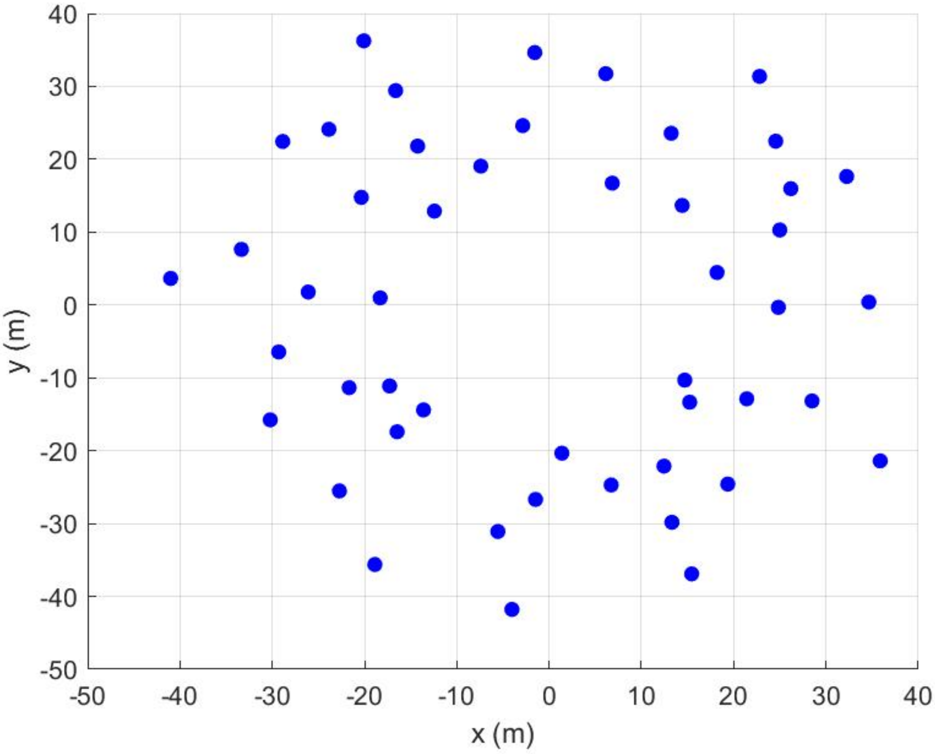


Figure 1.4: LOFAR-LBA outer station layout.

1.3. Motivation

1.3.1. Current LOFAR station calibration approach

Most low-frequency astronomical observatories possess several stations (or subarrays) distributed over a large area, each having receiver elements arranged in random or in patterns. The radio frequency (RF) path is specific for each element based on it's individual gain, direction of arrival of the source signal, position of the element in the station, mutual coupling between the receiver elements and atmospheric effects. The RF path calibration of each element is essential for accurate beamforming and imaging.

The goal of station calibration is to accurately estimate the station beam by estimating the antenna gains in the station accurately. A station beam is formed by assigning complex valued weights to the antenna beam patterns, as a function of frequency. The complex weights are assigned to coherently add the signals originating from sources in a desired direction while also suppressing radio frequency interference (RFI).

The complex valued weights, also regarded as the beamformer weights, should rectify the geometric delays over the array taking the antenna positions, antenna orientations, array positions and array orientation into account. The modified station beam should also suppress radio frequency interference (RFI), taking into account the side lobes and phase, gain variations. Finally, the system noise as a function of frequency has to be accurately estimated to improve the station beam [14].

Operation of LOFAR involves multiple stages of processing, each performing the aforementioned steps for calibration. An overview of the steps involved in LOFAR station calibration is given below.

Currently LOFAR station calibration is carried out in four steps [14]:

1. **Flagging of RFI affected data**

   The calibration data is compared to a source model consisting of celestial sources to detect RFI using Frobenius norm of the array covariance matrix. Spatial and spectral characteristics of subbands affected with RFI have a higher Frobenius norm than the surrounding RFI free subbands. This distinction is used to identify subbands containing RFI.

2. **Estimation of calibration parameters**

   The received signal at antenna j from equation 1.3a for multiple sources is calculated as a superposition of received signals from all the sources, given by equation 1.4. Furthermore, the antenna gains in LOFAR stations can be presented as a product of direction dependent and direction independent components as shown in equation 1.4 as opposed to the single gain parameter model established in equation 1.3a.

   $$x_j(t) = q_j \sum_{l=1}^{L} g_{lj} a_{lj}(t) s_l(t) + \sum_{l=1}^{L} n_{lj}(t) = q_j \sum_{l=1}^{L} g_{lj} a_{lj}(t) s_l(t) + n_j(t), \tag{1.4}$$

   In equation 1.4, q_j is the direction independent complex valued gain of receiver j, g_{lj} is the direction dependent gain for antenna j towards l^{th} source (L sources in total indicate L directions), $a_{lj} = e^{\frac{2i\pi}{\lambda}\mathbf{l}(t)\cdot\mathbf{p}_j}$ is the geometrical delay at the j^{th} receiver positioned at $\mathbf{p}_j$ looking towards source $s_l(t)$ at $\mathbf{l}(t)$ and $n_{lj}(t)$ is the receiver noise signal for antenna element j corresponding to received signal from $s_l(t)$. The noise component of $n_j(t)$ is a combination of thermal noise at antenna j and noise from all directions in the sky.

   Assuming the signals are narrowband and a single source $s(t)$ at $\mathbf{l}(t)$, the received signal vector $\mathbf{x}(t)$ from equation 1.5 can be obtained by stacking the J antenna signals as $\mathbf{x}(t) = [x_1(t), ..., x_J(t)]^T$, $\mathbf{Q} = diag([q_1, ...q_J]^T)$ as the $J \times J$ diagonal matrix with direction independent gains of the J elements along the diagonal, $\mathbf{g} = [g_1, ..., g_J]^T$ as the direction dependent gains towards the source, $\mathbf{n}(t) = [n_1(t), ..., n_J(t)]^T$ is the $J \times 1$ noise vector and the geometrical delays stacked in the $J \times 1$ array response vector $\mathbf{a}_l(t) = \mathbf{a}(t) = [a_1(t), ...a_J(t)]^T$.

   $$\mathbf{x}(t) = \mathbf{Q}(\mathbf{a}(t) \odot \mathbf{g}) s(t) + \mathbf{n}(t) \tag{1.5}$$

   The entry-wise product [15] indicated with $\odot$ in equation 1.5 can be removed by changing $\mathbf{A} = diag(\mathbf{a}(t))$ as the $J \times J$ array response matrix and $\mathbf{G} = diag(\mathbf{g})$ as the $J \times J$ direction dependent gain matrix, changing equation 1.5 to equation 1.6.

   $$\mathbf{x}(t) = \mathbf{QAG} s(t) + \mathbf{n}(t) \tag{1.6}$$

   Equation 1.6 is used to obtain the $J \times J$ array covariance matrix $\mathbf{R}$ given by equation 1.7 where $\mathscr{E}\{\cdot\}$ is the expectation operator.

   $$\mathbf{R} = \mathscr{E}\{\mathbf{x}(t)\mathbf{x}(t)^H\} \tag{1.7}$$

Equation 1.6 is also used as reference to the data model for the received signal vector in LOFAR stations for deriving and implementing the calibration parameters in the rest of this report.

The calibration problem is then formulated as a weighted least squares problem with the knowledge of the measured array covariance matrix $\hat{\mathbf{R}}$ at the station correlator as

$$\{\hat{\mathbf{g}}, \hat{\mathbf{q}}, \hat{\sigma}_n\} = \underset{\mathbf{g},\mathbf{q},\sigma_n}{argmin} ||\mathbf{W}(\hat{\mathbf{R}} - \mathbf{R})\mathbf{W}||_F^2, \tag{1.8}$$

where the weighting matrix $\mathbf{W}$ is the covariance matching matrix and σ_n is the real valued noise variance parameter from the noise covariance matrix $\mathbf{\Sigma}_n$ rewritten as vec($\mathbf{\Sigma}_n$)=$\mathbf{I}_s\sigma_n$, for $\mathbf{I}_s$ as a selection matrix such that $\mathbf{\Sigma}_n = \mathscr{E}\{\mathbf{n}(t)\mathbf{n}(t)^H\}$ [14]. The solution to the equation 1.8 can be obtained by the weighted alternating least squares(WALS) approach for current implementation model by neglecting embedded element pattern variations [16]. Note that the weighted least square solution is only applicable if the calibration model follows the assumption of identical EEPs.

3. **Identifying incorrect solutions**

   An outlier threshold is determined from data monitoring campaigns using signal processing algorithms. The determination and imposition of this threshold on the calibration results should be very accurate to ensure rejection of all erroneous results. Once the threshold is determined, the erroneous results are identified.

4. **Passband model fitting**

   The identified erroneous results are rejected and the subbands containing either the RFI or for which no acceptable calibration solution can be found are excluded from further processing. Correction factors are obtained from the calibration results left after solution based flagging from the previous step by fitting the left over subbands with a band pass model.

The calibration process assumes that the lines of sight of all the antennas in the array encounter the same ionosphere. This might be true for small array configurations over an even smaller duration of observation. Furthermore, the direction dependent gains are assumed not to succumb to mutual coupling among the antennas. This is based on the assumption that the embedded element patterns (EEPs) of the array elements are identical and it is not taken into account while estimating calibration factors.

These two assumptions cause significant systematic errors that increase as the array compactness increases. Ionospheric scintillation may affect the calibration accuracy drastically, especially at low frequencies. The assumption that the spectral characteristics of the diffuse emission model are same across the sky has also been identified to contribute to erroneous calibration results [17].

The differences in calibration results caused by these assumptions are described in the following subsection 1.3.2 [18].

1.3.2. Accuracy issues

Assumption of identical EEPs:

Equation 1.4 can also be mathematically described as shown in equation 1.9. The RF path output voltage x_j of antenna j positioned at $p_j(t)$ in the array, having overall complex valued gain g_j, antenna response vector $E_j(l,t)$ in l directions, as a response to a source $s_l(t)$ is mathematically given by equation 1.9 [18],

$$x_j = g_j \int E_j(l,t) \cdot s_l(t)\, e^{j\frac{-2\pi}{\lambda}\mathbf{l}(t)\cdot\mathbf{p}_j(t)}\, dl \tag{1.9}$$

The antenna response vector $E_j(l,t)$ from equation 1.9 is relaxed by assuming that the embedded element patterns (EEPs) between the receiving elements are identical and is modified as $\tilde{E}(l,t) = \frac{1}{J}\sum_{j=1}^{j=J} E_j(l,t)$. Equation 1.9 is then simplified to obtain equation 1.10.

$$x_j \;=\; g_j \int \tilde{E}(l,t) \cdot s_l(t)\, e^{j\frac{-2\pi}{\lambda}\mathbf{l}(t)\cdot\mathbf{p}_j(t)}\, dl \tag{1.10}$$

Although the relaxation of the term $E_j(l,t)$ simplifies the calibration problem, it also significantly increases the systematic calibration errors with increasing differences between the EEPs. The margin of errors as a result of this assumption is significant as demonstrated in [18] [19]. The relaxation is an indication that the current calibration approach does not take the mutual coupling between the receiving elements into account, assuming that the directional response between them is identical [20].

Inaccuracies in source model:

Assumptions made in the source model can lead to equally significant inaccuracies in calibration results. The diffuse emission model of the sky is treated to have the same structure as observed in the 408 MHz Haslam survey. Deviations in the flux ratio in different parts of the sky have identified this being incorrect, pointing out the erroneous assumption in the source model [17].

Similarly, the impact of ionospheric fluctuations on flux towards lowest frequencies has been underestimated as it has been observed to worsen particularly at frequencies below 100 MHz.

These assumptions and their effects indicate that there is a scope for significant improvement in LOFAR calibration in both instrument model and source model for single-station calibration while another potential solution could be in the form of calibration using multiple sources simultaneously.

1.4. Problem statement and outline of thesis

Considering that the assumption of identical EEPs introduces calibration errors in LOFAR-LBA and that these errors are prone to increase with increasing compactness and antenna elements of stations, which is evidently forthcoming in case of SKA-low, the following set of objectives are laid down to form the motivation of this thesis:

1. to identify computationally efficient calibration approaches that can incorporate the embedded element patterns (EEPs) for LOFAR-LBA calibration
2. to identify the computationally optimal methodology among the proposed calibration approaches
3. to compare the results from objective 5 with results from implementation of proposed methodology through simulation
4. to implement the proposed calibration strategy on a LOFAR-LBA measurement data
5. to compare the impact of the systemic errors with the calibration errors using LOFAR-LBA measurement data in order to validate the applicability of the proposed calibration approach

1.4.1. Proposed solutions

Few proposed improvements such as calibration of multiple stations simultaneously either by forming clusters of stations as superstations or by holography are innovative and theoretically have many advantages. These techniques displayed tolerance towards local RFI and have been observed to be fast and accurate, enabling easy data reduction. One other straightforward solution is to improve the calibration model of each station. Improving the calibration model involves corrections in both the instrument model and source model for each station individually, which is an arduous task.

The merits and demerits of these proposed improvements are prioritised and one of them is implemented on LOFAR data. It is essential to observe the underlying structure in the measurement model for LOFAR calibration described in section 2.1 before looking into feasibility of suggested improvements in the later part of chapter 2. Implementation model of computationally optimal approach among these proposed solutions based on requirements and merits is presented in chapter 3. Implementation of the model presented in chapter 3 through simulation and observations from the simulation are presented in chapter 4. Results from the implementation of the proposed approach at the end of chapter 3 on measurement data are reported in section 5. A validation of the results and analysis from chapter 5 is done with results from implementation on simulation model in chapter 4, followed by conclusions in chapter 6.

2
Station calibration and proposed improvements

LOFAR provides us with a tremendous opportunity as a pathfinder for the SKA-low system calibration. Before looking further into the proposed solutions to improve the calibration strategy for LOFAR and thereby for the SKA-low, it is essential to understand the components, layout and the signal processing pipeline for LOFAR.

A brief description of LOFAR system design is presented in section 2.1 followed by the SKA design in section 2.2. The similarities between LOFAR and the low-frequency system of SKA that allow us to use the improved calibration strategies from LOFAR in SKA-low are discussed in section 2.2.1.

The proposed solutions: LOFAR calibration by holography, calibration of LOFAR superstations and stand-alone singe station calibration with improved source model; are discussed in section 2.3.1, section 2.3.2 and section 2.3.3 respectively.

Section 2.4 describes the signal processing stages implemented for LOFAR operation. This allows an evaluation of the proposed solutions based on their computing demands to determine their compatibility.

A comparison of these proposed solutions with regards to compatibility with LOFAR using the LOFAR signal processing pipeline is done in section 2.5. The choice of an optimal strategy among the proposed solutions is made based on computational costs and development effort to conclude chapter 2. The choice made at the end of this chapter will be further studied for the remainder of the thesis.

2.1. LOFAR design

LOFAR is a multipurpose radio telescope employing phased array technology. It consists of an array of 38 dipole antenna stations distributed in the Netherlands and 14 similar international stations in UK, France, Sweden, Ireland, Poland and Germany. The operational range of LOFAR is from 10-240 MHz, split into two receiving systems: the low band antenna from 10-90 MHz and the high band antenna operating from 110-240 MHz.

The antenna arrays in LOFAR are used as radio interferometers [21] and with the all-sky coverage of these omni-directional antennas, LOFAR is capable of a large field-of-view (FoV). The antennas eliminate any moving parts and can virtually form a conventional telescope dish through beamforming.

Beams towards a specific direction can be formed by digital beamforming at all the antennas, making the system agile and allowing rapid repointing towards another direction. In this way the LOFAR system is capable of making observations in multiple directions simultaneously and can also form beams for multiple stations towards the same source.

2.1.1. LOFAR station layout

Out of the 38 stations distributed in the Netherlands, 24 stations are located within a 2km radius of a village called Exloo and are referred to as *core* stations. These core stations are optimally distributed to achieve specific key scientific goals of the LOFAR such as the epoch of reionization (EoR) observations and search for radio transients. 14 other stations are distributed within a 90km radius in the Netherlands and are referred to as *remote* stations.

Each of the 24 core stations has one LBA substation (having 96 LBA dipoles out of which 48 can be used at the same time) and two HBA substations (each with 24 tiles) as shown in figure 2.1 [11]. All the remote stations have 96 LBA dipoles and a single HBA subsystem with 48 tiles, unlike the two HBA subsystems with 24 tiles each in the core stations. Nevertheless, both the core and remote stations operate with a total of 96 receiver units (RCUs) and hence 96 digital signal paths.

In total, each station in the Netherlands has 96 signal paths and can be used to simultaneously process signals from either 48 dual-polarized or 96 single-polarized antennas.

Figure 2.1: The aerial view of the heart of LOFAR core site at Exloo, Netherlands, shows the circular island called the superterp. The superterp has 6 LOFAR stations with 1 LBA subsystem and 2 HBA subsystems each and is surrounded by 18 other *core* stations.

The 14 international LOFAR stations have 96 HBAs and 96 LBAs along with 192 RCUs. The increased number of RCUs in these international stations provide a total set of 192 digital signal paths allowing a full set of HBA tiles or LBA dipoles to operate for any given observation.

Each tile in HBA subsystems is optimized to operate in the 110-240 MHz range and consists of 4 x 4 element (dual polarized) antennas with built-in amplifiers. They are also equipped with an analog beamformer consisting of delay units and summators. The tiles in the HBA subsystems are regularly arranged unlike the irregularly placed aperture arrays in the LBA. The difference in mutual coupling effects among the antennas in the station are significant for the LBA subsystem which makes it reasonable to validate the calibration improvements for the LBA system before the HBA system.

Nevertheless, improvement in calibration strategy for the LBA subsystem corresponding to the assumption of mutual coupling between the antenna elements can also be applied to the HBA subsystem since it can be shown that mathematical expressions for calibration strategy for LBA and HBA are equivalent.

2.2. SKA design

Phase 1 of the Square Kilometre Array (SKA) is expected to be operational by 2025 and has an aperture of 1,000,000 sq.m operating over the 50 MHz - 20 GHz frequency range. The two sites of construction for the SKA are: the Karoo region of central Africa and the state of Western Australia respectively [22].

Each operating band within the range of 50 MHz- 20 GHz requires a different technological solution for almost all design parameters. In this context, similar to LOFAR, SKA is divided into SKA-low, proposed to operate in 50 MHz - 350 MHz range, and SKA-mid, proposed to operate in 350 MHz - 4 GHz [22].

The proposed station layout for the SKA-low system has an irregular aperture array configuration. The massive SKA-low system is being designed with 512, 35m diameter stations each having 256 log-periodic dipoles [23]. The most notable observation that can be made about the SKA system is that the SKA-low system is analogous to the LOFAR radio telescope in the Netherlands while the SKA-mid system is analogous to the Westerbork Synthesis Radio Telescope (WSRT) and the Very Large Array (VLA). Any calibration improvements with the LOFAR, WSRT and the VLA can likely be applied to SKA calibration [24] [25].

2.2.1. Similarities in LOFAR and SKA-low station layout

The emphasis in this context is more on the low band antenna (LBA) system of LOFAR. Similar to the array layout in the LOFAR-LBA, SKA-low system has sparse aperture arrays. Any improvements in the calibration strategy of LOFAR-LBA with regards to the assumption of identical EEPs are expected to lead to improvements in SKA-low station calibration. The improvements, in fact, can potentially be even more profound for the much more compact SKA-low stations compared to LOFAR-LBA system [18].

2.3. Approaches to improve station calibration

The proposed approaches in the form of holography, superstation calibration and standalone single station calibration calibration by improving the source model are briefly described below with the context of number of computations required to calculate the array covariance matrix for calibration following equation 1.8.

2.3.1. Holography

Holographic measurements use one or more reference stations for calibration of the station under test (s.u.t). The s.u.t. forms beams towards each direction in the sky. The received signals at the s.u.t. towards all B beams are correlated with the reference signal(s) [26] from the calibration beam of the reference station in holography unlike the auto-correlations in other calibration methods.

The B beam received signals (following the model given by equation 1.6) are correlated with the reference signal resulting in the array covariance vector (unlike the array covariance matrix from equation 1.7) used for calibration in equation 1.8.

In general, the array covariance matrix required for calibration of a station with J antennas beams requires ${}^{J}C_2$ [1] computations while it requires $B+1$ computations alone to calculate the the array covariance vector in case of holography for each subband.

With Δf_{signal} as the signal bandwidth, the number of computations required per station to generate the correlation data is hence $B\Delta f_{signal}$ and $BJ\Delta f_{signal}$ to calculate the beamformed data required per station for calibration.

2.3.2. Superstation Calibration

The AARTFAAC facility operating alongside LOFAR in real time can be considered as a superstation [27]. Signals from all the individual elements of all the stations form the superstation. The correlation matrix requires crosscorrelations between JR elements for the entire superstation for each subband. Superstation calibration thus has ${}^{JR}C_2$ array covariance matrix, J being the array size and R being the number of stations forming the superstation. This can, hence, be considered as stand-alone calibration of all the individual elements of all the stations.

The number of computations per station, hence required to calculate the array covariance matrix is given by $\frac{{}^{JR}C_2\Delta f_{signal}}{R}$.

2.3.3. Stand Alone Single station calibration by improving the source model

The diffuse emission of the sky can be meticulously modelled and included in the calibration model of each station, which also includes improving the instrument model by taking individual EEPs into account. The number of computations for the array covariance matrix is ${}^{J}C_2$ for an array of size J. However, accounting the embedded element pattern for each antenna into the instrument model for each station requires higher computations for the calculating the array covariance matrix.

The computational complexity further increases with the inclusion of the diffuse emission model of the sky. The number of computations per station required to calculate the array covariance matrix without the inclusion of individual EEPs and the diffuse emission model are ${}^{J}C_2\Delta f_{signal}$.

Radio telescopes such as the LOFAR have humongous systems that are distributed into stages to facilitate their operation. Implementing one of these proposed methodologies, along with the improvements needed to include individual EEPs has to be sporadically followed over the involved stages.

[1] ${}^{J}C_r = \frac{J!}{(J-r)!\,r!}$

2.4. LOFAR signal processing pipeline

The LOFAR digital signal processing pipeline has three operational stages as shown in figure 2.2 [11]. The flow of information through the three signal processing stages is used to calculate the computational costs involved in each of the alternatives considered for station calibration.

Figure 2.2: LOFAR signal processing schematic.

2.4.1. Station processing

LOFAR system has station processing facility at each station site to process the raw data received from each antenna or tile. The received signal for an antenna j in the station and received signal vector for each station while observing a single source for a single polarisation and a single narrowband frequency f_c can be mathematically modelled as in equation 2.1 and equation 2.2 respectively, similar to equation 1.4,

$$x_j(t) = q_j a_j g_j s(t) + n_j(t) \tag{2.1}$$

$$\mathbf{x}(t) = \mathbf{QGa}(\boldsymbol{\theta})s(t) + \mathbf{n}(t), \tag{2.2}$$

where $\mathbf{G} = diag[g_1, ..., g_J]^T$ is the $J \times J$ diagonal matrix with direction dependent antenna gains, $\mathbf{Q} = diag[q_1, ..., q_J]^T$ is the $J \times J$ diagonal matrix with direction independent antenna gains for J antennas, $s(t)$ is the time-dependent source signal, $\mathbf{a}(\boldsymbol{\theta})$ is the $J \times 1$ array response vector given by equation 2.3 and $\mathbf{n}(\mathbf{t}) = [n_1(t), ..., n_J(t)]^T$ is the $J \times 1$ vector representing uncorrelated noise at the station receivers.

