

Quel paramètre de régularisation pour une radio-goniométrie parcimonieuse avec la norme ℓ_0 ?

Alice DELMER^{1,2}, Anne FERRÉOL^{1,2}, Pascal LARZABAL²

¹Thales

4 avenue des Louvresses, 92230 Gennevilliers Cedex, France

²SATIE, CNRS, Université Paris-Saclay

61 avenue du Président Wilson, 94235 Cachan Cedex, France

nom.prenom@satie.ens-cachan.fr

Résumé – Cet article traite de la sélection du paramètre de régularisation d'un critère parcimonieux régularisé par la norme ℓ_0 pour l'estimation de directions d'arrivée à haute résolution. Ce paramètre est habituellement ajusté de manière empirique. Dans ce travail, l'observation est le vecteur propre associé à la valeur propre maximale de la matrice de covariance. Le degré de parcimonie est supposé connu. Dans cet article, nous établissons un intervalle permettant d'obtenir une bonne estimation, tant en mono-observation qu'en multi-observations. Les résultats numériques supportent notre étude théorique.

Abstract – This paper deals with the selection of the regularization parameter of a sparse ℓ_0 -regularized criterion for high-resolution DOA (Direction-of-Arrival) estimation. This parameter is generally empirically obtained. In this paper, the observation is the eigenvector associated with the maximal eigenvalue of the covariance matrix. The number of sources is assumed known. We establish a range allowing a good estimation for single or multi-observations. Numerical results support our theoretical investigations.

1 Introduction

L'estimation de directions d'arrivée est un problème classique de traitement d'antennes. Les méthodes conventionnelles que sont la formation de voies, la méthode de Capon [1], ou les méthodes à sous-espaces telles MUSIC [2], ont des performances limitées dans le cas d'erreurs de modèle, en présence de sources cohérentes, où quand le nombre d'instant de mesure est trop faible.

L'intérêt pour les méthodes parcimonieuses d'estimation des directions d'arrivées s'est intensifié ces dernières années, avec pour objectif d'améliorer les performances pour les scénarios complexes tels que ceux décrits plus haut. Le cadre de l'estimation parcimonieuse est en totale adéquation avec l'utilisation, dans les systèmes opérationnels, d'une table de calibration, qui contient les enregistrements des réponses du réseau dans des directions du champ d'observation.

Dans cet article, nous nous focalisons sur la minimisation d'une fonction de régularisation définie comme la somme d'un critère quadratique d'attache aux données et d'une pénalité ℓ_0 multipliée par un paramètre de régularisation λ . La problématique principale dans ce cadre est le choix de ce paramètre λ , qui est traditionnellement ajusté de façon empirique. Cet article étudie le comportement numérique de l'algorithme en fonction de λ . Cela nous permettra de définir une stratégie de sélection du paramètre λ permettant d'implémenter des algorithmes parcimonieux plus performants que les méthodes à haute résolu-

tion (HR) dans des scénarios sévères.

2 Formulation du problème

2.1 Modélisation du signal

On considère un réseau d'antennes composé de N capteurs, sur lequel on reçoit des signaux $\tilde{s}_m(t)$ provenant de M sources. Le vecteur des observations complexes est donnée par :

$$\mathbf{z}(t) = [z_1(t), \dots, z_N(t)]^T = \sum_{m=1}^M \mathbf{a}(\tilde{\theta}_m) \tilde{s}_m(t) + \mathbf{n}(t), \quad (1)$$

où $\tilde{\theta}_m$ est la direction d'arrivée de la m -ième source, $\mathbf{a}(\tilde{\theta}_m) \in \mathbb{C}^N$ est la réponse du réseau à cette direction d'arrivée, qui vérifie $\|\mathbf{a}(\tilde{\theta}_m)\|_2^2 = N$ et $\mathbf{n}(t) \in \mathbb{C}^N$ est un bruit additif.

