

Mise en oeuvre efficace d'une radiogoniométrie 2D au sens du maximum de vraisemblance via un algorithme parcimonieux quasi Gauss-Newton

Thomas AUSSAGUÈS¹ Anne FERRÉOL^{1,2} Alice DELMER¹ Pascal LARZABAL¹

¹Thales, 4 avenue des Louvresses, 92230 Gennevilliers, France

²SATIE, Université Paris-Saclay, UMR CNRS 8029, 4 avenue des Sciences, 91190 Gif-sur-Yvette, France

Résumé – La mise en œuvre opérationnelle d'une radiogoniométrie au sens du Maximum de Vraisemblance (MV) est très lourde voire impossible car elle nécessite l'optimisation d'un critère non-convexe et hautement non-linéaire. Dans cette communication, l'équivalence entre le MV et les méthodes parcimonieuses, obtenue après une étape de pré-blanchiment, est exploitée afin de proposer une implémentation à moindres coûts du MV en 2D. Plus particulièrement, nous montrons que l'exploitation des effets du blanchiment accélère substantiellement la convergence de l'algorithme parcimonieux tout en rendant équivalentes les dernières itérations de celui-ci à celles d'une méthode du second-ordre dont le coût computationnel serait bien plus élevé.

Abstract – Maximum Likelihood Direction-of-Arrival (DoA) estimation is impractical due to the requirement of highly non-convex multi-dimensional optimization. This work exploits equivalence between ML and sparse DoA estimators, obtained after a pre-whitening step, to propose a low cost 2D ML implementation. Specifically, the pre-whitening transform effects are leveraged to significantly improve the algorithm's speed of convergence. Furthermore, the last iterations of the algorithm are shown to be equivalent with a more computationally demanding second-order algorithm.

1 Introduction

La radiogoniométrie est un problème classique de traitement d'antenne pour lequel de nombreuses méthodes d'estimation ont été proposées comme MUSIC [1]. Cependant, cet estimateur voit ses performances se dégrader dans des scénarios difficiles à faible Rapport Signal à Bruit (RSB) ou faible nombre d'échantillons. Le Maximum de Vraisemblance (MV) [2] n'est quant à lui pas appliqué en raison de la complexité de sa mise œuvre.

Certaines limites précédemment évoquées peuvent être levées en utilisant le Modèle de la Matrice de Covariance Vectorisée (MMCV) [3] qui repose sur l'utilisation d'un réseau d'antennes virtuelles offrant de meilleures performances. Les directions d'arrivées des sources peuvent alors être estimées au travers d'un estimateur parcimonieux. À la suite d'une étape de pré-blanchiment, transformant l'erreur d'estimation sur le MMCV, initialement non-blanche, en un bruit blanc gaussien, nous avons montré l'équivalence [4] entre les estimateurs parcimonieux et le MV appliqué au MMCV. L'estimateur parcimonieux des directions d'arrivée peut être formulé comme la minimisation d'un critère quadratique pénalisé par la pseudo-norme ℓ_0 . La minimisation est ensuite effectuée au travers de l'algorithme *Forward-Backward Splitting* (FBS) [5]. Néanmoins, l'algorithme FBS est une méthode du premier ordre dont la vitesse de convergence est limitée par la valeur du pas utilisé dans l'étape de descente de gradient.

Cette communication se propose d'exploiter les effets du blanchiment [6] par le biais d'un algorithme de descente à pas variable proposé par Ablin [7]. Nous montrons que le blanchiment permet de choisir des pas largement supérieurs accélérant ainsi la convergence. De plus, nous montrons que les dernières étapes de l'algorithme sont équivalentes à celles

d'un algorithme du second ordre dont l'implémentation est bien plus coûteuse.

2 Formulation du problème

2.1 Modèle de signal

On considère un réseau de N antennes sur lequel sont reçus les signaux de M sources de directions $\boldsymbol{\theta} = \{\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M\}$ avec en 2D $\boldsymbol{\theta}_m = \{\theta_m, \Delta_m\}$ le couple formé par les angles en azimuth θ_m et en élévation Δ_m de la m -ième source. La sortie du réseau est :

$$\mathbf{x}(t) = \sum_{m=1}^M \mathbf{a}(\boldsymbol{\theta}_m) s_m(t) + \mathbf{n}(t) = \mathbf{A}(\boldsymbol{\theta}) \mathbf{s}_\theta(t) + \mathbf{n}(t) \quad (1)$$

