

Beam Shape Calibration for Multi-Beam Radio Astronomical Phased Arrays

Stefan J. Wijnholds

R&D Department, Netherlands Institute for Radio Astronomy

Dwingeloo, The Netherlands

wijnholds@astron.nl

Abstract—The ability of phased array systems to form multiple beams simultaneously holds great promise to revolutionise radio astronomy by enlarging the instantaneous sky coverage. Currently, measurements obtained by different beams are largely treated as separate observations. In this paper, a new approach to beam shape calibration, i.e., the calibration of direction dependent instrumental gains, is presented, in which the fact that these observations are done simultaneously is adequately exploited. It follows from imposing a novel spatial constraint on the direction dependent instrumental gain model. It may provide better calibration performance and may even allow beam shape calibration in scenarios in which the observations with the individual beams provide insufficient information to do so when they are treated as separate observations.

I. INTRODUCTION

The ability of phased array systems to form multiple beams simultaneously holds the potential to revolutionise radio astronomy by enlarging the instantaneous sky coverage [1], [2]. The Low Frequency Array (LOFAR, [3]) has a digital beamformer that can trade signal bandwidth per beam and number of beams. The low-frequency observing system of the Square Kilometre Array (SKA, [4]) will provide similar beamforming capability. Phased-array-fed reflector systems, like the Aperture Tile in Focus (APERTIF) system in the Netherlands [5], [6] and the Australian SKA Pathfinder (ASKAP) system in Australia [7], which are currently being commissioned, will even provide multiple beams with full signal bandwidth. This is also expected for the future Mid-Frequency Aperture Array (MFAA) system of the SKA [8], [9].

Calibration of direction dependent (instrumental) gain variations is a key prerequisite for high-quality imaging with these instruments [10], [11]. Current schemes to calibrate the beam shape, i.e., the direction dependent gain of the instrument, treat each beam independently [12], [13]. This is quite remarkable as beam errors in beams formed simultaneously by the same system are likely to be related. For example, a mechanical pointing error of the reflector will affect all beams formed by a phased array feed system. In this paper, a new direction dependent instrumental gain model is therefore introduced that assumes that the direction dependent response of the instrument over the compound field-of-view provided by all beams together can be decomposed into a local function that is the same for all beams, and a global function spanning the compound field-of-view. In the example of the mechanical pointing error, the local function would describe the beam

shape of the individual beams with a parameter to fit the common offset while the global function would describe the slow gain variation experienced by a beam that is moved across the compound field-of-view caused by the varying illumination of the reflector for each beam pointing.

This example also illustrates that the details of the local and global function and their parameterisations are strongly instrument dependent. After introducing the decomposition of the direction dependent instrumental gain into a local and a global function, this paper will therefore specialise to a simplified example that is representative for an aperture array with multi-beaming capability consisting of identical elements. The data model and resulting problem statement are described in Sec. II. In Sec. III, an alternating direction implicit (ADI) method is presented to estimate the free parameters if the local and global functions are each a linear superposition of a suitably chosen set of basis functions. Using this estimation method, the beam shape reconstruction performance of the proposed multi-beam approach is compared to that of the commonly used approach in which each beam is treated independently using simulated data. The simulations presented in Sec. IV show that the reduced number of degrees of freedom in the proposed multi-beam approach leads to better beam shape estimation performance and may even allow direction dependent gain calibration in scenarios where there is insufficient information available to calibrate the beams individually. The paper is concluded by summarising the main results.

II. DATA MODEL AND PROBLEM STATEMENT

The instruments mentioned in the introduction all consist of an array of receiving subsystems with multi-beaming capability. In the case of aperture arrays, these receiving subsystems are subarrays or *stations* consisting of a number of antennas. In the case of phased array feed systems, these receiving subsystems are phased-array-fed reflector dishes. If the narrowband condition holds, the signal with wavelength λ received by the n th beam of the p th station or dish located at ξ_p can be described as

$$x_{n,p}(t) = \sum_{q=1}^Q g_{n,p,q}^1 g_{p,q}^g s_q(t) e^{2\pi j l_q \cdot \xi_p / \lambda} + n_{n,p}(t), \quad (1)$$

where $g_{n,p,q}^1$ represents the local gain for the n th beam in the direction of the q th source and $g_{p,q}^g$ represents the global gain

towards the q th source. The source signal of the q th source is described by $s_q(t)$ and is coming from direction \mathbf{l}_q , which is a unit vector of direction cosines. Finally, $n_{n,p}(t)$ represents the additive noise.

