Robust calibration of radio interferometers in multi-frequency scenario
V Ollier, M. N. El Korso, A Ferrari, Remy Boyer, Pascal Larzabal

To cite this version:
V Ollier, M. N. El Korso, A Ferrari, Remy Boyer, Pascal Larzabal. Robust calibration of radio interferometers in multi-frequency scenario. 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Apr 2018, Calgary, Canada. 10.1109/icassp.2018.8461382 . hal-01851033

HAL Id: hal-01851033
https://hal-centralesupelec.archives-ouvertes.fr/hal-01851033
Submitted on 28 Jul 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
ROBUST CALIBRATION OF RADIO INTERFEROMETERS IN MULTI-FREQUENCY SCENARIO

V. Ollier*, M. N. El Korso‡, A. Ferrari*, R. Boyer*, P. Larzabal*

* SATIE, UMR 8029, ENS Paris-Saclay, Cachan, France
‡ L2S, UMR 8506, Université Paris-Sud, Gif-sur-Yvette, France
† LEME, EA 4416, Université Paris-Nanterre, Ville d’Avray, France
△ Laboratoire J.L. Lagrange, UMR 7293, Université Nice Sophia-Antipolis, France

ABSTRACT

This paper investigates calibration of sensor arrays in the radio astronomy context. Current and future radio telescopes require computationally efficient algorithms to overcome the new technical challenges as large collecting area, wide field of view and huge data volume. Specifically, we study the calibration of radio interferometry stations with significance dependent distortions. We propose an iterative robust calibration algorithm based on a relaxed maximum likelihood estimator for a specific context: i) observations are affected by the presence of outliers and ii) parameters of interest have a specific structure depending on frequency. Variation of parameters across frequency is addressed through a distributed procedure, which is consistent with the new radio synthesis arrays where the full observing bandwidth is divided into multiple frequency channels. Numerical simulations reveal that the proposed robust distributed calibration estimator outperforms the conventional non-robust algorithm and/or the mono-frequency case.

Index Terms— Robust calibration, distributed optimization, consensus, direction dependent distortions regime.

1. INTRODUCTION

With a resolution and sensitivity greater than any previous systems in the low observing frequencies, the low frequency array (LOFAR) and the square kilometre array (SKA) impose major challenges in terms of telescope design and data processing [1–3]. Among these challenges, the calibration step is crucial for the advanced phased array radio telescopes due to the large number of receivers and their wide field-of-view. These design features result in direction dependent effects [4], varying over the field-of-view, due, e.g., to ionospheric disturbances, and in gain differences from one receiving element to another. To avoid any calibration error preventing the exploitation of the full sensitivity potential and decreasing the high dynamic range performance of imaging [5], environmental and instrumental unknowns need to be corrected for.

In several studies [6, 7], array signal processing tools are used to perform calibration, especially the maximum likelihood (ML) estimator which is usually introduced under a Gaussian noise model assumption [8]. However, in our application, such noise modeling is not adapted: the sky model is composed of several bright known sources (the calibrator sources) but also many unknown unmodeled sources which give rise to incomplete sky models [9]. Furthermore, radio astronomical measurements are contaminated by radio interference [10]. All these effects, and others, lead to the presence of outliers in the data set. To take them into account, a robust calibration technique was introduced in [9] where the noise model is specifically described as a Student’s t with independent identically distributed entries. To improve calibration, we proposed, in a previous work [11], to adopt a compound-Gaussian modeling [12] which includes a broad range of different distributions and revealed to be more robust [13].

Furthermore, it should be noted that direction independent gains of each sensor and direction dependent perturbations associated to each source are estimated through the calibration process, where the latter parameters are frequency dependent. To exploit the known structure of variation w.r.t. frequency [14], calibration is reformulated as a constrained consensus problem and addressed with the alternating direction method of multipliers (ADMM) [15, 16]. To reduce the computational burden, we process the data thanks to a distributed architecture with a network of agents [17]. Decentralized and distributed strategies have already been applied for image reconstruction in radio astronomy [18] and also for calibration [19] in the Gaussian and non-structured case. In this work, we particularly focus on robust calibration of radio interferometry compact stations with direction dependent distortions, named as the 3DC calibration regime (structured case) [13].