où $\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\boldsymbol{\theta}_1), \dots, \mathbf{a}(\boldsymbol{\theta}_M)] \in \mathbb{C}^{N \times M}$ est la matrice de mélange des sources, $\mathbf{s}_\theta(t) = [s_1(t), \dots, s_M(t)] \in \mathbb{C}^M$ le vecteur formés par les signaux $s_m(t)$ émis par les M sources et $\mathbf{n}(t) \in \mathbb{C}^N$ un bruit complexe blanc gaussien, indépendant de $\mathbf{s}_\theta(t)$ ie. $\mathbb{E}[\mathbf{s}_\theta^H(t) \mathbf{n}(t)] = 0$ avec $\mathbb{E}[\cdot]$ l'espérance mathématique et de covariance $\mathbb{E}[\mathbf{n}(t) \mathbf{n}^H(t)] = \sigma^2 \mathbf{I}_N$. Enfin, sans perte de généralité, nous utiliserons les coordonnées $\{u_m = \cos(\theta_m) \cos(\Delta_m), v_m = \sin(\theta_m) \cos(\Delta_m)\}$ dans ce qui suit pour localiser une source.

La matrice de covariance de (1) est alors :

$$\mathbf{R}_x = \mathbb{E}[\mathbf{x}(t) \mathbf{x}^H(t)] = \mathbf{A}(\boldsymbol{\theta}) \mathbf{R}_s \mathbf{A}^H(\boldsymbol{\theta}) + \sigma^2 \mathbf{I}_N \quad (2)$$

où $\mathbf{R}_s = \mathbb{E}[\mathbf{s}_\theta(t) \mathbf{s}_\theta^H(t)]$ est la matrice de covariance des sources. N'étant pas accessible, \mathbf{R}_x est remplacée par l'estimée suivante obtenue en supposant les échantillons $\mathbf{x}(t_k)$, $1 \leq$

$k \leq K$, indépendant et identiquement distribués :

$$\widehat{\mathbf{R}}_x = \frac{1}{K} \sum_{k=1}^K \mathbf{x}(t_k) \mathbf{x}^H(t_k) = \mathbf{R}_x + \Delta \mathbf{R}_x \quad (3)$$

avec $\Delta \mathbf{R}_x$ une erreur d'estimation suivant une loi de Wishart pouvant être approximée par une loi complexe gaussienne lorsque le nombre d'échantillons est suffisant. En supposant les sources temporellement décorréllées, l'observation issue du MMCV s'obtient en appliquant l'opérateur de vectorisation par colonne $\text{vec}(\cdot)$ à (2) [3] :

$$\begin{aligned} \mathbf{r} &= \text{vec} \left(\widehat{\mathbf{R}}_x - \sigma^2 \mathbf{I}_N \right) = \sum_{m=1}^M \mathbf{b}(\boldsymbol{\theta}_m) \gamma_m + \text{vec}(\Delta \mathbf{R}_x) \\ &= \mathbf{B}(\boldsymbol{\theta}) \boldsymbol{\gamma} + \boldsymbol{\delta} \end{aligned} \quad (4)$$

avec $\mathbf{B}(\boldsymbol{\theta})$ la matrice formée par les vecteurs directeurs du réseau virtuel $\mathbf{b}(\boldsymbol{\theta}_m) = \mathbf{a}^*(\boldsymbol{\theta}_m) \otimes \mathbf{a}(\boldsymbol{\theta}_m)$ avec \otimes le produit de Kronecker, $\boldsymbol{\gamma} = \text{diag}(\mathbf{R}_s) = [\gamma_1, \dots, \gamma_M]^T$ le vecteur des puissances des sources et $\boldsymbol{\delta} = \text{vec}(\Delta \mathbf{R}_x)$ le vecteur associé à l'erreur d'estimation vectorisée de loi complexe gaussienne $\mathcal{CN}(\mathbf{0}_{N^2 \times 1}, \boldsymbol{\Gamma})$ avec :

$$\boldsymbol{\Gamma} = \mathbb{E} \left[\boldsymbol{\delta} \boldsymbol{\delta}^H \right] = \frac{1}{K} (\mathbf{R}_x^T \otimes \mathbf{R}_x) \quad (5)$$

