

Algorithme de calibration robuste dans un contexte de radio interférométrie

Virginie OLLIER^{1,3}, Mohammed Nabil EL KORSO², Rémy BOYER³, Pascal LARZABAL¹, Marius PESAVENTO⁴

¹SATIE/ENS Cachan/Université Paris-Saclay, 61 av. du Président Wilson, 94235 Cachan cedex, France

²LEME/Université Paris-Ouest, 50 rue de Sèvres, 92410 Ville d'Avray, France

³L2S/Université Paris-Sud, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette, France

⁴TUD/Communication Systems Group, Merckstrasse 25, 64283 Darmstadt, Allemagne

virginie.ollier@satie.ens-cachan.fr

Résumé – Ce travail s'inscrit dans le contexte de la radio interférométrie où la calibration consiste à estimer les perturbations subies par le signal depuis son émission par une source jusqu'à sa réception par le réseau d'antennes. Pour tenir compte de la présence de valeurs aberrantes contribuant au bruit, nous modélisons ce dernier par un processus non Gaussien, un PAIS (processus aléatoire à invariance sphérique). Nous comparons l'algorithme itératif proposé basé sur le maximum de vraisemblance, pour lequel des expressions analytiques permettent d'estimer les paramètres inconnus, avec des techniques de calibration traditionnelles.

Abstract – In radio interferometry, calibration amounts to estimate all the perturbation effects introduced along the signal propagation path until reception by the antenna array. In order to take into account outliers, we choose to model the noise by a non-Gaussian process, specifically by a SIRP (spherically invariant random process). We compare the proposed iterative algorithm based on the maximum likelihood estimator, where closed-form expressions are obtained, with state-of-the art calibration techniques.

1 Introduction

Les nouveaux radio interféromètres tels que le LOFAR (low frequency array) [1] et le futur projet SKA (square kilometre array) [2] sondent le ciel à l'aide d'un nombre important d'antennes à très large champ de vision et très sensibles. Parmi les défis à relever, l'étape de calibration est essentielle afin de pouvoir reconstruire des images à haute résolution. Elle consiste à estimer les différentes perturbations introduites le long du trajet du signal, qu'elles soient d'origine environnementale ou instrumentale. Nous modélisons ces effets perturbateurs à l'aide de matrices dites de Jones couramment utilisées par la communauté en radioastronomie [3, 4].

La plupart des algorithmes de calibration utilisent la méthode des moindres carrés en minimisant une fonction coût alternativement par rapport à chacun des paramètres inconnus [5]. Afin d'améliorer le coût calculatoire et la qualité d'estimation, l'algorithme EM (expectation maximization) [6], ainsi que SAGE (space alternating generalized expectation maximization) [7], ont été proposés [8]. Cependant, l'hypothèse classique de bruit Gaussien est considérée et ne correspond pas toujours à la réalité. En effet, des erreurs de modèle dues à la présence de sources de très faible intensité ou des interférences causées par des événements ponctuels ou des brouillages artifi-

ciels peuvent créer de manière aléatoire des valeurs aberrantes dans les données. À notre connaissance, seul l'algorithme proposé dans [9] traite le cas de la calibration robuste, en modélisant le bruit par une loi de Student. Ici, nous proposons d'avoir recours à une plus large gamme de distributions, qui inclut entre autres la loi de Student. Plus précisément, un VAIS (vecteur aléatoire à invariance sphérique) [10] s'écrit comme le produit d'une variable aléatoire positive appelée texture et d'un procédé aléatoire Gaussien, le speckle. En s'appuyant sur ce modèle, nous dérivons alors le maximum de vraisemblance, afin de pouvoir estimer les paramètres inconnus. Pour réduire les temps de calcul, la méthode proposée est itérative et permet d'obtenir directement des expressions analytiques pour l'estimation du bruit (texture et speckle), mais aussi pour les paramètres d'intérêt (matrices de Jones) à l'aide de l'algorithme BCD (block coordinate descent) [11].

