

BAYESIAN CALIBRATION USING DIFFERENT PRIOR DISTRIBUTIONS: AN ITERATIVE MAXIMUM A POSTERIORI APPROACH FOR RADIO INTERFEROMETERS

V. Ollier*[†], M. N. El Korso[‡], A. Ferrari[§], R. Boyer[†] and 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

In this paper, we aim to design robust estimation techniques based on the compound-Gaussian (CG) process and adapted for calibration of radio interferometers. The motivation beyond this is due to the presence of outliers leading to an unrealistic traditional Gaussian noise assumption. Consequently, to achieve robustness, we adopt a maximum a posteriori (MAP) approach which exploits Bayesian statistics and follows a sequential updating procedure here. The proposed algorithm is applied in a multi-frequency scenario in order to enhance the estimation and correction of perturbation effects. Numerical simulations assess the performance of the proposed algorithm for different noise models, Student's t , K , Laplace, Cauchy and inverse-Gaussian compound-Gaussian distributions *w.r.t.* the classical non-robust Gaussian noise assumption.

Index Terms— Bayesian calibration, compound-Gaussian distribution, robustness, maximum a posteriori estimation

1. INTRODUCTION

Robust calibration in radio astronomy amounts to estimating all environmental and instrumental perturbation effects which corrupt the signal of interest in a non-Gaussian environment. Indeed, radio interferometric data are affected by the occurrence of outliers due, *e.g.*, to man-made radio frequency interferences or unmodelled weak sources [1–3]. Therefore, estimation under a conventional Gaussian noise model does not perform optimally and robust calibration is required. Furthermore, unknown perturbation effects being highly variable in frequency [4], exploiting their variation *w.r.t.* frequency improves estimation performance in the calibration process while maintaining an acceptable computational cost [5, 6]. To this end, calibration can be reformulated as a distributed consensus optimization problem and solved thanks to the alternating direction method of multipliers (ADMM) [7].

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

To investigate non-Gaussian modeling, we consider a two-scale compound-Gaussian (CG) process, also referred to as a spherically invariant random variable in the radar community [8, 9], which is represented as the product of a positive texture term with a given a priori distribution, multiplied by a complex speckle component following a zero-mean Gaussian distribution. Such a model is suitable for practical scenario and widely used in signal processing applications [10, 11]. In our scenario, the statistical distribution of the texture parameter is unknown. Thus, we study different prior distributions for the texture, available in closed-form, and generating various non-Gaussian (heavy-tailed) noise models. In order to estimate perturbation effects and noise parameters, we devise an iterative algorithm where optimization is performed successively *w.r.t.* each unknown parameter, while fixing the others. A maximum a posteriori (MAP) approach is adopted to exploit the statistical information and incorporate the texture prior distribution, leading to the so-called iterative maximum a posteriori estimator (IMAPE).

In this paper, we use the following notation: symbols $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$ denote, respectively, the transpose, the complex conjugate and the Hermitian operator. The Kronecker product is represented by \otimes , $E\{\cdot\}$ denotes the expectation operator and $\text{diag}\{\cdot\}$ converts a vector into a diagonal matrix. The trace and determinant operators are, respectively, given by $\text{tr}\{\cdot\}$ and $|\cdot|$. Finally, the symbol \mathbf{I}_2 represents the 2×2 identity matrix, $\text{vec}(\cdot)$ stacks the columns of a matrix on top of one another, j is the complex number whose square equals -1 , $\Gamma(\cdot)$ is the gamma function and $\Psi(\cdot)$ the digamma function, such that, $\Psi(x) = \partial \ln(\Gamma(x)) / \partial x$.

2. DATA MODEL

2.1. Direction dependent regime

In radio astronomy, propagation of the i -th 2-dimensional incoming signal $\mathbf{s}_i^{[f]}$ at frequency $f \in \mathcal{F} = \{f_1, \dots, f_F\}$ generates the following voltage at antenna p [12]

$$\mathbf{v}_{i,p}^{[f]}(\boldsymbol{\theta}^{[f]}) = \mathbf{J}_{i,p}^{[f]}(\boldsymbol{\theta}^{[f]}) \mathbf{s}_i^{[f]} \quad (1)$$

where the so-called 2×2 Jones matrix $\mathbf{J}_{i,p}^{[f]}(\boldsymbol{\theta}^{[f]})$ is parametrized by unknown vector $\boldsymbol{\theta}^{[f]}$ and accounts for all perturbation effects along the propagation path [13].