In multi-beaming systems, the n th beam of each station or dish is pointed towards the same position on the sky. The signals produced by the corresponding beams from all stations or dishes are correlated with each other leading to covariance estimates for the n th beam,

$$\begin{aligned} R_{n,p_1 p_2} &= \mathcal{E} \{x_{n,p_1} \overline{x_{n,p_2}}\} \\ &= \sum_{q=1}^Q g_{n,p_1,q}^c g_{n,p_2,q}^c \sigma_q e^{2\pi j \mathbf{l}_q (\boldsymbol{\xi}_{p_1} - \boldsymbol{\xi}_{p_2}) / \lambda} \\ &\quad + \mathcal{E} \{n_{n,p_1} \overline{n_{n,p_2}}\}, \end{aligned} \quad (2)$$

where we have introduced the combined direction dependent gain $g_{n,p,q}^c = g_{n,p,q}^l g_{p,q}^g$ and σ_q is the power of the q th source. Using these covariance matrix estimates and a priori knowledge of the power and positions of the brightest sources in the sky (often referred to as calibration sources), the direction dependent gains $g_{n,p,q}^c$ can be estimated [10], [11], [14]. As $g_{n,p,q}^c$ is estimated independently for many frequency channels and relatively short time intervals of order 10 s, this parameter estimation problem has many degrees of freedom. Several methods have therefore been proposed to exploit the continuity of the estimated parameters over time, frequency and space [15]–[17]. The gain model proposed in this paper provides an excellent means to reduce the number of degrees of freedom in the spatial dimension and can be easily integrated in the fusion step of distributed algorithms such as proposed in [16], [17]. Depending on the implementation of the fusion center, the spatial constraint proposed here can be imposed as a hard or a soft constraint on the direction dependent gain solutions. In the latter case, the fusion center can also easily be extended to a Bayesian framework. These implementation details are outside the scope of this paper.

In the following description of the proposed approach, it will be assumed that estimates $\widehat{g}_{n,p,q}^c$ are already obtained using one of the available direction dependent calibration methods (see, e.g., [10], [11], [14]). That process effectively splits the problem into direction dependent gain estimates per station or dish. The following discussion therefore focuses on solving this station or dish based subproblem.

The local function describes a parameterised direction dependent gain that has the same shape for each beam, but with a different beam center, i.e., $g_{n,p,q}^l(\mathbf{l}_q) = f_p^l(\mathbf{l}_q - \mathbf{l}_{0,n}; \boldsymbol{\alpha}_p)$, where \mathbf{l}_q is the position of the q th source, $\mathbf{l}_{0,n}$ is the pointing center of the n th beam and $\boldsymbol{\alpha}_p$ is a parameter vector. The global function spans all beams and can thus be described w.r.t. the global coordinate center, i.e., $g_{p,q}^g(\mathbf{l}_q) = f_p^g(\mathbf{l}_q; \boldsymbol{\beta}_p)$, where $\boldsymbol{\beta}_p$ is a parameter vector. Thus, the direction dependent gain model is described by

$$g_{n,p,q}^c(\mathbf{l}_q; \boldsymbol{\alpha}_p, \boldsymbol{\beta}_p) = f_p^l(\mathbf{l}_q - \mathbf{l}_{0,n}; \boldsymbol{\alpha}_p) f_p^g(\mathbf{l}_q; \boldsymbol{\beta}_p). \quad (3)$$

Each of the Q calibration sources available in any of the beams provides an estimate $\widehat{g}_{n,p,q}^c$. All these estimates can be

stacked in a Q -element vector $\widehat{\mathbf{g}}_p^c$. For each of these positions, the local function can be evaluated given a specific value of parameter vector $\boldsymbol{\alpha}_p$. The results of these function evaluations can be stacked in a vector $\mathbf{f}_p^l(\boldsymbol{\alpha}_p)$. For the global function, the corresponding vector $\mathbf{f}_p^g(\boldsymbol{\beta}_p)$ can be constructed. With these conventions, the calibration problem can be formulated as the following least squares estimation problem:

$$\{\widehat{\boldsymbol{\alpha}}_p, \widehat{\boldsymbol{\beta}}_p\} = \underset{\boldsymbol{\alpha}_p, \boldsymbol{\beta}_p}{\operatorname{argmin}} \left\| \widehat{\mathbf{g}}_p^c - \mathbf{f}_p^l(\boldsymbol{\alpha}_p) \odot \mathbf{f}_p^g(\boldsymbol{\beta}_p) \right\|^2. \quad (4)$$

