

Estimation modale de signaux 2-D par approximation parcimonieuse simultanée

Souleymen SAHNOUN, El-Hadi DJERMOUNE, David BRIE

CRAN, Nancy-Université, CNRS, UMR 7039
Boulevard des Aiguillettes, BP 70239, 54506 Vandoeuvre, France
Prenom.Nom@univ-lorraine.fr

Résumé – Nous proposons une méthode d’approximation parcimonieuse pour l’estimation modale 2-D de signaux de grande taille. L’idée consiste à décomposer ce problème d’estimation 2-D en deux problèmes 1-D et à exploiter le principe de parcimonie simultanée. Le couplage des modes est également formulé comme un problème de minimisation avec contrainte de parcimonie. L’efficacité de la méthode proposée est démontrée en utilisant des simulations numériques et des signaux expérimentaux de RMN.

Abstract – We propose a sparse modal estimation approach for analyzing large 2-D signals. The idea consists in decomposing the 2-D modal estimation problem into two 1-D problems and to exploit the simultaneous sparse principle. The mode pairing is also formulated as a sparse minimization problem. The effectiveness of the proposed method is demonstrated on simulated signals and experimental NMR data.

1 Introduction

Plusieurs méthodes ont été proposées pour traiter le problème d’estimation modale bidimensionnelle, notamment les approches de sous-espaces telles que TLS-Prony [1], *matrix enhancement and matrix pencil* (MEMP) [2], ESPRIT 2-D [3], *improved multidimensional folding* (IMDF) [4], etc. Parmi ces méthodes, celles qui estiment séparément les modes de chaque dimension et utilisent ensuite une procédure de couplage pour recombinaison des modes 2-D sont les plus efficaces en terme de coût de calcul. Une des limitations de ces approches réside dans la nécessité de sélectionner *a priori* le nombre correct de modes à estimer. Cette information n’est que rarement disponible dans des applications réelles telles que la spectroscopie RMN.

Récemment, des méthodes exploitant le principe d’approximation parcimonieuse ont été proposées pour résoudre le problème d’estimation harmonique ou modale. Elles sont pour la plupart destinées aux signaux 1-D et offrent l’avantage de ne pas nécessiter forcément la connaissance préalable du nombre de modes. En revanche, leurs performances dépendent d’une bonne sélection de la grille permettant la construction du dictionnaire [5]. Dans le cas de l’analyse modale, il est nécessaire de sélectionner une grille fine des fréquences et des facteurs d’amortissement menant à un dictionnaire d’une grande taille. Ce problème est encore amplifié dans le cas des signaux 2-D qui nécessite une grille 4-D. Nous avons récemment proposé une méthode multigrille d’estimation modale [6] où l’estimation démarre avec un dictionnaire de faible dimension (correspondant à une grille grossière) qui est ensuite affinée itérativement par l’insertion de nouveaux éléments dans le dictionnaire (raffinement de la grille). Pour les signaux 1-D,

cette méthode permet d’atteindre une bonne résolution tout en conservant une taille du dictionnaire raisonnable. Toutefois, pour les signaux 2-D de grande taille comme en RMN, le coût de calcul reste très important. Nous proposons ici une méthode d’estimation modale 2-D parcimonieuse qui permet de traiter les signaux de grande dimension. Elle repose sur deux étapes principales : dans la première, la méthode estime séparément les modes selon chaque dimension en utilisant le principe d’approximation parcimonieuse simultanée [7] combiné à l’approche multigrille. Dans la deuxième étape, les modes estimés sont couplés. La nouvelle procédure de couplage est également posée comme un problème d’approximation parcimonieuse où le dictionnaire 2-D est construit à partir des modes 1-D estimés et n’a donc que peu de colonnes. Le papier est organisé comme suit. Dans la section 2, nous présentons le modèle des données. Dans la section 3, nous développons le problème d’estimation modale 2-D et sa formulation en deux problèmes d’approximation parcimonieuse simultanée couplés à une procédure d’appariement des modes 2-D. L’efficacité de la méthode proposée est démontrée en utilisant des simulations numériques et des signaux RMN expérimentaux dans la section 4. Les conclusions et perspectives sont données dans la section 5.