On note $\tilde{\Theta} = \{\tilde{\theta}_1, \dots, \tilde{\theta}_M\}$, et on définit la matrice de mélange des sources $\tilde{\mathbf{A}} \in \mathbb{C}^{N \times M}$ par :

$$\tilde{\mathbf{A}} = \mathbf{A}(\tilde{\Theta}) = [\mathbf{a}(\tilde{\theta}_1) \dots \mathbf{a}(\tilde{\theta}_M)]. \quad (2)$$

On a alors :

$$\mathbf{z}(t) = \tilde{\mathbf{A}} \tilde{\mathbf{s}}(t) + \mathbf{n}(t), \text{ avec } \tilde{\mathbf{s}}(t) = [\tilde{s}_1(t), \dots, \tilde{s}_M(t)]^T. \quad (3)$$

2.2 Représentation parcimonieuse du signal

Dans cet article, on suppose que les directions d'arrivée des sources appartiennent à une grille prédéfinie qui discrétise l'es-

pace de recherche. Cette grille a par exemple été utilisée pour mesurer la réponse du réseau, et obtenir ainsi une table de calibration. La matrice de mélange $\tilde{\mathbf{A}}$ est donc une sous-matrice de la table de calibration $\mathbf{A} \in \mathbb{C}^{N \times G}$, qui contient les réponses du réseau pour l'ensemble des directions $\Theta = \{\theta_1, \dots, \theta_G\}$, avec $G \gg M$:

$$\mathbf{A} = \mathbf{A}(\Theta) = [\mathbf{a}(\theta_1) \dots \mathbf{a}(\theta_G)]. \quad (4)$$

Les directions d'arrivée des sources vérifient $\tilde{\Theta} \subset \Theta$. On suppose $G \gg N$; le vecteur d'observation (1) peut alors se réécrire sous la forme sous-déterminée suivante :

$$\mathbf{z}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \quad (5)$$

où le vecteur $\mathbf{s}(t)$, de dimension $G \times 1$ est parcimonieux, avec M composantes non nulles correspondantes aux signaux $\tilde{s}_m(t)$: $s_{i_m}(t) = \tilde{s}_m(t)$. Dans le cadre de l'estimation parcimonieuse de directions d'arrivée, l'objectif est de déterminer le support $I_M = \{i_m, m = 1 \dots M\}$. Les colonnes de \mathbf{A} associées à ce support sont les réponses du réseau $\mathbf{a}(\tilde{\theta}_m)$, et indiquent donc les directions d'arrivée des sources $\{\tilde{\theta}_1, \dots, \tilde{\theta}_M\}$.

On se placera dans la suite dans un cadre plus général en considérant le modèle parcimonieux suivant :

$$\mathbf{y} = \mathbf{A}\mathbf{x}_0 + \mathbf{n}_0, \text{ avec } \mathbf{A}\mathbf{x}_0 = \tilde{\mathbf{A}}\tilde{\mathbf{x}}_0, \quad (6)$$

où le vecteur $\tilde{\mathbf{x}}_0$ de dimension $M \times 1$ contient les composantes non nulles du vecteur \mathbf{x}_0 . L'observation à un unique instant de mesure de (5) est directement de la forme du modèle général (6). Si l'on considère plusieurs instants de mesure, les vecteurs propres de la matrice de covariance associés à l'espace signal s'expriment également sous cette forme. Au lieu d'utiliser directement les observations mesurées à plusieurs instants, il est intéressant de travailler sur ces vecteurs propres pour réduire le temps de calcul [4]. Nous proposons d'utiliser uniquement le vecteur propre associé à la valeur propre maximale pour se rattacher directement au modèle (6).

2.3 Problèmes d'estimation parcimonieuse

On considère le modèle (6). On souhaite estimer le vecteur $\mathbf{x}_0 \in \mathbb{C}^G$ à partir de l'observation bruitée $\mathbf{y} \in \mathbb{C}^N$.

Le problème est sous-déterminé ($G \gg N$). Pour obtenir une unique solution, on exploite la parcimonie du vecteur \mathbf{x}_0 . Celle-ci est mesurée par la pseudo-norme ℓ_0 , définie par :

$$\|\mathbf{x}\|_0 = \text{Card} \{g \in \{1, \dots, G\} : x_g \neq 0\}, \quad (7)$$

où $\text{Card} \{\cdot\}$ est le cardinal.

Pour estimer \mathbf{x}_0 , on pose le problème de minimisation suivant :

$$\min_{\mathbf{x} \in \mathbb{C}^G} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2, \text{ s.c. } \|\mathbf{x}\|_0 \leq k, \quad (8)$$

où k représente le degré de parcimonie. Ce problème contraint est NP-difficile, et la recherche d'une solution optimale est nécessairement exhaustive.

