Signal Processing 153 (2018) 348–354

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Short communication

Robust distributed calibration of radio interferometers with direction dependent distortions[☆]

Virginie Ollier[a,b,*], Mohammed Nabil El Korso[c], André Ferrari[d], Rémy Boyer[b], Pascal Larzabal[a]

[a] ENS Paris-Saclay, SATIE UMR 8029, 61 avenue du Président Wilson, Cachan, 94235, France
[b] Paris-Sud University, L2S UMR 8506, 3 rue Joliot-Curie, Gif-sur-Yvette, 91192, France
[c] Paris-Nanterre University, LEME EA 4416, 50 rue de Sèvres, Ville d'Avray, 92410, France
[d] Nice Sophia-Antipolis University, Lab. J.L. Lagrange, UMR 7293, Parc Valrose, Nice, 06108, France

ARTICLE INFO

Article history:
Received 2 February 2018
Revised 12 July 2018
Accepted 30 July 2018
Available online 31 July 2018

Keywords:
Calibration
Robustness
Multi-frequency
Radio astronomy

ABSTRACT

In radio astronomy, accurate calibration is of crucial importance for the new generation of radio interferometers. More specifically, because of the potential presence of outliers which affect the measured data, robustness needs to be ensured. On the other hand, calibration is improved by taking advantage of these new instruments and exploiting the known structure of parameters of interest across frequency. Therefore, we propose in this paper an iterative robust multi-frequency calibration algorithm based on a distributed and consensus optimization scheme which aims to estimate the complex gains of the receivers and the directional perturbations caused by the ionosphere. Numerical simulations reveal that the proposed distributed calibration technique outperforms the conventional non-robust algorithm and per-channel calibration.

1. Introduction

In radio interferometry, the easiest approach to perform calibration is to consider one frequency bin at a time [1,2], with a single centralized agent processing the data, leading to suboptimality and computational limitations. However, considering the specific variation of parameters across frequency enables to add some modeling structure on the unknown parameters to estimate and therefore perform more accurately direction dependent calibration [3], notably for the new generation of radio interferometers where multiple frequency sub-bands are present [4]. Exploiting the frequency-dependent response has already been studied in the context of bandpass solutions [5], assuming a smooth response in frequency and also for direction dependent calibration with smooth polynomials [6,7]. As direction dependent perturbation effects are particularly significant in observations with new instruments, we focus in this paper on the regime where the receiving elements of the array have a large field-of-view and possibly long baselines, resulting in a direction dependent calibration problem in which receptors see different parts of the ionosphere [2]. We call this the direction dependent distortion regime.

On the one hand, multi-frequency calibration aims to take into account a whole frequency range, *e.g.*, between 30 and 240 MHz for the nominal operating bandpass of the Low Frequency Array (LOFAR) [8,9]. A computationally efficient way to handle the multiple sub-frequency bands in radio astronomy is to apply distributed and consensus algorithms with a decentralized strategy. In this context, we apply the Alternating Direction Method of Multipliers (ADMM) [6,10], which is well-suited for large-scale problems as in radio interferometry. This technique is based on decomposition and coordination tasks with a group of computational agents. Each of them has access to a part of the data and finds a solution to a local sub-problem, in a restricted frequency interval. Communication with a fusion center enforces consensus among all agents, the goal being to solve a global constrained optimization problem. By collecting and storing data in a distributed way among different computational agents, the global operational cost in the network is substantially reduced.

On the other hand, measurements are frequently affected by the presence of outliers due to interference or weak background sources. As a consequence, the noise can no longer be considered Gaussian as in [6]. In order to propose an alternative to [11], where the noise model is based on the Student's t distribution, in this work, we make use of a compound-Gaussian model which includes

☆ This work was supported by MAGELLAN (ANR-14-CE23-0004-01), ON FIRE project (Jeunes Chercheurs GDR-ISIS) and ANR ASTRID project MARGARITA (ANR-17-ASTR-0015).

* Corresponding author.
E-mail address: virginie.ollier@satie.ens-cachan.fr (V. Ollier).

https://doi.org/10.1016/j.sigpro.2018.07.024
0165-1684/

heavy-tailed distributions [12]. This is more general and reveals to be more robust [13]. Therefore, to take into account the presence of outliers and the variability of parameters across frequency, we propose here a Multi-frequency Robust Calibration Algorithm (MRCA) based on a compound-Gaussian distribution for modeling the noise and the consensus ADMM approach [10]. As in [14], we aim at estimating physical parameters appearing in the Jones terms [3,15].