2 Modèle du signal

Un signal RMN 2-D s’écrit comme la superposition de sinusoides amorties et d’un bruit additif :

$$\tilde{y}(m_1, m_2) = \sum_{i=1}^F c_i a_i^{m_1-1} b_i^{m_2-1} + e(m_1, m_2) \quad (1)$$

pour $m_1 = 1, \dots, M_1$ et $m_2 = 1, \dots, M_2$, où $a_i = e^{-\alpha_{a,i} + j2\pi f_{a,i}}$ sont les modes de la première dimension et $b_i = e^{-\alpha_{b,i} + j2\pi f_{b,i}}$ sont ceux de la deuxième dimension. Les $\{\alpha_{a,i}, \alpha_{b,i}\}_{i=1}^F$ sont les facteurs d'amortissement, $\{f_{a,i}, f_{b,i}\}_{i=1}^F$ sont les fréquences et $\{c_i\}_{i=1}^F$ sont les amplitudes complexes; $e(m_1, m_2)$ est un bruit additif. Le problème est d'estimer l'ensemble des paramètres $\{a_i, b_i, c_i\}_{i=1}^F$ à partir des données $\tilde{y}(m_1, m_2)$. La notation tilde (\sim) est utilisée pour représenter la version bruitée d'une donnée.

La matrice de données non bruitées \mathbf{Y} contenant les échantillons $y(m_1, m_2)$ s'écrit sous la forme :

$$\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{M_2}] \\ = \left[\sum_{i=1}^F c_i \mathbf{a}_i, \sum_{i=1}^F c_i b_i \mathbf{a}_i, \dots, \sum_{i=1}^F c_i b_i^{M_2-1} \mathbf{a}_i \right] \quad (2)$$

où $\mathbf{a}_i = [1, a_i, \dots, a_i^{M_1-1}]^T$. Les colonnes \mathbf{y}_{m_2} de \mathbf{Y} sont toutes des combinaisons linéaires des modes $a_i, i = 1, \dots, F$, mais ont des poids (amplitudes) différents. La matrice de données bruitées $\tilde{\mathbf{Y}}$ s'écrit alors :

$$\tilde{\mathbf{Y}} = \mathbf{A}\mathbf{H}_b + \mathbf{E} \quad (3)$$

où \mathbf{A} est une matrice de Vandermonde dont les colonnes sont $\{\mathbf{a}_i\}_{i=1}^F$, \mathbf{H}_b est une matrice de taille $F \times M_2$ définie par les éléments $h(i, m_2) = c_i b_i^{m_2-1}$ où l'indice b est introduit pour montrer que la matrice des poids dépend des modes b_i . \mathbf{E} est la matrice de bruit. De façon similaire, la transposée de la matrice $\tilde{\mathbf{Y}}^T$ s'écrit :

$$\tilde{\mathbf{Y}}^T = \mathbf{B}\mathbf{H}_a + \mathbf{E}^T \quad (4)$$

En fait, ces deux écritures résultent du modèle des données :

$$\tilde{\mathbf{Y}} = \mathbf{A}\mathbf{C}\mathbf{B}^T + \mathbf{E} \quad (5)$$

et de la façon dont la matrice $\mathbf{C} = \text{diag}(c_1, \dots, c_F)$ est associée à l'une des deux matrices \mathbf{A} ou \mathbf{B} . Notons que, pour réduire l'influence du bruit, $\tilde{\mathbf{Y}}$ peut être remplacée par son approximation de rang réduit [8].

3 Approximation parcimonieuse simultanée pour l'analyse modale 2-D

3.1 Estimation parcimonieuse simultanée

A partir de (2), il apparaît que les modes $\{a_i\}$ peuvent être obtenus à partir de toute colonne de \mathbf{Y} . On peut donc formuler le problème d'estimation modale, indépendamment pour chaque colonne $\tilde{\mathbf{y}}_{m_2}, m_2 = 1, \dots, M_2$ comme un problème d'approximation parcimonieuse :

$$\mathbf{x}_{m_2} = \min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{sous contrainte} \quad \|\tilde{\mathbf{y}}_{m_2} - \mathbf{Q}_a \mathbf{x}\|_2^2 \leq \epsilon \quad (6)$$

où \mathbf{Q}_a est un dictionnaire modal et \mathbf{x} est un vecteur parcimonieux contenant les coefficients des colonnes activées dans \mathbf{Q}_a . Le dictionnaire \mathbf{Q}_a est construit comme suit. Soit α_{\max} une limite supérieure sur $\{\alpha_{a,i}\}_{i=1}^F$ et soit P le nombre de points d'une grille uniforme couvrant l'intervalle de définition du facteur d'amortissement $[0, \alpha_{\max}]$. De même, soit K le nombre de