III. ESTIMATION OF FREE PARAMETERS

The problem formulated in (4) needs to be solved for each receiving subsystem in the array. Equation (4) formulates the problem in a generic way that does not make a priori assumptions on the vectors \mathbf{f}_p^l and \mathbf{f}_p^g and their parameterisation.

Assuming that the parameter vector $[\boldsymbol{\alpha}_p^T, \boldsymbol{\beta}_p^T]^T$ is identifiable, a general solver, like the Levenberg-Macquardt algorithm or the BFGS method, can be used. In many practical scenarios, the free parameters will be coefficients of suitably chosen basis functions, i.e., the vectors \mathbf{f}_p^l and \mathbf{f}_p^g are described by

$$\mathbf{f}_p^l = \boldsymbol{\Psi}^l \boldsymbol{\alpha}_p \quad (5)$$

$$\mathbf{f}_p^g = \boldsymbol{\Psi}^g \boldsymbol{\beta}_p, \quad (6)$$

where the matrices $\boldsymbol{\Psi}^l$ and $\boldsymbol{\Psi}^g$ are constructed by evaluating the basis functions selected for the local and global function respectively at the positions of the calibration sources.

This representation for the vectors \mathbf{f}_p^l and \mathbf{f}_p^g leads to a quadratic problem that can be solved iteratively using an alternating direction implicit (ADI) scheme:

- 1) *Initialisation*: Set the iteration counter $i = 1$ and set $\boldsymbol{\beta}_p^{[0]}$ to some suitable initial value and calculate $\mathbf{f}_p^g(\boldsymbol{\beta}_p^{[0]})$.
- 2) *Update $\widehat{\boldsymbol{\alpha}}_p$* by solving

$$\boldsymbol{\alpha}_p^{[i]} = \underset{\boldsymbol{\alpha}_p}{\operatorname{argmin}} \left\| \widehat{\mathbf{g}}_p^c - \operatorname{diag}(\mathbf{f}_p^g(\boldsymbol{\beta}_p^{[i-1]})) \boldsymbol{\Psi}^l \boldsymbol{\alpha}_p \right\|^2,$$

which is a linear problem that is straightforward to solve.

- 3) *Update $\widehat{\boldsymbol{\beta}}_p$* by solving

$$\boldsymbol{\beta}_p^{[i]} = \underset{\boldsymbol{\beta}_p}{\operatorname{argmin}} \left\| \widehat{\mathbf{g}}_p^c - \operatorname{diag}(\mathbf{f}_p^l(\boldsymbol{\alpha}_p^{[i-1]})) \boldsymbol{\Psi}^g \boldsymbol{\beta}_p \right\|^2,$$

which is another linear problem that is straightforward to solve.

- 4) *Check for convergence or maximum number of iterations*. Repeat steps 2-4 if the solution has not converged sufficiently and the maximum number of iterations has not been reached, otherwise stop.

This approach is used in the simulations in the next section, in which polynomial basis functions are assumed. By limiting polynomial order, a smoothness constraint is effectively adopted and the number of free parameters is reduced. This will be shown to improve calibration performance.

IV. SIMULATIONS

A. Description

To illustrate the advantages of the proposed multi-beam approach over calibrating the beam shapes of all beams individually, let us apply this idea to a simple example of a single one-dimensional aperture array station ($p = 1$ always). Assuming a perfectly calibrated, dense (antenna separation $\leq \lambda/2$) uniform linear array, the array factor will be a sinc function. If all antennas have the same element beam pattern, the actual beam shape will be determined by the array factor multiplied by the element beam pattern [18]. Around its maximum, the sinc function can be well approximated by a paraboloid. A second order polynomial is usually a good approximation around the main beam maximum of the broad element beam pattern as well. For the simulations, we will therefore use second order polynomials to describe the main beam area within the half power beamwidth for both the array factor (local function) and element beam (global function), i.e.,