In this paper, we use the following notation: The trace and determinant operators are, respectively, given by tr $\{\cdot\}$ and $|\cdot|$. The $B\times B$ identity matrix is referred by $\mathbf{I}_B$ and $||\cdot||_2$ denotes the l_2 norm. 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 while bdiag$\{\cdot\}$ is the block-diagonal operator. Finally, $\Re\{\cdot\}$ and $\Im\{\cdot\}$ are, respectively, the real and imaginary parts, $[\cdot]_k$ refers to the k-th entry of the considered vector, arg$\{\cdot\}$ is the argument of a complex number and j is the complex number whose square equals -1.

2. Model setup

2.1. Direction dependent distortion regime

An interferometer output consists of visibilities, *i.e.*, correlations of signals measured by two array elements along the corresponding baseline vector. In the noise free case, the measurements are given, for each frequency $f\in\mathcal{F}=\{f_1,\dots,f_F\}$, by

$$\mathbf{S}_{pq}^{[f]}(\boldsymbol{\theta}^{[f]})=\sum_{i=1}^{D}\mathbf{J}_{i,p}^{[f]}(\boldsymbol{\theta}^{[f]})\mathbf{C}_i^{[f]}\mathbf{J}_{i,q}^{[f]^H}(\boldsymbol{\theta}^{[f]}) \tag{1}$$

where $(p,q)\in\{1,\dots,M\}^2$ is a pair of receivers with $p<q$, M denotes the total number of array elements and D the number of calibrator sources, while $\mathbf{J}_{i,p}^{[f]}(\boldsymbol{\theta}^{[f]})$ refers to the so-called 2×2 Jones matrix [3,15], accounting for the perturbation effects along the path from the i-th source to the p-th receiver. We note $\boldsymbol{\theta}^{[f]}$ the unknown parameter vector of interest, whose elements are detailed further below. Finally, $\mathbf{C}_i^{[f]}$ is the known source brightness matrix of the i-th calibrator source, describing its polarization state. After vectorization of measurements (1) and considering the noise effect, we obtain the 4×1 data vector $\mathbf{x}_{pq}^{[f]}=\text{vec}(\mathbf{S}_{pq}^{[f]}(\boldsymbol{\theta}^{[f]}))+\mathbf{n}_{pq}^{[f]}$ where $\mathbf{n}_{pq}^{[f]}$ is the noise vector for baseline $(p,\,q)$ and $\text{vec}(\mathbf{S}_{pq}^{[f]}(\boldsymbol{\theta}^{[f]}))=\sum_{i=1}^{D}\mathbf{s}_{i,pq}^{[f]}(\boldsymbol{\theta}^{[f]})$ in which $\mathbf{s}_{i,pq}^{[f]}(\boldsymbol{\theta}^{[f]})=(\mathbf{J}_{i,q}^{[f]^*}(\boldsymbol{\theta}^{[f]})\otimes\mathbf{J}_{i,p}^{[f]}(\boldsymbol{\theta}^{[f]}))\mathbf{c}_i^{[f]}$ and $\mathbf{c}_i^{[f]}=\text{vec}(\mathbf{C}_i^{[f]})$. Finally, the $4B\times 1$ full measurement vector at frequency f, with $B=\frac{M(M-1)}{2}$ the total number of baselines, reads

$$\mathbf{x}^{[f]}=\left[\mathbf{x}_{12}^{[f]^T},\dots,\mathbf{x}_{(M-1)M}^{[f]^T}\right]^T=\sum_{i=1}^{D}\left[\mathbf{s}_{i,12}^{[f]^T}(\boldsymbol{\theta}^{[f]}),\dots,\mathbf{s}_{i,(M-1)M}^{[f]^T}(\boldsymbol{\theta}^{[f]})\right]^T+\left[\mathbf{n}_{12}^{[f]^T},\dots,\mathbf{n}_{(M-1)M}^{[f]^T}\right]^T. \tag{2}$$

In the direction dependent distortion regime, a particular decomposition of the Jones matrix is given by [15,16]

$$\mathbf{J}_{i,p}^{[f]}(\boldsymbol{\theta}^{[f]})=\mathbf{G}_p^{[f]}(\mathbf{g}_p^{[f]})\mathbf{H}_{i,p}^{[f]}\mathbf{Z}_{i,p}^{[f]}(\varphi_{i,p}^{[f]})\mathbf{F}_{i,p}^{[f]}(\vartheta_{i,p}^{[f]}). \tag{3}$$

Specifically, the complex electronic gain matrix is represented by $\mathbf{G}_p^{[f]}(\mathbf{g}_p^{[f]})=\text{diag}\{\mathbf{g}_p^{[f]}\}$ while $\mathbf{H}_{i,p}^{[f]}$ is an assumed known matrix gathering the geometric delay and beam pattern [3,15]. In addition, propagation through the ionosphere induces two effects. The first one is a phase delay given by the matrix $\mathbf{Z}_{i,p}^{[f]}(\varphi_{i,p}^{[f]})$ [3], and written as

$$\mathbf{Z}_{i,p}^{[f]}(\varphi_{i,p}^{[f]})=\exp\left(j\varphi_{i,p}^{[f]}\right)\mathbf{I}_2 \tag{4}$$

where $\varphi_{i,p}^{[f]}\propto\text{TEC}_{i,p}/f$ [17], with $\text{TEC}_{i,p}$ the Total Electron Content defined as the integrated electron density along line of sight $i-p$.