Nous précisons enfin les notations utilisées dans le papier : les opérateurs transposé, conjugué et transposé conjugué sont respectivement indiqués par les symboles $(\cdot)^T$, $(\cdot)^*$ et $(\cdot)^H$. Nous représentons le produit de Kronecker par \otimes , l'espérance par $E\{\cdot\}$ et une matrice diagonale par blocs par $\text{bdiag}\{\cdot\}$. La trace et le déterminant sont respectivement décrits par $\text{tr}\{\cdot\}$ et $|\cdot|$. Enfin, l'opérateur $\text{vec}(\cdot)$ permet de transformer une matrice en vecteur en accumulant les colonnes les unes sur les autres.

Ces travaux sont financés par le projet MAGELLAN (ANR-14-CE23-0004-01) et par le projet ON FIRE (Jeunes Chercheurs GDR-ISIS).

2 Modèle de données

Nous considérons les signaux provenant de D sources et un réseau de M antennes. Chaque signal incident se décompose selon deux directions de polarisation orthogonales (x, y) . La relation entre le signal émis par la i -ème source et la tension mesurée par la p -ème antenne est la suivante

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

où $\mathbf{s}_i = [s_{ix}, s_{iy}]^T$, $\bar{\mathbf{v}}_{i,p}(\boldsymbol{\theta}) = [v_{i,px}(\boldsymbol{\theta}), v_{i,py}(\boldsymbol{\theta})]^T$ et $\mathbf{J}_{i,p}(\boldsymbol{\theta})$ est la matrice de Jones de taille 2×2 , paramétrée par le vecteur inconnu $\boldsymbol{\theta}$. Cette matrice modélise toutes les perturbations sur le trajet depuis la source i jusqu'à l'antenne p . Sachant qu'il existe une matrice de Jones différente pour chaque paire $i-p$, le nombre total de matrices à estimer est donc DM .

Les mesures acquises par un radio interféromètre sont les corrélations des signaux mesurés par chaque paire d'antennes, décrite chacune par une distance relative spécifique. Notons $B = \frac{M(M-1)}{2}$ le nombre total de paires d'antennes (p, q) de sorte que l'expression de la corrélation $\tilde{\mathbf{V}}_{pq}(\boldsymbol{\theta}) = \mathbb{E}\{\bar{\mathbf{v}}_p(\boldsymbol{\theta})\bar{\mathbf{v}}_q^H(\boldsymbol{\theta})\}$ s'écrit comme suit dans le cas non bruité

$$\tilde{\mathbf{V}}_{pq}(\boldsymbol{\theta}) = \sum_{i=1}^D \mathbf{J}_{i,p}(\boldsymbol{\theta})\mathbf{C}_i\mathbf{J}_{i,q}^H(\boldsymbol{\theta}) \text{ avec } p < q, p, q \in \{1, \dots, M\}, \quad (2)$$

où $\mathbf{C}_i = \mathbb{E}\{\mathbf{s}_i\mathbf{s}_i^H\}$ décrit l'état de polarisation de la i -ème source, supposé connu (source de calibration). De manière équivalente, nous pouvons considérer $\tilde{\mathbf{v}}_{pq}(\boldsymbol{\theta}) = \text{vec}(\tilde{\mathbf{V}}_{pq}(\boldsymbol{\theta})) = \sum_{i=1}^D \mathbf{u}_{i,pq}(\boldsymbol{\theta})$ où $\mathbf{u}_{i,pq}(\boldsymbol{\theta}) = (\mathbf{J}_{i,q}^*(\boldsymbol{\theta}) \otimes \mathbf{J}_{i,p}(\boldsymbol{\theta}))\mathbf{c}_i$ et $\mathbf{c}_i = \text{vec}(\mathbf{C}_i)$. Dans un contexte réaliste, il est nécessaire d'introduire un vecteur bruit \mathbf{n}_{pq} qui prend éventuellement en compte la présence de valeurs aberrantes. En radioastronomie, les valeurs aberrantes peuvent être dues à des phénomènes d'interférence ou encore à la présence de sources de très faible intensité. Ainsi, le vecteur corrélation s'écrit finalement $\mathbf{v}_{pq} = \tilde{\mathbf{v}}_{pq}(\boldsymbol{\theta}) + \mathbf{n}_{pq}$. Toutes les corrélations bruitées mesurées par l'interféromètre sont alors regroupées dans un vecteur $\mathbf{x} = [\mathbf{v}_{12}^T, \mathbf{v}_{13}^T, \dots, \mathbf{v}_{(M-1)M}^T]^T \in \mathbb{C}^{4B \times 1}$ tel que