2.2 Pré-blanchiment

L'équivalence entre les estimateurs parcimonieux et le MV a été obtenue sous l'hypothèse d'un bruit blanc gaussien [4]. Ainsi, une étape de pré-blanchiment est nécessaire afin de traiter l'erreur d'estimation $\boldsymbol{\delta}$ non-blanche (4). Pour cela, l'observation \mathbf{r} (4) est pré-multipliée par la matrice de blanchiment $\widehat{\mathbf{W}} = \widehat{\boldsymbol{\Gamma}}^{-1/2} = \sqrt{K} \left(\widehat{\mathbf{R}}_x^{-T/2} \otimes \widehat{\mathbf{R}}_x^{-1/2} \right)$:

$$\mathbf{y} = \widehat{\mathbf{W}} \mathbf{r} = \widehat{\mathbf{W}} \mathbf{B}(\boldsymbol{\theta}) \boldsymbol{\gamma} + \widehat{\mathbf{W}} \boldsymbol{\delta} = \mathbf{B}_w(\boldsymbol{\theta}) \boldsymbol{\gamma} + \boldsymbol{\delta}_w \quad (6)$$

avec $\mathbf{B}_w(\boldsymbol{\theta}) = \widehat{\mathbf{W}} \mathbf{B}(\boldsymbol{\theta})$ le dictionnaire après blanchiment et $\boldsymbol{\delta}_w = \widehat{\mathbf{W}} \boldsymbol{\delta}$ l'erreur d'estimation blanchie telle que $\mathbb{E} \left[\boldsymbol{\delta}_w \boldsymbol{\delta}_w^H \right] = \mathbf{I}_{N^2}$.

2.3 Estimation parcimonieuse

Sous l'hypothèse que les directions des sources $\boldsymbol{\theta}$ soient contenues dans une grille $\boldsymbol{\varphi} = \{\boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_G\}$ de G directions avec $\boldsymbol{\varphi}_g = \{v_g, \nu_g\}$ la g -ième direction de la grille, une réécriture parcimonieuse de (6) peut être obtenue :

$$\mathbf{y} = \mathbf{B}_w(\boldsymbol{\varphi}) \boldsymbol{\gamma}_0 + \boldsymbol{\delta}_w \quad (7)$$

où $\mathbf{B}_w(\boldsymbol{\varphi}) = \widehat{\mathbf{W}} \mathbf{B}(\boldsymbol{\varphi})$ avec $\mathbf{B}(\boldsymbol{\varphi}) = [\mathbf{b}(\boldsymbol{\varphi}_1), \dots, \mathbf{b}(\boldsymbol{\varphi}_G)]$ est un dictionnaire de taille $N^2 \times G$, $G \gg N^2$ et $\boldsymbol{\gamma}_0$ un vecteur de parcimonie M dont les M composantes non-nulles sont associées aux directions des sources. Ainsi, l'estimation des directions repose sur celle de $\boldsymbol{\gamma}_0$. En utilisant l'*a priori* parcimonieux sur $\boldsymbol{\gamma}_0$, le problème d'estimation est formulé comme la minimisation du critère suivant :

$$\min_{\boldsymbol{\gamma} \in \mathbb{C}^G} \{ \mathcal{J}_{\ell_0}(\lambda, \boldsymbol{\gamma}) = \mathcal{J}_{LS}(\boldsymbol{\gamma}) + \lambda \|\boldsymbol{\gamma}\|_0 \} \quad (8)$$

avec $\mathcal{J}_{LS}(\boldsymbol{\gamma}) = \frac{1}{2} \|\mathbf{y} - \mathbf{B}_w(\boldsymbol{\varphi}) \boldsymbol{\gamma}\|_2^2$ et $\lambda > 0$ un hyperparamètre équilibrant la fidélité aux données par rapport à la parcimonie. Dans ce qui suit, λ est réglé selon [4] afin d'obtenir l'équivalence avec le MV *ie.* les critères MV et \mathcal{J}_{ℓ_0} possèdent le même minimum global.

3 Accélération de l'algorithme FBS par le blanchiment

3.1 Algorithme FBS accéléré

Le problème d'optimisation (8) peut être résolu par l'algorithme *Forward-Backward Splitting* (FBS) [5] décrit par Alg. 1 avec $\text{prox}_{\ell_0^{\lambda\beta}}(\cdot)$ l'opérateur proximal de la pseudo-norme ℓ_0 ,