points d'une grille uniforme couvrant l'intervalle de fréquences normalisées $[0, 1]$. Alors \mathbf{Q}_a est donné par :

$$\mathbf{Q}_a = [\mathbf{q}(0, 0), \dots, \mathbf{q}(0, (K-1)\delta_f), \mathbf{q}(\delta_\alpha, 0), \dots, \\ \mathbf{q}(\delta_\alpha, (K-1)\delta_f), \dots, \mathbf{q}((P-1)\delta_\alpha, (K-1)\delta_f)] \quad (7)$$

où $\mathbf{q}(\alpha, f) = \mathbf{a}(\alpha, f) / \|\mathbf{a}(\alpha, f)\|_2$ avec $\mathbf{a}(\alpha, f) = [1, e^{(-\alpha + j2\pi f)}, \dots, e^{(-\alpha + j2\pi f)(M_1-1)}]^T$, $\delta_\alpha = \alpha_{\max}/P$, et $\delta_f = 1/K$. En résumé, \mathbf{Q}_a est obtenu par discrétisation du plan (α, f) . Chaque point de la grille correspond à un mode hypothétique. Le nombre total des colonnes de \mathbf{Q}_a est $N = KP \gg F$, chacune d'entre elles est appelée atome. Au lieu de traiter chaque colonne de \mathbf{Y} indépendamment, il est possible de les utiliser conjointement. En effet, chaque vecteur $\tilde{\mathbf{y}}_{m_2}$ correspond à un signal 1-D engendré par les mêmes modes mais avec des amplitudes différentes; les positions des coefficients non nuls doivent être les mêmes pour tous les vecteurs \mathbf{x}_{m_2} . Soit alors la matrice \mathbf{X} définie par

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{M_2}], \quad (8)$$

la parcimonie de \mathbf{X} peut être mesurée en calculant la norme euclidienne de chaque ligne; celles qui ont des normes non nulles définissent les positions des atomes du dictionnaire \mathbf{Q}_a (c'est-à-dire les modes a_i) activés simultanément dans toutes les colonnes de \mathbf{X} . Par conséquent, on se retrouve avec un problème d'approximation parcimonieuse simultanée :

$$\min_{\mathbf{X}} \|\mathbf{X}\|_{2,0} \quad \text{sous contrainte} \quad \|\tilde{\mathbf{Y}} - \mathbf{Q}_a \mathbf{X}\|_f^2 \leq \epsilon \quad (9)$$

où

$$\|\tilde{\mathbf{Y}} - \mathbf{Q}_a \mathbf{X}\|_f^2 = \|\text{vec}(\tilde{\mathbf{Y}} - \mathbf{Q}_a \mathbf{X})\|_2^2 \quad (10)$$

$$\|\mathbf{X}\|_{2,0} = \left\| \left[\|\mathbf{X}[1, :]\|_2 \quad \dots \quad \|\mathbf{X}[N, :]\|_2 \right]^T \right\|_0 \quad (11)$$

et $\mathbf{X}[n, :]$ représente la n^{e} ligne de \mathbf{X} . L'intérêt d'une approche simultanée est de prendre en compte dans la phase d'estimation des modes l'ensemble des données disponibles ce qui conduit à une diminution de la variance d'estimation. Plusieurs travaux ont déjà été présentés pour résoudre le problème d'approximation parcimonieuse simultanée de plusieurs signaux. Un panorama de ces algorithmes est présenté dans [9]. Dans ce papier nous utilisons la méthode S-OMP (*Simultaneous Orthogonal Matching Pursuit*) [7]. Une fois les modes correspondant à la première dimension (a_i) sont estimés, ceux de la seconde (b_k) sont obtenus de la même manière par traitement de la matrice $\tilde{\mathbf{Y}}^T$ et utilisation d'un dictionnaire \mathbf{Q}_b . Enfin, a_i et b_k doivent être couplés pour former les modes 2-D. C'est le point traité dans le prochain paragraphe. Cependant, afin d'atteindre une bonne résolution, nous proposons d'associer à l'approximation parcimonieuse simultanée, l'approche multigrille d'adaptation du dictionnaire [6]. Cette dernière consiste à démarrer l'estimation avec une grille grossière, qui est ensuite affinée itérativement par l'insertion de nouveaux points dans la grille menant à une procédure qui s'apparente aux techniques multirésolution.