For each antenna pair (p, q) , the interferometer measures cross-correlations between voltages, given by

$$\begin{aligned} \mathbf{s}_{pq}^{[f]}(\boldsymbol{\theta}^{[f]}) &= \mathbb{E} \left\{ \left(\sum_{i=1}^D \mathbf{v}_{i,p}^{[f]}(\boldsymbol{\theta}^{[f]}) \right) \left(\sum_{i=1}^D \mathbf{v}_{i,q}^{[f]H}(\boldsymbol{\theta}^{[f]}) \right) \right\} \\ &= \sum_{i=1}^D \mathbf{J}_{i,p}^{[f]}(\boldsymbol{\theta}^{[f]}) \mathbf{C}_i^{[f]} \mathbf{J}_{i,q}^{[f]H}(\boldsymbol{\theta}^{[f]}) \end{aligned} \quad (2)$$

where D is the number of known calibrator sources, $(p, q) \in \{1, \dots, M\}^2$ with M the number of sensor elements and $\mathbf{C}_i^{[f]} = \mathbb{E}\{\mathbf{s}_i^{[f]} \mathbf{s}_i^{[f]H}\}$ stands for the known source coherency matrix [14]. We rewrite (2) as a 4×1 measurement vector and add the noise contribution, such that,

$$\mathbf{x}_{pq}^{[f]} = \sum_{i=1}^D \mathbf{s}_{i,pq}^{[f]}(\boldsymbol{\theta}^{[f]}) + \mathbf{n}_{pq}^{[f]} \quad (3)$$

in which $\mathbf{s}_{i,pq}^{[f]}(\boldsymbol{\theta}^{[f]}) = \text{vec} \left(\mathbf{J}_{i,p}^{[f]}(\boldsymbol{\theta}^{[f]}) \mathbf{C}_i^{[f]} \mathbf{J}_{i,q}^{[f]H}(\boldsymbol{\theta}^{[f]}) \right) = \left(\mathbf{J}_{i,q}^{[f]*}(\boldsymbol{\theta}^{[f]}) \otimes \mathbf{J}_{i,p}^{[f]}(\boldsymbol{\theta}^{[f]}) \right) \mathbf{c}_i^{[f]}$ and $\mathbf{c}_i^{[f]} = \text{vec}(\mathbf{C}_i^{[f]})$. If we note $B = \frac{M(M-1)}{2}$ the global number of antenna pairs, all cross-correlation measurements are given by the following $4B \times 1$ full observation vector at frequency f

$$\mathbf{x}^{[f]} = \left[\mathbf{x}_{12}^{[f]T}, \mathbf{x}_{13}^{[f]T}, \dots, \mathbf{x}_{(M-1)M}^{[f]T} \right]^T. \quad (4)$$

Each Jones matrix $\mathbf{J}_{i,p}^{[f]}(\boldsymbol{\theta}^{[f]})$ in (1) can be decomposed into a particular sequence of individual Jones terms [14]. Each of them corresponds to a particular effect along the signal propagation path. More specifically, we can write

$$\mathbf{J}_{i,p}^{[f]}(\boldsymbol{\theta}^{[f]}) = \mathbf{G}_p^{[f]}(\mathbf{g}_p^{[f]}) \mathbf{H}_{i,p}^{[f]} \mathbf{Z}_{i,p}^{[f]}(\varphi_{i,p}^{[f]}) \mathbf{F}_{i,p}^{[f]}(\vartheta_{i,p}^{[f]}) \quad (5)$$

