

Reconstruction d'impulsions de Dirac à partir de mesures linéaires bruitées

Laurent CONDAT

GIPSA-lab – CNRS, Université de Grenoble
11 rue des mathématiques, campus BP 46, 38402 St Martin d'Hères Cedex, France
Laurent.Condat@gipsa-lab.grenoble-inp.fr

Résumé – Le problème considéré est la reconstruction d'un nombre fini d'impulsions de Dirac de positions et amplitudes arbitraires, à partir d'échantillons passe-bas uniformes bruités. Il est montré que l'estimateur du maximum de vraisemblance des paramètres inconnus est obtenu en résolvant un problème d'approximation matricielle structurée de rang faible. Un nouvel algorithme itératif heuristique est proposé pour ce problème non convexe et réputé NP-difficile. En l'absence de convexité, nous ne disposons pas de preuve de convergence, mais l'algorithme converge en pratique vers une solution locale, et même vers la solution globale lorsque le niveau de bruit n'est pas trop élevé. Ainsi, on améliore la méthode classique de débruitage de Cadzow, pour essentiellement les mêmes facilité d'implémentation et rapidité.

Abstract – We consider the recovery of a finite stream of Dirac pulses, with arbitrary locations and amplitudes, from noisy lowpass-filtered uniform samples. It is shown that maximum-likelihood estimation of the unknown parameters amounts to solve a structured low rank matrix approximation problem. A new heuristic iterative optimization algorithm is proposed to solve this difficult, even believed NP-hard, problem. Although the algorithm comes, in absence of convexity, with no convergence proof, it converges in practice to a local solution, and even to the global solution of the problem, when the noise level is not too high. Thus, our method improves upon the classical Cadzow denoising method, for essentially the same implementation ease and speed.

1 Formulation du problème

La reconstruction de signaux appartenant à des espaces linéaires, comme les signaux à bande limitée ou les splines, a été étudiée en détail en théorie de l'échantillonnage. Récemment, le cadre de la reconstruction d'une fonction définie continûment, à partir d'échantillons discrets, a été étendu à une classe plus large de signaux ayant des propriétés de parcimonie, au-delà du paradigme classique enraciné dans les travaux de Shannon [1, 2, 3, 4]. Le problème emblématique de cette théorie en plein essor, auquel nous nous intéressons ici, est celui de la reconstruction d'une séquence finie d'impulsions de Dirac, à partir de mesures à bande limitée uniformes et bruitées [2, 5, 6].

Plus précisément, le signal inconnu s que nous voulons estimer consiste en une somme d'un nombre connu $K \geq 1$ d'impulsions dans un intervalle $[0, \tau]$, avec $\tau > 0$ un réel connu :

$$s(t) = \sum_{k=1}^K a_k \delta(t - t_k), \quad \forall t \in [0, \tau], \quad (1)$$

où $\delta(t)$ est la distribution de Dirac, les $\{t_k\}_{k=1}^K$ sont les positions inconnues et les $\{a_k\}_{k=1}^K$ sont les amplitudes réelles non nulles inconnues. Le but est d'estimer ces $2K$ paramètres inconnus, ce qui forme un problème d'estimation paramétrique non bayésien. L'information disponible sur s consiste seulement en un nombre $N \geq 1$ de mesures linéaires uniformes et bruitées, de la forme

$$v_n = \int_0^\tau s(t) \varphi\left(\frac{n\tau}{N} - t\right) dt + \varepsilon_n = \sum_{k=1}^K a_k \varphi\left(\frac{n\tau}{N} - t_k\right) + \varepsilon_n, \quad (2)$$

pour tout $n = 0, \dots, N - 1$, où $\varphi(t)$ est la fonction d'échantillonnage et les $\varepsilon_n \sim \mathcal{N}(0, \sigma^2)$ sont des réalisations aléatoires gaussiennes indépendantes. Il a été montré dans la littérature que la condition $N \geq 2K + 1$, supposée satisfaite par la suite, est nécessaire et suffisante pour que la reconstruction parfaite en l'absence de bruit soit possible [5, 7]. Pour simplifier les notations, nous supposons de plus que N est impair, de la forme $N = 2M + 1$. Pour ce travail, où l'accent est mis sur le traitement du bruit, nous nous concentrons sur le cas le plus simple où φ est la fonction de Dirichlet :

$$\varphi(t) = \frac{\sin(N\pi t/\tau)}{N \sin(\pi t/\tau)} = \frac{1}{N} \sum_{m=-M}^M e^{j2\pi m t/\tau}, \quad \forall t \in \mathbb{R}. \quad (3)$$