The second effect is the so-called Faraday rotation [16] in (3), given by

$$\mathbf{F}_{i,p}^{[f]}(\vartheta_{i,p}^{[f]})=\begin{bmatrix}\cos(\vartheta_{i,p}^{[f]}) & -\sin(\vartheta_{i,p}^{[f]})\\ \sin(\vartheta_{i,p}^{[f]}) & \cos(\vartheta_{i,p}^{[f]})\end{bmatrix} \tag{5}$$

where the unknown rotation angle $\vartheta_{i,p}^{[f]}\propto\text{RM}_{i,p}/f^2$ [3] and $\text{RM}_{\text{i,p}}$ is the Rotation Measure which depends on the magnetic field and the electron density along the path $i-p$.

Therefore, the $(2MD+2M)\times 1$ complex unknown parameter vector is given by $\boldsymbol{\theta}^{[f]}=[\boldsymbol{\epsilon}^{[f]^T},\mathbf{g}_1^{[f]^T},\dots,\mathbf{g}_M^{[f]^T}]^T$ in which $\boldsymbol{\epsilon}^{[f]}=[\vartheta_{1,1}^{[f]},\dots,\vartheta_{D,M}^{[f]},\exp(j\varphi_{1,1}^{[f]}),\dots,\exp(j\varphi_{D,M}^{[f]})]^T$ refers to the frequency dependent per-receiver and per-source ionospheric effects.

2.2. Noise modeling as a compound-Gaussian distribution

The presence of outliers has multiple causes in radio astronomy, such as errors in the sky model due to weak sources in the background [11] or man-made Radio Frequency Interference (RFI) [18], leading to statistics, with heavy-tailed distributions, different from the classical Gaussian case [1]. To ensure robustness in the proposed estimator, we adopt a two-scale compound-Gaussian noise modeling given for each baseline by

$$\mathbf{n}_{pq}^{[f]}=\sqrt{\tau_{pq}^{[f]}}\;\boldsymbol{\mu}_{pq}^{[f]}, \tag{6}$$

where the power factor $\tau_{pq}^{[f]}$ is a positive real random variable and the 4×1 vector $\boldsymbol{\mu}_{pq}^{[f]}$ follows a zero-mean complex circular Gaussian distribution, *i.e.*, $\boldsymbol{\mu}_{pq}^{[f]}\sim\mathcal{CN}(\mathbf{0},\boldsymbol{\Omega}^{[f]})$[1]. Therefore, calibration amounts to estimate for each frequency f the $(2MD+2M)\times 1$ vector $\boldsymbol{\theta}^{[f]}$ describing Jones matrices, as well as $B\times 1$ texture realizations $\boldsymbol{\tau}^{[f]}=[\tau_{12}^{[f]},\tau_{13}^{[f]},\dots,\tau_{(M-1)M}^{[f]}]^T$ and the 4×4 speckle covariance matrix $\boldsymbol{\Omega}^{[f]}$ (which must satisfy, *e.g.*, $\text{tr}\{\boldsymbol{\Omega}^{[f]}\}=1$ to avoid the ambiguity with the power factor, the choice of this constraint being arbitrary [19]). In the following, we assume independence of $\mathbf{n}_{pq}^{[f]}$ between baselines and frequencies and no specific prior structure exists *w.r.t. f*. Let us note that the algorithm can be adapted to perform an independent estimation of unknown parameters for each time interval. In order to consider different time scales for the unknown effects, a similar approach to the one proposed in this work for the multi-frequency scenario can be adopted, using specific time variation models or assuming smoothness across time.

3. Description of the proposed estimator

In this section, we introduce the Relaxed Maximum Likelihood (RML) method. Then, the ADMM algorithm is proposed to estimate the frequency dependent parameters in a distributed way. In the iterative procedure, each subset of parameters is updated alternatively, while fixing the remaining parameters.

3.1. Robust estimation of Jones matrices

Robust calibration is based on the model (2) and the compound-Gaussian noise model (6). Estimations are performed iteratively with the ML method similar to that of [19]. Specifically, here, we choose not to specify the probability density function (pdf) of the texture parameters which are assumed unknown and deterministic, leading to the RML. By doing so, we ensure more flexibility and robustness to any prior mismatch *w.r.t.* the unknown

[1] It is possible to consider a baseline dependent covariance matrix $\boldsymbol{\Omega}_{pq}^{[f]}$. In this case, the proposed algorithm requires a few modifications which are straightforward.