where $\mathbf{G}_p^{[f]}(\mathbf{g}_p^{[f]}) = \text{diag}\{\mathbf{g}_p^{[f]}\}$ stands for the complex electronic gain matrix and $\mathbf{H}_{i,p}^{[f]}$ is assumed known, including the geometric delay and beam pattern effects [13]. Due to large distances between antennas and wide field-of-view, propagation through the ionosphere leads to a per-antenna direction dependent phase delay $\mathbf{Z}_{i,p}^{[f]}(\varphi_{i,p}^{[f]}) = \exp(j\varphi_{i,p}^{[f]}) \mathbf{I}_2$ but also to a rotation effect [15], named Faraday rotation and written as $\mathbf{F}_{i,p}^{[f]}(\vartheta_{i,p}^{[f]}) = \begin{bmatrix} \cos(\vartheta_{i,p}^{[f]}) & -\sin(\vartheta_{i,p}^{[f]}) \\ \sin(\vartheta_{i,p}^{[f]}) & \cos(\vartheta_{i,p}^{[f]}) \end{bmatrix}$.

Thus, the unknown parameter vector of interest which describes all Jones matrices at frequency f reads $\boldsymbol{\theta}^{[f]} = [\vartheta_{1,1}^{[f]}, \dots, \vartheta_{D,M}^{[f]}, \exp(j\varphi_{1,1}^{[f]}), \dots, \exp(j\varphi_{D,M}^{[f]}), \mathbf{g}_1^{[f]T}, \dots, \mathbf{g}_M^{[f]T}]^T$.

2.2. Robust noise modeling

In order to propose a robust calibration algorithm, we study the flexible CG representation which encompasses a broad range of heavy-tailed distributions and can be written as [16, 17]

$$\mathbf{n}_{pq}^{[f]} = \sqrt{\tau_{pq}^{[f]}} \boldsymbol{\mu}_{pq}^{[f]}, \quad (6)$$

where $\tau_{pq}^{[f]}$ denotes the positive texture random variable and $\boldsymbol{\mu}_{pq}^{[f]} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Omega}^{[f]})$ with a uniform prior on the positive semidefinite cone for $\boldsymbol{\Omega}^{[f]}$. To avoid any ambiguity [18, 19], we impose an arbitrary constraint on the speckle covariance matrix $\boldsymbol{\Omega}^{[f]}$, e.g., $\text{tr}\{\boldsymbol{\Omega}^{[f]}\} = 1$. Therefore, in our context, calibration requires estimating not only parameters of interest $\boldsymbol{\theta}^{[f]}$ but also noise terms $\boldsymbol{\tau}^{[f]} = [\tau_{12}^{[f]}, \tau_{13}^{[f]}, \dots, \tau_{(M-1)M}^{[f]}]^T$ and $\boldsymbol{\Omega}^{[f]}$.

3. MAXIMUM A POSTERIORI ESTIMATION

3.1. Principle of the proposed algorithm

In this section, we investigate robust calibration from a Bayesian perspective. For this purpose, we focus on the texture realizations and derive MAP-based estimators.

Assuming independence of $\mathbf{n}_{pq}^{[f]}$ w.r.t. frequency channels f and antenna pairs (p, q) , the conditional log-likelihood function is written as

$$\begin{aligned} \mathcal{L}_C &= \ln p \left(\{\mathbf{x}^{[f]}\}_{f \in \mathcal{F}} \mid \{\boldsymbol{\tau}^{[f]}; \boldsymbol{\theta}^{[f]}, \boldsymbol{\Omega}^{[f]}\}_{f \in \mathcal{F}} \right) = - \sum_{f \in \mathcal{F}} \\ &\sum_{pq} \left(\frac{1}{\tau_{pq}^{[f]}} \mathbf{u}_{pq}^{[f]H}(\boldsymbol{\theta}^{[f]}) \boldsymbol{\Omega}^{[f]-1} \mathbf{u}_{pq}^{[f]}(\boldsymbol{\theta}^{[f]}) + \ln |\pi \tau_{pq}^{[f]} \boldsymbol{\Omega}^{[f]}| \right) \end{aligned} \quad (7)$$

in which $\mathbf{u}_{pq}^{[f]}(\boldsymbol{\theta}^{[f]}) = \mathbf{x}_{pq}^{[f]} - \sum_{i=1}^D \mathbf{s}_{i,pq}^{[f]}(\boldsymbol{\theta}^{[f]})$. Thanks to Bayes' law, the joint log-likelihood function, denoted as \mathcal{L}_J , reads

$$\mathcal{L}_J = \mathcal{L}_C + \sum_{f \in \mathcal{F}} \sum_{pq} \ln \left(p(\tau_{pq}^{[f]}; \varphi^{[f]}) \right) \quad (8)$$

in which $\varphi^{[f]}$ stand for the corresponding unknown hyperparameters which need to be estimated.