The notation used through this paper is the following: symbols $(\cdot)\top$, $(\cdot)\ast$, $(\cdot)\dagger$ denote, respectively, the transpose, the complex conjugate and the Hermitian transpose. The symbol $\otimes$ represents the Kronecker product, $\text{vec}(\cdot)$ stacks the columns of a matrix on top of one another, $\text{diag}(\cdot)$ converts a vector into a diagonal matrix and the trace is given by $\text{tr}\{\cdot\}$. The $B \times B$ identity matrix is referred by $I_B$. $\|\cdot\|_F$ is the Frobenius norm, while $\|\cdot\|_2$ denotes the $L_2$ norm. Finally, $j$ is the complex number whose square equals $-1$ and $[\cdot]_k$ refers to the $k$-th entry of the considered vector.

2. MODEL SET UP AND BACKGROUND ON ROBUST CALIBRATION ALGORITHM

2.1. Data model

We consider $D$ signal waves emitted by calibrator sources impinging on an array of $M$ antennas. The cross correlation between (noisy free) voltages measured by two antennas $p$ and $q$, is given by [20]

$$V_{pq}(\theta) = \sum_{i=1}^{D} J_{i,p}(\theta) C_i J_{i,q}(\theta) \text{ for } p < q, \quad (p, q) \in \{1, \ldots, M\}^2,$$  

(1)
where \( C_i \) is a known matrix describing the polarization state of the \( i \)-th calibrator source while \( J_{i,p}(\theta) \) stands for all the perturbations along the full corresponding signal path (from the \( i \)-th source to the \( p \)-th antenna) and is referred as a \( 2 \times 2 \) Jones matrix [21]. The aim of calibration is to estimate the parameter vector \( \theta \). We rewrite (1) as

\[
\tilde{v}_{pq}(\theta) = \text{vec} \left( V_{pq}(\theta) \right) = \sum_{i=1}^{D} u_{i,pq}(\theta)
\]

in which \( u_{i,pq}(\theta) = \left(J_{i,p}(\theta) \otimes J_{i,p}(\theta)\right) c_i \) and \( \text{vec}(C_i) \). The noisy output observation vector of the full array is \( x = \left[v_1^T, v_2^T, \ldots, v_{M-1,M}^T\right]^T \) t.i. \( v_{pq} = \tilde{v}_{pq}(\theta) + n_{pq} \). We note \( n_{pq} \) the noise sample at a particular antenna pair which takes into account background Gaussian noise but also the presence of outliers.

To deal with non-Gaussianity of the noise, we assume a compound-Gaussian noise model since it includes a broad range of heavy-tailed distributions [12, 22]. Its expression is given by

\[
\text{CN}(0, \Omega)
\]

where \( \tau_{pq} \) is the positive texture variable and \( \mu_{pq} \sim \text{CN}(0, \Omega) \) is the complex speckle part.

### 2.2. Robust estimation of Jones matrices

The ML method is used to estimate iteratively parameters of interest \( \theta \), the covariance matrix \( \Omega \) and all realizations \( \tau_{pq} \) for \( p < q \). \( (p, q) \in \{1, \ldots, M\}^2 \) are considered unknown in the algorithm, leading to a relaxed version of the ML [23]. Results are directly exposed in non-structured calibration algorithm (NSCA) and details can be found in [11, 13]. Let us denote

\[
a_{pq}(\theta) = v_{pq} - \tilde{v}_{pq}(\theta), \quad B = \frac{M(M-1)}{2}
\]

Each individual Jones term and its perturbation \( \Omega_t = \Omega_t + \Delta_t \) are considered unknown and determined in the algorithm, leading to a relaxed version of the ML [35]. Let us note that calibration amounts to estimate the parameter vector \( \theta \) along the full corresponding signal path (from the \( p \)-th antenna to the \( i \)-th calibrator source) and \( \Omega_t = \Omega_t + \Delta_t \) is the known position vector of the \( p \)-th antenna in wavelength units. Therefore, we deduce that \( \hat{V}_{i,p}^T \sim \mathcal{N}(0, \Omega) \).