On peut également formuler le problème sous une forme régularisée à l'aide d'une fonction coût $J : \mathbb{R} \times \mathbb{C}^G \rightarrow \mathbb{R}$:

$$\min_{\mathbf{x} \in \mathbb{C}^G} \left\{ J(\lambda, \mathbf{x}) = \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_0 \right\}, \quad (9)$$

où le paramètre de régularisation $\lambda > 0$ vise à équilibrer le poids relatif entre l'attache aux données $\frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2$ et la régularisation par la norme ℓ_0 . Le minimum dépend à la fois de λ et de \mathbf{x} . Pour un paramètre λ fixé, on notera J_λ la fonction :

$$J_\lambda : \begin{array}{ccc} \mathbb{C}^G & \longrightarrow & \mathbb{R} \\ \mathbf{x} & \longmapsto & J(\lambda, \mathbf{x}) \end{array} \quad (10)$$

L'algorithme proximal itératif de seuillage dur (IHT) [3] permet de trouver une valeur $\hat{\mathbf{x}}$ qui minimise localement la fonction J_λ . La difficulté réside dans la détermination d'une valeur de λ qui permet d'obtenir une bonne approximation de \mathbf{x}_0 . Si λ est très faible, la solution s'approche de celle des moindres carrés. Si λ est trop fort, la solution obtenue est le vecteur nul.

À notre connaissance, dans la littérature sur la radio-goniométrie, le paramètre de régularisation λ pour l'estimation de directions d'arrivée est déterminé de façon empirique. Dans cet article, nous effectuons une analyse théorique sur λ .

3 Sélection optimale du paramètre de régularisation

3.1 Étude théorique pour une observation

Notre approche est basée sur le travail [5]. Ce dernier établit, pour une seule observation, les liens entre les solutions du problème contraint (8) et du problème régularisé (9). La relation obtenue permet de déterminer un intervalle pour le paramètre de régularisation λ tel que les minimiseurs globaux de (9) soient les mêmes que le problème contraint par un degré de parcimonie M .

On définit la valeur optimale et l'ensemble des solutions de (8) par :

$$c_k = \inf \{ \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 \mid \mathbf{x} \in \mathbb{C}^G \text{ et } \|\mathbf{x}\|_0 \leq k \} \quad (11)$$

$$\hat{X}_c^k = \{ \mathbf{x} \in \mathbb{C}^G \text{ et } \|\mathbf{x}\|_0 \leq k \mid \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 = c_k \} \quad (12)$$

Notons qu'en présence de M sources, l'ensemble \hat{X}_c^M ne doit contenir qu'un seul élément afin de pouvoir estimer \mathbf{x}_0 sans ambiguïtés.

On définit ensuite la valeur optimale et l'ensemble des solutions de (9) pour un λ donné, par :

$$r_\lambda = \inf \{ J_\lambda(\mathbf{x}) \mid \mathbf{x} \in \mathbb{C}^G \} \quad (13)$$

$$\hat{X}_r^\lambda = \{ \mathbf{x} \in \mathbb{C}^G \mid J_\lambda(\mathbf{x}) = r_\lambda \} \quad (14)$$

On recherche les valeurs du paramètre λ tel que les solutions du problème contraint (8) et régularisé (9) coïncident, soit : $\hat{X}_r^\lambda = \hat{X}_c^k$.

Pour tout $k, \forall \hat{\mathbf{x}}_k \in \hat{X}_c^k$, on a :

$$J(\lambda, \hat{\mathbf{x}}_k) = \frac{1}{2} c_k + \lambda k. \quad (15)$$

Pour tout \mathbf{x}_k tel que $\|\mathbf{x}_k\|_0 \leq k$, on a $J_\lambda(\hat{\mathbf{x}}_k) \leq J_\lambda(\mathbf{x}_k)$.

On considère un exemple de détection de directions d'arrivée avec M sources. On appelle L le plus petit entier tel que c_L soit

nul : il vérifie $L \leq N$. Il est évident que la séquence c_k , $k = 1 \dots L$ vérifie $c_L \leq \dots \leq c_{M+1} \leq c_M \leq c_{M-1} \leq \dots \leq c_0$. Pour $\hat{\mathbf{x}}_M \in \hat{X}_c^M$, on souhaite trouver λ tel que $\forall k \neq M$ et $\hat{\mathbf{x}}_k \in \hat{X}_c^k$, $J(\lambda, \hat{\mathbf{x}}_M) \leq J(\lambda, \hat{\mathbf{x}}_k)$, soit :

$$\frac{1}{2}c_M + \lambda M < \frac{1}{2}c_k + \lambda k. \quad (16)$$

De (16), on déduit que λ doit vérifier $\lambda_M^- < \lambda < \lambda_M^+$, avec :

$$\lambda_M^+ = \min_k \left\{ \frac{c_k - c_M}{2(M-k)} \mid 0 \leq k < M \right\} \quad (17)$$

$$\lambda_M^- = \max_k \left\{ \frac{c_M - c_k}{2(k-M)} \mid M < k \leq L \right\} \quad (18)$$