In equation 2.3, $a_j(\theta) = e^{\frac{-2i\pi f c}{c} \mathbf{l}(t) \cdot \mathbf{p}_j}$ is the geometrical delay at the j^{th} antenna positioned at $\mathbf{p}_j$ looking towards source $s(t)$ at $\mathbf{l}(t)$ and c is the speed of light.

$$\mathbf{a}(\boldsymbol{\theta}) = [a_1(\theta), ..., a_J(\theta)]^T \tag{2.3}$$

The signal bandwidth Δf_{signal} after sampling at Nyquist rate generates twice the number of samples $2\Delta f_{signal}$ per second.

The raw received signals for each station modelled as in equation 2.2 are transmitted from the station processing facility to the next signal processing stage through high speed fibre optic channels either as:

- data for the J antenna dipoles forming B beams in case of holography. This generates B output streams of data to be transmitted from station facility to the next stage for each of the two polarisations and hence requires a streaming capacity of $4B\Delta f_{signal}$ samples per second.

- data from each antenna element in case of superstation calibration. This requires J output streams of data and hence requires a streaming capability of $4J\Delta f_{signal}$ samples per second for two polarisations.

- no correlations are streamed to the central signal processing stage since the crosscorrelations are calculated at station itself in case of stand-alone station calibration.

2.4.2. Central signal processing

The central signal processing (CSP) stage of the LOFAR signal processing pipeline is handled by the central processing facility. The data received from station processing facility is used to calculate desired correlations. The data processing and transmitting requirements for the central processing facility to calculate the correlation data of the array for each of the R stations in case of the proposed calibration strategies are described as follows:

- In case of calibration by holography, the correlations of B streams of received data corresponding to B beams for the s.u.t. generate the $B \times 1$ station covariance vector $\mathbf{r}$ for each of the R reference stations with the s.u.t.
- In case of superstation calibration, the $J \times 1$ raw received signal data from all the R stations involved form $JR \times JR$ correlations. These integrated correlations form a modest amount of data that needs to be transmitted from the CSP stage for further analysis.
- In case of stand-alone station calibration, if the correlations are already performed at the station site, all the available $\frac{J(J+1)}{2}$ crosscorrelations are streamed to the science data processing facility. In case the crosscorrelations are calculated at the CSP, all the J streams of crosscorrelations are transmitted to the science data processing facility from CSP.

2.4.3. Science data processing

Science data processing stage is hardly involved in calculating the desired calibration data and hence does not contribute in the study of comparison of computational requirements for the proposed improvements.

To further calculate the computational cost involved in calibration by superstation or holography or stand-alone station calibration, an analysis of data processing requirements for each of the proposed methods is done in the following section.

2.5. Comparison of proposed methods

The following criteria have been identified to compare the proposed SKA-low calibration approaches as shown in Table 2.1:

1. computational complexity
2. I/O to central processing facility data streaming
3. ease of incorporating EEP variations
4. compatibility with LOFAR and SKA
5. source model complexity
6. method in use.

Parameter	Holography	Superstation	Stand-alone
Computations required per station	$JB\Delta f_{signal}$	$({}^{JR}C_2\Delta f_{signal})/R$	${}^{J}C_2\Delta f_{signal}$
I/O to CPF streaming necessity	$4B\Delta f_{signal}$ sps	$4J\Delta f_{signal}$ sps	Negligible
EEP compensation	possible	possible	possible
Compatibility	compatible	compatible	compatible
Source model complexity	Low	Low	Very high
Method in use	in use	in use	in use

Table 2.1: Comparison of proposed approaches.

In table 2.1, the first measure compares the minimum number of computations required to obtain the calibration data for single station using each proposed approach. The I/O to CPF streaming necessity is a measure to compare data streaming capability from station processing facility to the central processing facility. The complexity of source model involved in both holography and superstation calibration is simple unlike the stand-alone calibration where the diffuse emission has to be meticulously modelled.

All the three proposed approaches have been explored and are compatible with both LOFAR and SKA-low systems. Finally, although all the three proposed approaches can incorporate individual EEPs in station calibration, both holography and superstation approaches present better features in terms of source model complexity.

Holography is chosen for further study in this thesis based on it's merit in terms of number of computations required for calibration of a single station compared to superstation calibration. The subsequent chapters of this report explore the calibration model for LOFAR by holography with and without the incorporation of individual EEPs and their impact on calibration results. The major contribution of this work is to investigate the impact of incorporating individual EEPs in the calibration model of LOFAR-LBA by holography.

3

LOFAR Holography

Holography uses a single reference signal per station for calibration. The received signal at the calibration beam of the reference station steered towards an identified source (calibration source) is used as reference signal for LOFAR calibration by holography. The calibration model for LOFAR stations with J antennas each, for a single polarisation and for a single narrowband channel at frequency f is described in the chapter.

Section 3.1 introduces the current holography model based on the implementation for LOFAR-HBA. Section 3.2 describes the the instrument model with the incorporation of individual EEPs in calibration by holography, followed by simulation model using individual EEPs in the instrument model. Note that the inclusion of individual EEPs is done only in the instrument model while calibration is still done with the model assuming identical EEPs.

3.1. Current implementation model

For each LOFAR station, the origin in (x, y, z) coordinate system is defined at the geometric mean of the station. The time delay $\tau_j(t)$ for a signal arriving at antenna j from a single source located at direction $\mathbf{l}(t)$ is given by equation 3.1,

$$\boldsymbol{\tau}_j(t) = \frac{\mathbf{p}_j \cdot \mathbf{l}(t)}{c}, \tag{3.1}$$

where '$\cdot$' represents inner product, c is the speed of light, $\mathbf{p}_j$ is the position of antenna with respect to origin and $\mathbf{l}(t)$ is the unit vector pointing from the phase reference centre of the station (at origin) to the source $s(t)$ located at $\mathbf{l}(t)$.

The location of the source is determined in the (l, m, n) coordinate system as a function of azimuth and zenith angle. The time delay at the j^{th} antenna for multiple sources can be obtained by stacking the geometric delays for each source into a vector.

The received signal $x_j(t)$ at the j^{th} antenna can then be modelled as in equation 3.2, as the product of antenna gain q_j, source power s(t), along with additive noise $n_j(t)$ at the j^{th} receiver. For the sake of simplicity, attenuation of source signal is neglected.

$$x_j(t) = s(t)\, q_j\, e^{-j2\pi f \tau_j(t)} + n_j(t) \tag{3.2}$$

The received signal vector is then obtained by stacking received signals from J antennas as $\mathbf{x}(t) = [x_1(t), ..., x_J(t)]^T$, to obtain equation 3.3.

$$\mathbf{x}(t) = \mathbf{Q}\,\mathbf{a}(l)s(t) + \mathbf{n}(t), \tag{3.3}$$

where $\mathbf{Q} = diag[q_1, ..., q_J]^T$ is a $J \times J$ diagonal matrix, $\mathbf{a}(l) = \exp(-2j\pi f \boldsymbol{\tau}(t))$ is the $J \times 1$ array response vector with $\boldsymbol{\tau}(t) = [\tau_1(t), ..., \tau_J(t)]^T$.

The array response vector represents the phase delays of the J received signals with respect to the phase reference centre of the station .

The received signal vector $\mathbf{x}(t)$ represents the signals from the source captured at all the antennas, which are coherently added using beamformer weights $\mathbf{w}$, that are unique to each station, position of the source and subband frequency. The purpose of a beamformer is to compensate the delays of the received signals in the station and steer the station towards the source. The output for the received signal vector using the classical beamformer is given by equation 3.4.

$$y(t) = \mathbf{w}^H \mathbf{x}(t) = \mathbf{a}^H(l)\ (\mathbf{Q}\mathbf{a}(l)s(t) + \mathbf{n}(t)) = \sum_{j=1}^{j=J} q_j s(t) + n_{bf}(t) \tag{3.4}$$

The beamformer hence amplifies the received source signal by $\sum_{j=1}^{j=J} q_j$ times. The side lobes in the beamformer output $n_{bf}(t)$ can be minimized if other beamformers such as the MVDR or the LMMSE or the Wiener filter are used instead of the classical beamformer. The beamformer specifications are not of much interest for this thesis and hence not further discussed.

The correlation of the received signal vector with the reference signal is a bit more complicated in practice. In general, each station in the LOFAR forms focused beams observing sections of the sky. The received signal vector is beamformed using B beams for each station.

The beamformed outputs for each station are used to obtain the beamformed signal and the reference signal correlations. Beamformer outputs, $y_b(t)$ and $y_r(t)$ respectively, for the b^{th} beam of the station under test (S.U.T) and the calibration beam of the reference station r are given by equation 3.5 and equation 3.6 respectively,

$$y_b(t) = \mathbf{w}_b^H \mathbf{x}(t) \tag{3.5}$$

$$y_r(t) = \mathbf{w}_r^H \mathbf{x_r}(t) = \mathbf{a}_r^H(l)\ (\mathbf{a}_r(l)s(t) + \mathbf{n}_r(t)) = Js(t), \tag{3.6}$$

where $\mathbf{w}_b$ is the beamformer for the b^{th} beam steering the S.U.T towards the source $s(t)$ and $\mathbf{w}_r$ is the beamformer for the calibration beam of the reference station. Since the reference station is known to have been pointed towards the source $s(t)$, the beamformer $\mathbf{w}_r = \mathbf{a}_r(l)$. The correlated signal v_b of the b^{th} beam of the S.U.T is calculated as the correlation between the beamformed outputs $y_b(t)$ and $y_r(t)$ as shown below:

$$\begin{aligned} v_b = \mathscr{E}\{y_b(t)\ y_r^H(t)\} &= \mathscr{E}\{\mathbf{w}_b^H \mathbf{x}(t)\ (Js(t))^H\} \\ &= \mathscr{E}\{\mathbf{w}_b^H \mathbf{Q}\mathbf{a}(l)|s(t)|^2 J\} = \mathbf{w_b}^H \mathbf{Q}\mathbf{a}(l)\sigma J \end{aligned}$$

$$v_b = \mathbf{w_b}^H \mathbf{Q}\mathbf{a}(l)\sigma J, \tag{3.7}$$

where $\sigma = |s(t)|^2$ is the expected power from the source and $\mathscr{E}\{.\}$ is the expectation operator.

The correlations between the beamformed outputs of all beams of the S.U.T and the calibration signal from reference station r are stacked into vector $\mathbf{v}_r$ to obtain the desired correlation data as shown in equation 3.8.

$$\mathbf{v}_r = \begin{bmatrix} \mathbf{w}_1^H \\ \vdots \\ \mathbf{w}_B^H \end{bmatrix} \mathbf{Qa}(l)\sigma J = J\sigma\mathbf{WQa}(l) \tag{3.8}$$

In equation 3.8, beamforming vectors for all the B beams in the station are stacked to obtain the beamformer matrix $\mathbf{W} = [\mathbf{w}_1, ..., \mathbf{w}_B]^H$. The obtained correlations are pre-processed for calibration. Pre-processing is done to reduce the data volume and should not change the expected value. Correlations for each station are time and sample averaged in the pre-processing stage. Pre-processing over a few minutes and for a few data samples improves the SNR and helps in suppression of interfering signals.

Current LOFAR calibration by holography implements a few final processing steps on the correlation vector after pre-processing. These final processing steps are done in stages and include normalisation using the received signal followed by time-averaging for each s.u.t., for each beam. The final correlation data towards each beam is obtained by station averaging. These processing steps are described in the following sections.

3.1.1. Pre-Processing

Time averaging is done in the pre-processing stage to allow early flagging of RFI. It is assumed that the power of the source and the gains to be calibrated are constant over the integration interval, such that the expected value of the measurement does not change.

If the interval for time-averaging in the pre-processing stage is M, then the total number of data samples in time are reduced by M-fold, reducing the data volume by as much. In radio astronomical perspective, the visibilities are expected to be constant over small time intervals after compensating for the rotation of the Earth, justifying time averaging.

In this description, sample averaging is not explicitly described since the pre-processing stage is excluded from implementation.

3.1.2. Final Processing

The correlation data for the s^{th} S.U.T after pre-processing still follows the model described in equation 3.8. The pre-processed correlation data $\mathbf{v}_r$ for the s^{th} S.U.T is normalised using correlated signal (v_{CB}) of the calibration beam CB of reference station r. The normalisation results in the $B \times 1$ normalised data vector $\tilde{\mathbf{v}}_r$ for the s^{th} S.U.T given by equation 3.9.

The normalised correlation vector $\tilde{\mathbf{v}}_r$ for s^{th} S.U.T can further be simplified to equation 3.10 by using $\mathbf{a}_s(l) = \mathbf{1} = [1, 1, ..., 1]^T$ as the array response vector for the S.U.T, towards a source at zenith. This is mathematically given by equation 3.10. Compensation for sources not at bore sight can be done by steering the stations towards the respective sources using beamforming.

$$\tilde{\mathbf{v}}_r = \frac{\mathbf{v}_r}{v_{CB}} = \frac{J\sigma\mathbf{W}_s\mathbf{Qa}_s(l)}{J\sigma} = \mathbf{W}_s\mathbf{Qa}_s(l) \tag{3.9}$$

In equation 3.9, $\mathbf{Q} = diag[\mathbf{q}] = diag[q_1, ...q_J]^T$ is the $J \times J$ antenna gain matrix and $\mathbf{W}_s$ is the $B \times J$ beamformer matrix for the s^{th} S.U.T.

$$\tilde{\mathbf{v}}_r = \mathbf{W}_s\mathbf{Q}\mathbf{a}_s(l) = \mathbf{W}_s\mathbf{Q}\mathbf{1} = \mathbf{W}_s\mathbf{q} \tag{3.10}$$

Equation 3.10 represents the model for normalized data for s^{th} S.U.T for a single time sample. Normalisation is then followed by time-averaging over the entire time for the cross-correlation data of the S.U.T with each reference station. Normalised and time-averaged correlations for all the beams of the s^{th} S.U.T with respect to all R reference stations can be stacked to form the $B \times R$ correlation matrix $\mathbf{V}$, given by equation 3.11.

$$\mathbf{V} = [\tilde{\mathbf{v}}_1, ..., \tilde{\mathbf{v}}_R] \tag{3.11}$$

The time-averaged cross-correlation values from the $B \times R$ matrix $\mathbf{V}$ are station averaged reducing the matrix into the $B \times 1$ correlation vector $\mathbf{v}$ given by equation 3.12. The correlation values after station averaging follow the final data model given by equation 3.12.

$$\mathbf{v} = [v_1, ..., v_B]^T = \mathbf{W}\mathbf{q}, \tag{3.12}$$

In equation 3.12, $\mathbf{W}$ is the final $B \times J$ final beamformer matrix.

The solution for antenna gains can be obtained from the model in equation 3.12 using least squares approach given by equation 3.13.

$$\mathbf{q} = (\mathbf{W^H W})^{-\mathbf{1}}\mathbf{W^H v} \tag{3.13}$$

3.2. Inclusion of EEPs in LOFAR Holography model

The inclusion of embedded element patterns in to the LOFAR data model can be done by using direction dependent gains $\mathbf{G}(l) = diag[g_1(l), ..., g_J(l)]$ as shown in equation 3.14 similar to the model from equation 2.2. This change mathematically acknowledges that the weights on antennas vary with direction because of the influence of unidentical embedded element patterns.

$$\mathbf{x}(t) = \mathbf{Q}\mathbf{G}(l)\mathbf{a}(l)s(t) + \mathbf{n}(t) \tag{3.14}$$

For the sake of simplicity and yet adhering to the principles of radio interferometry, the direction independent gains $\mathbf{Q}$ can be absorbed into the direction dependent gains $\mathbf{G}(l)$. The total gain for an antenna element j in the station is given by $\widetilde{g}_j(l) = q_j g_j(l)$. Similarly, antenna gain matrix $\widetilde{\mathbf{G}}(l)$ for the entire station is given by equation 3.15.

$$\widetilde{\mathbf{G}}(l) = \mathbf{G}(l) \odot \mathbf{Q} \tag{3.15}$$

In equation 3.15, $\odot$ represents entry wise matrix product [15]. After incorporating the EEPs into the calibration model, a further simplification in the form of station calibration towards a source at bore sight, similar to equation 3.8, is done to obtain a simplified correlation model in the form of equation 3.16.

Equating the array response vector $\mathbf{a}(l)$ to $\mathbf{1}$ essentially places the source and the central beam of the station at zenith. For other sources, corrections can be made by steering the beamformer weights in the appropriate staring direction.

$$\mathbf{v} = J\sigma\mathbf{W}\widetilde{\mathbf{G}}(l)\mathbf{a}(l) = J\sigma\mathbf{W}\widetilde{\mathbf{G}}(l)\mathbf{1} = J\sigma\mathbf{W}\widetilde{\mathbf{g}}(l) \tag{3.16}$$

where $\widetilde{\mathbf{g}}(l) = [\widetilde{g}_1(l), ..., \widetilde{g}_J(l)]^T$ are the combined direction dependent and independent antenna gains of the S.U.T that are to be estimated.

Following the pre-processing and final processing stages, the final model for correlation data in terms of the beamforming matrix $\mathbf{W}$ and gains $\widetilde{\mathbf{G}}(l)$ is given by equation 3.17.

$$\mathbf{v} = \mathbf{W}\widetilde{\mathbf{G}}(l) \tag{3.17}$$

The solution for gains from equation 3.17 can be estimated using the least squares approach similar to the solution given in equation 3.13 and is given by equation 3.18.

$$\widetilde{\mathbf{G}}(l) = (\mathbf{W}^{\mathbf{T}}\mathbf{W})^{-\mathbf{1}}\mathbf{W}^{\mathbf{T}}\mathbf{v} \tag{3.18}$$

3.3. Incorporating EEPs in LOFAR simulation model

The expected calibration errors from the assumption of identical EEPs cannot be observed if the system noise is very high. These errors are more obvious in cases of high SNR and depend on source power, array compactness, system noise and signal processing errors. To restrict this research towards analysing the impact of the assumption of identical EEPs on calibration errors, signal processing errors and system noise are ignored and a model for LOFAR calibration by holography is simulated.

The simulated model, compared to the measurement model, additionally has direction dependent embedded element patterns within the expression for received signal vector. The embedded element patterns are simulated based on an experimentally validated electromagnetic model. The received signal at each antenna element is modelled as a product of direction dependent gains, array response vector, direction independent gains, and attenuated source power along with additive zero mean white Gaussian noise.

The physical process of signal reception is in the aforementioned order. In case of simulation of LOFAR calibration by holography, embedded element patterns from electromagnetic simulations are used as direction dependent gains. Mathematically, received signal at the j^{th} antenna using the simulation model is given by equation 3.19.

$$x_j(t) = q_j a_j(l) g_j(l) s(t) + n_j(t), \tag{3.19}$$

Stacking the received signals for all J antennas in the array into a vector, the $J \times 1$ received signal vector is given by equation 3.20.

$$\mathbf{x}(t) = \mathbf{G}(l)[\mathbf{a}(l) \odot \mathbf{q}]s(t) + \mathbf{n}(t) = \widetilde{\mathbf{G}}(l)\mathbf{a}(l)s(t) + \mathbf{n}(t) \tag{3.20}$$

In equation 3.20, $\{\odot\}$ denotes Hadamard product or entry-wise matrix product [15], $\mathbf{G}(l)$ is the $J \times J$ diagonal matrix, $\mathbf{a}(l)$ is the $J \times 1$ array response vector, $\mathbf{q}$ is the $J \times 1$ direction independent gains vector and $\mathbf{n}(t)$ is the $J \times 1$ noise vector for the array of J antennas.

The classical beamformer $\mathbf{w}$ and array output $y(t)$ with the classical beamformer are given by expressions from equation 3.21.

$$\mathbf{w} = \mathbf{a}(l)$$

$$y(t) = \mathbf{w}^H \mathbf{x}(t) \tag{3.21}$$

The beamformer output in the actual simulation is a $B \times 1$ vector with outputs for the B beams formed by the J antennas of the array. The expressions for received signal, received signal vector and beamformer output hold good for all stations in both the LOFAR-LBA system and the LOFAR-HBA system. Holography is also currently being tested for the LOFAR-HBA system. The implementation, results and analysis of results from the simulation model are presented in chapter 4.

4

LOFAR-LBA holography simulation

The contents of this chapter focus on comparing results from a LOFAR-LBA holography simulation. Chapter 4 begins by showing variations in EEPs among antenna elements of LOFAR-LBA in section 4.1. Spherical wave coefficients for each antenna element are used to determine antenna radiation pattern towards the calibration source for LOFAR-LBA inner station. The theory involved in calculating these EEPs is not a part of this project and therefore not discussed further.

Section 4.2 and section 4.3 use antenna positions for the CS302 LBA station. The instrument model simulated to generate calibration data in section 4.2 assumes identical EEPs for all antenna elements while the instrument model in section 4.3 uses different embedded element pattern for each element. The distinction in the implementations in section 4.2 and section 4.3 lies in the instrument models used for generating calibration data alone. The final calibration by holography is done assuming that all the EEPs are identical.

The embedded element patterns used in section 4.3 are individually calculated using spherical wave functions at desired frequencies for all antenna elements in the CS302 LBA station towards each specified direction on the lm-plane. Further details of the simulations are described in the corresponding sections.

Chapter 4 is concluded in section 4.4 with a comparison of results from section 4.2 and section 4.3 followed by results showing the effect of SNR on calibration errors.

4.1. Variations in antenna radiation patterns

Embedded element pattern for each antenna element is calculated using spherical wave coefficients towards the specified direction. The MATLAB functions used to calculate spherical wave functions and thereby identifying the radiation pattern for a given antenna element are given in appendix A.3 and appendix A.2 respectively. The principles used in calculating the radiation patterns of antenna elements are based on the general theory of converting spherical harmonic wave functions into a series of plane wave expansions [28]. The theory behind the functions in not part of this thesis and hence not discussed further.

The θ and ϕ components of EEPs between two antennas, one at the center of the station while the other farther from center, in the LOFAR-LBA inner station are compared at 57 MHz based on the following considerations:

1. The antenna elements exhibit higher radiation patterns at their resonant frequency (57 MHz).

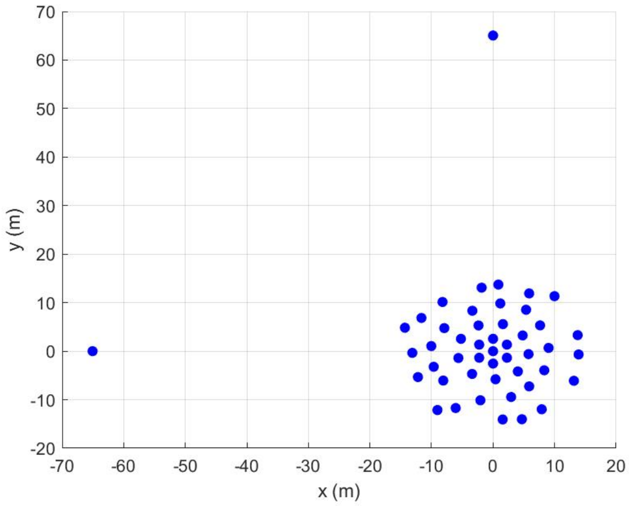


Figure 4.1: LOFAR-LBA inner station layout.

2. The mutual coupling effects that influence the radiation patterns are higher for LOFAR-LBA inner stations compared to the LOFAR-LAB outer stations.