$$\mathbf{x} = \sum_{i=1}^D \mathbf{u}_i(\boldsymbol{\theta}) + \mathbf{n} \quad (3)$$

où $\mathbf{u}_i(\boldsymbol{\theta}) = [\mathbf{u}_{i,12}^T(\boldsymbol{\theta}), \mathbf{u}_{i,13}^T(\boldsymbol{\theta}), \dots, \mathbf{u}_{i,(M-1)M}^T(\boldsymbol{\theta})]^T$ et $\mathbf{n} = [\mathbf{n}_{12}^T, \mathbf{n}_{13}^T, \dots, \mathbf{n}_{(M-1)M}^T]^T$.

3 Estimation par le maximum de vraisemblance

La présence de sources de très faible intensité, autres que les D sources de calibration supposées connues, est une des raisons pour lesquelles le bruit ne peut tout simplement pas être considéré comme Gaussien. Afin de proposer une calibration

robuste, nous considérons un modèle sous forme de PAIS [10], défini comme suit pour chaque paire d'antennes

$$\mathbf{n}_{pq} = \sqrt{\tau_{pq}} \mathbf{g}_{pq}, \quad (4)$$

où le facteur de puissance réel τ_{pq} est positif et aléatoire alors que le vecteur \mathbf{g}_{pq} suit une loi gaussienne de moyenne nulle, i.e., $\mathbf{g}_{pq} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Omega})$. Notons qu'une contrainte, purement arbitraire, est requise sur la matrice de covariance $\boldsymbol{\Omega}$ afin d'éviter tout problème d'ambiguïté, ici, nous supposons que $\text{tr}\{\boldsymbol{\Omega}\} = 1$. Avec ce modèle du bruit, l'expression de la vraisemblance, pour des mesures indépendantes entre elles, est la suivante

$$f(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\Omega}) = \prod_{pq} \frac{1}{|\pi\tau_{pq}\boldsymbol{\Omega}|} \exp\left\{-\frac{1}{\tau_{pq}} \mathbf{a}_{pq}^H(\boldsymbol{\theta})\boldsymbol{\Omega}^{-1}\mathbf{a}_{pq}(\boldsymbol{\theta})\right\}, \quad (5)$$

où le vecteur $\boldsymbol{\tau} = [\tau_{12}, \tau_{13}, \dots, \tau_{(M-1)M}]^T$ regroupe toutes les réalisations de la texture et $\mathbf{a}_{pq}(\boldsymbol{\theta}) = \mathbf{v}_{pq} - \tilde{\mathbf{v}}_{pq}(\boldsymbol{\theta})$. Ainsi, la log-vraisemblance s'écrit

$$\begin{aligned} \log f(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\Omega}) &= -4B \log \pi - 4 \sum_{pq} \log \tau_{pq} \\ &- B \log |\boldsymbol{\Omega}| - \sum_{pq} \frac{1}{\tau_{pq}} \mathbf{a}_{pq}^H(\boldsymbol{\theta})\boldsymbol{\Omega}^{-1}\mathbf{a}_{pq}(\boldsymbol{\theta}). \end{aligned} \quad (6)$$

La méthode proposée consiste à estimer de manière séquentielle les paramètres $\boldsymbol{\theta}$, $\boldsymbol{\tau}$ et $\boldsymbol{\Omega}$ en maximisant la log-vraisemblance par rapport à chacun d'entre eux, tout en fixant les paramètres restants. En outre, nous supposons dans notre algorithme que les paramètres τ_{pq} sont inconnus et déterministes, afin d'éviter d'avoir à spécifier leur densité de probabilité, assurant ainsi plus de flexibilité et de robustesse. Nous adoptons la méthodologie suivante pour estimer alternativement les paramètres inconnus [12] :