La figure 1 représente un exemple typique des fonctions affines en le paramètre λ de $J(\lambda, \mathbf{x} = \hat{\mathbf{x}}_k)$, pour $k = 0, \dots, 3$. Dans ce travail, on pose $M = 2$; par conséquent, il existe un intervalle $[\lambda_2^-, \lambda_2^+]$ tel que

$$\forall \lambda \in [\lambda_2^-, \lambda_2^+], \min_{\mathbf{x}} J(\lambda, \hat{\mathbf{x}}_k) = J(\lambda, \hat{\mathbf{x}}_2). \quad (19)$$

On peut noter que le caractère affine des fonctions $J(\lambda, \mathbf{x})$ est

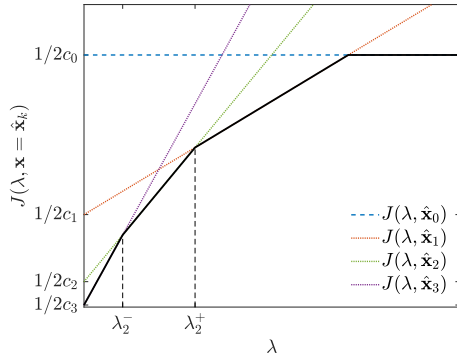


FIGURE 1 – Fonctions affines $J(\lambda, \mathbf{x})$ tracées en fonction de λ , pour $\mathbf{x} = \hat{\mathbf{x}}_k$, $k = 0 \dots 3$. λ_2^- et λ_2^+ délimitent l'intervalle où le minimum de $J(\lambda, \mathbf{x})$ (en noir) correspond à $J(\lambda, \hat{\mathbf{x}}_2)$.

exploité dans [6] pour proposer une autre approche de minimisation du problème régularisé par la norme ℓ_0 .

3.2 Stratégie de sélection de λ pour de multiples observations

Les paramètres λ_2^- et λ_2^+ dépendent des réalisations et sont des variables aléatoires. Comme il n'est pas envisageable de changer λ à chaque réalisation, il faut trouver un λ acceptable pour toutes les réalisations. Par conséquent, il s'agit de trouver une plage commune, avec une probabilité élevée, à de multiples réalisations. Selon les scénarios, il y aura une probabilité non nulle de recouvrement des statistiques de λ_2^- et λ_2^+ , comme on peut le voir sur la figure 2. S'il existe une telle plage, c'est qu'elle ne dépend pas des réalisations et pourra donc être calculée hors-ligne. Les statistiques de λ_2^- et λ_2^+ sont étudiées dans la partie numérique afin d'obtenir une plage commune qui pourra être utilisée.

4 Analyse numérique

On considère un réseau linéaire uniforme composé de $N = 10$ capteurs espacés d'une demi longueur d'onde. On souhaite estimer les directions d'arrivée de $M = 2$ sources provenant des vraies directions $\tilde{\theta}_1 = -6.9^\circ$ et $\tilde{\theta}_2 = 6.9^\circ$ qui appartiennent à la table de calibration, soit un écart de 1.2 largeur de lobe. Ces deux sources ont une corrélation temporelle de coefficient $c = 0.9$. On considère un bruit additif blanc gaussien, et on travaille à des rapports signal-sur-bruit (RSB) compris entre -20 et 0 dB. Il s'agit d'un scénario sévère pour la méthode MUSIC, dont on cherche à améliorer les performances. On a $K = 1000$ instants de mesure, et on utilise comme vecteur d'observation le vecteur propre associé à la valeur propre maximale.

Nous choisissons d'utiliser la pénalité CEL0 (*Continuous Exact ℓ_0*) récemment introduite [7], pour approximer la norme ℓ_0 . En effet leurs auteurs ont démontré, dans le cas d'un dictionnaire normalisé, que le critère CEL0 peut présenter moins de minima locaux que le critère régularisé par la norme ℓ_0 utilisé par IHT. Notre table de calibration vérifie cette hypothèse. Le critère CEL0 est minimisé par un algorithme proximal, appelé dans la suite FB-CEL0 (pour *Forward-Backward CEL0*).