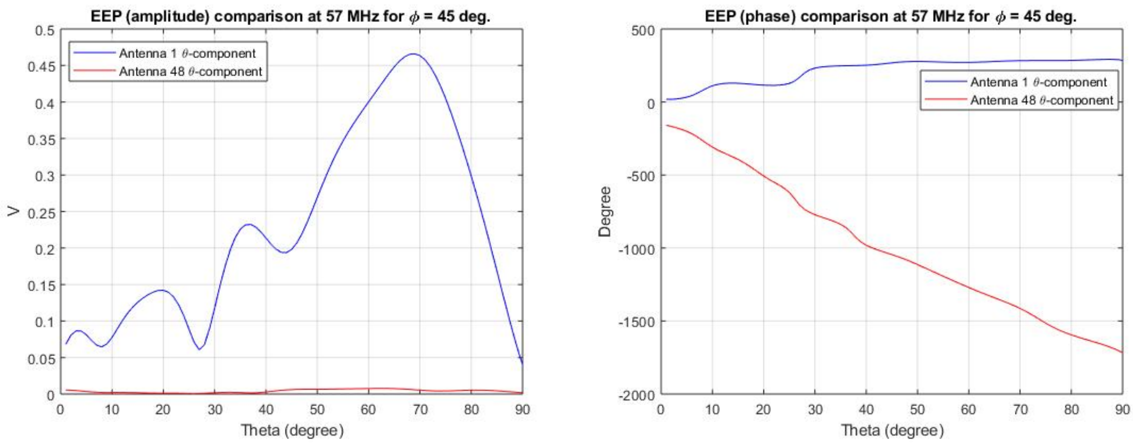
3. Figure 4.1 shows layout for the LOFAR-LBA inner station. It can be seen that two antenna elements are placed farther away from the actual cluster of the antennas. This is done to facilitate the imposition of restrictions over longer baselines.

4. Based on the previous observation, antenna at origin (designated as antenna 1 in this section) from figure 4.1 is expected to have large differences in θ and ϕ components of the embedded element patterns compared to antenna element at $(0,70)$ (designated as antenna 2 in this section) in figure 4.1.

Results for amplitude and phase of the embedded element patterns of antenna 1 and antenna 2 in the LOFAR-LBA inner station indicating differences in both θ and ϕ components at 57 MHz are shown in figure 4.2 and figure 4.3 respectively.

Figure 4.2 shows comparison of amplitude and phases of the EEPs for antenna 1 and antenna 2 with $\phi = 45°$ with varying azimuth angles. Figure 4.3 shows comparison of amplitude and phases of the EEPs for antenna 1 and antenna 2 with $\theta = 45°$ with varying zenith angles. It can be seen that the radiation patterns for the antenna elements 1 and 2 are clearly different the assumption of identical EEPs clearly leads to calibration errors.

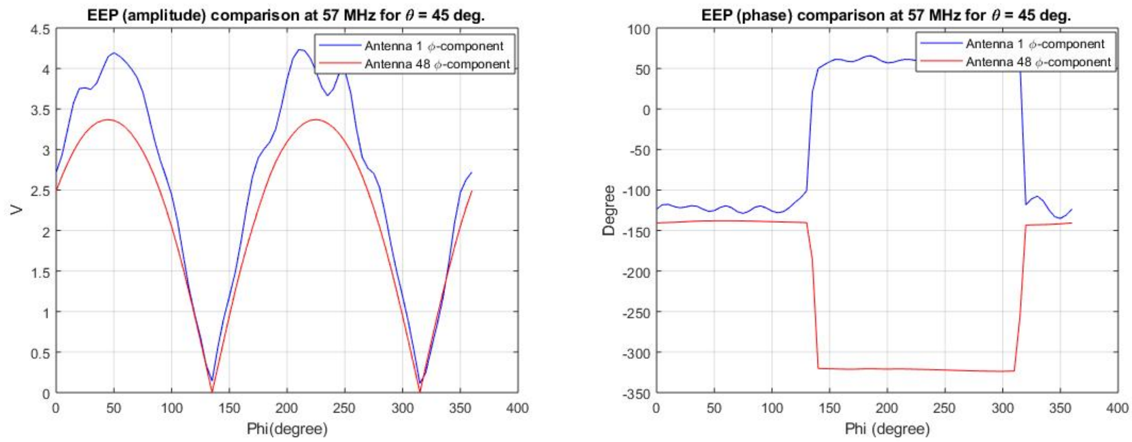
The individual embedded element patterns depend on the position of the antenna, operational frequency and direction of observations, all of which are unique to each antenna. Section 4.2 focuses on calibration by holography without including these individual embedded element patterns.



(a) EEP θ component amplitude differences between antenna 1 and antenna 2.

(b) EEP θ component phase differences between antenna 1 and antenna 2.

Figure 4.2: Amplitude and phase variations in θ component EEPs between antenna 1 and antenna 2 towards $\phi = 45$.



(a) EEP ϕ component amplitude differences between antenna 1 and antenna 2.

(b) EEP ϕ component phase differences between antenna 1 and antenna 2.

Figure 4.3: Amplitude and phase variations in ϕ component EEPs between antenna 1 and antenna 48 towards $\theta = 45$.

4.2. Holography without the inclusion of EEPs

The simulation model used in this section does not incorporate EEP variations. The mathematical expression to describe the received signal for the simulated antenna array with N_{ant} (48) elements positioned according to the LOFAR-LBA inner and outer layouts is given by equation 3.20.

MATLAB code used for the simulation of holography without inclusion of EEPs can be found in appendix B. The modelled source $\mathbf{S}_{ref}$ in the simulation is an unpolarised source signal sampled at Nyquist rate and has N_s samples. Since the source is assumed to be unpolarised, the source signal has both θ and ϕ components which are independently modelled using random distribution, making $\mathbf{S}_{ref}$ a $2 \times N_s$ matrix.

Sampling of the lm-plane is simulated with values for (l, m) coordinates ranging between $[-1, 1]$ each. The total number of beams B formed by the station elements is dependent on the number of sampling points along each direction b as $B = b^2$. The array response matrix $\mathbf{A}_{sut}$ is then a $48 \times B$ matrix, calculated using the geometric delays for antennas in the LOFAR-LBA inner/outer station antenna positions $\mathbf{P}$, with respect to each beam formed towards the lm-plane coordinates $\mathbf{l}$ (see equation 3.1).

The calculated $48 \times B$ array response matrix $\mathbf{A}_{sut}$ is used to determine the classical beamformer, while the geometric delays of the array elements towards the calibration beam c (source coordinates on the lm-plane) are used to construct the $N_{ant} \times 1$ array response vector $\mathbf{a}_{ref}$. Since the source signal is unpolarised and has θ and ϕ components, 2 identical $\mathbf{a}_{ref}$ array response vectors are used to construct the 48×2 matrix $\mathbf{A}_{ref} = [\mathbf{a}_{ref}, \mathbf{a}_{ref}]$ used in equation 4.1.

Additive zero mean white Gaussian noise $\mathbf{N}$ is modelled according to the desired SNR for the simulations. The SNR used in the simulations is the SNR per station. SNR per station is improved after beamforming and integration over the time for each subband. The term SNR refers to SNR per station for the rest of this chapter, unless specified.

Finally, antenna gains $\tilde{\mathbf{q}}$ are simulated according to Gaussian distribution to generate the received signals $\mathbf{X}_{ref}$ using the expression given by equation 4.1.

$$\mathbf{X}_{ref} = diag(\tilde{\mathbf{q}})\,\mathbf{A}_{ref}\,\mathbf{S}_{ref} + \mathbf{N} \tag{4.1}$$

The simulated received signals $\mathbf{X}_{ref}$ modelled according to equation 4.1 is a $N_{ant} \times N_s$ matrix. The beamforming operation using $\mathbf{W}_{sut} = \mathbf{A}_{sut}$ results in the $B \times N_s$ beamformed output matrix $\mathbf{V}_{sut}$ for the station, given by equation 4.2.

$$\mathbf{V}_{sut} = \mathbf{W}_{sut}^H\,\mathbf{X}_{ref} \tag{4.2}$$

$$\mathbf{v}_{ref} = \mathbf{w}_{ref}^H\,\mathbf{X}_{ref} \tag{4.3}$$

Similarly, the $1 \times N_s$ beamformed output vector $\mathbf{v}_{ref}$ towards the calibration beam is calculated using the classical beamformer $\mathbf{w}_{ref} = \mathbf{a}_{ref}$ as shown in equation 4.3. The final $B \times 1$ correlation vector $\mathbf{r}$ required for holography is obtained by correlating $\mathbf{V}_{sut}$ and $\mathbf{v}_{ref}$.

$$\mathbf{r} = \mathbf{V}_{sut}\,\mathbf{v}_{ref}^H \tag{4.4}$$

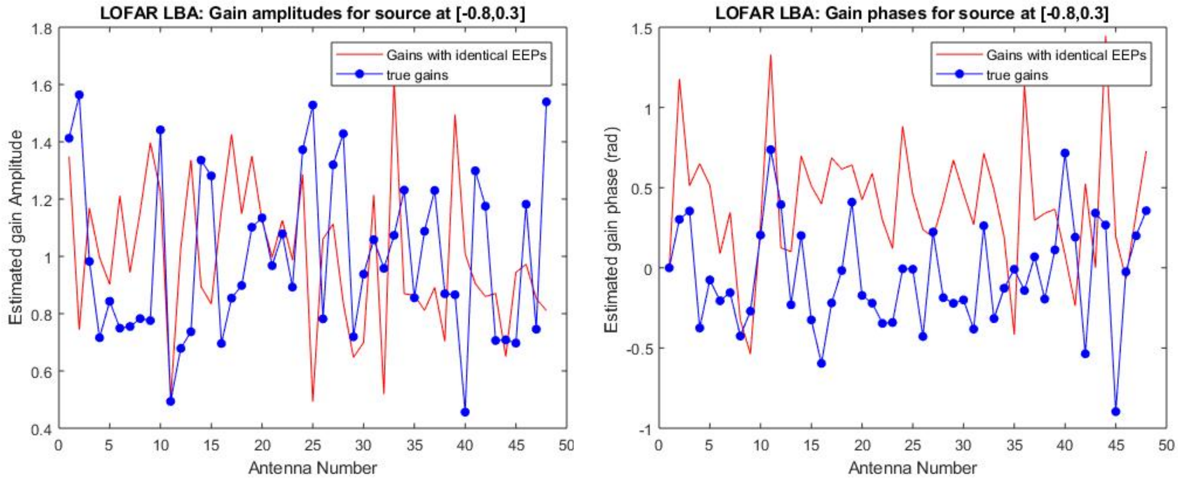
The obtained correlations from equation 4.4 are then normalised using the correlations corresponding to the calibration beam to obtain the normalised correlation vector $\mathbf{r}_{norm}$. The gain estimates $\mathbf{g}_{est}$ using the normalised correlations can be determined using the expression in equation 4.5, similar to the expression in equation 3.18.

$$\mathbf{g}_{est} = (\mathbf{W}_{sut}\mathbf{W}_{sut}^H)^{-1}(\mathbf{W}_{sut}\mathbf{r}_{norm}) \tag{4.5}$$

Results for the estimated gain amplitudes and phases for the 48 antennas positioned according to LOFAR-LBA inner and outer layout are shown in figure 4.4 and figure 4.5 respectively, towards a source positioned at [-0.8,0.3] on lm-plane having an SNR = 1000 (the noiseless case).

As evident from figure 4.4b and figure 4.5b, the antenna gain phases are calculated with respect to the phase of antenna 1 to facilitate comparison.

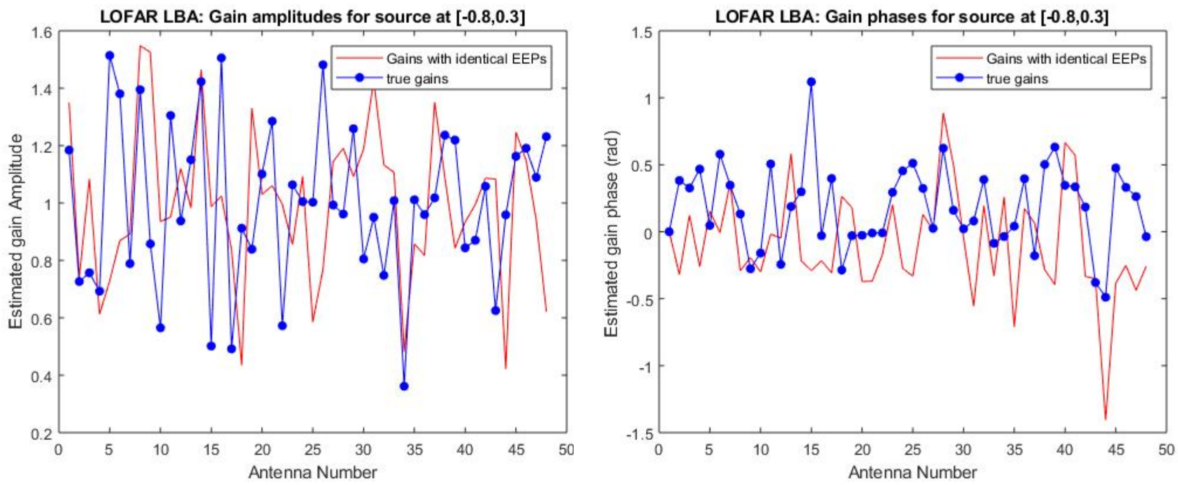
Figure 4.6 and figure 4.7 show the results from the simulation with the source positioned at (0,0) for LBA inner and outer layouts respectively. Even for the noiseless case with SNR = 1000, which is unlikely for the current LOFAR-LBA set up, estimated antenna gains show



(a) Estimated & true gain amplitudes vs antenna number.

(b) Estimated & true gain phases vs antenna number.

Figure 4.4: Estimated antenna gains for the LOFAR-LBA inner station at 35 MHz with the assumption of identical EEPs for SNR = 1000 with the source at (-0.8,0.3).



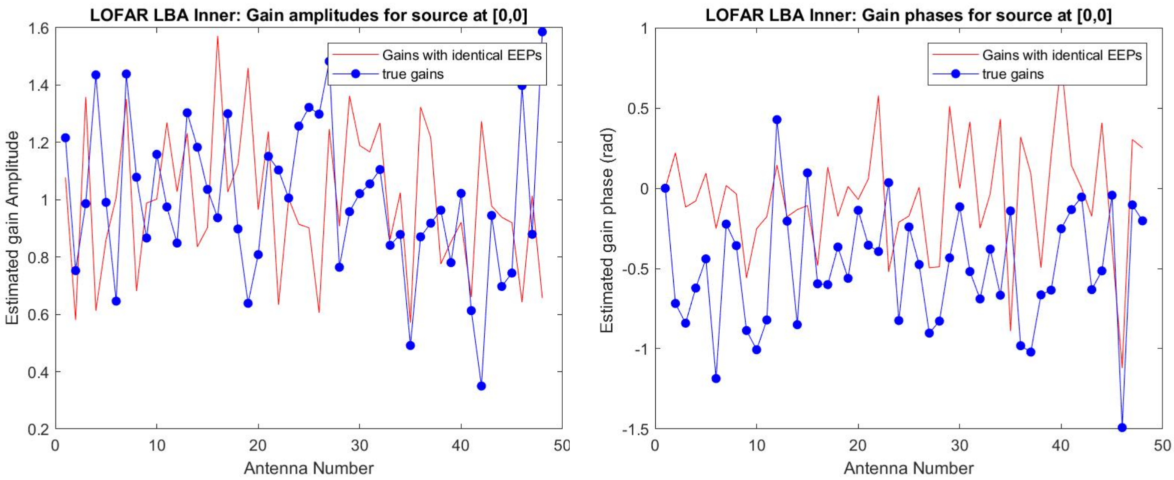
(a) Estimated & true gain amplitudes vs antenna number.

(b) Estimated & true gain phases vs antenna number.

Figure 4.5: Estimated antenna gains for the LOFAR-LBA outer station at 35 MHz with the assumption of identical EEPs for SNR = 1000 with the source at (-0.8,0.3).

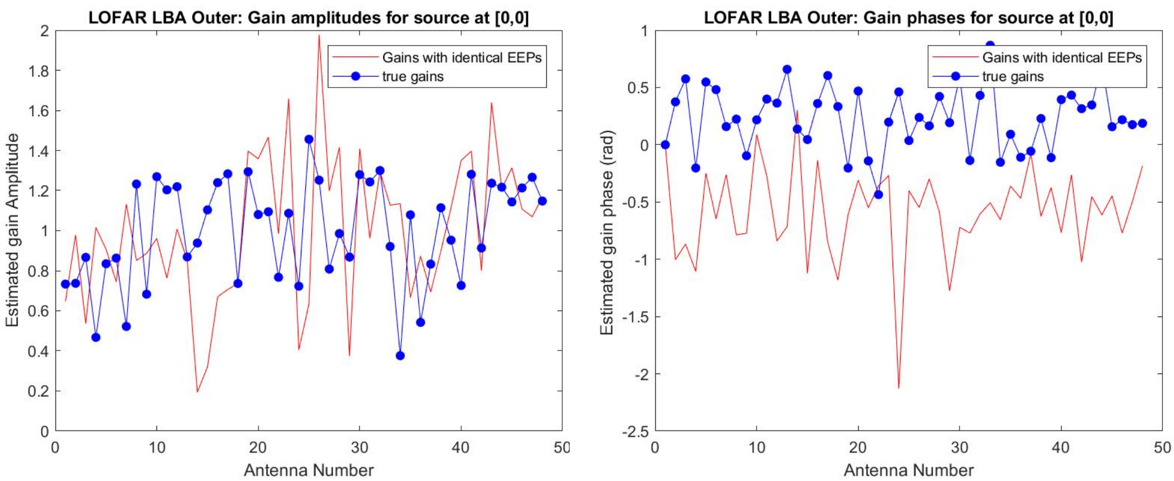
variations compared to the true gains as can be observed from figure 4.4, figure 4.5, figure 4.6 and figure 4.7.

Practically, the SNR for the LOFAR-LBA observations were identified to range from 0.01 to at most 4 in section 5.2. Section 4.2.1 compares the estimated antenna gains for varying SNR within this range. Section 4.2.2 shows the estimated antenna gains for the LOFAR-LBA operational frequencies $[35, 42, 54, 59, 65, 69, 75]$ MHz.



(a) Estimated & true gain amplitudes vs antenna number. (b) Estimated & true gain phases vs antenna number.

Figure 4.6: Estimated antenna gains for the LOFAR-LBA inner station at 35 MHz with the assumption of identical EEPs for SNR = 1000 with the source at (0,0).



(a) Estimated & true gain amplitudes vs antenna number. (b) Estimated & true gain phases vs antenna number.

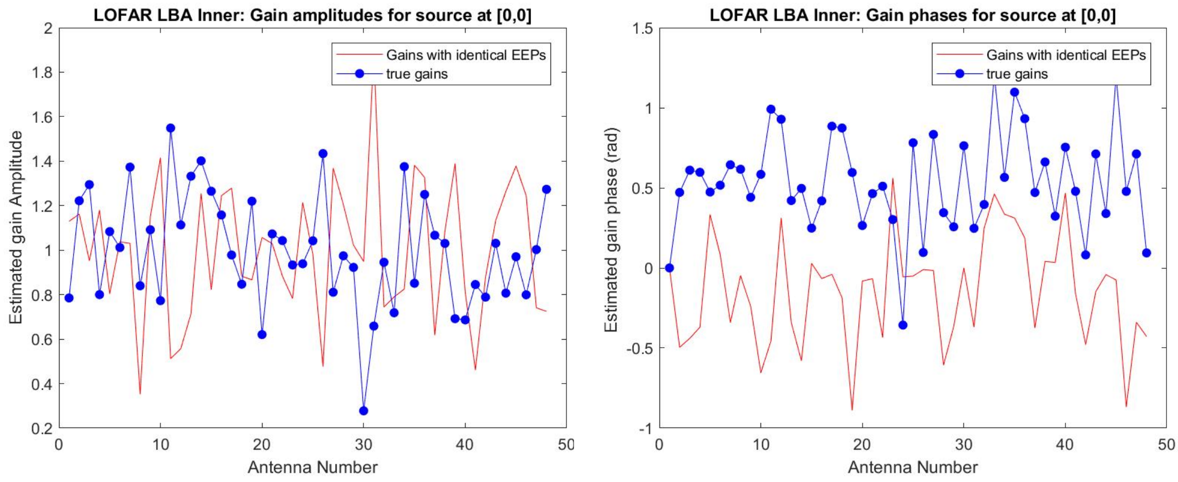
Figure 4.7: Estimated antenna gains for the LOFAR-LBA outer station at 35 MHz with the assumption of identical EEPs for SNR = 1000 with the source at (0,0).

4.2.1. Effect of SNR on estimation accuracy

The effect of SNR on gain estimation is predictable, given that the SNR of the observations influence the beamformed received signal output and thereby influence the gain estimates. The gain amplitude and phase estimates were observed to be accurate for the noiseless case (SNR = 1000) in previous section. It can be seen from figure 4.8, figure 4.9 and figure 4.10 that the gain estimates improve with increasing SNR.

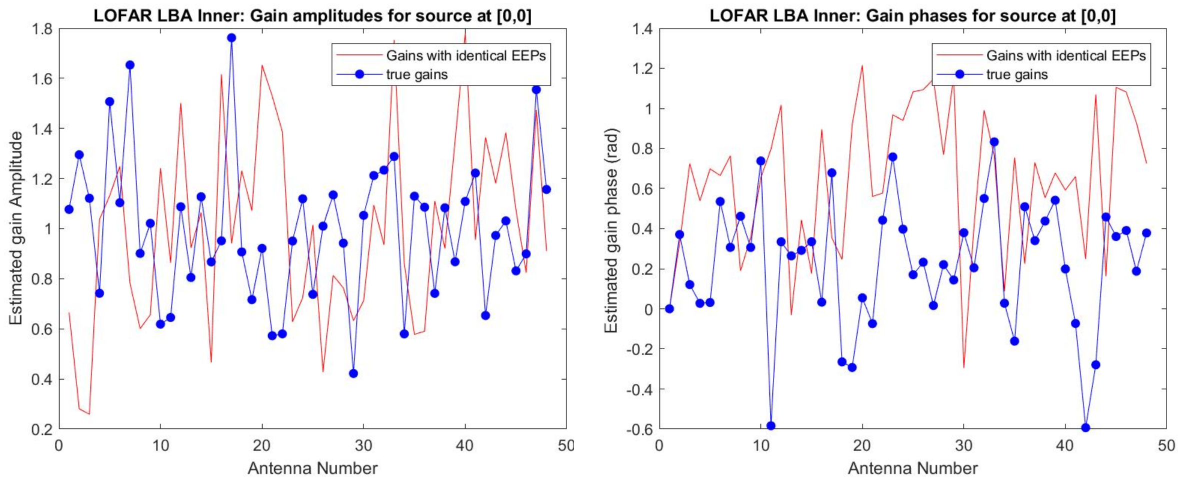
Gain estimates for the LOFAR LBA Outer layout exhibit the same pattern of improving accuracy with increasing SNR. Figure 4.13 shows estimated gain amplitudes and phases for the LOFAR-LBA outer station at 35 MHz and SNR = 4.0.

In general, calibration sources that offer an SNR of at least 0.001 are picked for reasonable accuracy. The improvement in the accuracy of gain estimates from SNR = 0.001 to SNR



(a) Estimated & true gain amplitudes vs antenna number. (b) Estimated & true gain phases vs antenna number.