$$g^l(l - l_{0,n}; \alpha_0, \alpha_1, \alpha_2) = \begin{cases} \alpha_0 + \alpha_1(l - l_{0,n}) + \alpha_2(l - l_{0,n})^2 & |l - l_{0,n}| \leq \frac{\Delta l}{2} \\ 0 & |l - l_{0,n}| > \frac{\Delta l}{2} \end{cases} \quad (7)$$

and

$$g^g(l; \beta_0, \beta_1, \beta_2) = \beta_0 + \beta_1 l + \beta_2 l^2. \quad (8)$$

Note that the direction cosine l is now a scalar as we are considering a one dimensional case. Also, the subscripts n and q have been omitted as those can be understood from $l_{0,n}$ and l respectively. The half power beam width of the array factor is represented by Δl . The discontinuity in the definition of the local function is not an issue as we will only consider the main beam area within the half power beamwidth in our simulations.

The simulations consider an array forming 9 beams simultaneously with the central beam pointing towards $l = 0$. Each beam has a half power beamwidth $\Delta l = 0.1$, which is assumed to be the separation in pointing between the beams as well. Fitting a paraboloid through the main beam peak and the two half power points gives $\alpha_0 = 1$, $\alpha_1 = 0$ and $\alpha_2 = -200$. Strictly speaking, these coefficients are valid when the fit is made in the power domain, while the gains defined in the previous section are signal domain gains. For this single station example, this does not matter as the gain phases cancel when the signal from a station is correlated with itself. This simplification thus avoids taking a square root in the analysis. For the element radiation pattern, it is assumed that it has unit gain towards $l = 0$ and zero gain at $l = \pm 1$. This gives $\beta_0 = 1$, $\beta_1 = 0$ and $\beta_2 = -1$. This model is shown in Fig. 1.

B. Scenario 1: many calibration sources

Current best practice is to calibrate the beam shape for each beam independently. This results in a problem similar to (4) for each beam in which $f_p^g = 1$ and one only needs to solve for α_p . If $f_p^l(\alpha_p)$ has N degrees of freedom, at

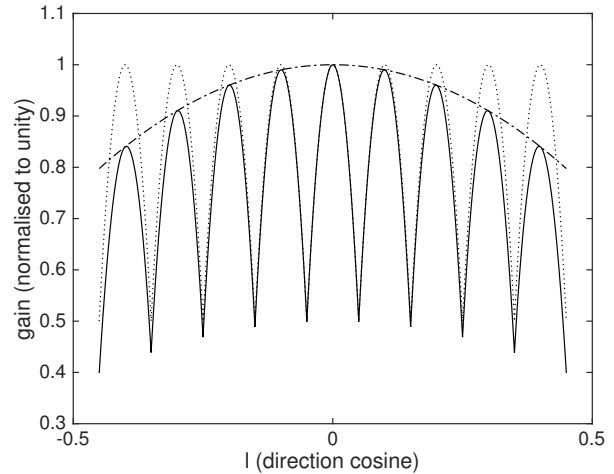


Fig. 1. Gain model for the 9-beam aperture array used in the simulations. The dotted curves represent the array factors of the 9 individual beams (described by the local function), the dash-dotted curve shows the element beam pattern (described by the global function) and the solid curve represents the combined result describing the actually realised beams.

least N calibration sources need to be available within each beam to make the problem identifiable. In this section, it will be demonstrated that the proposed multi-beam approach gives better beam shape estimates than beam shape calibration per beam even if there is a sufficiently high density of calibration sources. This will be done by comparing the standard deviation and bias across the compound field-of-view for three cases:

- 1) calibration using the proposed multi-beam approach with f_p^l and f_p^g both being second order polynomials (6 free parameters for all 9 beams);
- 2) calibration per beam using a second order polynomial (27 free parameters in total);
- 3) calibration per beam using a fourth order polynomial (45 free parameters in total).

As the combined gain of the element beam pattern and the array factor is a product of two second order polynomials, calibration using the third approach is expected to be bias free, but it requires estimation of a larger number of parameters than the first case. In the second case, the number of parameters to be estimated is reduced to three per beam, but the chosen model may not fit the actual data perfectly, resulting in a bias.