4.1 Initialisation

Le vecteur propre observation est une combinaison linéaire des deux vecteurs directeurs. Notre objectif est de fournir une initialisation à l'aide d'une première estimation de cette combinaison linéaire. Pour cela, nous mettons en œuvre un filtre adapté $\mathbf{x}_{ini}(k) = \mathbf{w}(k)^H \mathbf{y}$ recherché sous la forme suivante :

$$\min_{\mathbf{w}} \mathbf{w}^H \hat{\mathbf{R}}_{yy} \mathbf{w} \quad \text{s.c.} \quad \mathbf{w}^H \mathbf{a}(\theta_k) = 1 \quad (20)$$

Le résultat est donnée par la formule $\mathbf{w}(k) = \frac{\hat{\mathbf{R}}_{yy}^{-1} \mathbf{a}(\theta_k)}{\mathbf{a}^H(\theta_k) \hat{\mathbf{R}}_{yy}^{-1} \mathbf{a}(\theta_k)}$. Ce vecteur \mathbf{x}_{ini} complexe sert d'initialisation pour l'algorithme FB-CEL0. Cette initialisation "éclairée" permet d'optimiser la probabilité d'obtenir le minimum global de CEL0.

4.2 Comportement et plage de régularisation pour une observation donnée

On effectue un tirage aléatoire à $RSB = -14dB$, pour lequel on calcule λ_2^- et λ_2^+ donnés par (17) (18). On visualise sur la figure 3 le pseudo-spectre obtenu par FB-CEL0 pour quatre valeurs significatives de λ . Celui-ci est comparé au pseudo-spectre de MUSIC. Pour $\lambda < \lambda_2^-$, le spectre de FB-CEL0 contient beaucoup de lobes secondaires, et est donc très ambigu. Pour $\lambda = \lambda_2^-$, il n'y a plus qu'un seul petit lobe secondaire. À $\lambda = \lambda_2^+$, les deux uniques lobes restants, correspondant aux deux directions d'arrivées, sont plus étroits. Au delà de λ_2^+ , il n'y a d'abord plus qu'un seul lobe, avant d'obtenir comme solution le vecteur nul.

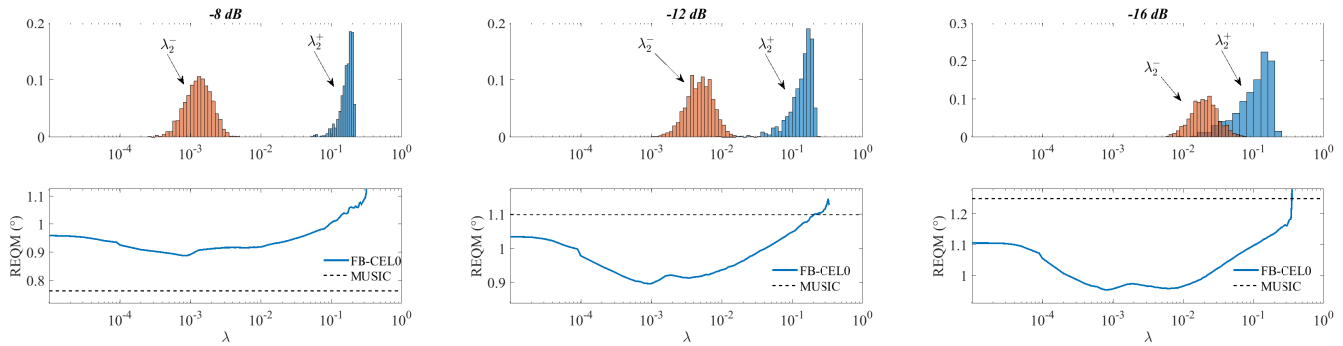


FIGURE 2 – Statistiques de λ_2^- et λ_2^+ (en haut) et racine de l’erreur quadratique moyenne (REQM) en fonction de λ (en bas) pour des rapports signaux-sur-bruit de -8dB (à gauche), -12dB (milieu), et -16dB (à droite).