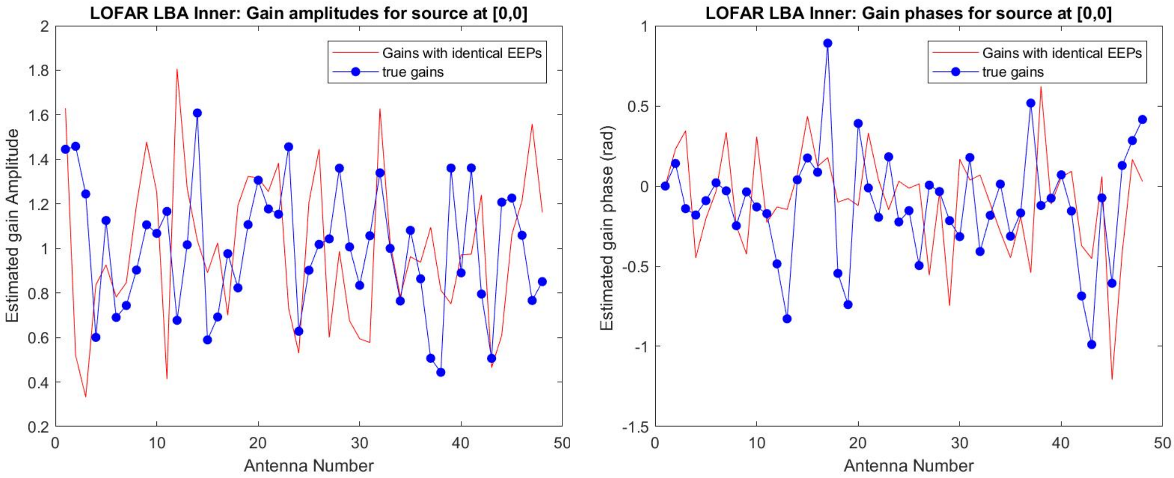
Figure 4.8: Estimated antenna gains for the LOFAR-LBA inner station at 35 MHz for SNR = 0.001.



(a) Estimated & true gain amplitudes vs antenna number. (b) Estimated & true gain phases vs antenna number.

Figure 4.9: Estimated antenna gains for the LOFAR-LBA inner station at 35 MHz for SNR = 2.

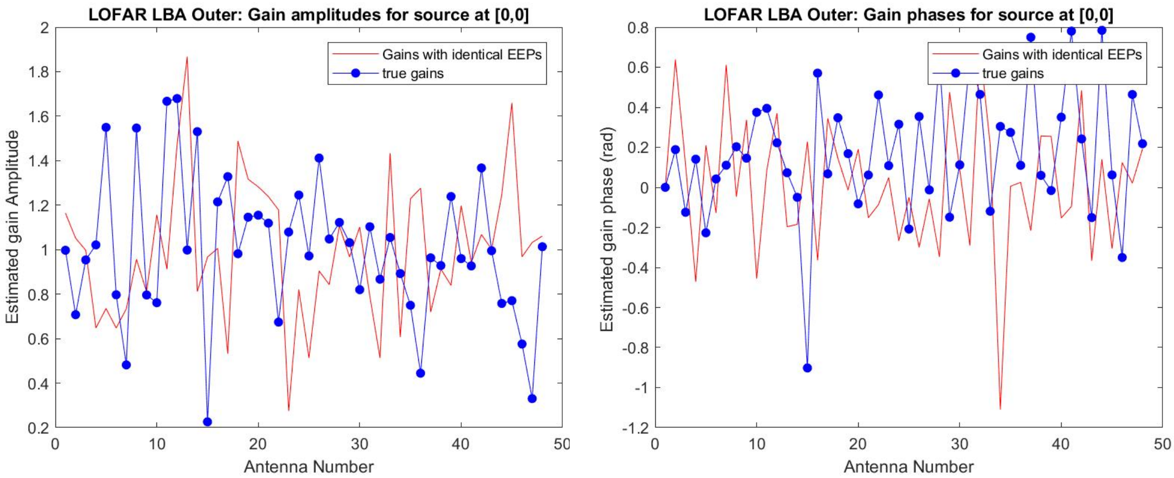
= 2 and to SNR = 4 can be observed from figure 4.11, figure 4.12 and figure 4.13 respectively for the LBA outer configuration.



(a) Estimated & true gain amplitudes vs antenna number.

(b) Estimated & true gain phases vs antenna number.

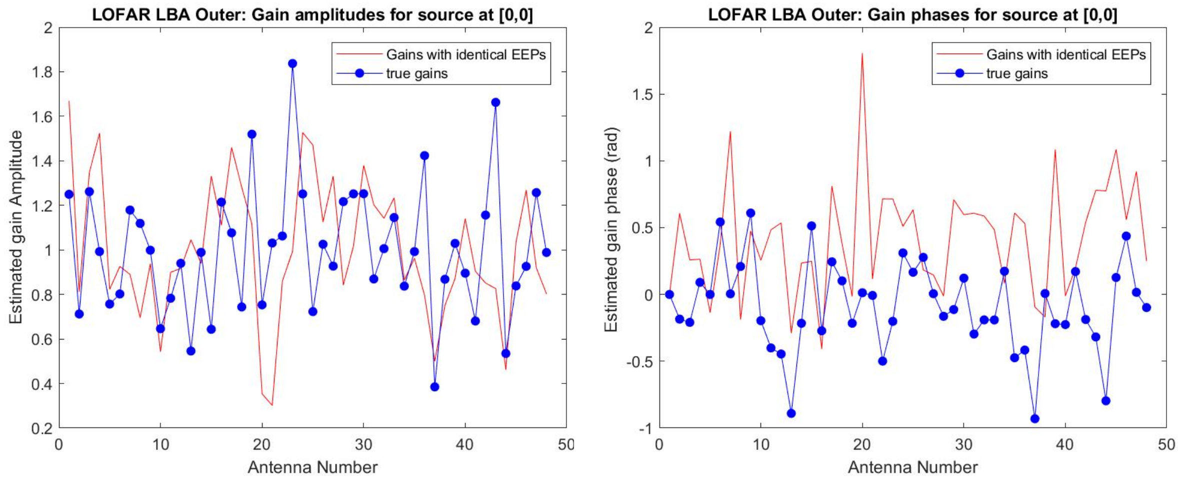
Figure 4.10: Estimated antenna gains for the LOFAR-LBA inner station at 35 MHz for SNR = 4.



(a) Estimated & true gain amplitudes vs antenna number.

(b) Estimated & true gain phases vs antenna number.

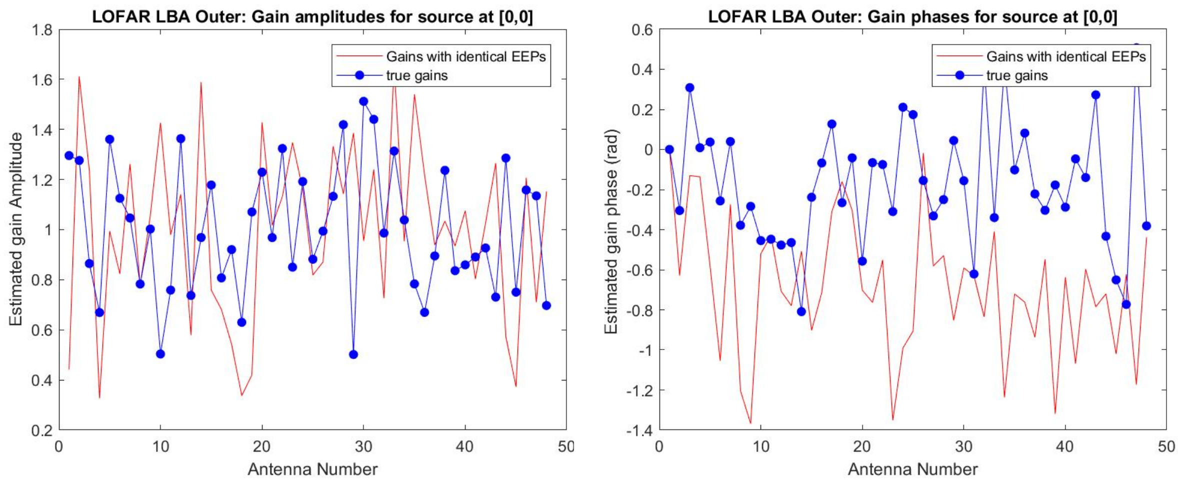
Figure 4.11: Estimated antenna gains for the LOFAR-LBA outer station at 35 MHz with the assumption of identical EEPs for SNR = 0.001.



(a) Estimated & true gain amplitudes vs antenna number.

(b) Estimated & true gain phases vs antenna number.

Figure 4.12: Estimated antenna gains for the LOFAR-LBA outer station at 35 MHz with the assumption of identical EEPs for SNR = 2.



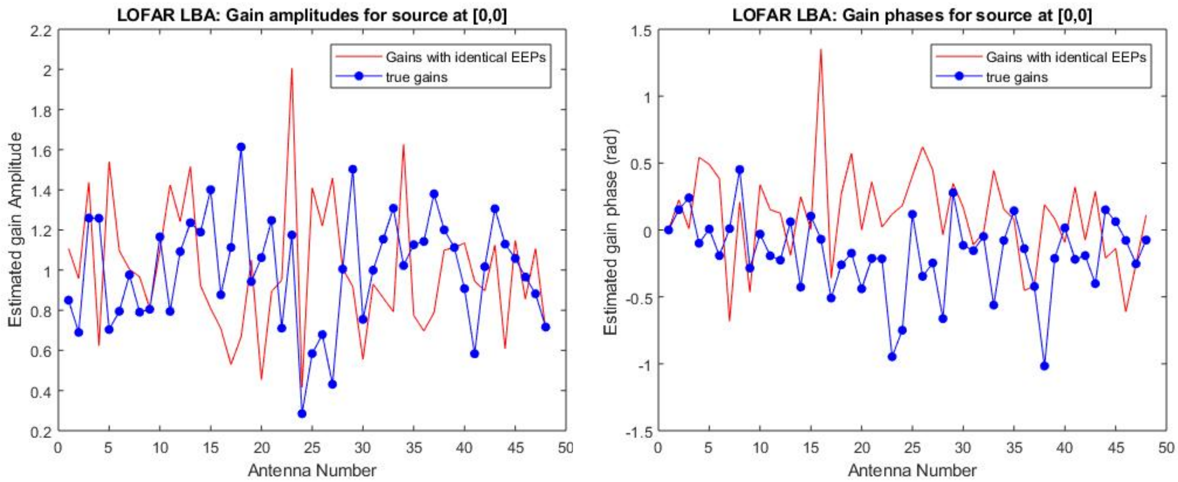
(a) Estimated & true gain amplitudes vs antenna number.

(b) Estimated & true gain phases vs antenna number.

Figure 4.13: Estimated antenna gains for the LOFAR-LBA outer station at 35 MHz with the assumption of identical EEPs for SNR = 4.

4.2.2. Gain estimates for the LOFAR-LBA operating frequencies

Measurement data for calibration of LOFAR-LBA by holography was initially recorded for 9 frequencies within its operational bandwidth. These frequencies were (19, 27, 35, 42, 54, 59, 65, 69, 75) MHz. It was observed from the measurement data that calibration of LOFAR-LBA by holography was impossible at 19 and 27 MHz frequencies (see section 5.1.1). For this reason, these frequencies are not considered in this chapter.



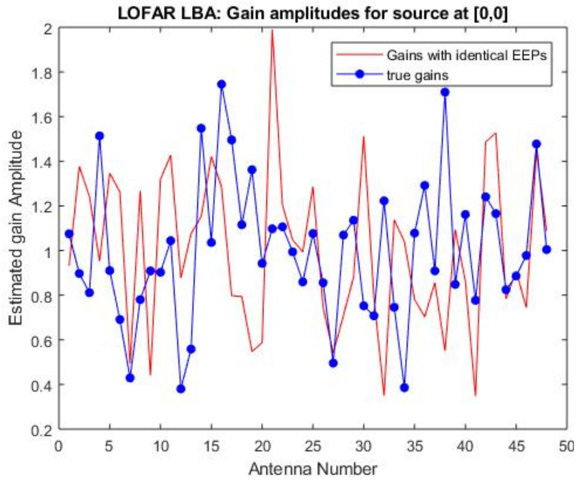
(a) Estimated & true gain amplitudes vs antenna number.

(b) Estimated & true gain phases vs antenna number.

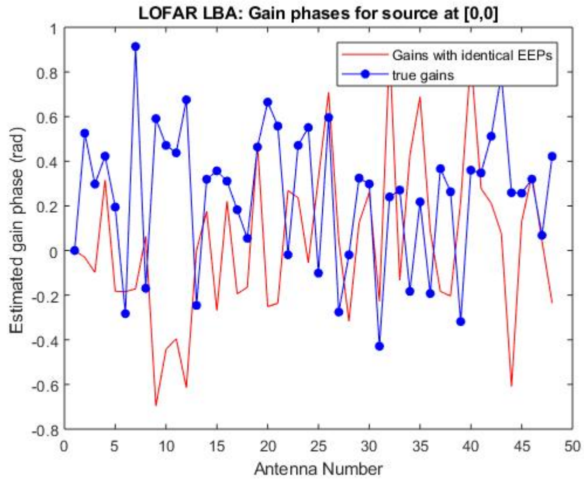
Figure 4.14: Estimated antenna gains for the LOFAR-LBA inner station at 59 MHz with the assumption of identical EEPs for SNR = 4.

Figure 4.14 and figure 4.15 show a comparison of the estimated gains with true gains for LBA inner and LBA outer respectively, at 59 MHz for SNR = 4. Similar observations were made for other frequencies within the operational bandwidth.

The notion of identifying gains towards a specific direction is trivial with identical EEPs in this section compared to gain estimation including the variations of EEPs. The impact of the position of the source is limited to calculating geometrical delays and for the purpose of normalisation in case of calibration with identical EEPs, unlike the case with calibration including EEP variations which is discussed in section 4.3.



(a) Estimated & true gain amplitudes vs antenna number.



(b) Estimated & true gain phases vs antenna number.

Figure 4.15: Estimated antenna gains for the LOFAR-LBA outer station at 59 MHz with the assumption of identical EEPs for SNR = 4.

4.3. Holography with the inclusion of EEPs

The mathematical model used to express the received signal vector in terms of antenna gains, array response matrix and the source signal is still the same as in equation 3.20. To incorporate the embedded element patterns into the received signal vector model, an experimentally validated electromagnetic model [18] is used to determine the radiation pattern of each antenna element as a function of frequency and position of the source in the lm-plane.

These electromagnetic models use spherical wave functions to calculate the EEPs for each antenna, with the phase centre of each antenna as the phase reference centre for the embedded element pattern variation. The electromagnetic model used to describe the interaction of antenna elements and their radiation patterns uses the general theory of converting spherical harmonic wave functions into a series of plane wave expansions [28].

The expression used in the simulation for received signal vector $\widetilde{\mathbf{X}}_{ref}$ incorporating the individual EEPs $\mathbf{E}_{ref}$ of the antennas towards the source is given by equation 4.6.

$$\widetilde{\mathbf{X}}_{ref} = diag(\mathbf{q})(\mathbf{E}_{ref} \odot \mathbf{A}_{ref})\mathbf{S}_{ref} + \mathbf{N} \tag{4.6}$$

Once the received signal vectors are calculated using equation 4.6, expressions from equation 4.7, equation 4.8 and equation 4.9 are used to determine the $B \times N_s$ beamformed received signal matrix $\widetilde{\mathbf{V}}_{sut}$, the $1 \times N_s$ beamformed reference signal vector $\widetilde{\mathbf{v}}_{ref}$ and the final $B \times 1$ correlation vector $\widetilde{\mathbf{r}}$ respectively, similar to the expressions used in section 4.2.

$$\widetilde{\mathbf{V}}_{sut} = \mathbf{W}_{sut}^{H}\widetilde{\mathbf{X}}_{ref} \tag{4.7}$$

$$\widetilde{\mathbf{v}}_{ref} = \mathbf{w}_{ref}^{H}\widetilde{\mathbf{X}}_{ref} \tag{4.8}$$

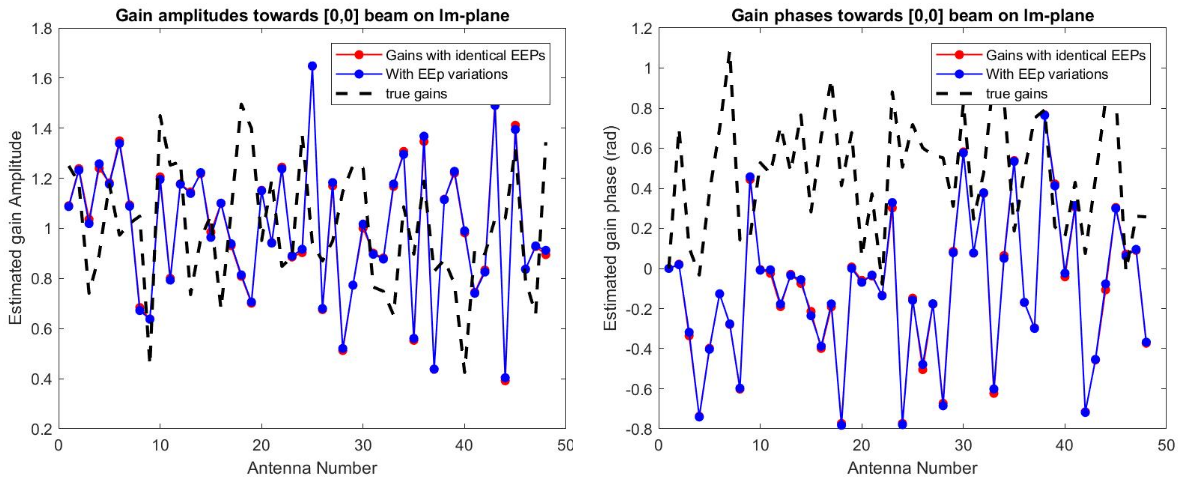
$$\widetilde{\mathbf{r}} = \widetilde{\mathbf{V}}_{sut}\widetilde{\mathbf{v}}_{ref}^{H} \tag{4.9}$$

The obtained correlations from equation 4.9 are normalised using the correlation signal corresponding to the calibration beam, to determine the $B \times 1$ normalised vector $\widetilde{\mathbf{r}}_{norm}$.

The gain estimates $\mathbf{g}_{eep}$ after including individual EEPs in the received signal model are determined using the least squares approach given by equation 4.10.

$$\mathbf{g}_{eep} = (\mathbf{W}_{sut}\mathbf{W}_{sut}^{H})^{-1}(\mathbf{W}_{sut}\widetilde{\mathbf{r}}_{norm}) \tag{4.10}$$

MATLAB code using the above expressions for the simulations is given in appendix C. Figure 4.16 shows a comparison of gain estimates at 35 MHz while the source is at bore sight for a LOFAR-LBA inner station. The received signals are generated with an SNR = 4 for the simulation results in this section. It was observed that the gain estimates were similar at all the operating frequencies.



(a) Amplitude estimates with source (SNR = 4) at bore sight.

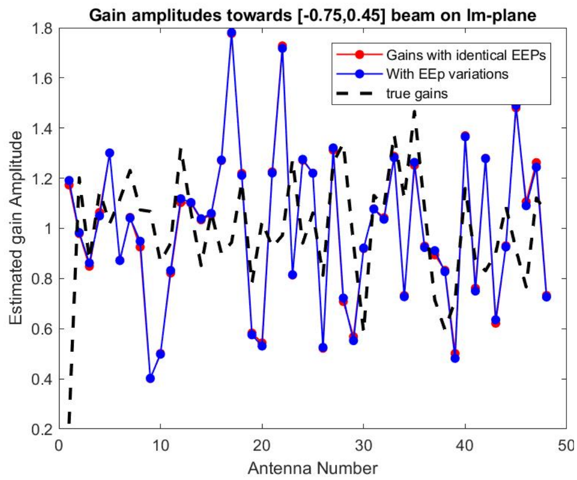
(b) Phase estimates with source (SNR = 4) at bore sight.

Figure 4.16: Comparison of LOFAR-LBA inner estimated gains with and without individual EEPs.

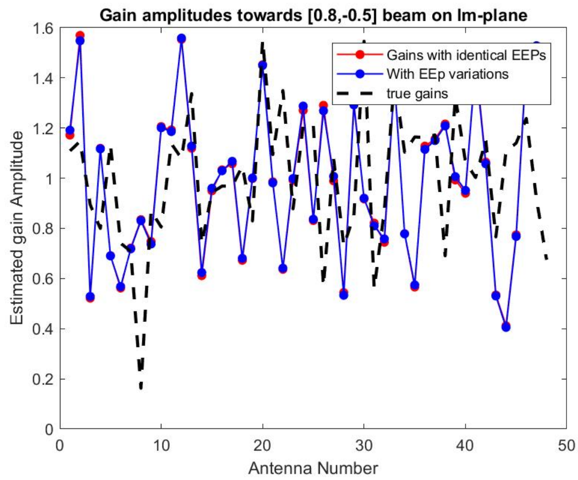
In figure 4.16a, gain amplitude estimates towards bore sight with and without the inclusion of individual EEPs match with the each other while displaying some minor errors with respect to the simulated true gains. The phase estimates with and without the inclusion of individual EEPs display similar errors as the amplitude estimates, as shown in figure 4.16b.

Calculation of EEPs is done using spherical wave functions, typically in polar coordinate system. For the case when the source is at bore sight ((0,0) on lm-plane), it corresponds to $(\theta,\phi) = (0,0)$ in polar coordinates. Calculating EEPs for antennas using spherical wave functions with $(\theta,\phi) = (0,0)$ is undefined. For this reason, EEPs at bore sight in Cartesian coordinate system are approximated to EEPs at $(\theta,\phi) = (0.01,0)$ in polar coordinates.

Figure 4.17a and figure 4.17b show a comparison of the gain amplitude estimates for the LOFAR-LBA inner station when the calibration beam is pointed towards $(-0.75, 0.45)$ and $(0.80, -0.50)$ respectively. The gain estimates after including individual EEPs in the instrument model can be observed to be indistinguishable from the gain estimates with the assumption of identical EEPs. The errors in the gain amplitude estimates with respect to the true gains, however, are relatively very low compared and are expected to be swamped by observation noise.



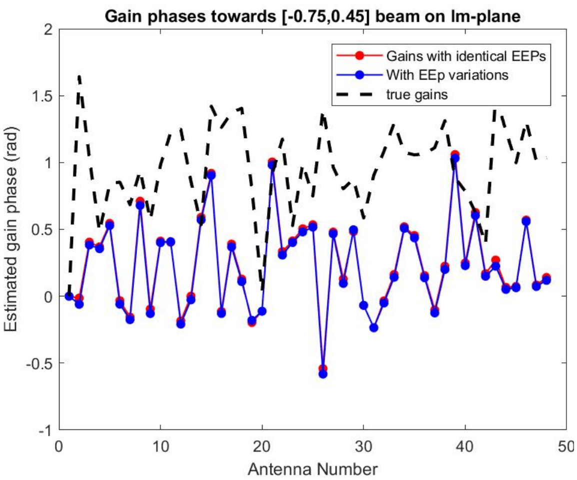
(a) Amplitude estimates with source (SNR = 4) at (-0.75,0.45).



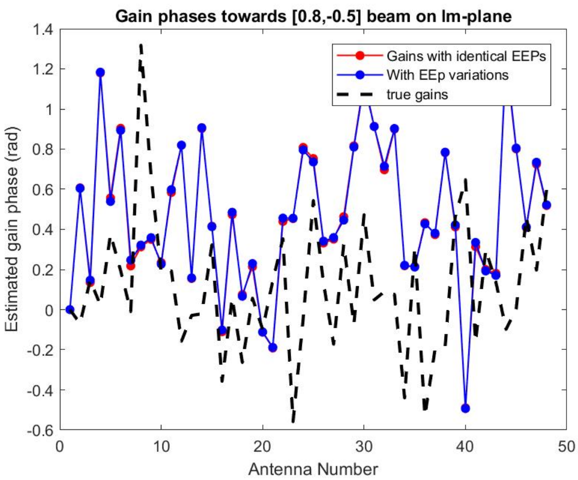
(b) Amplitude estimates with source (SNR = 4) at (0.80,-0.50).