Alg. 1 : Algorithme FBS [5]

```

1  $i = 0$ 
2 tant que critère d'arrêt pas satisfait faire
3    $\nabla \mathcal{J}_{LS}(\boldsymbol{\gamma}^{(i)}) = \mathbf{B}_w^H(\boldsymbol{\varphi})(\mathbf{B}_w(\boldsymbol{\varphi})\boldsymbol{\gamma}^{(i)} - \mathbf{y})$  (9)
4    $\tilde{\boldsymbol{\gamma}}^{(i+1)} = \boldsymbol{\gamma}^{(i)} - \beta \nabla \mathcal{J}_{LS}(\boldsymbol{\gamma}^{(i)})$  (10)
5    $\boldsymbol{\gamma}^{(i+1)} = \text{prox}_{\ell_0^{\lambda\beta}}(\tilde{\boldsymbol{\gamma}}^{(i+1)})$  (11)
6    $i \leftarrow i + 1$ 
7 fin

```

$\nabla \mathcal{J}_{LS}$ le gradient et β le pas de descente. La convergence de l'algorithme FBS vers un minimum local de \mathcal{J}_{ℓ_0} est assurée dès lors que le pas satisfait $\beta \leq 1/\mathcal{L}$ avec $\mathcal{L} = \lambda_{max}(\mathcal{H}_w)$ la plus grande valeur propre de la hessienne de \mathcal{J}_{LS} :

$$\mathcal{H}_w = \mathbf{B}_w^H(\boldsymbol{\varphi}) \mathbf{B}_w(\boldsymbol{\varphi}) \quad (12)$$

D'après [7, 8], connaissant $\boldsymbol{\gamma}^{(i)}$, l'étape de descente de gradient (10) de Alg. 1 peut se reformuler sous la forme de la minimisation d'un majorant $\mathcal{Q}_{\mathcal{L}}$ de la forme quadratique \mathcal{J}_{LS} . Le majorant $\mathcal{Q}_{\mathcal{L}}$ est obtenu en procédant à un développement limité au second ordre de \mathcal{J}_{LS} au voisinage de $\boldsymbol{\gamma}^{(i)}$:

$$\begin{aligned} \mathcal{J}_{LS}(\boldsymbol{\gamma}) &= \mathcal{J}_{LS}(\boldsymbol{\gamma}^{(i)}) + (\nabla \mathcal{J}_{LS}(\boldsymbol{\gamma}^{(i)}))^H (\boldsymbol{\gamma} - \boldsymbol{\gamma}^{(i)}) \\ &\quad + \frac{1}{2} (\boldsymbol{\gamma} - \boldsymbol{\gamma}^{(i)})^H \mathcal{H}_w (\boldsymbol{\gamma} - \boldsymbol{\gamma}^{(i)}) \end{aligned} \quad (13)$$

puis en majorant le terme quadratique de (13) :

$$\begin{aligned} \mathcal{J}_{LS}(\boldsymbol{\gamma}) &\leq \mathcal{Q}_{\mathcal{L}}(\boldsymbol{\gamma}) = \mathcal{J}_{LS}(\boldsymbol{\gamma}^{(i)}) + (\nabla \mathcal{J}_{LS}(\boldsymbol{\gamma}^{(i)}))^H (\boldsymbol{\gamma} - \boldsymbol{\gamma}^{(i)}) \\ &\quad + \frac{\mathcal{L}}{2} \|\boldsymbol{\gamma} - \boldsymbol{\gamma}^{(i)}\|_2^2 \end{aligned} \quad (14)$$

Enfin, en choisissant le pas $\beta = 1/\mathcal{L}$, $\tilde{\boldsymbol{\gamma}}^{(i+1)}$ est obtenu en minimisant $\mathcal{Q}_{\mathcal{L}}$ ce qui donne lieu à (10). Remarquons que dans le cas particulier $\beta = 1/\mathcal{L} = 1/\lambda_{max}(\mathcal{H}_w)$, (10) correspond à une itération de l'algorithme de Gauss-Newton [2] appliqué au majorant $\mathcal{Q}_{\mathcal{L}}$ dont la hessienne est $\mathcal{L} \times \mathbf{I}_G$.

Ablin [7] a proposé une version accélérée de Alg. 1 en exploitant la parcimonie de $\boldsymbol{\gamma}^{(i)}$ à chaque itération. Dans [7], il est montré que le pas β peut être remplacé par un pas adaptatif $\beta_S^{(i)} = 1/\mathcal{L}_S^{(i)}$, calculé à chaque itération, en utilisant $\mathcal{L}_S^{(i)} = \lambda_{max}(\mathcal{H}_{w,S}^{(i)})$ la plus grande valeur propre de la hessienne $\mathcal{H}_{w,S}^{(i)} = \mathbf{B}_{w,S}^{(i)H}(\boldsymbol{\varphi}) \mathbf{B}_{w,S}^{(i)}(\boldsymbol{\varphi})$ (12) obtenue en considérant la restriction $\mathbf{B}_{w,S}^{(i)}(\boldsymbol{\varphi})$ du dictionnaire $\mathbf{B}_w(\boldsymbol{\varphi})$ aux colonnes dont