Autrement dit, le signal s est périodisé avant convolution avec un sinus cardinal, afin de le rendre à bande limitée, puis un échantillonnage ponctuel uniforme est appliqué. L'extension à d'autres fonctions à bande limitée φ peut être traitée sans difficulté [5].

Notons rapidement que le problème considéré a de multiples applications, de la communication ultra-large-bande [3] à la détection de signaux biomédicaux impulsionnels [5].

2 Reformulation de l'estimation du maximum de vraisemblance comme un problème matriciel

Définissons les coefficients de Fourier

$\hat{v}_m = \sum_{n=0}^{N-1} v_n e^{-j2\pi mn/N}$, pour $-M, \dots, M$, que l'on obtient en appliquant la transformée de Fourier discrète (TFD) à la séquence $\{v_n\}_{n=0}^{N-1}$. Puisque la TFD est unitaire, à une constante près, on montre aisément que l'estimation du maximum de vraisemblance des paramètres revient à résoudre le problème des moindres carrés non linéaire suivant :

$$\underset{\substack{\{t'_k\}_{k=1}^K \in [0, \tau]^K, \{a'_k\}_{k=1}^K \in \mathbb{R}^K}}{\text{minimiser}} \sum_{m=-M}^M \left| \hat{v}_m - \sum_{k=1}^K a'_k e^{-j2\pi m t'_k / \tau} \right|^2. \quad (4)$$

Ainsi, nous nous retrouvons avec un problème d'estimation spectrale [8], réputé très difficile, car la fonctionnelle à minimiser a une forme oscillante avec de nombreux minima locaux [9]. De nombreuses méthodes d'estimation spectrale existent, comme les méthodes du périodogramme, MUSIC, ou ESPRIT ; soit elles sont sous-optimales, en résolvant un autre problème plus simple que (4) ou en ajoutant des hypothèses de séparation minimale des $\{t_k\}$ ou de type $N \gg K$; soit elles ne trouvent que le minimum local de la fonction coût le plus proche d'une première estimation du résultat, dont la détermination est problématique. L'approche proposée, elle, s'affranchit de ces limitations. Elle consiste à *débruiter* les coefficients de Fourier $\{\hat{v}_m\}$ avant d'appliquer la méthode classique d'*annihilation* de De Prony [10].

Afin de comprendre la méthode classique d'annihilation, plaçons-nous temporairement dans le cas non bruité. Alors, la convolution de la séquence $\{\hat{v}_m\}_{m=-M}^M$ avec la séquence $\{h_k\}_{k=0}^K$ vaut identiquement zéro : $\sum_{k=0}^K h_k \hat{v}_{m-k} = 0$, pour tout $m = -M+K, \dots, M$, où le filtre annihilateur est défini, à une constante près, par $\sum_{k=0}^K h_k z^k = \prod_{k=1}^K (z - e^{j2\pi t_k / \tau})$. Cela se traduit sous forme matricielle par

$$\underbrace{\begin{pmatrix} \hat{v}_{-M+K} & \cdots & \hat{v}_{-M} \\ \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ \hat{v}_M & \cdots & \hat{v}_{M-K} \end{pmatrix}}_{\mathbf{T}_K} \begin{pmatrix} h_0 \\ \vdots \\ h_K \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \quad (5)$$

Fixons maintenant un entier P entre K et M et formons, à partir des valeurs $\{\hat{v}_m\}$, la matrice Toeplitz (ayant des valeurs constantes sur chacune de ses diagonales) \mathbf{T}_P , de taille $N - P \times P + 1$; \mathbf{T}_K est représentée en (5). Alors, l'existence d'un filtre annihilateur de taille $K + 1$ pour la séquence $\{\hat{v}_m\}_{m=-M}^M$ est complètement équivalente à la propriété que \mathbf{T}_P est de rang au plus K .