Figure 4.17: Results for LOFAR-LBA inner estimated gain amplitudes for varying source direction.

The estimated gains after incorporating individual EEPs display similar phase variations as the amplitude estimates. The gain phase estimates with and without the assumption of identical EEPs are indiscernible from each other while show visible yet negligible errors compared to the simulated gains. No improvement in the gain estimates can be observed with changes in source positions as shown in figure 4.18a and figure 4.18b for the sources at $(-0.75, 0.45)$ and $(0.80, -0.50)$ respectively. The gain phase estimates in figure 4.18a and figure 4.18b trace each other while showing minimal errors compared to the true gain phases.



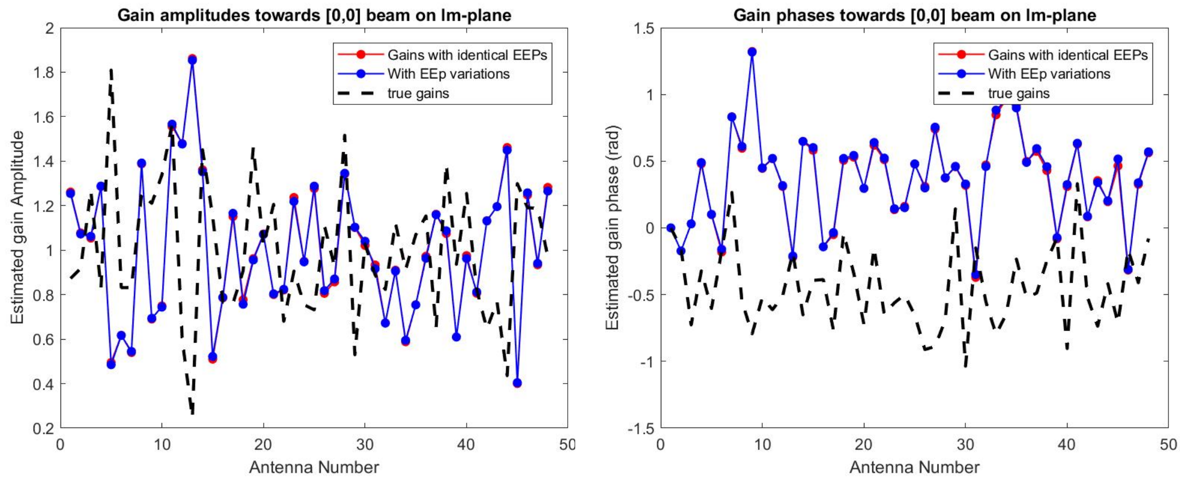
(a) Phase estimates with source (SNR = 4) at (-0.75,0.45).



(b) Phase estimates with source (SNR = 4) at (0.80,-0.50).

Figure 4.18: Results for the LOFAR-LBA inner estimated gain phases for varying source direction.

Similar observations can be made for the LOFAR-LBA outer station. Amplitude estimates, with and without the assumption of identical EEPs, when the station is steered towards a source at bore sight in figure 4.19a match accurately with each other with low variations compared to the true gain amplitudes. Figure 4.19 indicates similar behaviour in gain phase estimates.

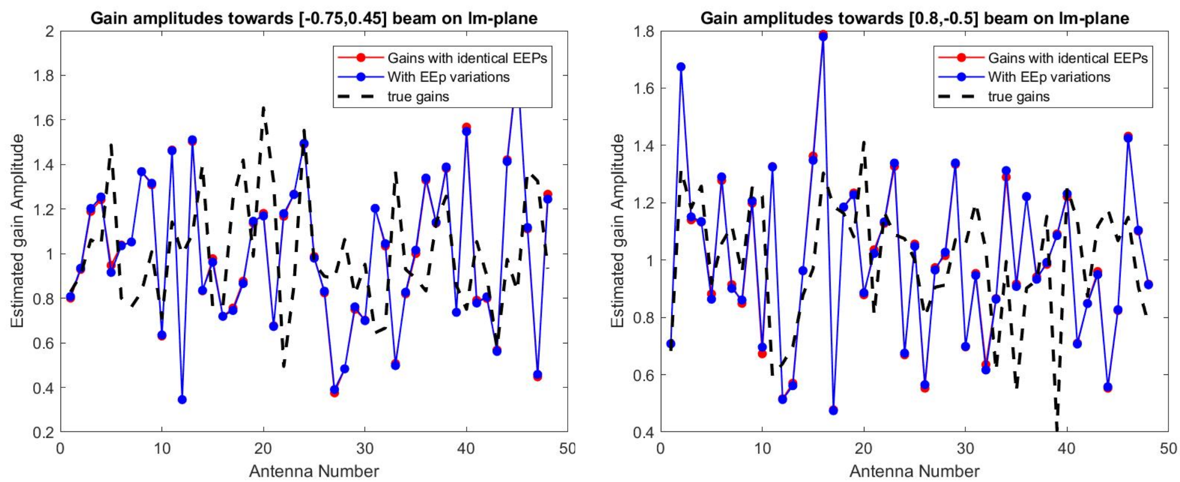


(a) Amplitude estimates with source (SNR = 4) at bore sight.

(b) Phase estimates with source (SNR = 4) at bore sight.

Figure 4.19: Comparison of the LOFAR-LBA outer estimated gains with and without individual EEPs.

The effect of changes in source direction, unusually, is observed to have no impact on the gain estimates. The gain estimates with the inclusion of individual EEPs are primarily influenced by the source direction and the fact that no appreciable improvement in the estimated gains can be observed can be attributed to the low SNR of the observations. The high noise levels as a result of the low SNR render the inclusion of individual EEPs ineffective. This can be further observed from the results for gain phase and amplitude estimates in figure 4.20 and figure 4.21 respectively with the source position varying from $(-0.75, 0.45)$ to $(0.80, -0.50)$ for LBA outer layout.

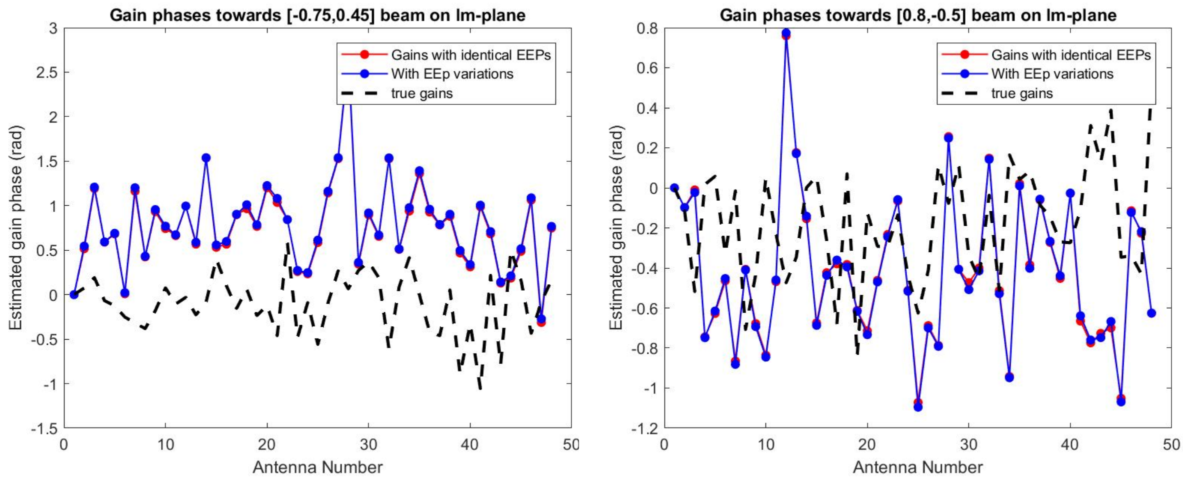


(a) Amplitude estimates with source (SNR = 4) at (-0.75,0.45).

(b) Amplitude estimates with source (SNR = 4) at (0.80,-0.50).

Figure 4.20: Results for LOFAR-LBA outer estimated gain amplitudes for varying source direction.

The results from this section indicate that the gain amplitude and phase estimates with and without the assumption of identical EEPs are identical, it does not quantitatively identify if the errors in these estimates compared to true gains are strong enough to be distinguished from system noise. The errors in gain estimates compared to the simulated gains for varying source directions and for different SNR are calculated in section 4.4.



(a) Phase estimates with source (SNR = 4) at (-0.75,0.45).

(b) Phase estimates with source (SNR = 4) at (0.80,-0.50).

Figure 4.21: Results for LOFAR-LBA outer estimated gain phases for varying source direction.

4.4. Impact of SNR on calibration errors

As mentioned earlier, the effect of noise was ignored in section 4.3 to observe the impact of individual EEPs on gain estimates. This section is focused on identifying the impact of individual EEPs on gain estimates with a simulation model close to the realistic set up of LOFAR-LBA.

The source SNR in the LOFAR-LBA measurements was observed to range between 0.01 to at most 4.0 in section 5.2. Simulation results for SNR within this range are considered in this section at a frequency of 35 MHz.

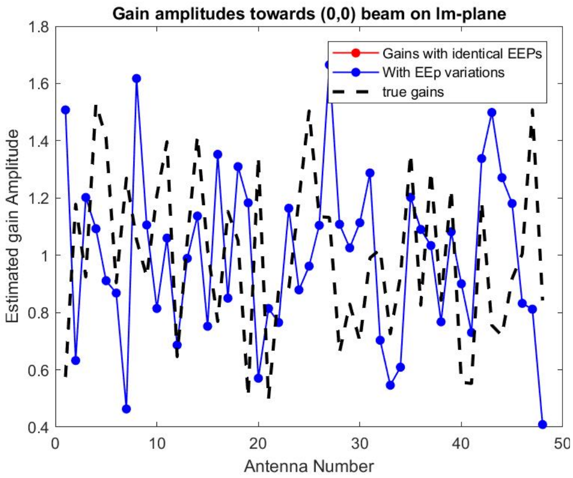
Additionally, random noise following standard normal distribution is modelled in the simulation to achieve the desired SNR per element per sample within the aforementioned range. To further investigate the effect of SNR, magnitudes and phases of the estimated gains with and without incorporating individual EEPs are compared for different SNR and with the source positioned at bore sight and at $(-0.50, -0.85)$ in case of LOFAR-LBA inner layout as shown in figure 4.22 and figure 4.23 respectively.

From figure 4.22, gain amplitude estimates were observed to agree with each other irrespective of SNR even for changes in source positions. This indicates that for LOFAR-LBA observations with calibration sources having low SNR (between 0.001 to 4 as in this case), estimated gains with and without the inclusion of individual EEPs are identical. This argument can be further strengthened with the comparison of phase estimates with the sources at bore sight and at $(-0.50, -0.85)$ in figure 4.23 for LOFAR-LBA inner layout.

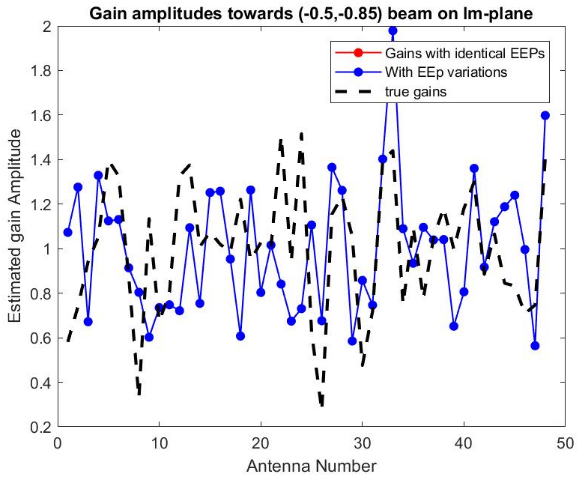
Increasing the SNR within the observed range of 0.001 to 4 was observed to not have any impact on the gain estimates, nor did the variations in source directions.

A common observation that can be made based on the results showing the impact of changing source position and SNR on gain estimates is that the gain estimates after the inclusion individual EEPs are at least as close to the true gains as the gains with the assumption of identical EEPs.

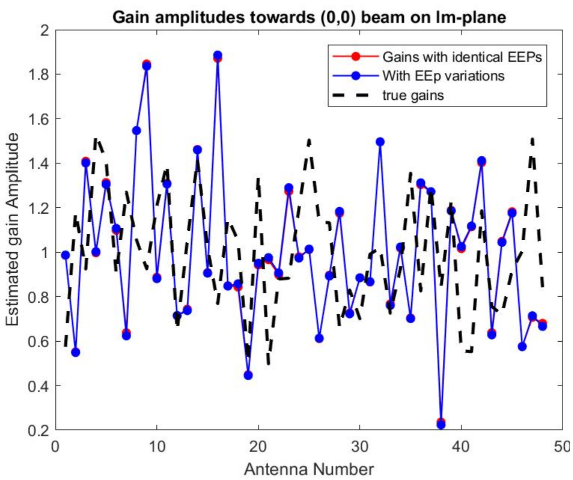
Similar observations can be made from results for the LBA outer configuration shown in figure 4.24 and figure 4.25.



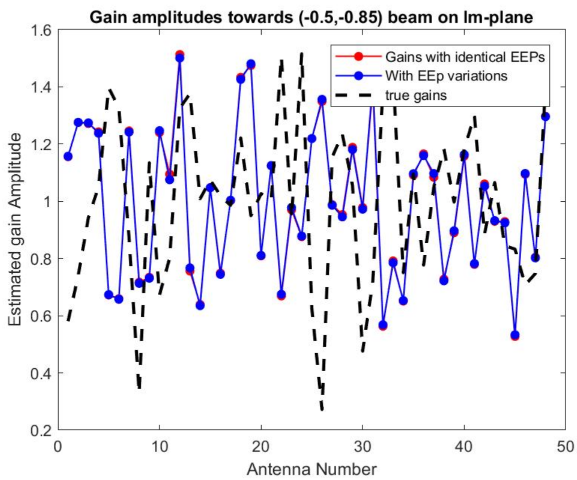
(a) SNR = 0.001, source at $(0,0)$.



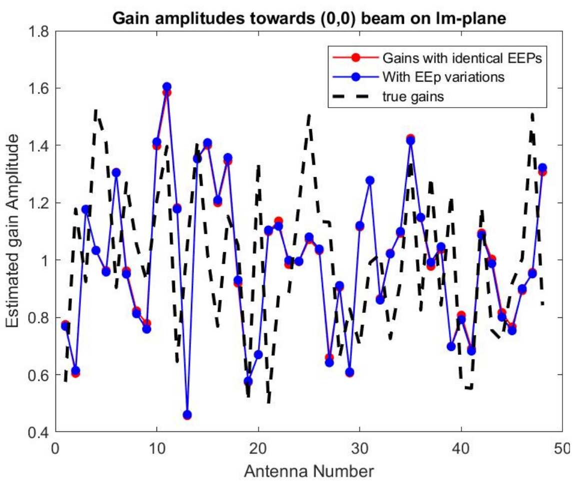
(b) SNR = 0.001, source at $(-0.5,-0.85)$



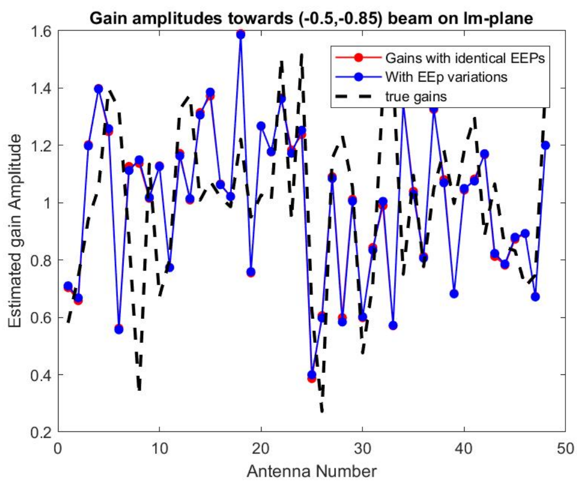
(c) SNR = 2, source at $(0,0)$.



(d) SNR = 2, source at $(-0.5,-0.85)$

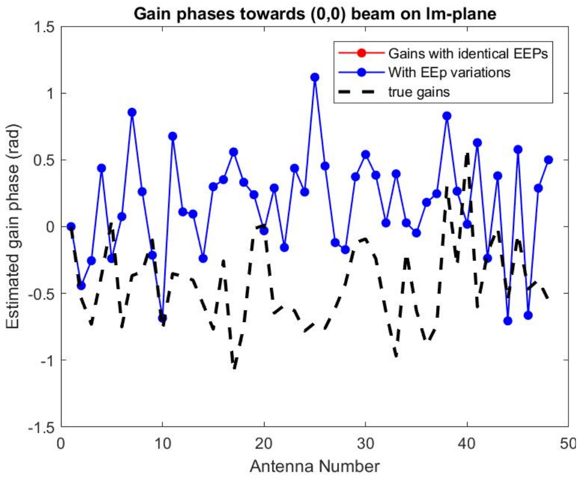


(e) SNR = 4, source at $(0,0)$.

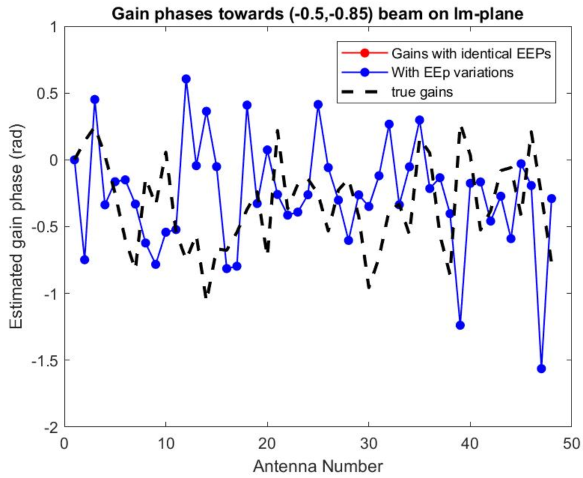


(f) SNR = 4, source at $(-0.5,-0.85)$

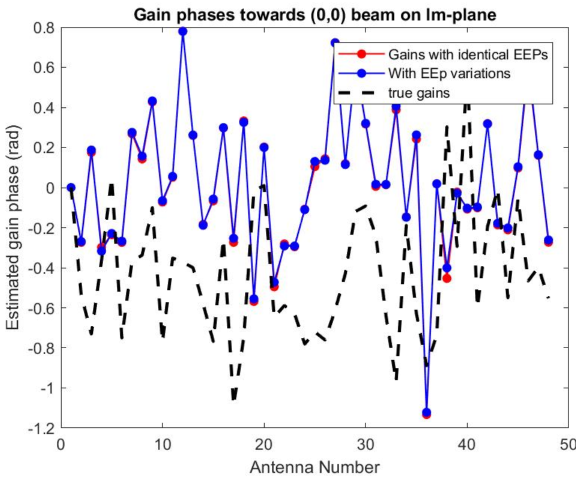
Figure 4.22: Estimated gain amplitude comparison with varying SNR and source direction for LBA inner layout at 65 MHz.



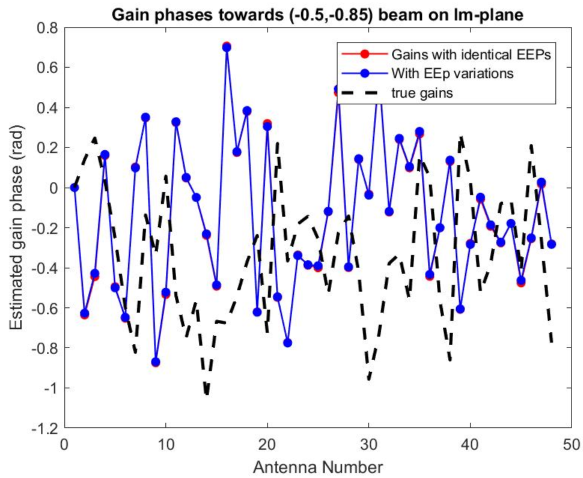
(a) SNR = 0.001, source at (0,0).



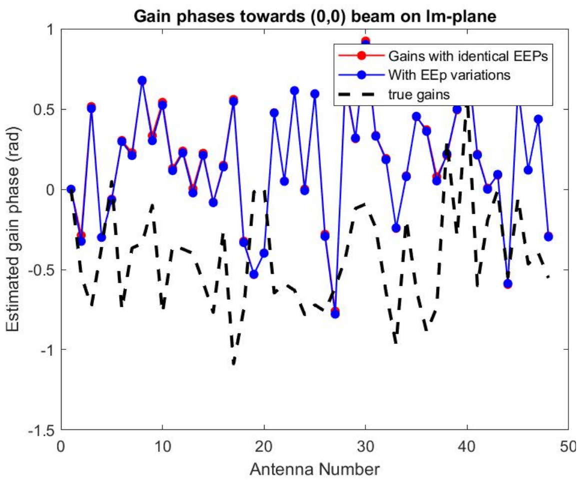
(b) SNR = 0.001, source at $(-0.5, -0.85)$



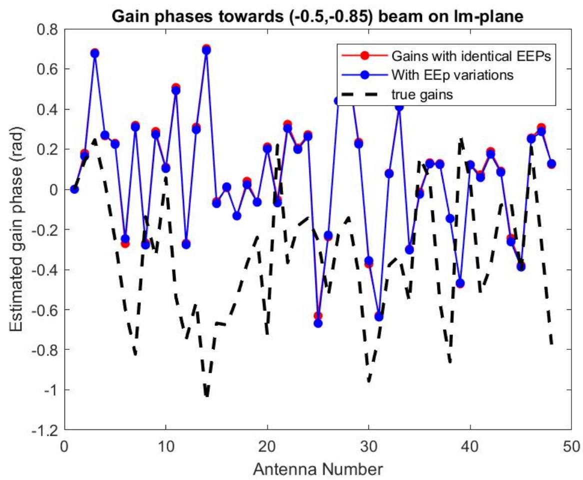
(c) SNR = 2, source at (0,0).



(d) SNR = 2, source at $(-0.5, -0.85)$



(e) SNR = 4, source at (0,0).



(f) SNR = 4, source at $(-0.5, -0.85)$

Figure 4.23: Estimated gain phase comparison with varying SNR and source direction for LBA inner layout at 65 MHz.

From the simulation results, it can be conclusively said that the gain estimates with the assumption of identical EEPs agree closely with true gains with the best possible accuracy within the observed SNR range. For the cases of lower SNR, no appreciable improvement can be observed in the gain amplitude or gain phase estimates after including individual EEPs compared to the gains with the assumption of identical EEPs that can surpass the noise in the observations. These observations can be further validated by comparing the error in gain estimates with varying SNR and source directions.

Table 4.1: Gain amplitude variations for varying SNR towards different source directions for LBA inner at 65 MHz.

Direction	Estimated Gains w/o EEPs			True Gains		
(l,m)	0.001	2	4	0.001	2	4
$(0,0)$	0.01%	0.65%	1.13%	35.60%	34.10%	35.73%
$(0.1,0.1)$	0.02%	1.14%	1.34%	46.52%	39.76%	42.56%
$(-0.4,0.7)$	0.01%	0.70%	0.90%	43.79%	39.31%	42.03%
$(0.3,-0.9)$	0.02%	0.66%	1.12%	43.74%	36.10%	42.55%
$(0.7,0.7)$	0.01%	0.57%	1.20%	36.27%	33.50%	40.85%

The average percentage in error of the gain amplitude estimates after including individual EEPs with respect to true antenna gains and with respect to gain amplitude estimates with the assumption of identical EEPs for SNR = 0.001, SNR = 2 and SNR = 4 with changing source directions are shown in table 4.1 for the LBA inner configuration at 65 MHz. Table 4.2 shows the average errors in gain phase estimates, in radians, for LBA inner layout for the same parameters.