The MAP estimator consists in maximizing (8) w.r.t. each unknown individual parameter. The expressions being mutually dependent, we adopt an iterative procedure leading to the IMAPE. More specifically, we optimize alternately w.r.t. $\boldsymbol{\theta}^{[f]}$, $\boldsymbol{\tau}^{[f]}$, $\varphi^{[f]}$ and $\boldsymbol{\Omega}^{[f]}$ in a step-wise approach. At each step, the remaining parameters are assumed fixed and known from the previous iteration. The overall procedure of the proposed IMAPE is shown in Table 1 hereafter.

Estimation of $\{\hat{\boldsymbol{\theta}}^{[f]}\}_{f \in \mathcal{F}}$ in step 1 of the IMAPE, for a given $\{\boldsymbol{\tau}^{[f]}\}_{f \in \mathcal{F}}$ and $\{\boldsymbol{\Omega}^{[f]}\}_{f \in \mathcal{F}}$, is detailed in [20, 21].

Table 1: IMAPE Iterative MAP estimator

	input : $D, M, B, \{\{C_i^{[f]}\}_{i=1, \dots, D}, \mathbf{x}^{[f]}\}_{f \in \mathcal{F}}$
	output : $\{\hat{\boldsymbol{\theta}}^{[f]}\}_{f \in \mathcal{F}}$
	initialize : $\{\hat{\boldsymbol{\Omega}}^{[f]} \leftarrow \boldsymbol{\Omega}_{\text{init}}^{[f]}\}_{f \in \mathcal{F}},$ $\{\hat{\tau}_{pq}^{[f]} \leftarrow \tau_{pq\text{init}}^{[f]}\}_{f \in \mathcal{F}, p < q, (p,q) \in \{1, \dots, M\}^2}$
	while <i>stop criterion unreached</i> do
1	$\{\hat{\boldsymbol{\theta}}^{[f]}\}_{f \in \mathcal{F}} = \underset{\{\boldsymbol{\theta}^{[f]}\}_{f \in \mathcal{F}}}{\text{argmax}} \{\mathcal{L}_C\}$, see [20, 21]
2	Estimation of $\boldsymbol{\varphi}^{[f]}$ for $f \in \mathcal{F}$
3	$\hat{\boldsymbol{\Omega}}^{[f]} = \frac{1}{B} \sum_{pq} \frac{\mathbf{u}_{pq}^{[f]}(\hat{\boldsymbol{\theta}}^{[f]}) \mathbf{u}_{pq}^{[f]H}(\hat{\boldsymbol{\theta}}^{[f]})}{\hat{\tau}_{pq}^{[f]}}$ $\hat{\boldsymbol{\Omega}}^{[f]} = \frac{\hat{\boldsymbol{\Omega}}^{[f]}}{\text{tr}\{\hat{\boldsymbol{\Omega}}^{[f]}\}}$ for $f \in \mathcal{F}$
4	Estimation of $\boldsymbol{\tau}^{[f]}$ for $f \in \mathcal{F}$

These parameters of interest vary specifically with frequency such that we can reformulate the problem as a constrained optimization scheme and solve it distributedly with a network of agents and the consensus ADMM procedure [22]. Let us note that $\partial \mathcal{L}_J = \partial \mathcal{L}_C$ if we take the derivative *w.r.t.* $\boldsymbol{\theta}^{[f]}$ or $\boldsymbol{\Omega}^{[f]}$ and as mentioned previously, the normalization step $\hat{\boldsymbol{\Omega}}^{[f]} = \frac{\hat{\boldsymbol{\Omega}}^{[f]}}{\text{tr}\{\hat{\boldsymbol{\Omega}}^{[f]}\}}$ is added in step 3 of the IMAPE for proper identifiability. Consequently, in what follows, we derive and study for each prior distribution $p(\tau_{pq}^{[f]}; \boldsymbol{\varphi}^{[f]})$ the corresponding IMAPE and highlight how the estimation of Jones matrices' parameters is affected in the numerical simulations section. Thus, we focus on steps 2 to 4 of the IMAPE. Furthermore, since the reasoning is equivalent for each frequency f , we omit the frequency dependence in the rest of the paper for sake of clarity.