3.2 Couplage des modes 2-D

Le modèle (1) du signal 2-D peut être écrit sous la forme vectorielle suivante :

$$\tilde{\mathbf{y}} = (\mathbf{A} \odot \mathbf{B})\mathbf{c} + \mathbf{e} \quad (12)$$

où $\tilde{\mathbf{y}} = \text{vec}(\tilde{\mathbf{Y}}^T)$, $\mathbf{A} \in \mathbb{C}^{M_1 \times F}$ et $\mathbf{B} \in \mathbb{C}^{M_2 \times F}$ sont deux matrices de Vandermonde dont les colonnes sont \mathbf{a}_i et \mathbf{b}_k , respectivement. Le vecteur $\mathbf{c} = [c_1, \dots, c_F]^T$ regroupe les amplitudes des modes 2-D et \odot représente le produit de Khatri-Rao. La méthode que nous proposons ici pour appairer les modes 1-D repose sur l'approche d'estimation modale parcimonieuse des signaux R-D proposée dans [6]. Soit F_a et F_b le nombre de modes \hat{a}_i et \hat{b}_k estimés. On définit les deux dictionnaires $\hat{\mathbf{Q}}_a$ et $\hat{\mathbf{Q}}_b$ tels que :

$$\hat{\mathbf{Q}}_a = [\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \dots, \hat{\mathbf{a}}_{F_a}], \quad \hat{\mathbf{Q}}_b = [\hat{\mathbf{b}}_1, \hat{\mathbf{b}}_2, \dots, \hat{\mathbf{b}}_{F_b}] \quad (13)$$

où $\hat{\mathbf{a}}_i = [1, \hat{a}_i, \dots, \hat{a}_i^{M_1-1}]^T$ pour $i = 1, \dots, F_a$ et $\hat{\mathbf{b}}_k = [1, \hat{b}_k, \dots, \hat{b}_k^{M_2-1}]^T$ pour $k = 1, \dots, F_b$. Ensuite, en utilisant le produit de Kronecker, on forme le dictionnaire $\hat{\mathbf{Q}}$ défini par :

$$\hat{\mathbf{Q}} = \hat{\mathbf{Q}}_a \otimes \hat{\mathbf{Q}}_b. \quad (14)$$

L'appariement des couples de modes 1-D consiste à sélectionner les $F \leq F_a \cdot F_b$ modes 2-D présents dans le signal $\tilde{\mathbf{y}}$ en résolvant le problème d'approximation parcimonieuse :

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \|\tilde{\mathbf{y}} - \hat{\mathbf{Q}}\mathbf{x}\|^2 \leq \epsilon. \quad (15)$$

Ce problème est résolu en utilisant une méthode glouton telle que OMP ou SBR [10]. Notons que cette méthode de couplage est efficace même en présence de modes alignés dans l'une ou l'autre des deux dimensions.

4 Algorithme et performances

4.1 Algorithme

L'algorithme complet peut être résumé comme suit :

1. Calculer la SVD de $\tilde{\mathbf{Y}}$ et prendre son approximation de rang réduit.
2. Appliquer l'algorithme multigrille couplé à la méthode S-OMP sur la matrice $\tilde{\mathbf{Y}}$ pour obtenir les modes a_i de la première dimension.
3. Refaire l'étape 2 en utilisant $\tilde{\mathbf{Y}}^T$ pour estimer les modes b_i de la deuxième dimension.
4. Former les modes 2-D en utilisant la procédure d'appariement.

4.2 Simulations

Les performances de la méthode proposée sont évaluées par des simulations numériques et comparées à celles obtenues avec ESPRIT 2-D, qui est l'une des méthodes d'estimation modale 2-D les plus efficaces. Le signal simulé de taille 30×30 échantillons contient trois modes :

$$\begin{aligned} (f_{1,1}, \alpha_{1,1}; f_{1,2}, \alpha_{1,2}) &= (0.164, 0.080; 0.342, 0.075), \\ (f_{2,1}, \alpha_{2,1}; f_{2,2}, \alpha_{2,2}) &= (0.310, 0.075; 0.110, 0.050), \\ (f_{3,1}, \alpha_{3,1}; f_{3,2}, \alpha_{3,2}) &= (0.412, 0.050; 0.090, 0.243). \end{aligned}$$