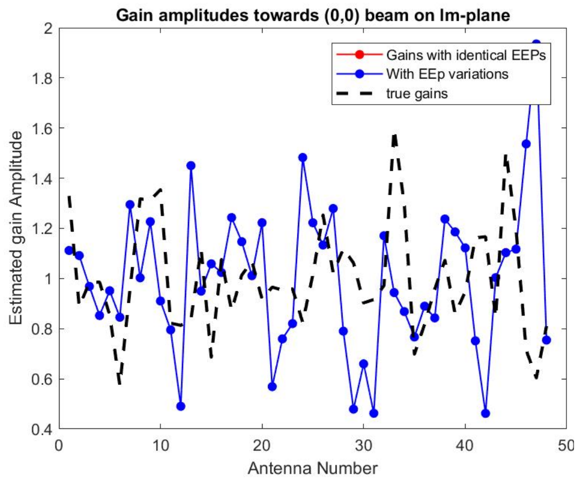
Table 4.2: Gain phase variations (in radians) for varying SNR towards different source directions for LBA inner at 65 MHz.

Direction	Estimated Gains w/o EEPs			True Gains		
(l,m)	0.001	2	4	0.001	2	4
$(0,0)$	2.2×10^{-4}	0.01	0.01	0.57	0.91	0.90
$(0.1,0.1)$	3×10^{-4}	0.01	0.02	0.71	0.57	0.34
$(-0.4,0.7)$	1.6×10^{-4}	0.01	0.02	0.56	0.46	0.40
$(0.3,-0.9)$	1×10^{-4}	0.007	0.01	1.20	0.91	0.54
$(0.7,0.7)$	2×10^{-4}	0.09	0.02	0.55	0.80	0.57

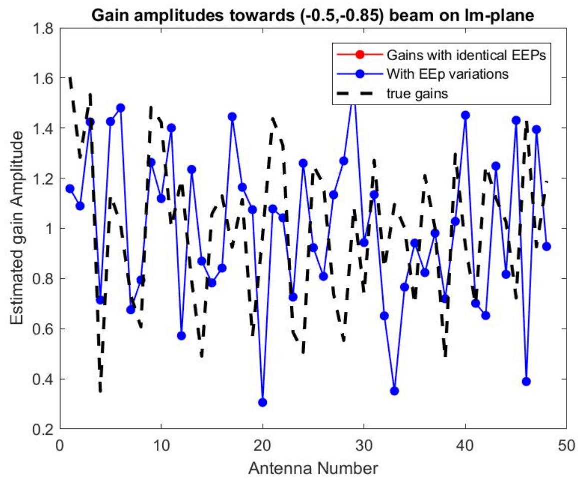
Table 4.3 and table 4.4 show comparison of amplitude and phase estimates after the inclusion of individual EEPs respectively, with respect to gain estimates with the assumption of identical EEPs and with respect to true gains for the LBA outer layout for the same parameters of SNR = (0.001, 2, 4) with changing source directions at 65 MHz.

Based on the results from this section, the following observations can be made.

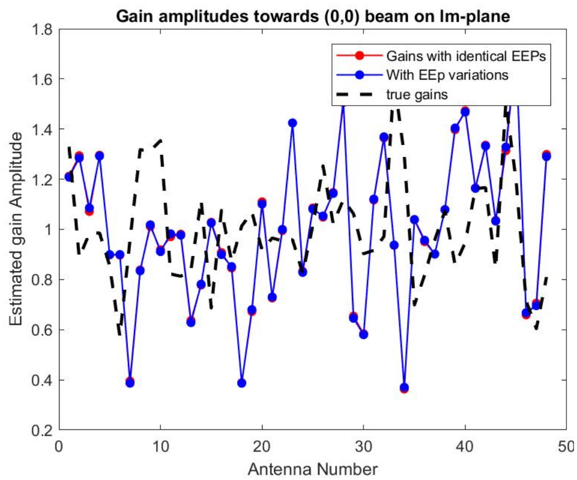
- Results from table 4.1, table 4.2, table 4.3 and table 4.4 indicate very low variations in gain estimates with varying SNR and source directions. The average error in the estimated gain phases with respect to true gains and with respect to gain phase estimates with the assumption of identical EEPs towards the source positions for varying SNR is not greater than 1.2 radians.



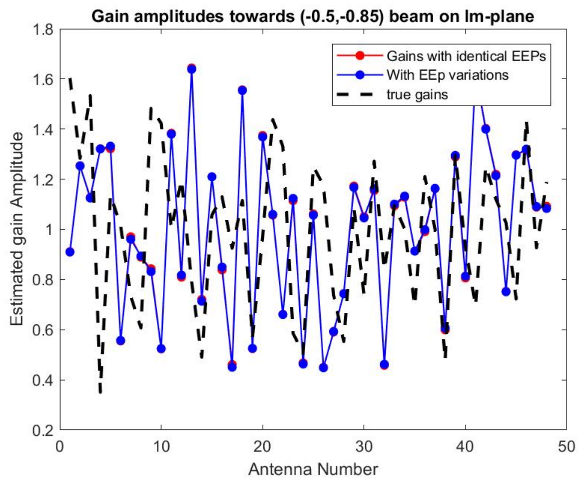
(a) SNR = 0.001, source at bore sight



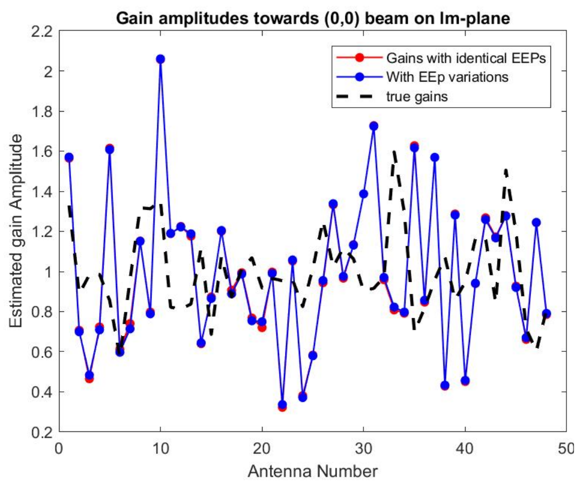
(b) SNR = 0.001, source at $[-0.5, -0.85]$



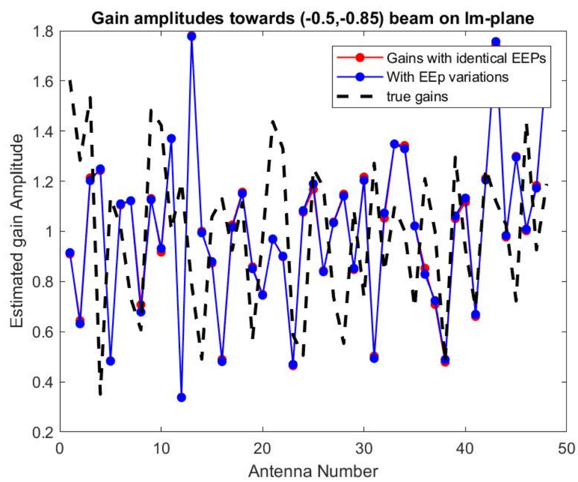
(c) SNR = 2, source at bore sight



(d) SNR = 2, source at $[-0.5, -0.85]$

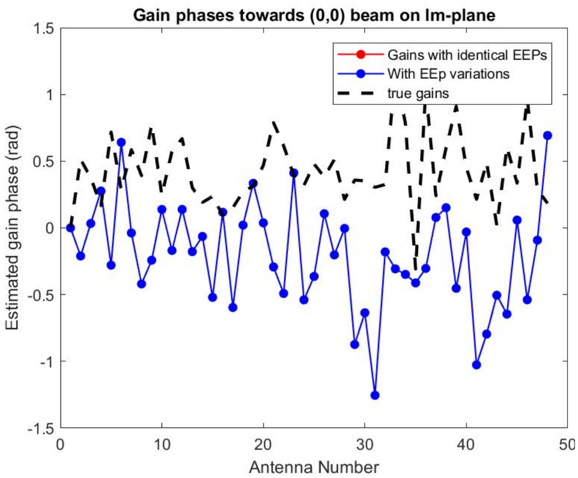


(e) SNR = 4, source at bore sight

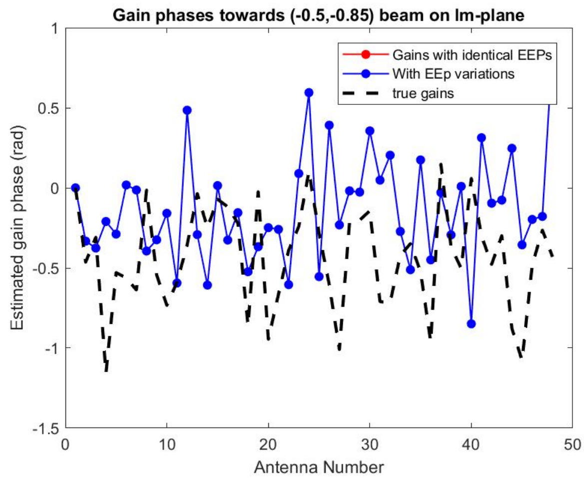


(f) SNR = 4, source at $[-0.5, -0.85]$

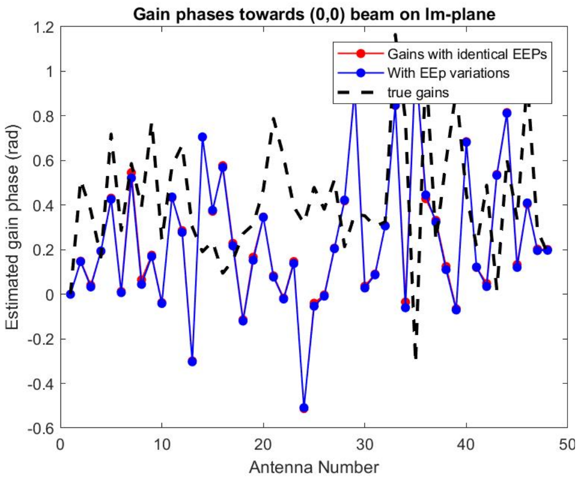
Figure 4.24: Estimated gain amplitude comparison with varying SNR and source direction for LBA outer layout at 65 MHz.



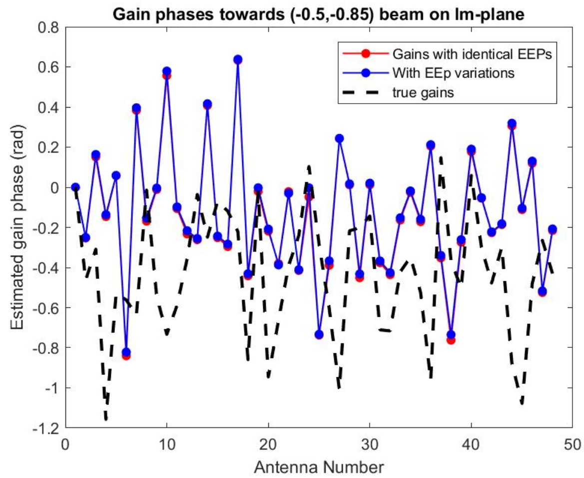
(a) SNR = 0.001, source at bore sight



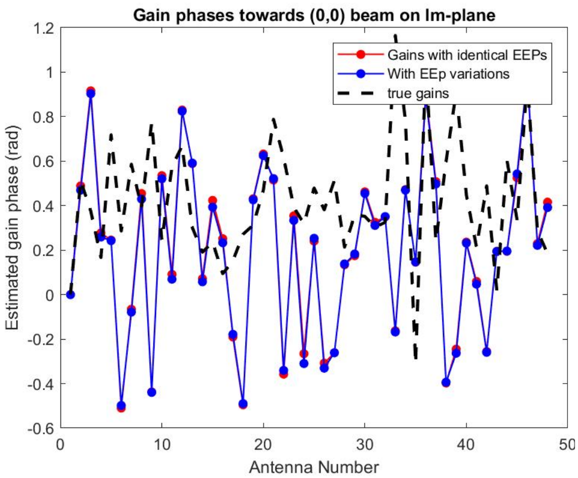
(b) SNR = 0.001, source at $[-0.5, -0.85]$



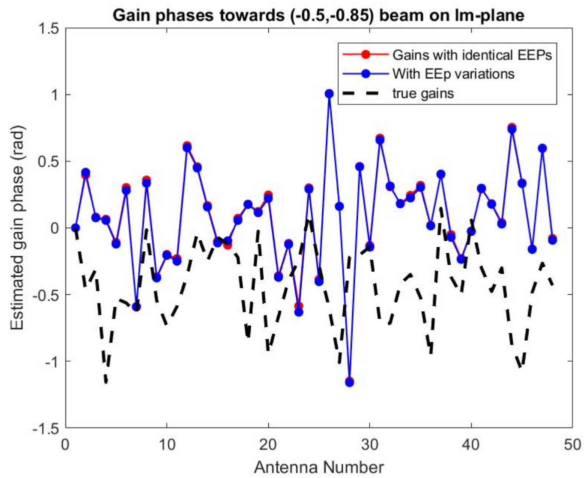
(c) SNR = 2, source at bore sight



(d) SNR = 2, source at $[-0.5, -0.85]$



(e) SNR = 4, source at bore sight



(f) SNR = 4, source at $[-0.5, -0.85]$

Figure 4.25: Estimated gain phase comparison with varying SNR and source direction for LBA outer layout at 65 MHz.

Table 4.3: Gain amplitude variations for varying SNR towards different source directions for LBA outer at 65 MHz.

Direction	Estimated Gains w/o EEPs			True Gains		
(l,m)	0.001	2	4	0.001	2	4
$(0,0)$	0.01%	0.58%	0.83%	42.41%	39.35%	39.18%
$(0.1,0.1)$	0.02%	0.97%	1.61%	35.71%	45.21%	48.60%
$(-0.4,0.7)$	0.02%	0.93%	1.40%	35.67%	33.02%	41.80%
$(0.3,-0.9)$	0.02%	1.00%	1.25%	33.02%	38.86%	36.60%
$(0.7,0.7)$	0.01%	0.58%	0.87%	37.10%	36.47%	35.83%

Table 4.4: Gain phase variations (in radians) for varying SNR towards different source directions for LBA outer at 65 MHz.

Direction	Estimated Gains w/o EEPs			True Gains		
(l,m)	0.001	2	4	0.001	2	4
$(0,0)$	1.6×10^{-4}	0.47	0.01	0.00	0.52	0.57
$(0.1,0.1)$	2.1×10^{-4}	0.42	0.01	0.01	0.33	0.42
$(-0.4,0.7)$	2.4×10^{-4}	0.01	0.01	0.33	0.54	0.57
$(0.3,-0.9)$	3.5×10^{-4}	0.01	0.01	0.67	0.83	0.72
$(0.7,0.7)$	1.4×10^{-4}	0.00	0.01	0.56	0.57	0.62

- The gain amplitude estimates indicate very low average percentage errors among the gain estimates with and without the inclusion of individual EEPs. The errors in gain amplitude estimates with respect to true gains, although display higher average percentage errors, are swamped by the noise in the measurement.

- The errors that are introduced as a result of the assumption of identical EEPs do not pose a threat to station calibration as calibration errors since they are likely to be swamped by the errors introduced as a result of low SNR and ionospheric variability.

- For the calibration of LOFAR-LBA by holography using sources with low SNR (ranging between 0.001 to 4), gain estimates with and without the assumption of identical EEPs are identical. In other words, for sources with such low SNR, calibration by holography with the assumption of identical EEPs is computationally optimal compared to the implementation with individual EEPs.

5

Results from Measurement data

Chapter 4 focused simulations of gain estimates with the inclusion of individual EEPs and the impact of SNR on calibration errors. This chapter is based on calibration of LOFAR-LBA using holography from measured data. Section 5.1 describes the data of the LOFAR-LBA measurement. Section 5.2 is focused on presenting calibration results from holography on LOFAR-LBA concluding chapter 5.

5.1. Data specifications

Each LOFAR-LBA station is accommodated with 48 antennas capable of signal reception in X and Y polarisations. Observed data for the $CS302$ LOFAR-LBA station has correlations at 9 frequencies in it's operational bandwidth, with $RS205LBA$, $RS306LBA$, $RS310LBA$, $RS407LBA$, $RS503LBA$ and $RS509LBA$ as reference stations. Correlations with each reference station have 100 samples of data at 9 frequencies of 19 MHz, 27 MHz, 35 MHz, 42 MHz, 55 MHz, 60 MHz, 65 MHz, 70 MHz and 75 MHz in both XX, XY, YX and YY polarisations.

The four polarisations (XX, XY, YX, YY) are the result of two cross-correlation components XY and YX as result of correlations between X and Y components besides the XX and YY components.

The position of the source in lm-coordinate system is recorded over the 100 second interval. constituting the 100 samples of data, along with the time stamp of the observation. These recorded positions of the source for 100 seconds facilitate the compensation of Earth rotation.

As described in chapter 4, each station forms beams observing sections of the sky. The $CS302$-LBA s.u.t has 188, 353, 478, 481, 481, 481, 481, 481 and 481 beams at the 9 frequencies in the aforementioned order. Each of these beams has the 100 samples of data for all 4 polarisations as a result of the correlations with each of the 6 reference stations.

The calibration source in this observation is 3C196 and the corresponding beam pointing towards this source is the beam 241. The sampling frequency in the measurements is 195 kHz and the system source equivalent flux density is 40kJy at 75 MHz [29].

5.1.1. Issue with measurement data

Based on current work that is being done on holography for LOFAR-HBA, a problem with measurement data has been identified during this project. The s.u.t. had formed fewer beams for observations at low frequencies. Since the number of independent beams formed by the stations at frequencies of 19 MHz and 27 MHz are quite small (183 and 353 beams respectively), they could not be used to obtain accurate calibration solutions for the 48 antenna elements in 2 polarisations in the current implementation. For this reason they have been excluded for further analysis.

5.1.2. SNR from measurement data

The SNR for the correlations with the 6 calibration beams over the 7 operational frequencies displayed undefined values at 75 MHz. For the 6 remaining frequencies, SNR for the correlations of the s.u.t. with the calibration beams of the 6 reference stations are plotted to facilitate comparison of calibration errors against the measurement uncertainty caused by noise levels.

Figure 5.1 and figure 5.2 show SNR vs frequency plots for the 6 LBA remote stations in XX and YY polarisations respectively. The plots indicate that the average SNR for the stations over all frequencies range between 0.01 and 2.4.

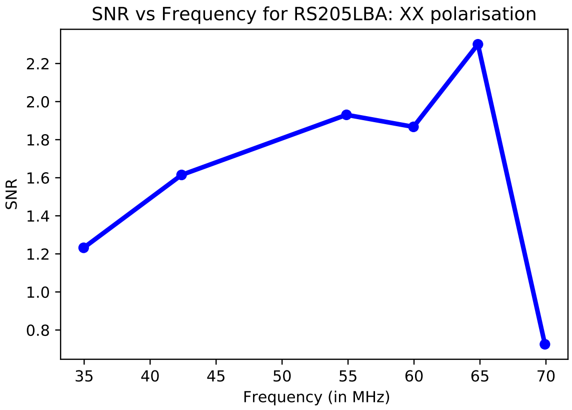
Correlation values over the 100 second interval for $RS509LBA$ station at 35 MHz are frequency shown by figure 5.3 and figure 5.4. The values can be observed to drift away over the 100-s interval as shown in the magnitude and phase correlation plots in figure 5.3 and figure 5.4 respectively. The high spread in these statistics result in an overestimated SNR, explaining the low SNR values in figure 5.1 and figure 5.2. Similar behaviour was observed for correlations with all the 6 reference stations at all frequencies.

In order to improve the accuracy in SNR estimation, the magnitude of the correlations are alone used to determine the SNR. The determined SNR for all 4 polarisations and at all 6 frequencies are averaged over the 6 reference stations. The results are tabulated in table 5.1.

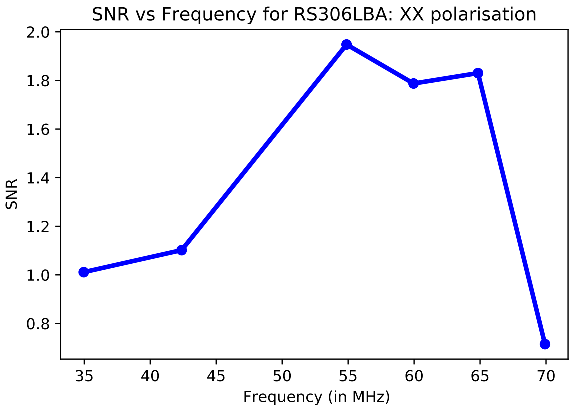
Frequency (MHz)	XX	XY	YX	YY
35	2.55	1.93	2.04	2.60
42	3.27	1.88	1.94	3.08
55	3.65	1.87	1.98	3.83
60	3.87	1.98	1.95	3.69
65	3.34	1.93	2.00	3.55
70	2.68	1.98	2.02	2.71

Table 5.1: Average SNR over the 6 reference stations for XX, XY,YX, YY polarisations

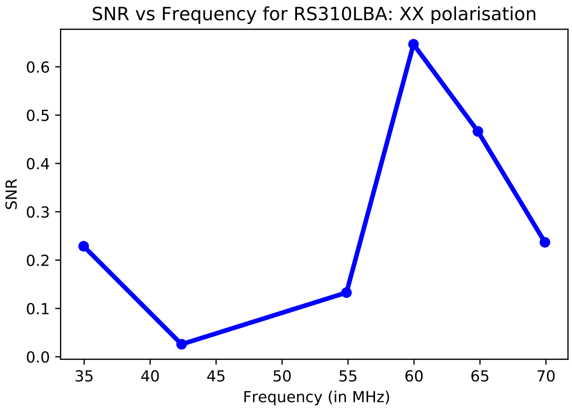
It can be seen from table 5.1 that the average SNR for all the frequencies is about 4. The range of 0.001 to 4 is chosen based on the premise that the least possible SNR for a source to be a calibration source is 0.001 and the observed average SNR from table 5.1 is 4. Table 5.1 shows unbiased average SNR estimates for each frequency and polarisation, unlike the SNR values suggested from figure 5.1 and figure 5.2. These values were used for reference in the simulations from chapter 4.



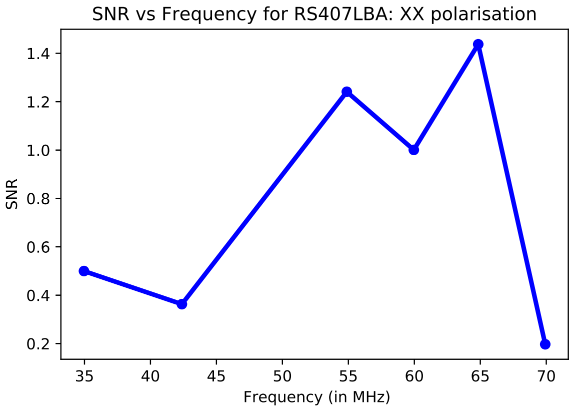
(a) With $RS205LBA$ reference station.



