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://centralesupelec.hal.science/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

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 significant direction 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 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 the measurement vector model, calibration is formulated 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 [13].

The notation used through this paper is the following: symbols $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^{H}$ denote, respectively, the transpose, the complex conjugate and the Hermitian transpose. The symbol $\otimes$ represents the Kronecker product, vec$(\cdot)$ stacks the columns of a matrix on top of one another, 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.

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 distortions, named as the 3DC calibration regime [13].

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 propose, 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].

The notation used through this paper is the following: symbols $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^{H}$ denote, respectively, the transpose, the complex conjugate and the Hermitian transpose. The symbol $\otimes$ represents the Kronecker product, vec$(\cdot)$ stacks the columns of a matrix on top of one another, 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]

$$V_{pq}(\theta) = \sum_{i=1}^{D} J_{i,p}(\theta) C_{i,q}(\theta)$$

for $p < q$, $(p, q) \in \{1, \ldots, M\}^2$,

(1)
where $C_l$ is a known matrix describing the polarization state of the $i$-th calibrator source while $J_{1,p}(\theta)$ 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

$$\hat{v}_{pq}(\theta) = \text{vec}(V_{pq}(\theta)) = \sum_{i=1}^{D} u_{i,pq}(\theta)$$

in which $u_{i,pq}(\theta) = (J_{i,p}(\theta) \otimes J_{i,p}(\theta)) c_i$ and $c_i = \text{vec}(C_l)$. The noisy output observation vector of the full array is $x = [v_{12}^T, v_{13}^T, \ldots, v_{(M-1)M}^T]^T$. $v_{pq} = \hat{v}_{pq}(\theta) + n_{pq}$. 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

$$n_{pq} = \sqrt{\tau_{pq}} h_{pq},$$

where $\tau_{pq}$ is the positive texture variable and $\mu_{pq} \sim \mathcal{CN}(0, \Omega)$ 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 $p < q$, $(p, q) \in \{1, \ldots, M\}^2$ which are considered as non-deterministic 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 $a_{pq}(\theta) = v_{pq} - \hat{v}_{pq}(\theta)$, $B = \frac{M}{4(M-1)}$ is the number of antenna pairs and we impose $\text{tr}(\Omega) = 1$ to remove scaling ambiguities [24]. Estimation of $\theta$ in step 2 of NSCA can be performed numerically or thanks to expectation maximization (EM) [25] and block coordinate descent (BCD) algorithms [26]. Such vector refers here to the entries of all $DM$ Jones matrices (non-structured case) [8, 27].

**NSCA: Non-structured calibration algorithm**

<table>
<thead>
<tr>
<th>input: $D, M, B, {C_l}_{i=1}^{D}, x$</th>
<th>output: $\hat{\theta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>initialize: $\Omega \leftarrow \Omega_{\text{init}}, {\tau_{pq} \leftarrow \tau_{pq_{\text{null}}}}_{p &lt; q, (p, q) \in {1, \ldots, M}^2}$</td>
<td></td>
</tr>
<tr>
<td>while stop criterion unreached do</td>
<td></td>
</tr>
<tr>
<td>$\hat{\theta} = \text{argmin} \left{ \sum_{p&lt;q, (p, q) \in {1, \ldots, M}^2} \left</td>
<td>a_{pq}(\theta) \right</td>
</tr>
<tr>
<td>$\hat{\Omega} = \frac{4}{\text{tr}(\Omega)} \sum_{pq} a_{pq}(\theta) a_{pq}^H(\theta) \Omega$</td>
<td></td>
</tr>
<tr>
<td>$\hat{\tau}<em>{pq} = \frac{4}{\text{tr}(\Omega)} a</em>{pq}(\hat{\theta}) \Omega^{-1} a_{pq}(\hat{\theta})$</td>
<td></td>
</tr>
<tr>
<td>end</td>
<td></td>
</tr>
</tbody>
</table>