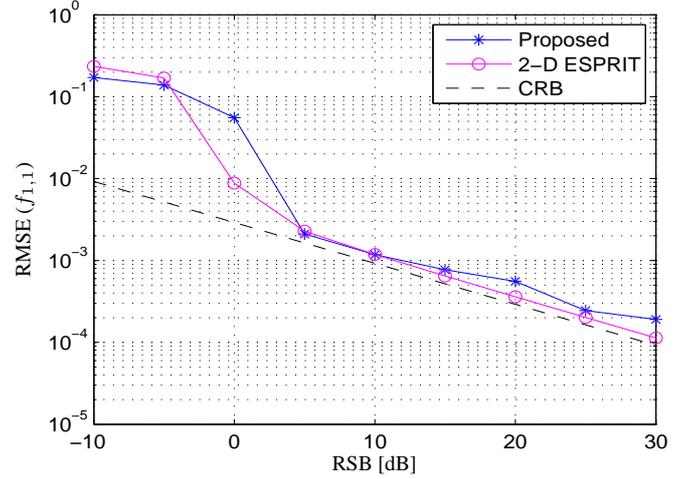


FIGURE 1 – Résultat de simulation : erreur type de $f_{1,1}$.

Toutes les amplitudes sont mises à 1. Nous effectuons des simulations de Monte Carlo avec différents niveaux de bruit. Pour la méthode proposée, la grille initiale utilisée pour construire le dictionnaire est la même pour les deux dimensions; elle contient 40 points en fréquence répartis uniformément sur l'intervalle $[0, 1]$, et 4 facteurs d'amortissement $\alpha \in \{0, 0.025, 0.05, 1\}$. On effectue 30 étages de résolution pour chaque estimation. L'erreur type d'estimation (RMSE) de la fréquence $f_{1,1}$ est représentée sur la figure 1. L'approche proposée donne des résultats très satisfaisants puisqu'elle atteint des performances d'estimation comparables à celles d'ESPRIT 2-D pour un coût de calcul similaire à celui de TLS-Prony 2-D, autorisant ainsi le traitement de signaux RMN 2-D de grande taille.

4.3 Analyse de signaux RMN 2-D

La spectroscopie RMN 2-D est une technique particulièrement bien adaptée à l'analyse de molécules ayant des structures complexes telles que celles qui composent les biomatériaux. Le signal analysé est de dimension 64×2024 . Avant d'appliquer l'approche multigrille combinée à l'algorithme S-OMP, on décompose le signal en sous-bandes selon la démarche présentée dans [11]. On utilise la même grille initiale et le même nombre d'étages de résolution que précédemment. Le nombre d'étages de résolution est fixé à 30. La figure 2 montre la décomposition finale en sous-bandes du signal RMN 2-D sur la plage spectrale $[-0.25, 0.25] \times [-0.25, 0.25]$ ainsi que le nombre de modes estimés dans chaque sous-bande par la méthode proposée. Les résultats obtenus dans deux sous-bandes sont montrés sur la figure 3 où les positions des modes estimés sont indiquées par des cercles. On constate que la méthode donne de bons résultats dans le cas de modes proches et/ou alignés. En outre, la méthode s'avère plus performante que TLS-Prony 2-D [12, 13] en terme de précision de l'estimation et comparable en terme de complexité numérique (temps de calcul).

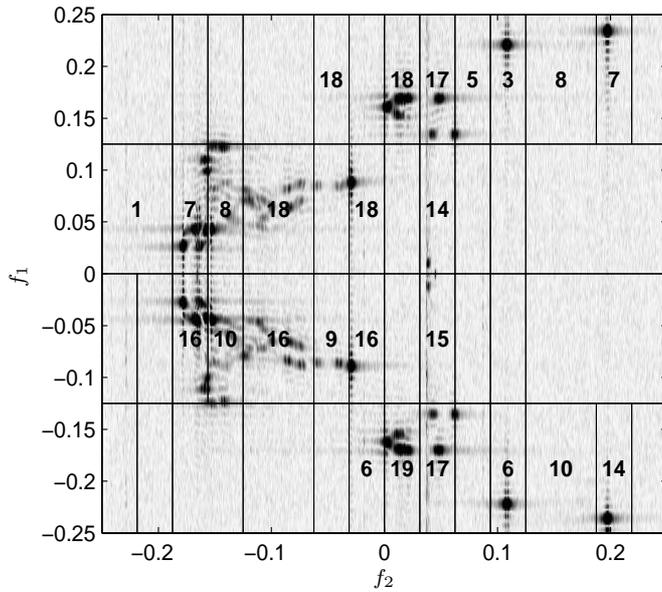


FIGURE 2 – Résultats obtenus par l’approche parcimonieuse dans la gamme spectrale $[-0.25, 0.25] \times [-0.25, 0.25]$ du signal RMN 2-D. Le nombre de modes estimés est indiqué pour chaque sous-bande.