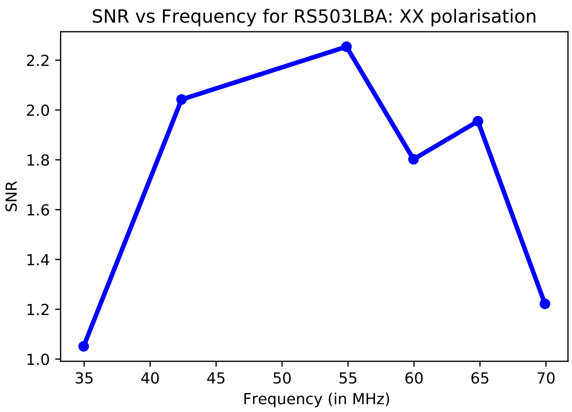
(b) With $RS306LBA$ reference station.



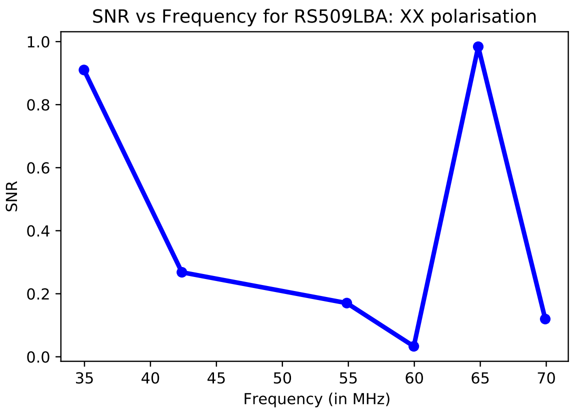
(c) With $RS310LBA$ reference station.



(d) With $RS407LBA$ reference station.



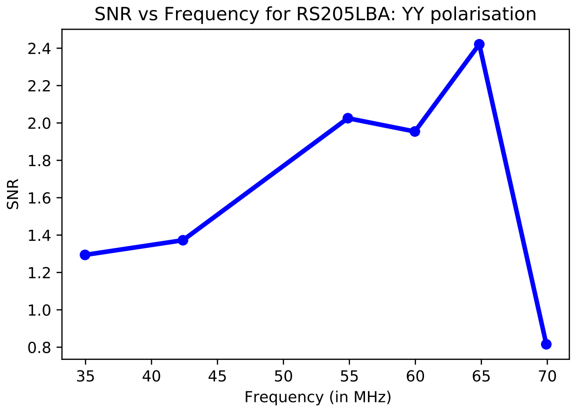
(e) With $RS503LBA$ reference station.



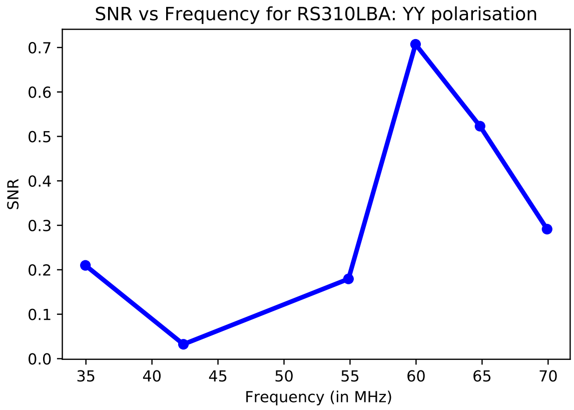
(f) With $RS509LBA$ reference station.

Figure 5.1: SNR of the correlations of CS302 with each reference station vs frequency for XX polarisation.

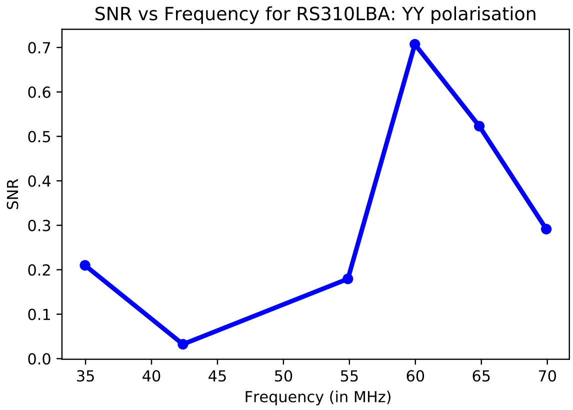
For observations with such low SNR, the noise levels in the correlation signals dominate the variations introduced in estimated gains with the inclusion of EEP variations. In other words, improvement in estimated gains with the inclusion of EEP variations is unnoticed, as also previously noticed from results in section 4.4. The reason for this observation is unfortunately ambiguous and could not be conclusively identified owing to the low SNR of the signal correlations.



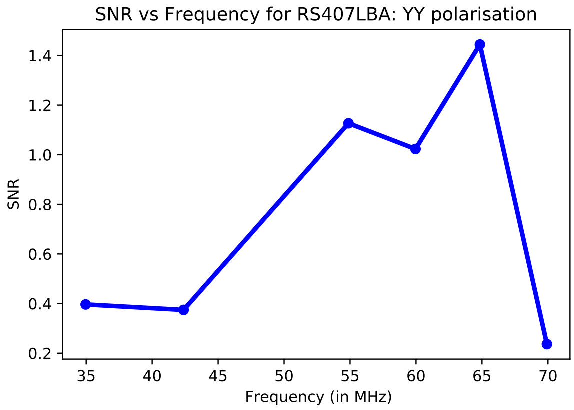
(a) With $RS205LBA$ reference station.



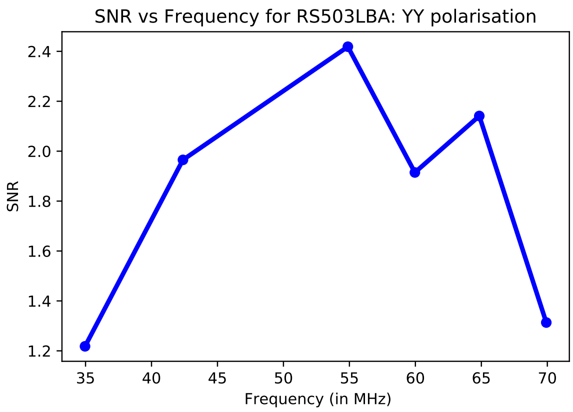
(b) With $RS306LBA$ reference station.



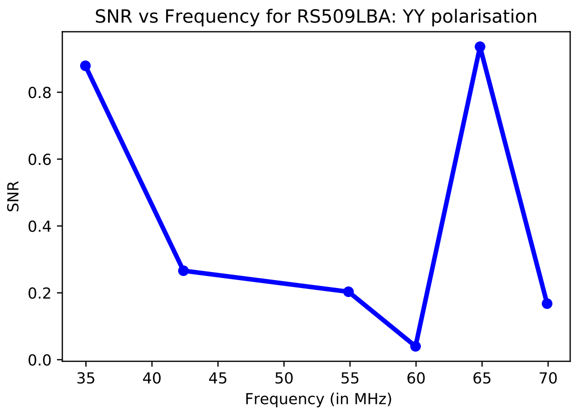
(c) With $RS310LBA$ reference station.



(d) With $RS407LBA$ reference station.



(e) With $RS503LBA$ reference station.



(f) With $RS509LBA$ reference station.

Figure 5.2: SNR of the correlations of CS302 with each reference station vs frequency for YY polarisation.

5.2. Results from measurement data

Holography was applied to CS302 LOFAR-LBA station based on the current implementation available for LOFAR-HBA. During this project, several issues were solved that prevented the holography code currently in use for the LOFAR HBA system from working with LOFAR LBA data. Even after attending to the issues at hand, no conclusive results for calibration of LOFAR-LBA by holography could be made yet.

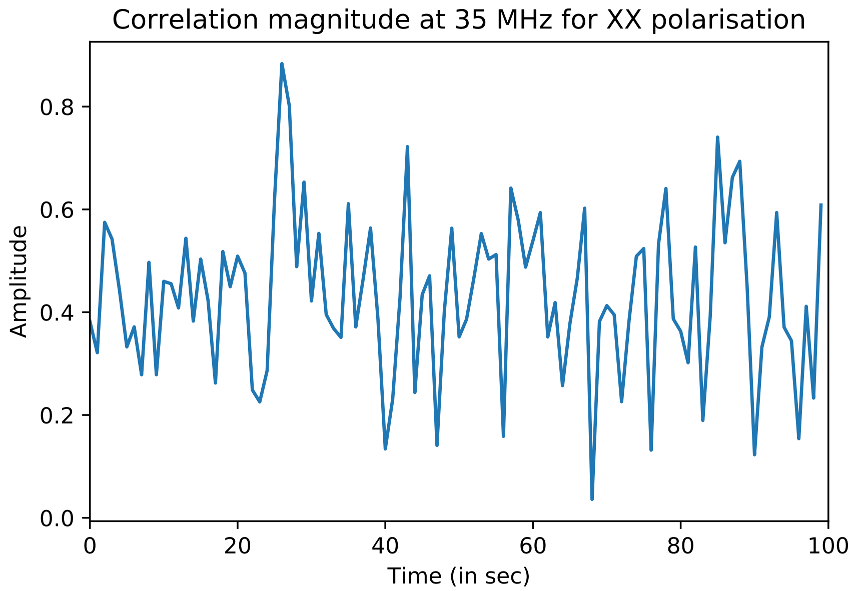


Figure 5.3: XX polarisation correlation magnitude for RS509 LBA station over the 100 second interval.

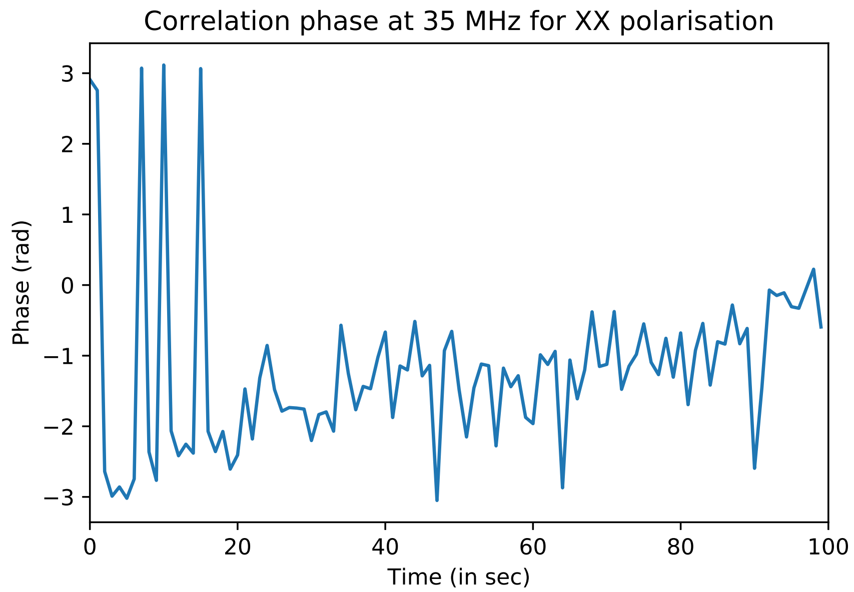


Figure 5.4: XX polarisation correlation phase for RS509 LBA station over the 100 second interval.

6
Conclusions

Holography is a computationally efficient strategy to incorporate embedded element patterns in the calibration model compared to superstation calibration and stand-alone station calibration.

Holography was implemented on LOFAR-LBA as a simulation. Individual EEPs were included in the instrument model during the simulation, while calibration was still done based on the model derived with the assumption of identical EEPs. Following conclusions were made based on the results from the simulation:

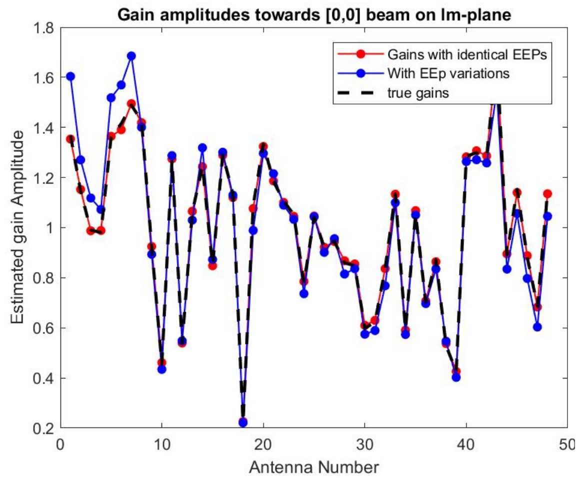
- For calibration of LOFAR-LBA by holography using sources with such low SNR ranging from 0.001 to 4, the gain estimates with and without the assumption of identical EEPs are identical and display indiscernible errors compared to true gains. This essentially means that calibration by holography with the assumption of identical EEPs is computationally optimal compared to the implementation with individual EEPs.

- Gain estimates with and without the assumption of identical EEPs closely trace each other irrespective of SNR and source direction variations.

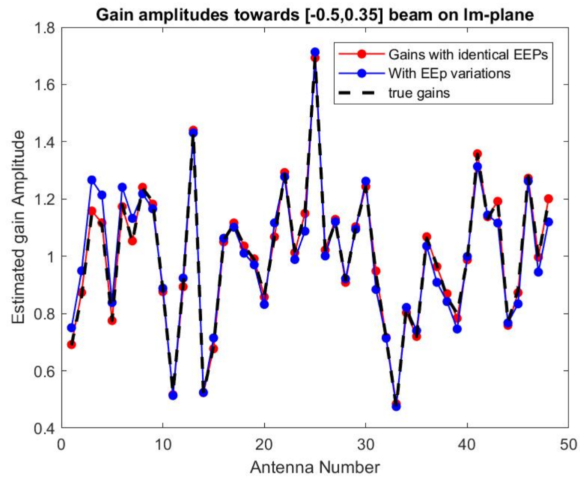
Holography was implemented on the CS302 LOFAR-LBA measurement data. During this implementation, the holography code used for LOFAR-HBA was improved to be able to work on LOFAR-LBA, issues with LOFAR-LBA measurement data were identified and addressed. The impact of gain estimate variations from calibration results using measurement data could not be successfully assessed owing to the low SNR in the observations and ionospheric variability.

The SNR of the calibration source measured from the correlations with the 6 remote reference stations determines the impact of including individual EEPs in the instrument model. If correlations with much higher SNR could be made, the calibration errors introduced because of the assumption of identical EEPs could be noticed and may even need to be corrected for. SNR of the calibration source can be improved by using more reference stations and hence more calibration beams. Counteracting the ionospheric variations could further improve the SNR of the calibration source.

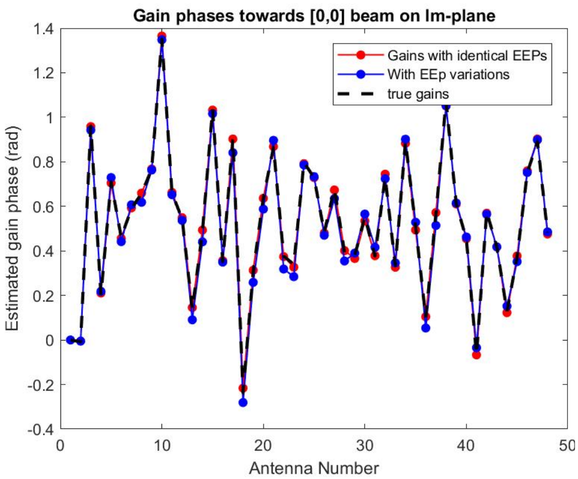
Through the simulation developed during this project, observations with an SNR of 50 dB or above, per station, was observed to improve calibration results, with both the gain estimates agreeing with the simulated true gains. Figure 6.1 shows gain phase and amplitude estimates for the LOFAR-LBA inner layout for 70 dB SNR towards the source at bore sight and at $(-0.5, 0.35)$. Similar results were observed for 70 dB SNR towards the same sources with the LOFAR-LBA outer layout as shown in figure 6.2. These results illustrate that, with observations having an SNR as high as 70 dB, calibration of LOFAR-LBA by holography with the assumption of identical EEPs performs at least as well as calibration of LOFAR-LBA by holography with the inclusion of individual EEPs.



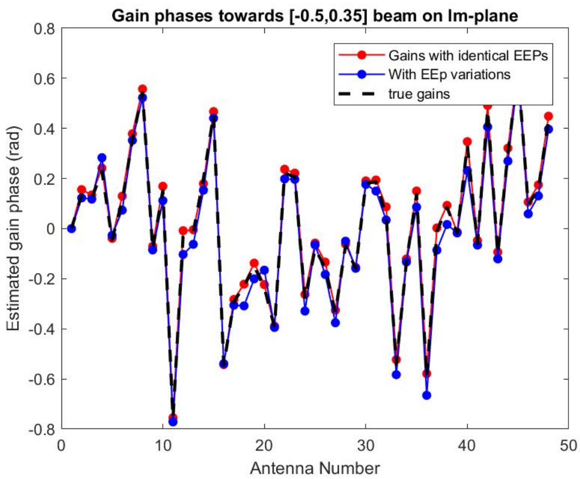
(a) Amplitude estimates with the source at bore sight.



(b) Amplitude estimates with the source at $(-0.5, 0.35)$

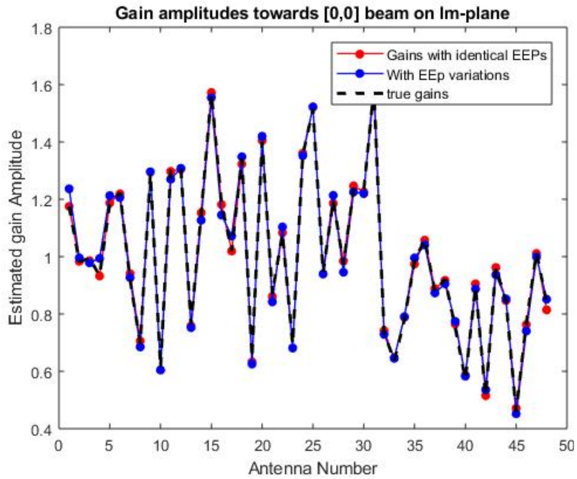


(c) Phase estimates with the source at bore sight.

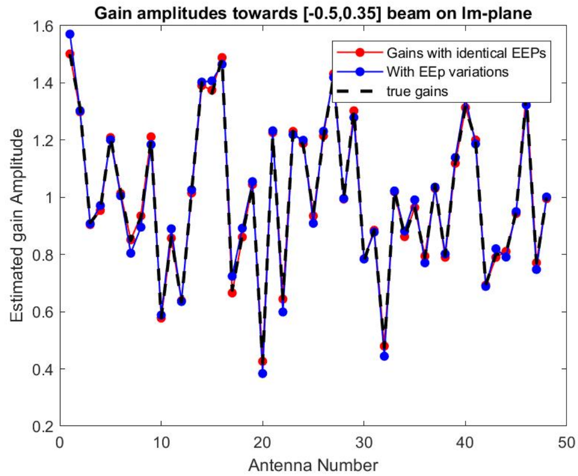


(d) Phase estimates with the source at $(-0.5, 0.35)$.

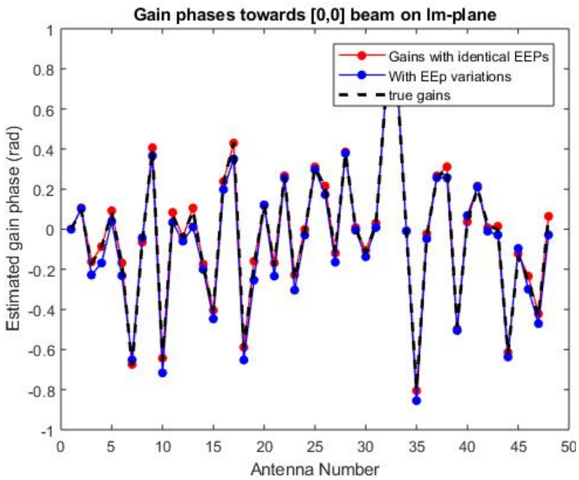
Figure 6.1: Comparison of gain estimates for SNR = 70 dB with varying source direction for LBA inner layout at 35 MHz.



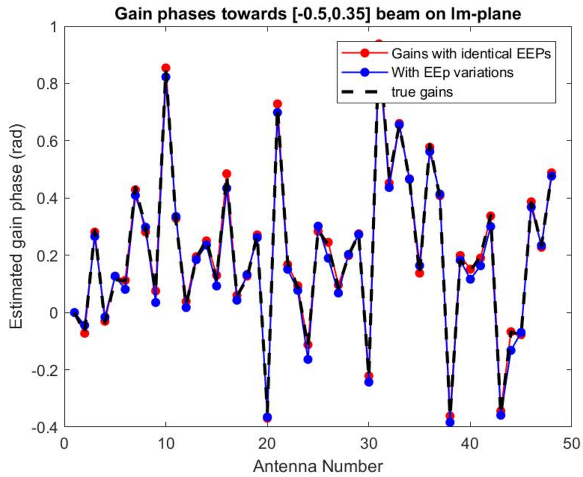
(a) Amplitude estimates with the source at bore sight.



(b) Amplitude estimates with the source at $(-0.5, 0.35)$



(c) Phase estimates with the source at bore sight.



(d) Phase estimates with the source at $(-0.5, 0.35)$.

Figure 6.2: Comparison of gain estimates for SNR = 70 dB with varying source direction for LBA outer layout at 35 MHz.