3.2. Study of different texture priors

In the Bayesian approach, the expression of the joint log-likelihood in (8) depends on $p(\tau_{pq}; \boldsymbol{\varphi})$, which describes the statistics of the texture variable. In this section, we study different kinds of heavy-tailed CG distributions obtained with different texture priors, available in closed-form [8, 9]. Then, the corresponding estimators for $\boldsymbol{\varphi}$, $\boldsymbol{\Omega}$ and $\boldsymbol{\tau}$ are deduced.

3.2.1. K-distribution

The noise vector \mathbf{n}_{pq} is said to follow the K-distribution [8] if the texture τ_{pq} is generated as

$$p(\tau_{pq}; a, b) = \frac{1}{\Gamma(a)b^a} \tau_{pq}^{a-1} \exp\left(-\frac{\tau_{pq}}{b}\right) \quad (9)$$

which is a gamma distribution characterized by the shape parameter a and the scale parameter b , such that, $\boldsymbol{\varphi} = [a, b]$.

Table 2: IMAPE K-distributed noise

	2 Obtain \hat{a} solving $-B\Psi(a) + \sum_{pq} \ln(\hat{\tau}_{pq}) - B \ln\left(\frac{\sum_{pq} \hat{\tau}_{pq}}{Ba}\right) = 0$ Obtain \hat{b} with (12)
	3 Obtain $\hat{\boldsymbol{\Omega}}$ with $\hat{\boldsymbol{\Omega}} = \frac{2}{B} \sum_{pq} \frac{\mathbf{u}_{pq}(\hat{\boldsymbol{\theta}}) \mathbf{u}_{pq}^H(\hat{\boldsymbol{\theta}})}{(\hat{a}-5)\hat{b} + \left((\hat{a}-5)^2\hat{b}^2 + 4\hat{b}\mathbf{u}_{pq}^H(\hat{\boldsymbol{\theta}})\hat{\boldsymbol{\Omega}}^{-1}\mathbf{u}_{pq}(\hat{\boldsymbol{\theta}})\right)^{1/2}}$ $\hat{\boldsymbol{\Omega}} = \frac{\hat{\boldsymbol{\Omega}}}{\text{tr}\{\hat{\boldsymbol{\Omega}}\}}$
	4 Obtain $\hat{\boldsymbol{\tau}}$ with (11)

Inserting (9) into (8) leads to

$$\mathcal{L}_J = \mathcal{L}_C + (a-1) \sum_{pq} \ln(\tau_{pq}) - \frac{\sum_{pq} \tau_{pq}}{b} - B \ln(\Gamma(a)) - Ba \ln(b). \quad (10)$$

Solving $\partial \mathcal{L}_J / \partial \tau_{pq} = 0$ gives the following estimate for the texture

$$\hat{\tau}_{pq} = \frac{(a-5)b + \left((a-5)^2b^2 + 4b\mathbf{u}_{pq}^H(\boldsymbol{\theta})\boldsymbol{\Omega}^{-1}\mathbf{u}_{pq}(\boldsymbol{\theta})\right)^{1/2}}{2}. \quad (11)$$

From $\partial \mathcal{L}_J / \partial \boldsymbol{\varphi} = 0$, we deduce the following analytical expression for the scale

$$\hat{b} = \frac{\sum_{pq} \tau_{pq}}{Ba}. \quad (12)$$

Finally, the shape parameter has to satisfy the following equation

$$-B\Psi(\hat{a}) + \sum_{pq} \ln(\tau_{pq}) - B \ln(b) = 0 \quad (13)$$

from which the shape parameter can be easily calculated numerically.