### 3. CALIBRATION IN 3DC REGIME

#### 3.1. 3DC calibration regime

We consider calibration of radio interferometry stations where antennas are enclosed in a compact array and direction dependent effects are dominant [28]. If we intend to calibrate the data at different frequencies, then a particular decomposition of the Jones matrix is [29, 30]

$$J_{1,p}(\theta_{1,p}) = G_p(g_p) H_{1,p}(\theta_{1,p}) Z_{1,p}(\phi_{1,p})) F_{1,p}(\phi_{1,p})$$

for $i \in \{1, \ldots, D\}, p \in \{1, \ldots, M\}$, $f \in \mathcal{F} = \{f_1, \ldots, f_F\}$ and $\theta_{1,p} = [\phi_{1,p}, g_{1,p}]^T$. Each individual Jones term and its corresponding effect is described in the following:

#### 3.1.1. Ionospheric effects

While travelling through a charged medium as the ionosphere, the incoming wave is affected by a phase delay due to refraction, and written as [31]

$$Z_{1,p}(\phi_{1,p}) = \exp \left( j \phi_{1,p} \right) I_2$$

where $\phi_{1,p} = \eta_{1,p}^f + \xi_{1,p}^f$ in which $\alpha_{1,p} = [\eta_{1,p}^f, \xi_{1,p}^f]^T$ and $f = \frac{1}{\sqrt{f^2 - \frac{1}{2}}}$ [32] stands for the apparent position shift in the source location and $\eta_{1,p}^f = [\eta_{1,p}^f, \xi_{1,p}^f]^T$ is the known position vector of the $p$-th antenna in wavelength units. Therefore, we deduce that $\phi_{1,p} \propto f^{-1}$ [33]. Propagation through the upper part of the atmosphere also results in a rotation of the polarization plane, called Faraday rotation [34].

$$F_{1,p}(\phi_{1,p}) = \begin{bmatrix} \cos(\phi_{1,p}) & -\sin(\phi_{1,p}) \\ \sin(\phi_{1,p}) & \cos(\phi_{1,p}) \end{bmatrix}$$

where the Faraday rotation angle $\phi_{1,p} \propto f^{-2}$ [29] is the same for all antennas due to the compact geometry of the array [28].

#### 3.1.2. Instrumental effects

Receiver electronics introduce direction and frequency independent effects [29], leading to a complex sensor gain matrix, noted $G_p(g_p) = \text{diag}\{g_p\}$.

Finally, let us note that $H_{1,p}(\theta_{1,p})$ is a known matrix given by electromagnetic modeling and a priori information (calibrator sources and antenna positions) [21]. Let us recall that the combined effect of all individual perturbations along a particular signal path is represented by a global Jones matrix, as shown in (2).

#### 3.2. Mono-frequency case

In 3DC regime, calibration amounts to estimate the physical parameters $\theta_{1,p}$ which appear in the structured Jones matrices (2). In the mono-frequency case, we use an alternating least squares approach to estimate unknowns sequentially [35], by optimizing a cost function w.r.t. one parameter while fixing the others. The global frame of the structured calibration algorithm (SCA) is exposed hereafter, for one given $i$ and $p$, and details can be found in [13]. Let us note that $\hat{\phi}_{i} = [\hat{\phi}_{i1}, \ldots, \hat{\phi}_{iM}]^T$ and $\Lambda = [r_1, \ldots, r_M]$.

#### 3.3. Multi-frequency calibration algorithm

The aim of calibration in a multi-frequency scenario is to estimate the parameter vector of interest $e = [\epsilon_1^f, \ldots, \epsilon_D^f]^T$, $g = [g_1^T, \ldots, g_M^T]^T$ where $\epsilon^f = [\phi_{1,p}^f], [\phi_{2,p}^f], [\phi_{3,p}^f], \ldots, [\phi_{D,p}^f]^T$ and $g = [g_1^T, \ldots, g_M^T]^T$.

To do so, we introduce the following cost function

$$l^f(i, \epsilon^f) = \sum_{i=1}^{D} l^f_i(i, \epsilon^f)$$


where $L$.