1) Estimation de τ_{pq} : Dériver (6) par rapport à τ_{pq} et égaliser à 0 conduit à

$$\hat{\tau}_{pq} = \frac{1}{4} \mathbf{a}_{pq}^H(\boldsymbol{\theta})\boldsymbol{\Omega}^{-1}\mathbf{a}_{pq}(\boldsymbol{\theta}). \quad (7)$$

2) Estimation de $\boldsymbol{\Omega}$: De manière similaire, nous obtenons

$$\hat{\boldsymbol{\Omega}} = \frac{1}{B} \sum_{pq} \frac{1}{\tau_{pq}} \mathbf{a}_{pq}(\boldsymbol{\theta})\mathbf{a}_{pq}^H(\boldsymbol{\theta}). \quad (8)$$

Puisque nous adoptons une procédure itérative, nous insérons (7) dans (8), ce qui donne

$$\hat{\boldsymbol{\Omega}}^{h+1} = \frac{4}{B} \sum_{pq} \frac{\mathbf{a}_{pq}(\boldsymbol{\theta})\mathbf{a}_{pq}^H(\boldsymbol{\theta})}{\mathbf{a}_{pq}^H(\boldsymbol{\theta})(\hat{\boldsymbol{\Omega}}^h)^{-1}\mathbf{a}_{pq}(\boldsymbol{\theta})} \quad (9)$$

où h fait référence à la h -ème itération. Enfin, n'oublions pas d'imposer la contrainte spécifiée, i.e.,

$$\hat{\boldsymbol{\Omega}}^{h+1} = \frac{\hat{\boldsymbol{\Omega}}^{h+1}}{\text{tr}\{\hat{\boldsymbol{\Omega}}^{h+1}\}}. \quad (10)$$

3) Estimation de $\boldsymbol{\theta}$: Estimer $\boldsymbol{\theta}$ pour un $\boldsymbol{\tau}$ et un $\boldsymbol{\Omega}$ fixés revient à effectuer l'opération suivante

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\text{argmin}} \left\{ \sum_{pq} \frac{1}{\tau_{pq}} \mathbf{a}_{pq}^H(\boldsymbol{\theta})\boldsymbol{\Omega}^{-1}\mathbf{a}_{pq}(\boldsymbol{\theta}) \right\}. \quad (11)$$

L'expression de θ dépend de la structure des matrices de Jones, qui peuvent se décomposer de différentes façons [13]. Dans ce qui suit, nous considérons le cas particulier où les matrices de Jones sont non structurées, i.e., estimer θ revient à estimer les entrées de toutes les matrices de Jones, évitant ainsi d'éventuelles erreurs de modélisation.

4 Estimation de θ pour des matrices de Jones non structurées

Réaliser l'étape de minimisation en (11) peut se révéler particulièrement coûteux. Remarquons que le vecteur d'intérêt θ se décompose comme suit

$$\theta = [\theta_1^T, \dots, \theta_D^T]^T = [\theta_{1,1}^T, \dots, \theta_{1,M}^T, \dots, \theta_{D,1}^T, \dots, \theta_{D,M}^T]^T, \quad (12)$$

où le vecteur $\theta_{i,p} \in \mathbb{R}^{8 \times 1}$ paramétrise le trajet i - p , i.e., $\mathbf{J}_{i,p}(\theta) = \mathbf{J}_{i,p}(\theta_{i,p})$.