To populate the data vector $\hat{\mathbf{g}}^c$, a regular grid of calibration sources with a SNR of 10 and a separation between sources of $\delta l = 0.01$ was assumed, resulting in 10 sources per beam. This unrealistic scenario ensures that all beams are equally well sampled, such that any difference between the three approaches can be attributed to the susceptibility of each method to measurement noise and not to their susceptibility to an unfavourable distribution of sources. Figures 2 and 3 show the standard deviation and bias across all nine beams obtained in a Monte Carlo simulation of 1000 runs.

Figure 2 shows that calibration per beam with a fourth order polynomial gives a higher variability between the estimated

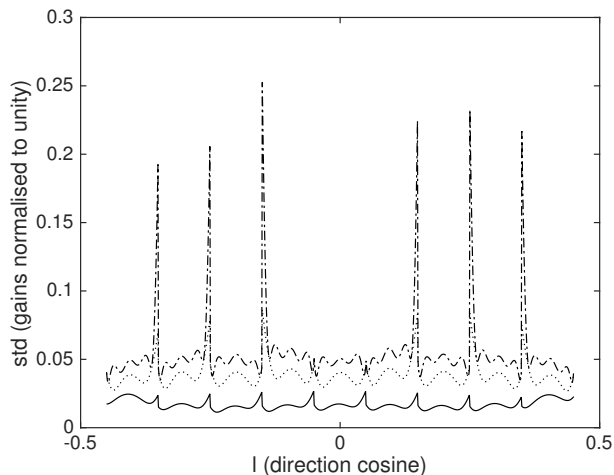


Fig. 2. Standard deviation of the estimated beam shape model across the compound field-of-view formed by nine beams obtained in a Monte Carlo simulation of 1000 runs. The results for multi-beam calibration (case 1), calibration per beam with a second order polynomial (case 2) and calibration per beam with a fourth order polynomial (case 3) are represented by the solid, dotted and dash-dotted curve respectively.

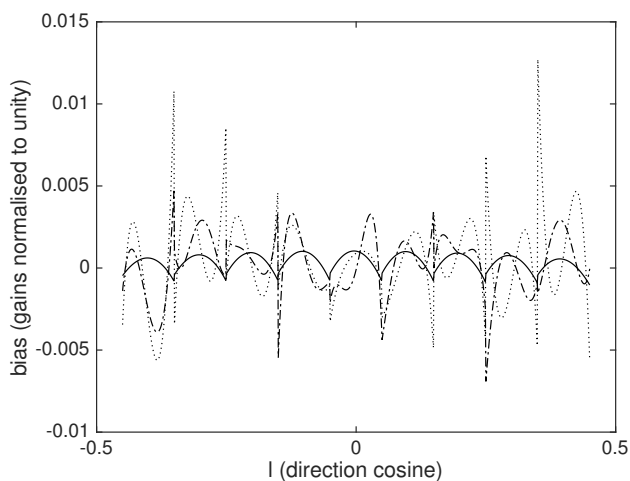


Fig. 3. Bias of the estimated beam shape model across the compound field-of-view formed by nine beams obtained in a Monte Carlo simulation of 1000 runs. The results for multi-beam calibration (case 1), calibration per beam with a second order polynomial (case 2) and calibration per beam with a fourth order polynomial (case 3) are represented by the solid, dotted and dash-dotted curve respectively.

beam shapes than calibration per beam with a second order polynomial, which in turn gives a higher variability between the estimated beam shapes than the proposed multi-beam calibration approach. The largest variations occur at the beam edges, which is typical behaviour for a polynomial fit at the edge of the fitting interval. Quantitatively, the standard deviation appears proportional to the square root of the number of degrees of freedom. This behaviour is expected for independent parameters and independent measurements. Figure 3 shows that the bias is effectively negligible being about an