Ainsi, en revenant au cas où du bruit est présent dans les données, on peut réécrire (4) sous forme matricielle, comme un problème d'*approximation structurée de rang faible* (SLRA) :

Trouver $\tilde{\mathbf{T}}_P \in \arg \min_{\mathbf{T}' \in \mathbb{C}^{N-P \times P+1}} \|\mathbf{T}' - \mathbf{T}_P\|_w^2$
 tq \mathbf{T}' est Toeplitz et $\text{rang}(\mathbf{T}') \leq K$, (6)

où la norme de Frobenius pondérée au carré d'une matrice $\mathbf{A} = \{a_{i,j}\} \in \mathbb{C}^{N-P \times P+1}$ est $\|\mathbf{A}\|_w^2 = \sum_{i=1}^{N-P} \sum_{j=1}^{P+1} w_{i,j} |a_{i,j}|^2$, avec des poids $w_{i,j}$ égaux à l'inverse de la taille de la diagonale de la matrice passant par la position (i, j) (voir la formule dans la version étendue de cet article [11, eq. (16)]).

Ainsi, le problème (6) consiste à *débruiter* la matrice \mathbf{T}_P , ou de manière équivalente les données $\{v_n\}$. Une fois résolu ce problème, la procédure pour obtenir les paramètres estimés, décrite dans [2], est simplement basée sur la propriété d'annihilation décrite plus haut. Premièrement, on réarrange les éléments de la matrice Toeplitz débruitée $\tilde{\mathbf{T}}_P$ en une matrice Toeplitz $\tilde{\mathbf{T}}_K$ de taille $N - K \times K + 1$. Deuxièmement, on calcule le vecteur singulier à droite $\tilde{\mathbf{h}} = \{\tilde{h}_k\}_{k=0}^K$ de $\tilde{\mathbf{T}}_K$, correspondant à la valeur singulière 0 (en pratique, la valeur singulière la plus faible). Troisièmement, on calcule les racines $\{\tilde{z}_k\}_{k=1}^K$ du polynôme $\sum_{k=0}^K \tilde{h}_k z^k$. Les positions estimées $\{\tilde{t}_k\}_{k=1}^K$ sont alors $\tilde{t}_k = \frac{\tau}{2\pi} \arg_{[0, 2\pi[}(\tilde{z}_k)$. Enfin, les amplitudes estimées $\{\tilde{a}_k\}_{k=1}^K$ sont obtenues en résolvant le système linéaire $\tilde{\mathbf{U}}^H \tilde{\mathbf{U}} \tilde{\mathbf{a}} = \tilde{\mathbf{U}}^H \tilde{\mathbf{v}}$, où $\tilde{\mathbf{v}} = [\tilde{v}_{-M} \cdots \tilde{v}_M]^T$, \cdot^H indique la transposée hermitienne, et

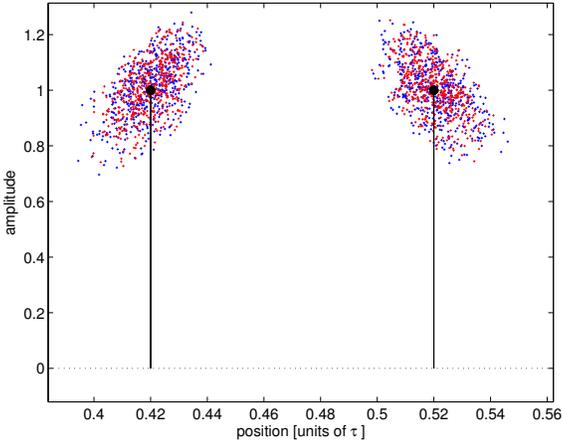
$$\tilde{\mathbf{U}} = \begin{pmatrix} e^{j2\pi M \tilde{t}_1 / \tau} & \cdots & e^{j2\pi M \tilde{t}_K / \tau} \\ \vdots & \ddots & \vdots \\ e^{-j2\pi M \tilde{t}_1 / \tau} & \cdots & e^{-j2\pi M \tilde{t}_K / \tau} \end{pmatrix}. \quad (7)$$