4.1 L'algorithme EM

L'algorithme EM permet de trouver les estimations des paramètres par le maximum de vraisemblance, à plus faible coût grâce à l'itération de deux étapes. Dans notre cas, l'étape E se résume à calculer l'espérance conditionnelle de données dites complètes sachant les données observées et les paramètres estimés. Puis l'étape M est une étape d'optimisation qui peut être réalisée numériquement, par exemple avec l'algorithme de Levenberg-Marquardt, ou à l'aide d'expressions analytiques si possible. Le problème (11) se décompose alors en plusieurs sous-problèmes de plus faible complexité. En effet, nous verrons par la suite que l'optimisation se fera selon $\theta_i \in \mathbb{C}^{4M \times 1}$ au lieu de $\theta \in \mathbb{C}^{4DM \times 1}$. La méthodologie de l'EM est la suivante :

1) **Étape E** : Nous introduisons le vecteur de données complètes $\mathbf{w} = [\mathbf{w}_1^T, \dots, \mathbf{w}_D^T]^T$, où $\mathbf{w}_i = \mathbf{u}_i(\theta_i) + \mathbf{n}_i$ tel que $\mathbf{x} = \sum_{i=1}^D \mathbf{w}_i$ et $\mathbf{n} = \sum_{i=1}^D \mathbf{n}_i$. Nous supposons les vecteurs bruit $\mathbf{n}_i \sim \mathcal{CN}(\mathbf{0}, \beta_i \Psi)$ statistiquement indépendants avec $\sum_{i=1}^D \beta_i = 1$ où Ψ est la matrice de covariance de \mathbf{n} telle que $\Psi = \text{bdiag}\{\tau_{12}\Omega, \dots, \tau_{(M-1)M}\Omega\}$. Nous pouvons directement obtenir l'expression de l'espérance conditionnelle caractéristique de l'étape E de l'EM, i.e.,

$$\begin{aligned} \hat{\mathbf{w}}_i &= \mathbb{E}\{\mathbf{w}_i | \mathbf{x}; \theta, \tau, \Omega\} \\ &= \mathbf{u}_i(\theta_i) + \beta_i \left(\mathbf{x} - \sum_{l=1}^D \mathbf{u}_l(\theta_l) \right). \end{aligned} \quad (13)$$

2) **Étape M** : Une fois les \mathbf{w}_i estimés pour $i \in \{1, \dots, D\}$, le but est d'estimer les θ_i en maximisant la vraisemblance suivante pour des \mathbf{w}_i indépendants

$$\begin{aligned} f(\hat{\mathbf{w}} | \theta, \tau, \Omega) &= \prod_{i=1}^D \frac{1}{|\pi \beta_i \Psi|} \\ &\exp \left\{ - \left(\hat{\mathbf{w}}_i - \mathbf{u}_i(\theta_i) \right)^H (\beta_i \Psi)^{-1} \left(\hat{\mathbf{w}}_i - \mathbf{u}_i(\theta_i) \right) \right\}. \end{aligned} \quad (14)$$

Pour la i -ème source, la fonction coût à minimiser est donc $\phi_i(\theta_i) = \left(\hat{\mathbf{w}}_i - \mathbf{u}_i(\theta_i) \right)^H (\beta_i \Psi)^{-1} \left(\hat{\mathbf{w}}_i - \mathbf{u}_i(\theta_i) \right)$. Afin de réduire le coût calculatoire de l'algorithme, nous proposons d'optimiser la fonction $\phi_i(\theta_i)$ par rapport à chaque $\theta_{i,p}$ de manière alternée grâce à l'algorithme BCD [11], ce qui conduit à des expressions analytiques.

4.2 L'algorithme BCD

Alternativement, la fonction $\phi_i(\theta_i)$ est minimisée par rapport à $\theta_{i,p} \in \mathbb{C}^{4 \times 1}$, tout en fixant les autres $\theta_{i,q}$ pour $q \neq p$. Par manque de place, nous donnons directement ci-dessous les résultats obtenus

$$\hat{\theta}_{i,p} = \begin{cases} (\Sigma_i^H \mathbf{A}_{i,p} \Sigma_i + \Upsilon_i^H \tilde{\mathbf{A}}_{i,p} \Upsilon_i)^{-1} (\Sigma_i^H \mathbf{A}_{i,p} \mathbf{w}_{i,p} + \Upsilon_i^H \tilde{\mathbf{A}}_{i,p} \tilde{\mathbf{w}}_{i,p}) & \text{pour } 1 < p < M \\ (\Sigma_i^H \mathbf{A}_{i,p} \Sigma_i)^{-1} \Sigma_i^H \mathbf{A}_{i,p} \mathbf{w}_{i,p} & \text{pour } p = 1 \\ (\Upsilon_i^H \tilde{\mathbf{A}}_{i,p} \Upsilon_i)^{-1} \Upsilon_i^H \tilde{\mathbf{A}}_{i,p} \tilde{\mathbf{w}}_{i,p} & \text{pour } p = M \end{cases} \quad (15)$$