5 Conclusion

Dans ce travail, nous avons proposé une approche d’approximation parcimonieuse pour l’estimation modale 2-D de signaux de grande dimension. L’idée consiste à exploiter le principe de parcimonie simultanée pour séparer ce problème d’estimation conjoint en deux problèmes 1-D. Les modes sont ensuite couplés par une nouvelle méthode, également basée sur la minimisation d’un critère parcimonieux. Les tests effectués sur des signaux RMN 2-D ont montré le potentiel des méthodes parcimonieuses multigrille. Comme perspective, cette méthode sera étendue au cas des signaux R -D ($R > 2$).

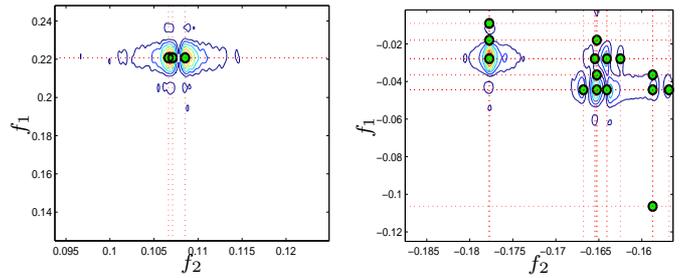
Références

[1] J. J. Sacchini, W. M. Steedly, and R. L. Moses. Two-dimensional Prony modeling and parameter estimation. *IEEE Trans. Signal Process.*, 41(11):3127–3137, 1993.

[2] Y. Hua. Estimating two-dimensional frequencies by matrix enhancement and matrix pencil. *IEEE Trans. Signal Process.*, 40(9):2267–2280, 1992.

[3] S. Rouquette and M. Najim. Estimation of frequencies and damping factors by two-dimensional ESPRIT type methods. *IEEE Trans. Signal Process.*, 49(1):237–245, 2001.

[4] J. Liu and X. Liu. An eigenvector-based approach for multidimensional frequency estimation with improved identifiability. *IEEE Trans. Signal Process.*, 54(12):4543–4556, décembre 2006.



(a) Sous-bande $(0.12, 0.24) \times (0.095, 0.12)$ (b) Sous-bande $(-0.12, 0.0) \times (-0.175, 0.155)$

FIGURE 3 – Résultats obtenus dans deux sous-bandes particulières du signal RMN. (o) Modes estimés. (· · ·) Coordonnés des modes.

[5] P. Stoica and P. Babu. Sparse estimation of spectral lines: grid selection problems and their solutions. *IEEE Trans. Signal Process.*, 60(2):962–967, 2012.

[6] S. Sahnoun, E.-H. Djermoune, C. Soussen, and D. Brie. Sparse multidimensional modal analysis using a multigrid dictionary refinement. *EURASIP J. Adv. Signal Process.*, March 2012.

[7] J.A. Tropp, A.C. Gilbert, and M.J. Strauss. Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit. *Signal Processing*, 86(3):572–588, 2006.

[8] D. Malioutov, M. Cetin, and A. S. Willsky. A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE Trans. Signal Process.*, 53(8):3010–3022, 2005.

[9] A. Rakotomamonjy. Surveying and comparing simultaneous sparse approximation (or group-Lasso) algorithms. *Signal processing*, 91(7):1505–1526, 2011.

[10] C. Soussen, J. Idier, D. Brie, and J. Duan. From Bernoulli-Gaussian deconvolution to sparse signal restoration. *IEEE Trans. Signal Process.*, 56(10):4572–4584, 2011.

[11] E. H. Djermoune, G. Kasalica, and D. Brie. Estimation of the parameters of two-dimensional NMR spectroscopy signals using an adapted subband decomposition. In *Proc. ICASSP*, pages 3641–3644, Las Vegas, USA, 2008.

[12] E. H. Djermoune, D. Brie, and M. Tomczak. A sub-band algorithm for estimating the parameters of two-dimensional exponential signals. In *Proc. European Signal Process. Conf.*, pages 1032–1036, Glasgow, Scotland, 2009.

[13] E. H. Djermoune, M. Tomczak, and D. Brie. NMR data analysis: A time-domain parametric approach using adaptive subband decomposition. *OGST – Rev. IFP Energies Nouvelles. Special issue on Advances in Signal Processing and Image Analysis for Physico-Chemical, Analytical Chemistry and Chemical Sensing*, 2013.