### 3.3. Multi-frequency calibration algorithm

The aim of calibration in a multi-frequency scenario is to estimate the vector parameter of interest \( \epsilon = [\epsilon_i]_{i=1}^D \) t.i. \( \epsilon(f) \) and \( \epsilon(f) \) where \( \epsilon_i = [\epsilon_i(f_1), \ldots, \epsilon_i(f_F)]^T \) and \( \epsilon_i = [\epsilon_i(f_1), \ldots, \epsilon_i(f_F)]^T \) are frequency-varying. To do so, we introduce the following cost function

\[
i_i^T(\epsilon_i) = \sum_{f=1}^{F} i_i^T(\epsilon_i)
\]
where \( B \) is the known frequency model, \( z_i \) is an unknown associated global variable, independent \( w.r.t. \) frequency and shared by all agents, and \( z = [z_1, \ldots, z_D] \). To solve (4), we use a consensus optimization scheme as in the ADMM procedure [16]. Instead of considering the original objective function (3), we study the following augmented Lagrangian

\[
\begin{align*}
\min_{\hat{e}^{[f]} \in F, \hat{z}} & \sum_{f \in F} \| \hat{e}^{[f]}(e^{[f]}(\hat{e}^{[f]}(z_i, y_i \hat{y}^{[f]})) \|_2^2 \\
\text{s.t.} & \quad e^{[f]}(z_i, y_i \hat{y}^{[f]}) = B^{[f]} z_i, \quad i \in \{1, \ldots, D\}, \quad f \in F
\end{align*}
\]

where \( B^{[f]} = \frac{1}{B} I_3 \) is the known frequency model, \( z_i \) is an unknown associated global variable, independent \( w.r.t. \) frequency and shared by all agents, and \( z = [z_1, \ldots, z_D] \). To solve (4), we use a consensus optimization scheme as in the ADMM procedure [16]. Instead of considering the original objective function (3), we study the following augmented Lagrangian

\[
L(e^{[f]}_1, \ldots, e^{[f]}_D, z, y^{[f]}_1, \ldots, y^{[f]}_D) = \sum_{f \in F} \sum_{i=1}^D L_f(e^{[f]}_i, z_i, y^{[f]}_i) = \sum_{f \in F} \sum_{i=1}^D L_f(e^{[f]}_i, z_i, y^{[f]}_i)
\]

where \( L_f(e^{[f]}_i, z_i, y^{[f]}_i) = \| e^{[f]}_i(e^{[f]}(z_i, y^{[f]})) - B^{[f]} z_i \|_2^2 + \| e^{[f]}(z_i, y^{[f]} - B^{[f]} z_i) \|_2^2 \). We note that \( y^{[f]}_i = [y_i^{[f]}(1), \ldots, y_i^{[f]}(D)] \) is the associated Lagrange parameter (or dual variables) and \( \rho > 0 \) a penalty factor. We notice separability of the Lagrangian \( w.r.t. \) source direction but above all, separability \( w.r.t. \) frequency, meaning that each agent solves a subproblem locally at a given frequency. The ADMM consists in updating sequentially the three following quantities:

\[
\bigg( e^{[f]}_i(t^{i+1}) \bigg) = \arg \min_{e^{[f]}_i} L_f(e^{[f]}_i, (\hat{z}_i)^t, (y^{[f]}_i)^t) \quad \text{performed locally by each agent for } i \in \{1, \ldots, D\}
\]

\[
\bigg( y^{[f]}_i(t^{i+1}) \bigg) = \bigg( y^{[f]}_i(t^i) + \rho \bigg( e^{[f]}_i(t^{i+1}) - B^{[f]} (\hat{z}_i)^t \bigg) \bigg) \quad \text{performed locally by each agent for } i \in \{1, \ldots, D\}
\]

Minimization (5) is addressed iteratively. To this end, we compute the gradient of \( L_f(e^{[f]}_i, z_i, y^{[f]}_i) \) \( w.r.t. \) \( e^{[f]}_i \), which induces

\[
\frac{\partial L_f(e^{[f]}_i, z_i, y^{[f]}_i)}{\partial e^{[f]}_i} = \begin{bmatrix} 0 \ldots 0 \end{bmatrix} + \rho \begin{bmatrix} 0 \ldots 0 \end{bmatrix}
\]

likewise, we have

\[
\frac{\partial L_f(e^{[f]}_i, z_i, y^{[f]}_i)}{\partial \hat{z}_i} = \begin{bmatrix} 0 \ldots 0 \end{bmatrix} + \rho \begin{bmatrix} 0 \ldots 0 \end{bmatrix}
\]

leading to the following estimate of each complex gain element for \( k \in \{1, 2\} \)

\[
\hat{g}_{i,p}[k] = \left( \sum_{f \in F} \sum_{i=1}^D \| W_{i,p}^{[f]} \|_2^2 \right)^{-1} \sum_{f \in F} \sum_{i=1}^D \| X_{i,p}^{[f]} \|_k^2
\]