où $\mathbf{w}_{i,p} = [\mathbf{w}_{i,p(p+1)}^T, \dots, \mathbf{w}_{i,pM}^T]^T$, $\tilde{\mathbf{w}}_{i,p} = [\mathbf{w}_{i,1p}^{*T}, \dots, \mathbf{w}_{i,(p-1)p}^{*T}]^T$, $\Sigma_i = [\Sigma_{i,p+1}^T, \dots, \Sigma_{i,M}^T]^T$, $\Upsilon_i = [\Upsilon_{i,1}^{*T}, \dots, \Upsilon_{i,p-1}^{*T}]^T$, $\mathbf{A}_{i,p} = \text{bdiag}\{\beta_i \tau_{p(p+1)} \Omega, \dots, \beta_i \tau_{pM} \Omega\}^{-1}$ et $\tilde{\mathbf{A}}_{i,p} = \text{bdiag}\{\beta_i \tau_{1p} \Omega^*, \dots, \beta_i \tau_{(p-1)p} \Omega^*\}^{-1}$. Nous avons également utilisé les notations suivantes :

$$\Sigma_{i,q} = \begin{bmatrix} \alpha_{i,q} & \beta_{i,q} & 0 & 0 \\ 0 & 0 & \alpha_{i,q} & \beta_{i,q} \\ \gamma_{i,q} & \rho_{i,q} & 0 & 0 \\ 0 & 0 & \gamma_{i,q} & \rho_{i,q} \end{bmatrix} \quad \Upsilon_{i,q} = \begin{bmatrix} \lambda_{i,q} & \mu_{i,q} & 0 & 0 \\ \nu_{i,q} & \xi_{i,q} & 0 & 0 \\ 0 & 0 & \lambda_{i,q} & \mu_{i,q} \\ 0 & 0 & \nu_{i,q} & \xi_{i,q} \end{bmatrix}$$

avec $\alpha_{i,q} = q_{i1}^* c_{i1} + q_{i2}^* c_{i3}$, $\beta_{i,q} = q_{i1}^* c_{i2} + q_{i2}^* c_{i4}$, $\gamma_{i,q} = q_{i3}^* c_{i1} + q_{i4}^* c_{i3}$, $\rho_{i,q} = q_{i3}^* c_{i2} + q_{i4}^* c_{i4}$, $\lambda_{i,q} = q_{i1} c_{i1} + q_{i2} c_{i2}$, $\mu_{i,q} = q_{i1} c_{i3} + q_{i2} c_{i4}$, $\nu_{i,q} = q_{i3} c_{i1} + q_{i4} c_{i2}$ et $\xi_{i,q} = q_{i3} c_{i3} + q_{i4} c_{i4}$. Nous notons $\mathbf{c}_i = [c_{i1}, c_{i2}, c_{i3}, c_{i4}]^T$, $\mathbf{J}_{i,p}(\theta_{i,p}) = \begin{bmatrix} p_{i1} & p_{i2} \\ p_{i3} & p_{i4} \end{bmatrix}$,

$\mathbf{J}_{i,q}(\theta_{i,q}) = \begin{bmatrix} q_{i1} & q_{i2} \\ q_{i3} & q_{i4} \end{bmatrix}$, i.e., $\theta_{i,p} = [p_{i1}, p_{i2}, p_{i3}, p_{i4}]^T$ et $\theta_{i,q} = [q_{i1}, q_{i2}, q_{i3}, q_{i4}]^T$. L'algorithme que nous proposons est nommé Algorithme 1 et décrit ci-après.