For the K-distributed noise model, steps 2 to 4 of the IMAPE are given in Table 2. Let us note that (12) was substituted into (13) and likewise, (11) was inserted into the expression of $\hat{\boldsymbol{\Omega}}$ due to our iterative approach.

In the following sections, using the same technique, we only give the results for the estimates of the unknown noise parameters and deduce similar procedures as in Table 2.

3.2.2. Student's t distribution

If the texture parameter follows an inverse gamma function, given by

$$p(\tau_{pq}; a, b) = \frac{b^a}{\Gamma(a)} \tau_{pq}^{-a-1} \exp\left(-\frac{b}{\tau_{pq}}\right), \quad (14)$$

the corresponding noise model is a Student's t distribution [23]. With (14), (8) and the derivative *w.r.t.* τ_{pq} , we obtain

$$\hat{\tau}_{pq} = \frac{b + \mathbf{u}_{pq}^H(\boldsymbol{\theta})\boldsymbol{\Omega}^{-1}\mathbf{u}_{pq}(\boldsymbol{\theta})}{a + 5} \quad (15)$$

while hyperparameters can be deduced from $\hat{b} = \frac{Ba}{\sum_{pq} \frac{1}{\tau_{pq}}}$ and

$$-B\Psi(a) - \sum_{pq} \ln(\tau_{pq}) + B \ln(b) = 0. \quad (16)$$

From now on, we introduce texture prior distributions with only one hyperparameter to estimate.

3.2.3. Cauchy distribution

Using the inverse gamma prior distribution (14) with fixed $a = 1$ leads to Cauchy noise modeling for \mathbf{n}_{pq} [9]. This is a particular case of section 3.2.2 where (16) becomes $\hat{a} = 1$.

3.2.4. Laplace distribution

Let us consider the following exponential distribution as prior: $p(\tau_{pq}; \lambda) = \lambda \exp(-\lambda\tau_{pq})$ characterized by rate parameter λ , and corresponding to Laplace noise [9]. After some calculus, the texture estimate reads

$$\hat{\tau}_{pq} = \frac{-4 + \left(16 + 4\lambda\mathbf{u}_{pq}^H(\boldsymbol{\theta})\boldsymbol{\Omega}^{-1}\mathbf{u}_{pq}(\boldsymbol{\theta})\right)^{1/2}}{2\lambda} \quad (17)$$

while the single parameter λ is given by $\hat{\lambda} = \frac{B}{\sum_{pq} \tau_{pq}}$.

3.2.5. Inverse Gaussian compound-Gaussian (IG-CG) distribution

Finally, we consider the case when the texture variable is modeled by an inverse Gaussian distribution, also known as the Wald distribution, with shape λ and assumed unit mean [8]. Thus, its probability density function is of the form

$$p(\tau_{pq}; \lambda) = \left(\frac{\lambda}{2\pi}\right)^{1/2} \tau_{pq}^{-3/2} \exp\left(\frac{-\lambda(\tau_{pq}-1)^2}{2\tau_{pq}}\right).$$

The resulting CG is called an IG-CG distribution and the corresponding noise parameters are given by

$$\hat{\tau}_{pq} = \frac{-11 + (121 + 4\lambda(\lambda + 2\mathbf{u}_{pq}^H(\boldsymbol{\theta})\boldsymbol{\Omega}^{-1}\mathbf{u}_{pq}(\boldsymbol{\theta})))^{1/2}}{2\lambda} \quad (18)$$

$$\text{and } \hat{\lambda} = \frac{B}{\sum_{pq} \frac{(\tau_{pq}-1)^2}{\tau_{pq}}}.$$

4. NUMERICAL SIMULATIONS

We aim to compare the estimation performance of the IMAPE for selected examples of CG-based noise models, in a realistic