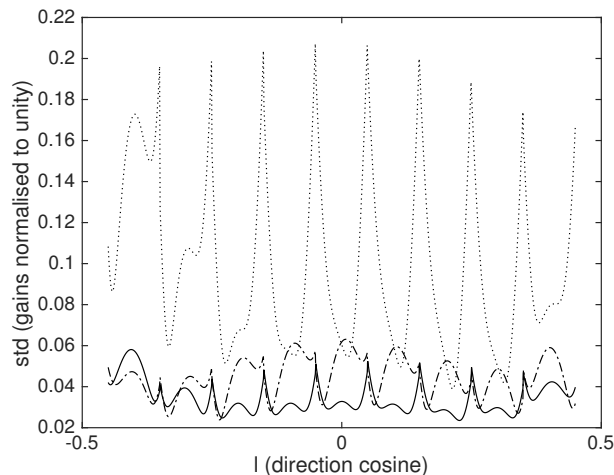


Fig. 4. Standard deviation of the estimated beam shape model across the compound field-of-view formed by nine beams obtained in a Monte Carlo simulation of 1000 runs. The results for multi-beam calibration with 9, 18 and 27 sources randomly distributed across the compound field-of-view are presented by the dotted, dash-dotted and solid curve respectively.

order of magnitude smaller than the standard deviation. As expected, the largest peaks occur for calibration per beam with a second order polynomial.

C. Scenario 2: few calibration sources

In the second, more realistic scenario, 9, 18 or 27 sources with a SNR of 10 are randomly distributed over the compound field-of-view of all beams. This implies that there are just one, two or three sources on average per beam, but due to their random placement, it is unlikely that three sources are available in each beam in the latter case. As a result, calibration per beam will fail as this poses an unidentifiable problem. Therefore, only the results for the proposed multi-beam calibration approach will be presented.

Figures 4 and 5 show the standard deviation and bias across all nine beams obtained in a Monte Carlo simulation of 1000 runs. The results for 18 and 27 sources are already very similar to those found for the first scenario with 10 sources uniformly distributed across each beam. The results obtained for 9 randomly distributed sources across the compound field-of-view are much worse. This is caused by the fact that in some runs, the spatial distribution of sources may be very unfavourable. If the number of sources increases, the more likely a reasonably homogenous spatial coverage becomes. Despite the obvious impact of the spatial distribution and SNR of the calibration sources, these results show that the proposed multi-beam calibration approach makes beam shape calibration in these scenarios viable by ensuring an identifiable calibration problem, whereas the traditional beam-by-beam calibration approach is faced with an unidentifiable problem for at least some of the beams.

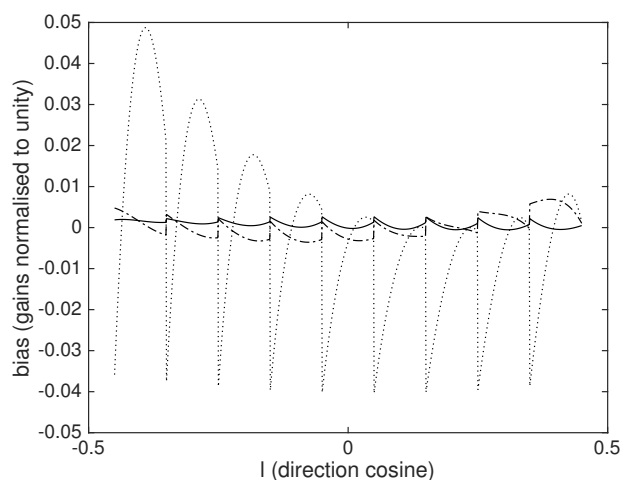


Fig. 5. Bias of the estimated beam shape model across the compound field-of-view formed by nine beams obtained in a Monte Carlo simulation of 1000 runs. The results for multi-beam calibration with 9, 18 and 27 sources randomly distributed across the compound field-of-view are presented by the dotted, dash-dotted and solid curve respectively.

V. CONCLUSIONS

The new generation of radio astronomical instrumentation based on phased array technology has the ability to increase the instantaneous sky coverage by forming multiple beams simultaneously. As those beams are formed by the same system at the same time, beam errors in those beams are likely to be strongly related. In this paper, a multi-beam calibration approach to direction dependent instrumental gain (beam shape) calibration is presented that exploits that by introducing a novel direction dependent instrumental gain model. Simulation results indicate that the proposed multi-beam calibration approach outperforms the traditional approach in which the direction dependent gains are solved for each beam individually. The proposed method provides higher beam shape estimation accuracy and works in scenarios in which the spatial density of suitable calibration sources is too low for the traditional method to work.

VI. ACKNOWLEDGMENT

This work is funded by the Netherlands Organisation for Scientific Research.