5 Simulations numériques

Nous souhaitons dans un premier temps évaluer les performances de l'Algorithme 1, lorsque le modèle de données suit effectivement un PAIS comme décrit en (4). Puis, nous comparons l'algorithme proposé avec l'approche exposée en [9] basée sur une loi de Student et le cas classique Gaussien [8] dans un contexte réaliste avec présence de valeurs aberrantes (D' sources de faible intensité).

Sur la figure 1(a), l'EQM (erreur quadratique moyenne) de la partie réelle d'un paramètre donné est tracée en fonction du RSB (rapport signal sur bruit), de même que la BCR (borne de Cramér-Rao), dont l'expression de l'inverse est donnée dans [14], notamment pour la loi de Student. Nous remarquons que l'EQM est proche de la BCR mais un léger décalage subsiste.

Algorithme 1 : Calibration robuste basée sur un PAIS

entrée : $D, M, B, \mathbf{C}_i, \beta_i, \mathbf{x}$
sortie : $\hat{\theta}$
initialisation : $\hat{\Omega} \leftarrow \Omega_{\text{init}}, \hat{\tau} \leftarrow \tau_{\text{init}}, \hat{\theta} \leftarrow \theta_{\text{init}}$

- 1 **tant que** critère d'arrêt non atteint **faire**
- 2 **tant que** critère d'arrêt non atteint **faire**
- 3 Étape E : $\hat{\mathbf{w}}_i$ donné par (13), $i = 1, \dots, D$
- 4 Étape M : $\hat{\theta}_i, i = 1, \dots, D$
- 5 **tant que** critère d'arrêt non atteint **faire**
- 6 $\hat{\theta}_{i,p}$ donné par (15), $p = 1, \dots, M$
- 7 **fin**
- 8 **fin**
- 9 $\hat{\Omega}$ obtenu par (9) puis (10),
- 10 $\hat{\tau}$ obtenu par (7)
- 11 **fin**

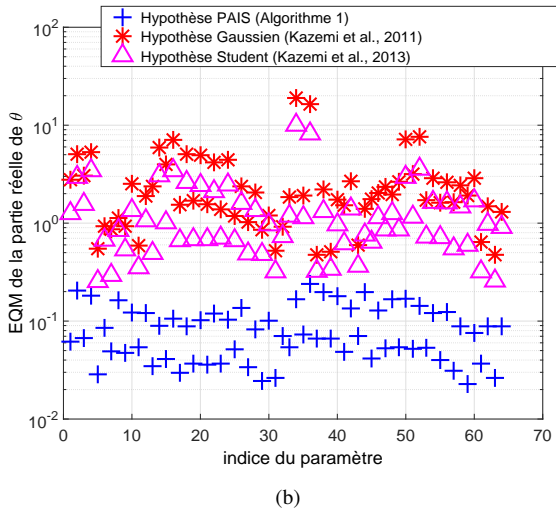
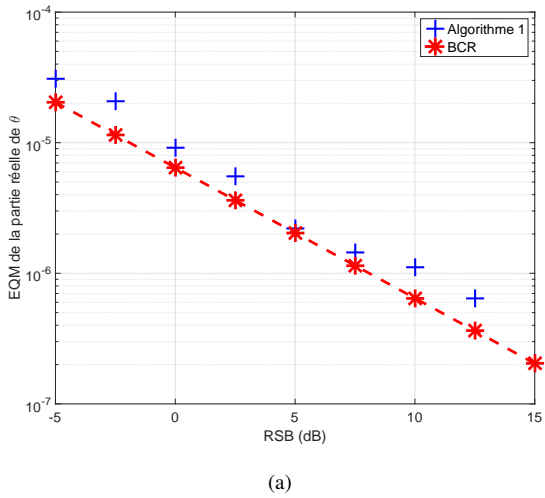


FIGURE 1 – (a) EQM d'un paramètre donné et BCR en fonction du RSB, (b) EQM de la partie réelle des 64 paramètres inconnus pour un RSB fixé, avec $D = 2$, $M = 8$ et $D' = 8$.

En effet, nous avons supposé des paramètres de texture déterministes et inconnus dans l'Algorithme 1 alors qu'il s'agit en réalité de variables aléatoires selon une loi inverse gamma dans les simulations.