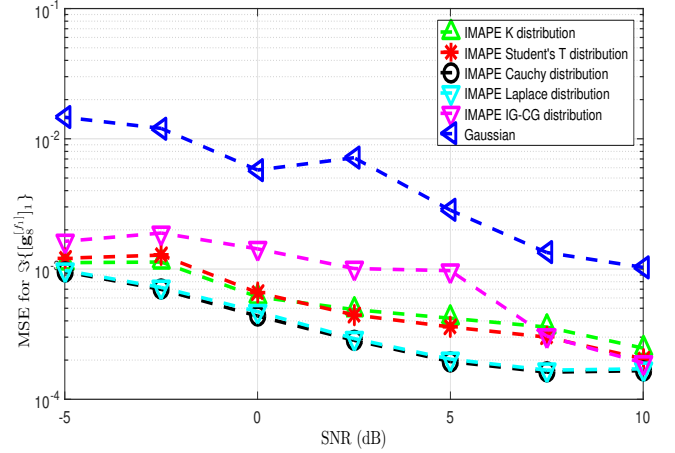


Fig. 1. Evolution of the MSE for the imaginary part of a given complex gain as a function of the SNR for different kinds of CG distributions.

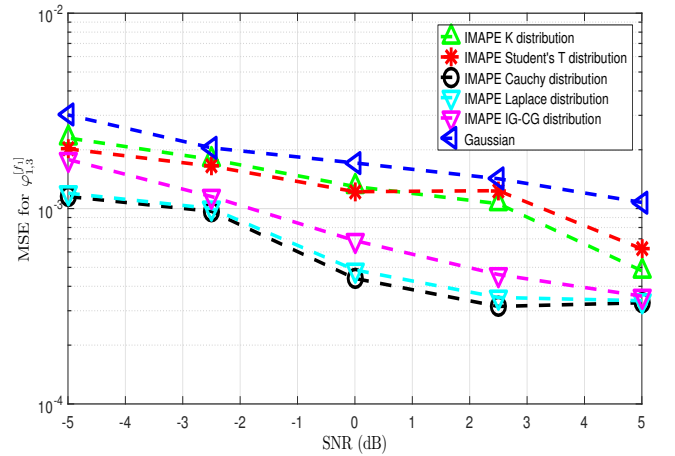


Fig. 2. Evolution of the MSE for a given ionospheric phase delay as a function of the SNR for different kinds of CG distributions.

situation for radio astronomy. Let us consider $M = 8$ antennas in a 2-dimensional sensor array, $D = 2$ bright calibrator sources and $D' = 4$ weak background sources (at least 10 times weaker than the calibrator sources) which are spatially randomly distributed following a uniform distribution. Perturbation effects are randomly generated thanks to the measurement equation in (5) and white Gaussian background noise is added to the data. Due to the presence of faint unknown sources, the noise can no longer be considered Gaussian. To highlight the robustness of the proposed Bayesian method, we compare with traditional calibration based on the minimization of a least squares cost function, *i.e.*, by assuming independent Gaussian noise with no Bayesian approach.

In Figures 1 and 2, we plot the mean square error (MSE) of the imaginary part of one given complex gain and one

ionospheric phase delay, respectively, as a function of the signal-to-noise ratio (SNR) for different noise modeling. The SNR is defined as the ratio of the normalized power of the D calibrator sources over the sum of normalized power of the D' background sources and a noise factor. The commonly used Gaussian noise assumption has poor performance as expected while we notice some dissimilarities for the IMAPE depending on which $p(\tau_{pq}; \varphi)$ is considered. We note that the IMAPE based on Cauchy distribution provides the best performance which can be selected as a competitive noise model for Bayesian calibration of radio interferometers.

5. CONCLUSION

In this paper, we adopted a Bayesian approach for robust calibration of radio interferometers where robustness needs to be ensured due to the presence of outliers. To address this problem, we derived specific MAP-based estimators which exploit the statistical distribution of the texture in the CG representation. Such family covers a wide variety of non-Gaussian models, obtained with different prior distributions. Application to simulated data shows that the choice of the prior actually affects the estimation of the perturbation effects and reveals to be more or less accurate to model the noise contribution in our application. This leads us to select the Cauchy distribution as a suitable model for the noise. Additional comparison with realistic data simulations in the software Meqtrees and real data will be performed in future work.