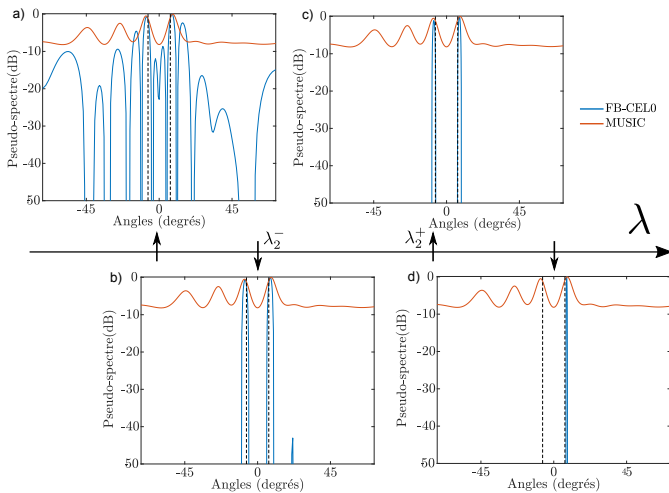


FIGURE 3 – Pseudo-spectres de FB-CELO obtenus selon le paramètre λ , pour a) $\lambda < \lambda_2^-$, b) $\lambda = \lambda_2^-$, c) $\lambda = \lambda_2^+$, d) $\lambda > \lambda_2^+$, comparés au pseudo-spectre de MUSIC. Les vraies directions d’arrivées sont indiquées par les traits verticaux en pointillés.

4.3 Analyse statistique dans le cas multi-observations

Les statistiques de λ_2^- et λ_2^+ sont représentées sur la figure 2 (en haut) pour différents RSB. Pour des RSB assez élevés, les statistiques ne se recouvrent pas, et on peut donc voir une grande plage de valeurs de λ qui permettent théoriquement d’obtenir la solution correctement régularisée. Pour des RSB plus faibles, les statistiques se rejoignent (-12dB), puis se chevauchent (-16dB). La racine des erreurs quadratiques moyennes (REQM) est calculée pour une plage de λ s’étendant de part et d’autre de l’intervalle théorique pour les trois mêmes scénarios. Pour des valeurs trop faibles de λ , la REQM tend vers une valeur limite équivalente à l’erreur des moindres carrés. Elle est ensuite à son plus bas pour des valeurs du paramètre correspondantes à des valeurs possibles de λ_2^- . La REQM augmente à nouveau à partir de valeurs correspondantes à λ_2^+ . Pour -12 et -16dB, qui sont ici les scénarios les plus sévères, on constate que FB-CELO fait toujours mieux que la méthode MUSIC pour peu que λ soit inférieure à la valeur maximale de λ_2^+ . De manière

générale, dès que λ dépasse cette valeur, l’algorithme parcimonieux dysfonctionne.

5 Conclusion

Dans cet article, nous nous sommes intéressés au réglage du paramètre de régularisation pour une radio-goniométrie parcimonieuse utilisant un critère régularisé avec la norme ℓ_0 . Nous avons donné un intervalle théorique pour ce paramètre λ pour une observation donnée, et vérifié la concordance avec les simulations. Nous avons également effectué une étude des statistiques de l’intervalle pour trouver une plage commune à de multiples observations, et mis en évidence l’existence d’une borne haute au dessus de laquelle l’algorithme échoue. Cette étude numérique permet de faire un choix hors-ligne pour le paramètre λ . Dans les scénarios draconiens, les réglages ainsi obtenus permettent d’obtenir de bien meilleurs résultats que la méthode MUSIC ; on peut noter que la méthode proposée fonctionne également dans le cas de sources totalement cohérentes. Dans un travail futur, on s’attachera à calculer analytiquement les statistiques pour éviter ensuite tout calcul combinatoire.

Références

- [1] J. Capon, “High-Resolution Frequency-Wavenumber Spectrum Analysis,” *Proceedings of the IEEE*, vol. 57, 1969.
- [2] G. Bienvenu and L. Kopp, “Optimality of High Resolution Array Processing Using the Eigensystem Approach,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1983.
- [3] T. Blumensath and M. E. Davies, “Normalized iterative hard thresholding : Guaranteed stability and performance,” *IEEE J. on Selected Topics in Signal Processing*, 2010.
- [4] D. Malioutov, M. Çetin, and A. S. Willsky, “A Sparse Signal Reconstruction Perspective for Source Localization With Sensor Arrays,” *IEEE Transactions on Signal Processing*, vol. 53, no. 8, 2005.
- [5] M. Nikolova, “Relationship between the optimal solutions of least squares regularized with ℓ_0 -norm and constrained by k -sparsity,” *Applied and Computational Harmonic Analysis*, vol. 41, pp. 237–265, 2016.
- [6] J. Duan, C. Soussen, D. Brie, and J. Idier, “A Continuation Approach to Estimate a Solution Path of Mixed L2-L0 Minimization Problems,” 2009.
- [7] E. Soubies, L. Blanc-Féraud, and G. Aubert, “A Continuous Exact L0 penalty (CEL0) for least squares regularized problem,” *SIAM J. Imaging Sciences*, vol. 8, 2015.