Sur la figure 1(b), nous traçons l'EQM de la partie réelle de tous les paramètres inconnus, à RSB fixé pour i) l'Algorithme 1, ii), l'algorithme de calibration robuste proposé dans [9] et iii) le cas Gaussien basé sur la résolution d'un moindres carrés classique. Nous remarquons de meilleures performances d'estimation pour l'Algorithme 1. En effet, contrairement à [9], nous n'avons pas à préciser de distribution particulière grâce au PAIS et les entrées du vecteur bruit ne sont pas supposées indépendantes.

6 Conclusion

L'algorithme de calibration robuste proposé consiste à estimer de manière itérative chaque paramètre inconnu, en dérivant le maximum de vraisemblance basé sur un PAIS. Afin d'accélérer les calculs, nous avons introduit différents algorithmes permettant d'obtenir des expressions analytiques. En comparant l'algorithme nommé Algorithme 1 à d'autres techniques de calibration, il se révèle être davantage robuste.

Références

- [1] M. P. Van Haarlem *et al.*, "LOFAR : The LOW-Frequency ARray," *Astronomy & Astrophysics*, vol. 556, no. A2, 2013.
- [2] P. E. Dewdney, P. J. Hall, R. T. Schilizzi, and T. J. L. Lazio, "The Square Kilometre Array," *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1482–1496, 2009.
- [3] J. P. Hamaker, J. D. Bregman, and R. J. Sault, "Understanding radio polarimetry. I. Mathematical foundations," *Astronomy and Astrophysics Supplement Series*, vol. 117, no. 1, pp. 137–147, 1996.
- [4] O. M. Smirnov, "Revisiting the radio interferometer measurement equation. I. A full-sky Jones formalism," *Astronomy & Astrophysics*, vol. 527, no. A106, 2011.
- [5] S. J. Wijnholds, "Fish-eye observing with phased array radio telescopes," Ph.D. dissertation, Technische Universiteit Delft, Delft, The Netherlands, 2010.
- [6] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.
- [7] J. A. Fessler and A. O. Hero, "Space-alternating generalized expectation-maximization algorithm," *IEEE Transactions on Signal Processing*, vol. 42, no. 10, pp. 2664–2677, 1994.
- [8] S. Kazemi, S. Yatawatta, S. Zaroubi, P. Lampropoulos, A. de Bruyn, L. Koopmans, and J. Noordam, "Radio interferometric calibration using the SAGE algorithm," *Monthly Notices of the Royal Astronomical Society*, vol. 414, no. 2, pp. 1656–1666, 2011.
- [9] S. Kazemi and S. Yatawatta, "Robust radio interferometric calibration using the t-distribution," *Monthly Notices of the Royal Astronomical Society*, vol. 435, no. 1, pp. 597–605, 2013.
- [10] K. Yao, "Spherically invariant random processes : Theory and applications," in *Communications, Information and Network Security*, V. Bhargava *et al.*, Eds. Dordrecht, The Netherlands : Kluwer Academic : Springer, 2003, pp. 315–331.
- [11] M. Hong, M. Razaviyayn, Z.-Q. Luo, and J.-S. Pang, "A unified algorithmic framework for block-structured optimization involving big data : With applications in machine learning and signal processing," *IEEE Signal Processing Magazine*, vol. 33, no. 1, pp. 57–77, 2016.
- [12] X. Zhang, M. N. El Korso, and M. Pesavento, "Maximum likelihood and maximum a posteriori direction-of-arrival estimation in the presence of SIRP noise," in *International Conference on Acoustics, Speech and Signal Processing*, Shanghai, China, 2016.
- [13] J. E. Noordam, "The measurement equation of a generic radio telescope, AIPS++ implementation note nr 185," ASTRON, Tech. Rep., 1996.
- [14] O. Besson and Y. I. Abramovich, "On the Fisher information matrix for multivariate elliptically contoured distributions," *IEEE Signal Processing Letters*, vol. 20, no. 11, pp. 1130–1133, 2013.