REFERENCES

- [1] J. Raza, A.-J. Boonstra, and A.-J. Van der Veen, "Spatial filtering of RF interference in radio astronomy," *IEEE Signal Processing Letters*, vol. 9, no. 2, pp. 64–67, 2002.
- [2] S. Yatawatta and S. Kazemi, "Robust radio interferometric calibration," in *International Conference on Acoustics, Speech and Signal Processing (IEEE ICASSP)*, Florence, Italy, 2014, pp. 5392–5396.
- [3] P. Zarka et al., "NenUFAR: Instrument description and science case," in *International Conference on Antenna Theory and Techniques (IEEE ICATT)*, Kharkiv, Ukraine, 2015.
- [4] J. E. Noordam and O. M. Smirnov, "The MeqTrees software system and its use for third-generation calibration of radio interferometers," *Astronomy & Astrophysics*, vol. 524, pp. A61, 2010.
- [5] S. Yatawatta, "Distributed radio interferometric calibration," *Monthly Notices of the Royal Astronomical Society*, vol. 449, no. 4, pp. 4506–4514, 2015.
- [6] M. Brossard, M. N. El Korso, M. Pesavento, R. Boyer, P. Larzabal, and S. J. Wijnholds, "Parallel multi-wavelength calibration algorithm for radio astronomical arrays," *Elsevier Signal Processing Journal*, vol. 145, pp. 258–271, 2018.
- [7] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.
- [8] E. Ollila, D. E. Tyler, V. Koivunen, and H.V. Poor, "Complex elliptically symmetric distributions survey, new results and applications," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5597–5625, 2012.
- [9] E. Jay, *Détection en Environnement non Gaussien*, Ph.D. thesis, Université de Cergy Pontoise, 2002.
- [10] E. Conte and M. Longo, "Characterisation of radar clutter as a spherically invariant random process," in *IEE Proceedings F (Communications, Radar and Signal Processing)*, 1987, vol. 134, pp. 191–197.
- [11] F. Gini and A. Farina, "Vector subspace detection in compound-Gaussian clutter. part I: survey and new results," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 4, pp. 1295–1311, 2002.
- [12] L. M. Ker, *Radio AGN evolution with low frequency radio surveys*, Ph.D. thesis, The University of Edinburgh, 2012.
- [13] A. R. Thompson, J. M. Moran, and G. W. Swenson Jr., *Interferometry and synthesis in radio astronomy*, 2nd ed., John Wiley & Sons, New York, 2001.
- [14] O. M. Smirnov, "Revisiting the radio interferometer measurement equation. I. A full-sky Jones formalism," *Astronomy & Astrophysics*, vol. 527, no. A106, 2011.
- [15] K. Davies, *Ionospheric radio*, Peter Peregrinus, London, 1990.
- [16] K. Yao, "Spherically invariant random processes: Theory and applications," in *Communications, Information and Network Security*, V.K. Bhargava et al., Eds., pp. 315–331. Springer, Dordrecht, The Netherlands: Kluwer Academic, 2003.
- [17] J. Wang, A. Dogandzic, and A. Nehorai, "Maximum likelihood estimation of compound-gaussian clutter and target parameters," *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3884–3898, 2006.
- [18] X. Zhang, M. N. El Korso, and M. Pesavento, "MIMO radar target localization and performance evaluation under SIRP clutter," *Signal Processing Journal*, vol. 130, pp. 217–232, 2017.
- [19] V. Ollier, M. N. El Korso, R. Boyer, P. Larzabal, and M. Pesavento, "Robust calibration of radio interferometers in non-Gaussian environment," *IEEE Transactions on Signal Processing*, vol. 65, no. 21, pp. 5649–5660, 2017.
- [20] V. Ollier, *Contributions to robust calibration methods in radio astronomy*, Ph.D. thesis, ENS Paris-Saclay, 2018.
- [21] V. Ollier, M. N. El Korso, A. Ferrari, R. Boyer, and P. Larzabal, "Robust calibration of radio interferometers in multi-frequency scenario," in *International Conference on Acoustics, Speech and Signal Processing (IEEE ICASSP)*, Calgary, Canada, 2018.
- [22] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: numerical methods*, Prentice Hall, 1989.
- [23] K. L. Lange, R. J. A. Little, and J. M. G. Taylor, "Robust statistical modeling using the t distribution," *Journal of the American Statistical Association*, vol. 84, no. 408, pp. 881–896, 1989.