In this section, we compare the performance of the proposed MSCA with the mono-frequency case, \( i.e., \) SCA in which frequency diversity is not taken into account. We recall that radio astronomy observations are affected by the presence of outliers. Thus, we also compare our robust approach with an algorithm based on a classical Gaussian noise assumption [8], which amounts to solve a non-linear least squares problem. First, in Fig. 1, we plot the mean square error (MSE) of \( y^{[f]}_i \) as a function of the signal-to-noise ratio (SNR), the
MSCA: Multi-frequency structured calibration algorithm

\[ \text{input : } D, M, F, r_p, H_p, \mathbf{J}_p, \text{ as output of NSCA for } i \in \{1, \ldots, D\}, p \in \{1, \ldots, M\} \text{ and } f \in F \]

\[ \text{output : } \hat{\epsilon} \]

\[ \text{initialize: } \hat{\epsilon} \leftarrow \epsilon_{\text{init}}, \hat{z} \leftarrow z_{\text{init}}, \{ \hat{y}_{f}^{[i]} \} \leftarrow y_{\text{init}}^{[f]} f \in F \]

\[ \text{while stop criterion unreached do} \]

\[ \text{while stop criterion unreached do} \]

\[ \text{Obtain } \{ \hat{g}_i^{[f]} \}_{i=1}^{D} \text{ locally with (9)} \]

\[ \text{Obtain } \{ \hat{h}_i^{[f]} \}_{i=1}^{D} \text{ locally with (10)} \]

\[ \text{Obtain } \{ \hat{c}_i^{[f]} \}_{i=1}^{D} \text{ locally with (11)} \]

\[ \text{Obtain } \{ \hat{z}_i \}_{i=1}^{D} \text{ globally with (8)} \]

\[ \text{Obtain } \{ \hat{y}_i^{[f]} \}_{i=1}^{D} \text{ locally with (7)} \]

\[ \text{Obtain } \{ \hat{g}_p \}_{p=1}^{M} \text{ with (12)} \]

\[ \text{end} \]

\[ \text{end} \]

behavior being the same for any other parameter of \( \epsilon \). We compare the estimation performance for different number of frequencies \( F \) and notice better statistical performance when multi-frequency robust calibration is performed. Robust calibration based on the Student’s [9] is not exposed in the simulations due to a different model which is not adapted to the 3DC regime.

In the following figures, we use Meqtrees [29] to generate the data model, the observations and compare its least squares solver to MSCA. Here, we choose to correct for Faraday rotation matrices, which are the only introduced perturbations in the observations. We consider \( M = 7 \) antennas (KAT-7 instrument), \( D = 1 \), \( D' = 16 \) weak realistic background sources taken from the SUMSS survey using a spectral index of 0.7. The full duration of the observation is 12 hours, for 60 seconds integration time per data sample. After calibration and subtraction of the bright calibrator source, a dirty image, namely the corrected residual, is constructed with Meqtrees using `lwimag`er. Fig. 2 gives the corrected residual image at 895 MHz in a small area surrounding the calibrator, whose position corresponds to the red cross. Fig. 3 gives the recovered flux for one of the \( D' \) weak outlier sources. Therefore, we notice better flux estimation of weak background sources and better calibrator removal using joint frequency dependent calibration with MSCA compared to a frequency independent calibration.

5. CONCLUSION

This paper introduces a robust ML based calibration technique in the context of radio interferometry. Robustness is addressed thanks to a compound-Gaussian noise modeling and a particular scenario is studied, i.e., the 3DC calibration regime. Variation of parameters w.r.t. frequency imposes additional constraints which are considered in an ADMM-based distributed algorithm. We show in the simulations the advantages of the proposed algorithm in regards to standard mono-frequency and/or non-robust scenario.

6. REFERENCES