REFERENCES

- [1] A. van Ardenne, J. D. Bregman, W. A. van Cappellen, G. W. Kant, and J. G. bij de Vaate, "Extending the Field of View With Phased Array Techniques: Results of European SKA Research," *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1531–1542, Aug. 2009.
- [2] S. J. Wijnholds, W. A. van Cappellen, J. G. bij de Vaate, and A. van Ardenne, "Phased-array antenna system development for radio-astronomy applications," *IEEE Antennas and Propagation Magazine*, vol. 55, no. 6, pp. 293–308, Dec. 2013.
- [3] M. P. van Haarlem *et al.*, "LOFAR: The LOw Frequency ARray," *Astronomy & Astrophysics*, vol. 556, no. A2, pp. 1–53, Aug. 2013.
- [4] P. E. Dewdney, P. J. Hall, R. T. Schilizzi, and T. J. L. W. Lazio, "The Square Kilometre Array," *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1482–1496, Aug. 2009.
- [5] W. A. van Cappellen and L. Bakker, "APERTIF: Phased array feeds for the Westerbork Synthesis Radio Telescope," in *IEEE International Symposium on Phased Array Systems and Technology (ARRAY)*, Waltham (MA), USA, 12–15 Oct. 2010.
- [6] B. Hut, R. H. van den Brink, and W. A. van Cappellen, "Status update on the system validation of APERTIF, the phased array feed system for the Westerbork Synthesis Radio Telescope," in *European Conference on Antennas and Propagation (EUCAP)*, Paris, France, 19–24 Mar. 2017.
- [7] A. E. T. Schinkel and D. C. J. Bock, "The Australian SKA Pathfinder: project update and initial operations," in *Proceedings SPIE Astronomical Telescopes and Instrumentation*, vol. 9906, Edinburgh, UK, Aug. 2016.
- [8] W. A. van Cappellen *et al.*, "MANTIS: The Mid-Frequency Aperture Array Transient and Intensity-Mapping System," <https://arxiv.org/abs/1612.07917>, Dec. 2016, MANTIS white paper.
- [9] A. Faulkner, "Mid-Frequency AA Technology," in *3rd MIDprep / AA-mid workshop*, Cape Town, South Africa, 7–9 Mar. 2016.
- [10] S. J. Wijnholds, S. van der Tol, R. Nijboer, and A. J. van der Veen, "Calibration Challenges for the Next Generation of Radio Telescopes," *IEEE Signal Processing Magazine*, vol. 27, no. 1, pp. 30–42, Jan. 2010.
- [11] O. M. Smirnov, "Revisiting the radio interferometer measurement equation – III. Addressing direction-dependent effects in 21 cm WSRT observations of 3C 147," *Astronomy & Astrophysics*, vol. 527, no. A108, pp. 1–12, 2011.
- [12] J. Landon *et al.*, "Phased Array Feed Calibration, Beamforming and Imaging," *The Astronomical Journal*, vol. 139, no. 3, Feb. 2010.
- [13] M. V. Ivashina, O. Iupikov, R. Maaskant, W. A. van Cappellen, and T. Oosterloo, "An Optimal Beamforming Strategy for Wide-Field Surveys With Phased-Array-Fed Reflector Antennas," *IEEE Transactions on Antennas and Propagation*, vol. 59, no. 6, pp. 1864–1875, Jun. 2011.
- [14] S. Kazemi *et al.*, "Radio Interferometric Calibration Using the SAGE Algorithm," *Monthly Notices of the Royal Astronomical Society*, vol. 414, no. 2, pp. 1656–1666, Jun. 2011.
- [15] S. Yatawatta, "Radio interferometric calibration using a Riemannian manifold," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Vancouver, Canada, 26–31 May 2013.
- [16] —, "Distributed radio interferometric calibration," *Monthly Notices of the Royal Astronomical Society*, vol. 449, no. 4, pp. 4506–4514, Jun. 2015.
- [17] M. Brossard *et al.*, "Parallel multi-wavelength calibration algorithm for radio astronomical arrays," *Signal Processing*, vol. 145, pp. 258–271, Apr. 2018.
- [18] C. A. Balanis, *Antenna Theory - Analysis and Design*, 2nd ed. John Wiley & Sons, 1997.