les indices appartiennent au support $\mathcal{S}^{(i)}$ de $\gamma^{(i)}$ à la i -ième itération que nous définissons comme suit :

$$\mathcal{S}^{(i)} = \text{supp}(\gamma^{(i)}) = \{g \mid \gamma_g^{(i)} \neq 0\}, i > 0 \quad (15)$$

$$\mathcal{S}^{(0)} = \{1, \dots, G\}, i = 0$$

Si la condition [7] :

$$\text{supp}(\gamma^{(i+1)}) \subset \mathcal{S}^{(i)} \quad (\diamond)$$

où $\gamma^{(i+1)}$ est obtenu en minimisant le majorant $\mathcal{Q}_{\mathcal{L}_S}^{(i)}$ calculé en utilisant $\mathcal{L}_S^{(i)}$ dans (13), est vérifiée alors $\beta_S^{(i)} = 1/\mathcal{L}_S^{(i)}$ est un pas valide et $\mathcal{Q}_{\mathcal{L}_S}^{(i)} \leq \mathcal{Q}_{\mathcal{L}}$.

Ainsi, la méthode [7] décrite par Alg. 2 minimise $\mathcal{Q}_{\mathcal{L}_S}^{(i)}$ qui est un meilleur majorant de (13) que $\mathcal{Q}_{\mathcal{L}}$ tout en bénéficiant d'un pas $\beta_S^{(i)} = 1/\mathcal{L}_S^{(i)}$ supérieur au pas initial $\beta = 1/\mathcal{L}$ permettant donc d'accélérer la convergence de Alg. 1 vers un minimiseur local de \mathcal{J}_{ℓ_0} . La FIGURE 1 présente les majorants

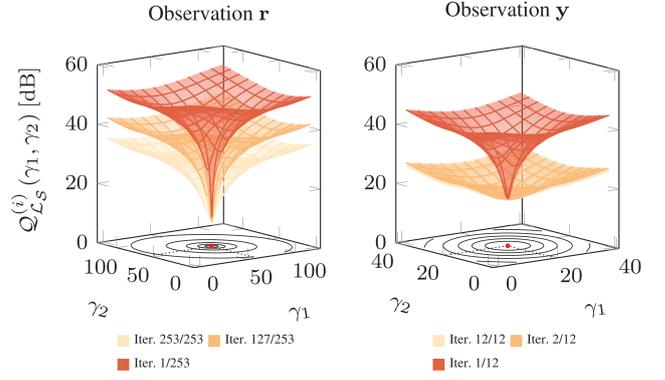


FIGURE 1 – Projections des majorants $\mathcal{Q}_{\mathcal{L}_S}^{(i)}$ dans les directions des sources avant et après blanchiment. Le minimum global du critère est représenté par un point rouge pour chaque critère. Pour cet exemple, $M = 2$ sources de directions $\{\theta_1, \theta_2\} = \{(0.5, -0.30), (-0.75, 0.29)\}$ donnant lieu à une corrélation spatiale $r_A^2(\theta_1, \theta_2) = 0.3$ sont considérées avec $K = 400$ échantillons, $\text{RSB} = 10$ dB ainsi que le réseau défini dans la section 4.

Alg. 2 : Algorithme FBS avec pas adaptatif [7]

```

1  i = 0
2  tant que critère d'arrêt pas satisfait faire
3  |   Calculer  $\mathcal{S}^{(i)}$  ainsi que  $\beta_S^{(i)}$ 
4  |   Obtenir  $\gamma_{temp}^{(i+1)}$  via (9), (10), (11) en utilisant  $\beta_S^{(i)}$ 
5  |   si  $\text{supp}(\gamma^{(i+1)}) \subset \mathcal{S}^{(i)}$  ( $\diamond$ ) alors
6  |   |    $\gamma^{(i+1)} \leftarrow \gamma_{temp}^{(i+1)}$ 
7  |   sinon
8  |   |   Obtenir  $\gamma^{(i+1)}$  via (9), (10), (11) en utilisant  $\beta$ 
9  |   fin
10 |   i ← i + 1
11 fin

```

$\mathcal{Q}_{\mathcal{L}_S}^{(i)}$ obtenus avant (observation r (4)) et après blanchiment (observation y (6)) pour plusieurs itérations de Alg. 2. Les majorations obtenues après blanchiment sont plus fines que celles obtenues sans cette transformation. De plus, remarquons qu'après blanchiment, les surfaces sont substantiellement plus plates. Ainsi, des pas plus grands peuvent être utilisés pour l'étape de descente de gradient (10) permettant donc d'accélérer l'algorithme.