A

MATLAB code: EEP variations

A.1. MATLAB code for EEP variations in LOFAR-LBA

```
% Description: Embedded element pattern calculation
%       for each antenna
% Calculation for 1. XX polarisation (pol = 1)
%               2. Inner station LBA configuration
%
%               Inputs:
% Frequency    = 57 MHz, resonant frequency.
%              higher EEPs at resonant frequency
% phase_res    = False (EEPs calculated with
% respective antenna as phase reference centre
% [theta,phi] = all-sky coordinates in polar form
% LOFAR-LBA inner antenna positions
%               Output:
% eep_th_phi   = EEPs for each antenna towards
%                   each [theta,phi].
%    Has: theta-component , phi-component
% Functions used= reconstruct: to calculate EEPs


freq        = 57*1e6;
phase_res   = false;
Pol         = 1;

ants        = [1,48];
N_ant       = length(ants);

theta       = 1:1:90;
phi         = 0:5:360;
eeps_th_phi = zeros([length(theta), length(phi),N_ant,2]);

%   use El = El +48 for LBA outer
```

```
for El       = 1:N_ant
    EEPs = reconstruct(ants(El),freq./1e6,theta,phi,Pol,phase_res);
    theta_comp      = EEPs(:,:,1);
    phi_comp        = EEPs(:,:,2);
    eeps_th_phi(:,:,(El),1) = squeeze(theta_comp);
    eeps_th_phi(:,:,(El),2) = squeeze(phi_comp);
end
% [ph,th] = meshgrid(phi,theta);

e1      = squeeze(eeps_th_phi(:,:,1,:));
elast   = squeeze(eeps_th_phi(:,:,2,:));

figure;
plot(theta,abs(e1(:,phi==0,2)),'b')
grid on
hold on
plot(theta,abs(elast(:,phi==0,2)),'r')
xlabel('Theta (degree)')
ylabel('V')
title(['EEP (amplitude) comparison at ' num2str(freq/1e6) ' MHz for \phi = 0 deg.'
legend('Antenna 1 \theta-component','Antenna 48 \theta-component')

figure;
plot(theta,unwrap(angle(e1(:,phi==0,2)))./pi.*180,'b')
grid on
hold on
plot(theta,unwrap(angle(elast(:,phi==0,2)))./pi.*180,'r')
xlabel('Theta (degree)')
ylabel('Degree')
title(['EEP (phase) comparison at ' num2str(freq/1e6) ' MHz for \phi = 0 deg.'])
legend('Antenna 1 \theta-component','Antenna 48 \theta-component')

figure;
plot(phi,abs(e1(theta==90,:,1)),'b')
grid on
hold on
plot(phi,abs(elast(theta==90,:,1)),'r')
xlabel('Phi(degree)')
ylabel('V')
title(['EEP (amplitude) comparison at ' num2str(freq/1e6) ' MHz for \theta = 90 deg
legend('Antenna 1 \phi-component','Antenna 48 \phi-component')

figure;
plot(phi,unwrap(angle(e1(theta==90,:,1)))./pi.*180,'b')
grid on
hold on
```

```
plot(phi,unwrap(angle(elast(theta==90,:,1)))./pi.*180,'r')
xlabel('Phi (degree)')
ylabel('Degree')
title(['EEP (phase) comparison at ' num2str(freq/1e6) ' MHz for \theta = 90 deg.'])
legend('Antenna 1 \phi-component','Antenna 48 \phi-component')
```

A.2. Calculation of EEPs for each antenna element

```
function Erecon=reconstruct(El,Freq,Theta,Phi,Pol,Phase_restore)
%
% Erecon=reconstruct(El,Freq,Pol,Phase_restore)
%    Contents of loaded files
%
%    Freq0: Frequencies for which the coefficients
%           are calculated(MHz)
%    Q81:   matrix containing the spherical wave coefficients.
%           Format: coefficient index x frequency index.
%
%    LBA.P: x-coordinate of the positions of each element
%    LBA.Q: y-coordinate of the positions of each element
%
% Format:
%
%   Input:
%     El:             element number
%     Freq:           Frequency (MHz)
%     Theta:          theta-angles for which the EEP
%                     will be calculated (degrees)
%     Phi:            phi-angles for which the EEP
%                     will be calculated (degrees)
%     Pol:            Polarization, 1=dipole in phi=45 deg. plane,
%                    2=dipole in phi=135 deg. plane
%     Phase_restore: true if phase reference of
%                     the EEPs should be at
%                     the origin of the coordinate system
%                     false if the phase reference is at
%                     position of each antenna element.
%
%   Output:
%     Erecon:         the calculated embedded element pattern.
%                     l-index x m-index x component
%                     (1=theta-component, 2=phi-component)

% The function F4far_new evaluates the spherical wave functions
% at the specified angles.

load (['Q81_el' num2str(El) '_pol' num2str(Pol) '.mat'])
```

```
Q=Q81(:,Freq0==Freq);
ThetaMat=Theta.'*ones(1,length(Phi));
PhiMat=ones(length(Theta),1)*Phi;
ThetaLin=ThetaMat(:).';
PhiLin=PhiMat(:).';

beta=2.*pi.*Freq.*1e6./3e8;
jn=1;
Nmax=21;
F=zeros(2.*length(ThetaLin),2*Nmax*(Nmax+2));
for n=1:Nmax
    for m=-n:n
        for s=1:2
            [~,q2,q3]=F4far_new(s,m,n,ThetaLin,PhiLin,beta);
            F(:,jn)=[q2.' ; q3.'];
            jn=jn+1;
        end
    end
end
Erecon=F*Q;
Erecon=reshape(Erecon,length(Theta),length(Phi),2);

if Phase_restore
    K=2.*pi.*Freq.*1e6./3e8;
    load CS302_coords
    Positions=[LBA.P LBA.Q];
        Erecon(:,:,1)=Erecon(:,:,1).* ...
            exp(1i.*K.*((ThetaMat).*
            Positions(El,1)+ (PhiMat).*Positions(El,2)));
        Erecon(:,:,2)=Erecon(:,:,2).* ...
            exp(1i.*K.*((ThetaMat).*
            Positions(El,1)+ (PhiMat).*Positions(El,2)));
else
    return
end
```

A.3. Function to evaluate spherical wave functions

```
function [q1,q2,q3]=F4far_new(s,m,n,theta,phi,beta)
if s==1
    [q1,q2,q3]=F41(m,n,theta,phi,beta);
elseif s==2
    [q1,q2,q3]=F42(m,n,theta,phi,beta);
end
end

function q=P(m,n,x)
    q=legendre(n,x);
    if n>0
        q=q(abs(m)+1,:);
    end
    if m<0
        q=-1.^-m.*factorial(n+m)./factorial(n-m).*q;
    end
end

function q=Pacc(m,n,z)
q=(-(n+m).*(n-m+1).*sqrt(1-z.^2).*
    P(m-1,n,z)-m.*z.*P(m,n,z))./(z.^2-1);
end

function [q1,q2,q3]=F41(m,n,theta,phi,beta)
if m~=0
    Const=beta.*sqrt(60).*1./sqrt(n.*(n+1)).*(-m./abs(m)).^m;
else
    Const=beta.*sqrt(60).*1./sqrt(n.*(n+1));
end
q1=zeros(1,length(theta));

q2=Const.*(-1i).^(-n-1)./beta.*1i.*m./(sind(theta)).*
        sqrt((2.*n+1)./2.*factorial(n-abs(m))./factorial(n+abs(m))).*
        P(abs(m),n,cosd(theta)).*exp(1i.*m.*phi./180.*pi);

q3=Const.*(-1i).^(-n-1)./beta.*sqrt((2.*n+1)./2.*factorial(n-abs(m))
        ./factorial(n+abs(m))).*Pacc(abs(m),n,cosd(theta)).*
        sind(theta).*exp(1i.*m.*phi./180.*pi);
end

function [q1,q2,q3]=F42(m,n,theta,phi,beta)
if m~=0
    Const=beta.*sqrt(60).*1./sqrt(n.*(n+1)).*(-m./abs(m)).^m;
else
    Const=beta.*sqrt(60).*1./sqrt(n.*(n+1));
```

```
end
q1=zeros(1,length(theta));

q2=-Const.*(-1i).^-n./beta.*sqrt((2.*n+1)./2.*factorial(n-abs(m)).
        /factorial(n+abs(m))).*Pacc(abs(m),n,cosd(theta)).*
        sind(theta).*exp(1i.*m.*phi./180.*pi);

q3=Const.*(-1i).^-n./beta.*1i.*m./sind(theta).*sqrt((2.*n+1)./2.*
        factorial(n-abs(m))./factorial(n+abs(m))).*
        P(abs(m),n,cosd(theta)).*exp(1i.*m.*phi./180.*pi);
end
```

B

MATLAB simulation: Holography without EEP variations

```
%% Comparison of estimated gains from
% holography with and without EEP variations

% SIMULATION VARIABLES:

% tau                 :   integation time in seconds
% fs                  :   LOFAR subband sampling
%                         frequency in Hz at Nyquist rate
% Ts                  :   Sampling time in seconds
% Ns                  :   Number of samples =  length of Ts
% SNR                 :   SNR of the correaltions in dB
% freq                :   [35, 42, 54, 59, 65, 69, 75] MHz
% [l,m]               :   lm-plane coordiantes range: -1 to 1
% [l_ref,m_ref]       :   Reference-central beam steering coordinates
% pos                 :   CS302 LOFAR-LBA station coordiantes
% g_est               :   estiamted gains with identical EEP assumption
% w_sut               :   Beamforming Matrix
% s_ref               :   Unpolarised source signal
%                         (s_theta- theta component, s_phi- phi component)
% a_ref               :   Array response vector

c           = 3e8;
tau         = 1e-3;
fs          = 195312.5;
Ts          = (0:1/fs:tau);
Ns          = length(Ts);
SNR         = 1e3;

freq        = 35e6;
l           = -1:0.05:1;
m           = -1:0.05:1;
```

```matlab
[lgrid, mgrid] = meshgrid(l, m);
%%  Calculate EEPs, load antenna positions

load('antpos_CS302_LBA_INNER.mat')
% load('antpos_CS302_LBA_OUTER.mat')

N_ant       = length(pos);
x_stat      = pos(:,1);
y_stat      = pos(:,2);


%% lm-plane coordinates

l_ref       = 0;
m_ref       = 0;

l_index     = find(abs(l-l_ref)<1e-6);
m_index     = find(abs(m-m_ref)<1e-6);
ind         = sub2ind(size(lgrid),l_index,m_index);

s_theta     = randn(1, Ns) + 1i * randn(1, Ns);
s_phi       = randn(1, Ns) + 1i * randn(1, Ns);

s_ref       = [s_theta ; s_phi];

tau_ref     = (x_stat * l_ref + y_stat * m_ref) / c;
a_ref       = exp(-2i * pi * tau_ref * freq);
w_ref       = a_ref;

tau_sut     = (x_stat * lgrid(:).' + y_stat * mgrid(:).') / c;
a_sut       = exp(-2i * pi * tau_sut * freq);

w_sut       = a_sut';

gains       = ones(N_ant, 1) +
            0.3 * (randn(N_ant, 1) + 1i * randn(N_ant, 1));


%% With identical EEP

x_ref       = diag(gains) * [a_ref, a_ref] * s_ref;

noise       = (randn(N_ant, Ns) + 1i * randn(N_ant, Ns))
            /(sqrt(SNR/Ns)/N_ant);
x_ref       = x_ref + noise;
```

```
v_ref       = w_ref' * x_ref;

v_out_sut   = w_sut * x_ref;
r           = v_out_sut * v_ref';


%% Perform holography: Estimate gains

r_norm      = r / r(ind); % normalize measurements on central beam

g_est       = (w_sut' * w_sut) \ w_sut' * r_norm;
%%  Compare gains

% correct for source that is not in bore sight
g_est = g_est ./ a_ref;

% use first element as phase reference to facilitate comparison
g_est = g_est ./ (g_est(1) / abs(g_est(1)));
gains = gains ./ (gains(1) / abs(gains(1)));

% normalise on average gain to facilitate comparison
g_est = g_est / mean(abs(g_est));
gains = gains / mean(abs(gains));

%%
figure;
plot(1:N_ant, abs(g_est), 'r-', ...
     1:N_ant, abs(gains), 'mo');
legend('Gains with identical EEPs',  'true gains')
xlabel('Antenna Number');
ylabel('Estimated gain Amplitude');
title(['LOFAR LBA Inner: Gain amplitudes for source at
        [', num2str(l_ref) ',' num2str(m_ref) ']']);

figure;
plot(1:N_ant, angle(g_est), 'r-', ...
     1:N_ant, angle(gains), 'mo');
legend('Gains with identical EEPs', 'true gains')
xlabel('Antenna Number');
ylabel('Estimated gain phase ');
title(['LOFAR LBA Inner: Gain phases for source at
        [', num2str(l_ref) ',' num2str(m_ref) ']']);
```

C

MATLAB simulation: Holography with EEP variations

C.1. Comparison of gain estimates with and without the incorporation of EEPs

```
%% Comparison of estiamted gains from holography
%       with and without EEP variations

% SIMULATION VARIABLES:

% tau                 :   integation time in seconds
% fs                  :   LOFAR subband sampling frequency
%                         in Hz at Nyquist rate
% Ts                  :   Sampling time in seconds
% Ns                  :   Number of samples =  length of Ts
% SNR                 :   integrated SNR per station
%                          (from the correlations) in linear scale
% freq                :   [35, 42, 54, 59, 65, 69, 75] MHz
% [l,m]               :   lm-plane coordiantes range: -1 to 1
% [l_ref,m_ref]       :   Reference-central beam steering coordinates
%                         abs maximum (l,m) or (m,l) ~ abs[0.95,0.3]
% pos                 :   CS302 LOFAR-LBA
%                         inner/outer station coordinates
% g_est               :   estimated gains with identical EEP assumption
% g_est_eep           :   estimated gains with EEP variations
% w_sut               :   Beamforming Matrix
% s_ref               :   Unpolarised source signal
%                         (s_theta- theta component, s_phi- phi component)
% a_ref               :   Array response vector
% eeps                :   EEPs for antennas towards each beam on
%                         lm_plane
% eep_ref             :   EEP for antennas towards [l_ref,m_ref]
%                         has both [theta,phi] components
```

```
% %  g_est cannot be estimated if (l_ref^2+m_ref^2) > 1
%    for (l_ref^2+m_ref^2) > 1, theta- imaginary


c            = 3e8;
tau          = 1e-3;
fs           = 195312.5;
Ts           = (0:1/fs:tau);
Ns           = length(Ts);
SNR          = 1;

freq         = 35e6;
l            = -1:0.05:1;
m            = -1:0.05:1;

[lgrid, mgrid] = meshgrid(l, m);
%%  Calculate EEPs, load antenna positions

load('antpos_CS302_LBA_INNER.mat')
% load('antpos_CS302_LBA_OUTER.mat')

N_ant        = length(pos);
x_stat       = pos(:,1);
y_stat       = pos(:,2);
phase_res    = false;
Pol          = 1;

theta        = 0.01:5:90;
phi          = 0:5:360;
eeps_th_phi = zeros([length(theta), length(phi),N_ant,2]);

%   use El = El +48 for LBA outer
for El       = 1:N_ant
    Erecon = reconstruct(El,freq./1e6,theta,phi,Pol,phase_res);
    theta_comp       = Erecon(:,:,1);
    phi_comp         = Erecon(:,:,2);
    eeps_th_phi(:,:,El,1) = squeeze(theta_comp);
    eeps_th_phi(:,:,El,2) = squeeze(phi_comp);
end

%% lm-plane coordinates

l_ref        = 0;
m_ref        = 0;

l_index      = find(abs(l-l_ref)<1e-6);
```

```
m_index      = find(abs(m-m_ref)<1e-6);
ind          = sub2ind(size(lgrid),l_index,m_index);

s_theta      = randn(1, Ns) + 1i * randn(1, Ns);
s_phi        = randn(1, Ns) + 1i * randn(1, Ns);

s_ref        = [s_theta ; s_phi];

tau_ref      = (x_stat * l_ref + y_stat * m_ref) / c;
a_ref        = exp(-2i * pi * tau_ref * freq);
w_ref        = a_ref;

tau_sut      = (x_stat * lgrid(:).' + y_stat * mgrid(:).') / c;
a_sut        = exp(-2i * pi * tau_sut * freq);
w_sut        = a_sut;

gains        = ones(N_ant, 1)
             + 0.3 * (randn(N_ant, 1) + 1i * randn(N_ant, 1));

% obtain eep_ref by interpolation

theta_ref    = max(theta(1),(asind(sqrt((l_ref).^2 + (m_ref).^2))));
if (ceil(l_ref) == ceil(m_ref))
    phi_ref      =  atan2d(m_ref, l_ref) ;
elseif (ceil(l_ref) > ceil(m_ref))
    phi_ref      =  360 - atan2d(abs(m_ref),l_ref) ;
elseif (ceil(m_ref) > ceil(l_ref))
    phi_ref      =  180 - atan2d(m_ref,abs(l_ref)) ;
else
    phi_ref      =  180 + atan2d(abs(m_ref),abs(l_ref)) ;
end

eep_ref      = zeros(N_ant, 2);

for El = 1:N_ant
    eep_ref(El, 1) = interp2(phi, theta, eeps_th_phi(:, :, El, 1),
             phi_ref, theta_ref);
    eep_ref(El, 2) = interp2(phi, theta, eeps_th_phi(:, :, El, 2),
             phi_ref, theta_ref);
end

%% With identical EEP

x_ref        = diag(gains) * [a_ref, a_ref] * s_ref;

noise        =(randn(N_ant, Ns) + 1i * randn(N_ant, Ns)) /
```

```
              (sqrt(SNR/Ns)/N_ant);
x_ref         = x_ref + noise;
v_ref         = w_ref' * x_ref;

v_out_sut     = w_sut' * x_ref;
r             = v_out_sut * v_ref';



%% With EEP variations

x_ref         = diag(gains) * (eep_ref .* [a_ref, a_ref]) * s_ref
              + noise;
v_ref         = w_ref' * x_ref;

v_out_sut     = w_sut' * x_ref;
r_eep         = v_out_sut * v_ref';

%% Perform holography: Estimate gains

r_norm        = r / r(ind); % normalize measurements on central beam
r_norm_eep    = r_eep / r_eep(ind);

g_est         = (w_sut * w_sut') \ w_sut * r_norm;
g_est_eep     = (w_sut * w_sut') \ w_sut * r_norm_eep;

%%  Compare gains

% correct for source that is not in bore sight
g_est = g_est ./ a_ref;
g_est_eep = g_est_eep ./ a_ref;


g_est = g_est ./ (g_est(1) / abs(g_est(1)));
g_est_eep = g_est_eep ./ (g_est_eep(1) / abs(g_est_eep(1)));
gains = gains ./ (gains(1) / abs(gains(1)));

%% normalise on average gain to facilitate comparison
g_est = g_est / mean(abs(g_est));
g_est_eep = g_est_eep / mean(abs(g_est_eep));
gains = gains / mean(abs(gains));

%%
figure;
p1 = plot(1:N_ant, abs(g_est), 'r.-', ...
     1:N_ant, abs(g_est_eep), 'b.-', ...
```

```
     1:N_ant, abs(gains), 'k--');
p1(1).MarkerSize = 20;
p1(2).MarkerSize = 20;
p1(3).MarkerSize = 20;
p1(1).LineWidth  = 1;
p1(2).LineWidth  = 1;
p1(3).LineWidth  = 2;
legend('Gains with identical EEPs',
        'With EEp variations', 'true gains')
xlabel('Antenna Number');
ylabel('Estimated gain Amplitude');
title(['Gain amplitudes towards [', num2str(l_ref) ','
        num2str(m_ref) '] beam on lm-plane']);

figure;
p2 = plot(1:N_ant, angle(g_est), 'r.-', ...
     1:N_ant, angle(g_est_eep), 'b.-', ...
     1:N_ant, angle(gains), 'k--');
p2(1).MarkerSize = 20;
p2(2).MarkerSize = 20;
p2(3).MarkerSize = 20;
p2(1).LineWidth  = 1;
p2(2).LineWidth  = 1;
p2(3).LineWidth  = 2;
legend('Gains with identical EEPs',
        'With EEp variations', 'true gains')
xlabel('Antenna Number');
ylabel('Estimated gain phase ');
title(['Gain phases towards [', num2str(l_ref) ','
        num2str(m_ref) '] beam on lm-plane']);
```

Bibliography

[1] Astronomy log. Milkyway in 21cm. URL `https://astronomylog.wordpress.com/radio-astronomy/milkyway-in-21cm/`.

[2] A.J. van der Veen and S. J. Wijnholds. *Handbook of Signal Processing Systems*. Springer-Verlag, New York, NY, USA, 2013.

[3] G. Ellis. Cyclotron Radiation From Jupiter. *Australian Journal of Physics*, 15(3):344, 1962. doi: 10.1071/ph620344.

[4] R. Wilson. History of the Discovery of the Cosmic Microwave Background Radiation. *Physica Scripta*, 21(5):599–605, 1980. doi: 10.1088/0031-8949/21/5/001.

[5] D. S. Heeschen and N. H. Dieter. Extragalactic 21-CM Line Studies. *Proceedings of the IRE*, 46(1):234–239, 1958.

[6] N. Kassim, R. Perley, W. Erickson and K. Dwarakanath. Subarcminute resolution imaging of radio sources at 74 MHz with the Very Large Array. *The Astronomical Journal*, 106:2218, 1993. doi: 10.1086/116795.

[7] P. Wilkinson. The Hydrogen Array. *International Astronomical Union Colloquium*, 131:428–432, 1991. doi: 10.1017/s0252921100013774.

[8] P. Dewdney, P. Hall, R. Schilizzi and T. Lazio. The Square Kilometre Array. *Proceedings of the IEEE*, 97(8):1482–1496. doi: 10.1109/jproc.2009.2021005.

[9] J. Dowell, G. Taylor, F. Schinzel, N. Kassim and K. Stovall. The LWA1 Low Frequency Sky Survey. *Monthly Notices of the Royal Astronomical Society*, 469(4):4537–4550, 2017. doi: 10.1093/mnras/stx1136.

[10] C. Trott, S. Tingay and R. Wayth. Prospects for detection of Fast Radio Bursts with the Murchison Widefield Array. *The Astrophysical Journal*, 776(1):L16, 2013. doi: 10.1088/2041-8205/776/1/l16.

[11] M.P. van Haarlem and et al. LOFAR: The LOw-Frequency ARray. *Astronomical instrumentation*, 556:1–53, july 2013. doi: 10.1051/0004-6361/201220873. URL `https://doi.org/10.1051/0004-6361/201220873`.

[12] NCRA website. Introducing GMRT. URL `http://www.ncra.tifr.res.in/ncra/gmrt/about-gmrt/introducing-gmrt-1/introducing-gmrt`.

[13] Lofar.org. About LOFAR | LOFAR, 2019. URL `http://www.lofar.org/about-lofar/about-lofar`.

[14] S. J. Wijnholds. Overview of the Initial Production Version of the Station Calibration Pipeline. LOFAR-LOFAR-MEM-264, 2011.

[15] G. Styan. Hadamard products and multivariate statistical analysis. *Linear Algebra and its Applications*, 6:217–240, 1973. doi: 10.1016/0024-3795(73)90023-2.

[16] S. Wijnholds and A. van der Veen. Multisource Self-Calibration for Sensor Arrays. *IEEE Transactions on Signal Processing*, 57(9):3512–3522, 2009. doi: 10.1109/tsp.2009.2022894.

[17] K. Maeda, H. Alvarez, J. Aparici, J. May, and P. Reich. A 45-MHz continuum survey of the northern hemisphere. *Astronomy Astrophysics Supplement Series*, 140(2):145–154, Dec. 1999.

[18] S. J. Wijnholds, M. Arts, P. Bolli, P. Di Ninni and G. Virone. Using Embedded Element Patterns to Improve Aperture Array Calibration. *International Conference on Electromagnetics in Advanced Applications (ICEAA)*, pages 1–6, 2019. doi: 10.1109/ICEAA.2019.8878949.

[19] I. Gupta and A. Ksienski. Effect of mutual coupling on the performance of adaptive arrays. *IEEE Transactions on Antennas and Propagation*, 31(5):785–791, 1983. doi: 10.1109/tap.1983.1143128.

[20] S. Chiarucci and S. J. Wijnholds. Blind calibration of radio interferometric arrays using sparsity constraints and its implications for self-calibration. *Monthly Notices of the Royal Astronomical Society*, 474(1):1028–1040, Feb. 2018.

[21] R. A. Thompson, J. M. Moran and G. W. Swenson Jr. *Interferometry and Synthesis in Radio Astronomy*. Springer, USA, 2017.

[22] P. E. Dewdney, P. J. Hall, R. T. Schilizzi and T. J. L. W. Lazio. The Square Kilometre Array. *IEEE*, 97(8):1482–1496, Aug. 2009. doi: 10.1109/JPROC.2009.2021005.

[23] SKA.org. SKA. URL `https://www.skatelescope.org/`.

[24] Astron.nl. History of WSRT | Astron. URL `https://www.astron.nl/telescopes/history-wsrt`.

[25] National Radio Astronomy Observatory. Very Large Array - National Radio Astronomy Observatory. URL `https://public.nrao.edu/telescopes/vla/`.

[26] Bennett, J., Anderson, A., McInnes, P. and Whitaker, A. Microwave holographic metrology of large reflector antennas. *IEEE Transactions on Antennas and Propagation*, 24(3):295–303, 1976.

[27] P. Prasad, S. J. Wijnholds, F. Huizinga, and R. A. M. J. Wijers. Real-time calibration of the AARTFAAC array for transient detection. *Astronomy Astrophysics*, 568(A48):1–18, Aug. 2014.

[28] R. MacPhie and Ke-Li Wu. A plane wave expansion of spherical wave functions for modal analysis of guided wave structures and scatterers. *IEEE Transactions on Antennas and Propagation*, 51:2801–2805, 2003.

[29] NRAO.edu. Calibrating the flux density scale. URL `https://science.nrao.edu/facilities/vla/docs/manuals/oss/performance/fdscale`.