Instead of considering the original objective function (3), we study the following augmented Lagrangian in which $z$ shared by all agents, and $y_i$.

\[
\begin{align*}
\phi_{i,p} &= \text{argmin}_{\varphi_{i,p}} \left\{ \sum_{j=1}^{M} \left( |J_{i,p} - G_p H_{i,p} Z_{i,p} F_i \varphi_{i,p}|^2 \right) \right\} \\
\bar{z}_i &= \sum_{p=1}^{P} \beta B_i^T B_i \left( \sum_{l \in F} (y_i^{(l)} + \rho \epsilon_i^{(l)}) \right) .
\end{align*}
\]

Minimization (5) is addressed iteratively. To this end, we compute the gradient of $L_i^{(l)}(\epsilon_i^{(l)}, z_i, y_i^{(l)})$ w.r.t. $\epsilon_i^{(l)}$, which induces

\[
\frac{\partial L_i^{(l)}(\epsilon_i^{(l)}, z_i, y_i^{(l)})}{\partial \epsilon_i^{(l)}} = [y_i^{(l)}]_1 + \rho(\hat{y}_i^{(l)} - \frac{1}{f} [z_i]_1) + \sum_{p=1}^{M} \text{tr} \left\{ S_{i,p} + S_{i,p}^T \right\} .
\]

Likewise, we have

\[
\frac{\partial L_i^{(l)}(\epsilon_i^{(l)}, z_i, y_i^{(l)})}{\partial z_i} = [y_i^{(l)}]_2 + \rho(\hat{y}_i^{(l)} - \frac{1}{f} [z_i]_2) + \sum_{p=1}^{M} \text{tr} \{ D_{i,p} + D_{i,p}^T \} .
\]

Estimation of the gains $g_p$ is done as minimization of the following least squares cost function

\[
\kappa(g_p) = \sum_{f \in F} \sum_{r=1}^{D} \left| J_{i,p} - G_p H_{i,p} Z_{i,p} F_i \right|^2
\]

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

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

\textbf{4. NUMERICAL SIMULATIONS}

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 $n_i^{(l)}$ as a function of the signal-to-noise ratio (SNR), the
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 $t$ 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 twinager. 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


**MSCA: Multi-frequency structured calibration algorithm**

```
input : $D, M, F, r_p^{[f]}, H_{h, p}^{[f]}$, $J_{h, p}^{[f]}$ as output of NSCA for $i \in \{1, \ldots, D\}, p \in \{1, \ldots, M\}$ and $f \in F$
output : $\hat{\epsilon}$
initialize: $\epsilon \leftarrow \epsilon_{\text{init}}, \hat{z} \leftarrow z_{\text{init}}, \{\hat{y}^{[f]}\}_{f \in F}$
while stop criterion unreached do
    while stop criterion unreached do
        Obtain $\{\hat{g}^{[f]}_i\}_{i=1,\ldots,D}$ locally with (12)
        Obtain $\{\hat{\epsilon}^{[f]}_i\}_{i=1,\ldots,D}$ locally with (11)
        Obtain $\{\hat{\epsilon}^{[f]}_i\}_{i=1,\ldots,D}$ locally with (10)
        Obtain $\{\hat{\epsilon}^{[f]}_i\}_{i=1,\ldots,D}$ locally with (9)
        Obtain $\{\hat{g}^{[f]}_i\}_{i=1,\ldots,D}$ globally with (8)
        Obtain $\{\hat{g}^{[f]}_i\}_{i=1,\ldots,D}$ locally with (7)
    end
    end
end
```

Fig. 1. MSE of $\eta^{[f]}_i$ vs. SNR, for $D = 2$ bright signal sources, $M = 8$ antennas and 4 weak outlier sources.

Fig. 2. Recovered flux for (a) ideal case, (b) MSCA and (c) Meqtrees solver at 895 MHz. The red cross corresponds to the location of the calibrator source.

Fig. 3. Recovered flux of one weak outlier source after calibration for $F = 9$.