3.2 Justification de l'accélération après blanchiment

Distance entre la surface \mathcal{J}_{LS} et son majorant $\mathcal{Q}_{\mathcal{L}_S}^{(i)}$ Dans cette partie, nous allons montrer comment le blanchiment donne lieu à des majorations plus fines de (13) et donc permet d'accélérer la convergence de Alg. 2. Plaçons nous dans le cas particulier de $M = 2$ sources et supposons que le support de $\gamma^{(i)}$, $\mathcal{S}^{(i)}$ (15), contienne uniquement les indices associés aux directions des sources *ie.* $\gamma = [\gamma_1 \ \gamma_2]^T$. Sous l'hypothèse que $\beta_S^{(i)}$ soit un pas valide [7] *ie* (\diamond) est satisfaite, nous avons $\text{supp}(\gamma) \subset \mathcal{S}^{(i)}$ (15). La distance $\varepsilon_1^{(i)}$ entre le terme quadratique intervenant dans (13) et son majorant est :

$$\begin{aligned} \varepsilon_1^{(i)} &= \left| \frac{\mathcal{L}_S^{(i)}}{2} \|\gamma - \gamma^{(i)}\|_2^2 - \frac{1}{2} (\gamma - \gamma^{(i)})^H \mathcal{H}_{w,S}^{(i)} (\gamma - \gamma^{(i)}) \right| \\ &= \frac{r_{\mathbf{B}_w}(\theta_1, \theta_2)}{1 + r_{\mathbf{B}_w}(\theta_1, \theta_2)} \times |\gamma_2 - \gamma_2^{(i)}|^2 \end{aligned} \quad (16)$$

en associant la source de direction θ_1 à l'indice 1 dans $\mathcal{S}^{(i)}$ ainsi qu'à la plus grande valeur propre de $\mathcal{H}_{w,S}^{(i)}$. Réciproquement, la source de direction θ_2 est reliée à l'indice 2 ainsi qu'à la plus faible valeur propre.

Preuve Une décomposition en valeurs propres de $\mathcal{H}_{w,S}^{(i)}$ suivie de l'utilisation des expressions des valeurs propres $\mathcal{H}_{w,S}^{(i)}$ dans le cas $M = 2$ sources donne (16). ■

La proposition (16) connecte $\varepsilon_1^{(i)}$ à la corrélation spatiale entre les sources $r_{\mathbf{B}_w}(\theta_1, \theta_2) = \frac{\mathbf{b}_w^H(\theta_1)\mathbf{b}_w(\theta_2)}{\|\mathbf{b}_w(\theta_1)\|_2\|\mathbf{b}_w(\theta_2)\|_2}$ calculée pour les vecteurs $\mathbf{b}_w(\theta_1), \mathbf{b}_w(\theta_2)$. À la suite du blanchiment $r_{\mathbf{B}_w}(\theta_1, \theta_2) \rightarrow 0$ d'où $\varepsilon_1^{(i)} \rightarrow 0$, ainsi, la surface majorante se rapproche donc du critère jusqu'à se confondre avec celui-ci. Avant blanchiment, la corrélation spatiale est importante, ce qui donne lieu à une plus grande distance $\varepsilon_1^{(i)}$.

Équivalence avec un algorithme du second ordre En plus du résultat (16), il est possible de montrer que les dernières itérations de Alg. 2 après blanchiment sont équivalentes à celles de l'algorithme de Gauss-Newton [2]. En reprenant les hypothèses du paragraphe 3.2, une itération de l'algorithme de Gauss-Newton s'écrit :

$$\gamma^{(i+1)} = \gamma^{(i)} - (\mathcal{H}_{w,S}^{(i)})^{-1} \nabla \mathcal{J}_{LS,S}^{(i)}(\gamma^{(i)}) \quad (17)$$

avec $\nabla \mathcal{J}_{LS,S}^{(i)}$ le gradient et $\mathcal{H}_{w,S}^{(i)}$ la hessienne (12) obtenus en se restreignant aux indices associés aux directions des sources. Une itération de la partie gradient (9) de Alg. 2 est :

$$\gamma^{(i+1)} = \gamma^{(i)} - \frac{1}{\lambda_{\max}(\mathcal{H}_{w,S}^{(i)})} \mathbf{I}_2 \nabla \mathcal{J}_{LS,S}^{(i)}(\gamma^{(i)}) \quad (18)$$

Les éq. (18) et (17) sont similaires à la différence près que l'inverse de la hessienne est remplacée par $\frac{1}{\lambda_{\max}(\mathcal{H}_{w,S}^{(i)})} \mathbf{I}_2$ dans (18). Ainsi, l'équivalence entre les éq. (18) et (17) s'appuie sur la distance entre les matrices $\frac{1}{\lambda_{\max}(\mathcal{H}_{w,S}^{(i)})} \mathbf{I}_2$ et $(\mathcal{H}_{w,S}^{(i)})^{-1}$

au sens de la norme spectrale $\|\cdot\| = \lambda_{\max}(\cdot)$. Nous pouvons alors déduire le résultat suivant :

$$\varepsilon_2^{(i)} = \left\| \left(\mathcal{H}_{w,S}^{(i)} \right)^{-1} - \frac{1}{\lambda_{\max}(\mathcal{H}_{w,S}^{(i)})} \mathbf{I}_2 \right\| = \frac{2r_{\mathbf{B}_w}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)}{(1+r_{\mathbf{B}_w}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2))^2} \quad (19)$$

Preuve En effectuant un développement limité de $(\mathcal{H}_{w,S}^{(i)})^{-1}$ puis en injectant l'expression de $\lambda_{\max}(\mathcal{H}_{w,S}^{(i)})$ [9], (19) est obtenue. ■

Le blanchiment ayant pour effet d'orthogonaliser les vecteurs directeurs dans les directions des sources [6, 9] *ie.* $r_{\mathbf{B}_w}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \approx 0$, il vient que $\varepsilon_2^{(i)} \rightarrow 0$. Ainsi, les dernières itérations de l'algorithme après blanchiment sont équivalentes à celles de l'algorithme de Gauss-Newton [2].

4 Résultats numériques

4.1 Paramètres des simulations

Un réseau à 6 antennes dont 5 distribuées sur un cercle de rayon $r = 0.8\lambda_0$ avec λ_0 la longueur d'onde et une antenne centrale est utilisé. $M = 2$ sources de directions $\{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2\} = \{(0.5, -0.30), (-0.75, 0.29)\}$ donnant lieu à une corrélation spatiale $r_A^2(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = 0.3$ sont considérées avec $K = 400$ échantillons et RSB = 10 dB. Une grille oracle avec un pas de 0.1 en u et en v contenant les estimées du MV est employée afin de supprimer le biais lié à l'erreur de grille. La pénalité ℓ_0 est remplacée par la fonctionnelle Continuous Exact ℓ_0 penalty (CEL0) [10] qui préserve le minimiseur global de ℓ_0 tout en supprimant de nombreux minima locaux. λ est respectivement choisi via [11] pour l'observation non-blanchie \mathbf{r} (4) et selon [4] après blanchiment (observation \mathbf{y} (6)). Le pas $\beta_S^{(i)} = 1/\mathcal{L}_S^{(i)}$ de Alg. 2 est calculé en utilisant la méthode des puissances afin de réduire le coût calculatoire. Enfin, les algorithmes Alg. 1 et Alg. 2 sont initialisés avec $\boldsymbol{\gamma}^{(0)} = \mathbf{0}_{G \times 1}$.

4.2 Analyse numérique de la convergence

La FIGURE 2 présente les courbes de convergence obtenues avant et après blanchiment pour l'algorithme FBS standard Alg. 1 [5] et sa version accélérée Alg. 2 [7]. Les pas de l'étape de gradient (10) correspondants sont également représentés. Premièrement, remarquons que le blanchiment améliore légèrement la vitesse de convergence en réduisant le nombre d'itérations nécessaires pour atteindre le critère d'arrêt, $\|\boldsymbol{\gamma}^{(i+1)} - \boldsymbol{\gamma}^{(i)}\|_2 / \|\boldsymbol{\gamma}^{(i)}\|_2 \leq 10^{-6}$, de 1257 itérations à 759 itérations. L'utilisation de la version accélérée après blanchiment Alg. 2 permet une convergence extrêmement rapide et stable en seulement 12 itérations. Au contraire, avant blanchiment, l'algorithme accéléré est très instable et converge en 253 itérations. Le TABLEAU 1 confirme l'amélioration de la vitesse de convergence obtenue en combinant l'algorithme FBS accéléré au blanchiment. En effet, pour l'algorithme accéléré, seulement 12 itérations sont nécessaires pour converger au lieu des 118,5 requises sans blanchiment.

5 Conclusion

Cette communication propose d'exploiter un algorithme parcimonieux à pas variable afin de mettre en œuvre une radio-

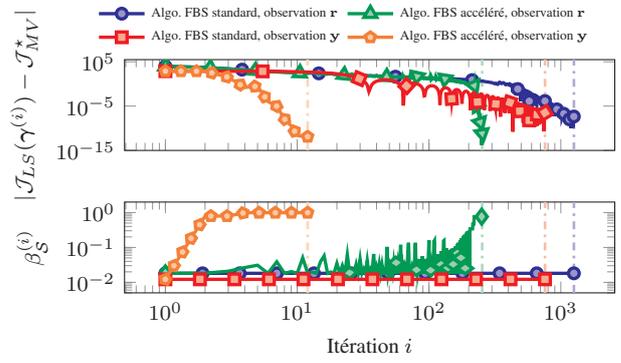


FIGURE 2 – Haut : Courbes de convergence avant et après blanchiment avec \mathcal{J}_{MV}^* le minimum global de J_{LS} . Bas : pas à chaque itération.

Algorithme	Observation \mathbf{r} (4)	Observation \mathbf{y} (6)
FBS standard [5]	1380,0	633,5
FBS accéléré [7]	118,5	11,3

TABLEAU 1 – Nombre moyen d'itérations sur 100 réalisations de bruit.

goniométrie 2D au sens du Maximum de Vraisemblance à moindres coûts. Plus particulièrement, nous montrons comment l'étape de blanchiment accélère significativement l'algorithme en rendant ses dernières itérations équivalentes à celles d'une méthode du second ordre dont l'implémentation est bien plus coûteuse. Des simulations numériques témoignent du gain en terme de vitesse de convergence.

Références

- [1] G. BIENVENU et L. KOPP : Optimality of high resolution array processing using the eigensystem approach. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 31(5):1235–1248, 1983.
- [2] B. OTTERSTEN, M. VIBERG, P. STOICA et A. NEHORAI : *Exact and Large Sample Maximum Likelihood Techniques for Parameter Estimation and Detection in Array Processing*, pages 99–151. Springer Berlin Heidelberg, Berlin, Heidelberg, 1993.
- [3] Joseph S. PICARD et Anthony J. WEISS : Direction finding of multiple emitters by spatial sparsity and linear programming. *In 2009 9th International Symposium on Communications and Information Technology*, pages 1258–1262, 2009.
- [4] Thomas AUSSAGUÈS, Anne FERRÉOL, Alice DELMER et Pascal LARZABAL : Looking for equivalence between maximum likelihood and sparse doa estimators. *In 2024 32th European Signal Processing Conference (EUSIPCO)*, 2024.
- [5] Patrick L. COMBETTES et Jean-Christophe PESQUET : Proximal splitting methods in signal processing, 2010.
- [6] Thomas AUSSAGUÈS, Anne FERRÉOL, Alice DELMER et Pascal LARZABAL : Whitening effects for ml-doa estimation using a sparse representation of array covariance. *In ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2025.
- [7] Pierre ABLIN, Thomas MOREAU, Mathurin MASSIAS et Alexandre GRAMFORT : Learning step sizes for unfolded sparse coding. *In Advances in Neural Information Processing Systems*, volume 32, 2019.
- [8] Amir BECK et Marc TEOULLE : A fast iterative shrinkage-thresholding algorithm with application to wavelet-based image deblurring. *In 2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 693–696, 2009.
- [9] Thomas AUSSAGUÈS, Anne FERRÉOL, Alice DELMER et Pascal LARZABAL : MI-doa estimation using a sparse representation of array covariance with a non-standard noise. *Submitted to Signal Processing*, 2025.
- [10] Emmanuel SOUBIES, Laure BLANC-FÉRAUD et Gilles AUBERT : A Continuous Exact ℓ_0 penalty (CEL0) for least squares regularized problem. *SIAM Journal on Imaging Sciences*, 8(3):pp. 1607–1639 (33 pages), juillet 2015.
- [11] Alice DELMER, Anne FERRÉOL et Pascal LARZABAL : On regularization parameter for ℓ_0 -sparse covariance fitting based doa estimation. *In ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4552–4556, 2020.