Notons qu'en l'absence de bruit, $\tilde{\mathbf{T}}_P = \mathbf{T}_P$ et les paramètres sont parfaitement estimés. Dans le cas bruité, l'avantage de la formulation matricielle (6) par rapport au problème (4) est qu'il n'y a pas de problème d'initialisation : un algorithme itératif pour résoudre (6) peut procéder avec la matrice bruitée \mathbf{T}_P comme estimée initiale de la solution débruitée $\tilde{\mathbf{T}}_P$.

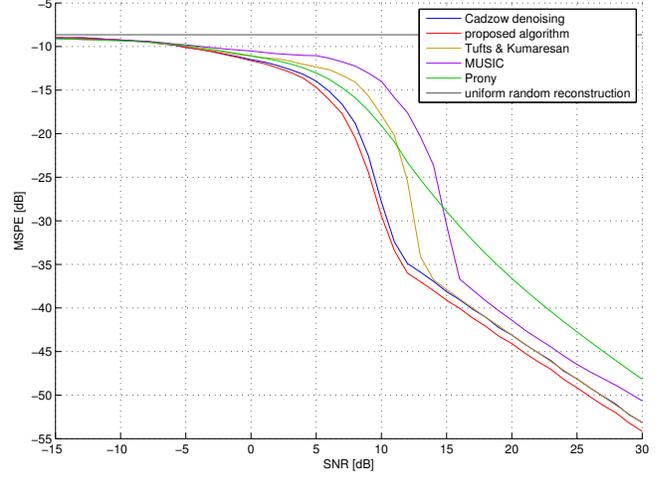
Il faut tout de même mentionner que la procédure ci-dessus ne délivre des estimées solution du problème (4) que si les racines $\{\tilde{z}_k\}_{k=1}^K$ sont distinctes et toutes sur le cercle complexe unité. Comme les matrices sont centro-hermitiennes, cela est le cas en pratique, excepté si le niveau de bruit est trop élevé. Dans ce cas, il peut arriver que deux racines forment une paire $(\tilde{z}_{k_1}, \tilde{z}_{k_2} = 1/\tilde{z}_{k_1}^*)$ des deux côtés du cercle unité, donnant deux positions $\tilde{t}_{k_1} = \tilde{t}_{k_2}$ identiques.

3 État de l'art pour les problèmes SLRA

Les problèmes SLRA, dont (6) est l'exemple typique, ont un large domaine d'application [13]. Ils sont par contre très difficile, et même supposés NP-difficiles. Ainsi, on devra se satisfaire d'une solution locale, c'est-à-dire une matrice Toeplitz de rang K qui soit un minimiseur local de la fonction coût $\|\cdot - \mathbf{T}_P\|_w^2$. La bonne nouvelle est que la matrice à approcher \mathbf{T}_P n'est pas n'importe quelle matrice, mais une version perturbée d'une matrice Toeplitz de rang K . Par conséquent, pour un niveau de bruit pas trop élevé, la solution locale trouvée par un algorithme sera bien la solution globale, comme on l'observe en pratique. Quelques algorithmes ont été proposés pour trouver une solution locale au problème (6) dans la communauté de l'algèbre numérique [14, 15] ; ceux-ci sont difficiles à implémenter et très coûteux. À notre connaissance, le code le plus performant, et disponible publiquement, est celui en cours de développement par Ivan Markovsky [16] ; cependant, ce code ne traite que des données réelles, alors que nos coefficients de Fourier sont complexes. Notons que l'idée de remplacer le rang par la norme nucléaire dans la formulation du problème (6), afin de le rendre convexe, ne semble pas donner de bons résultats, car deux



(a)



(b)

FIGURE 1 – Le signal à estimer, à partir de $N = 11$ mesures bruitées, consiste en $K = 2$ impulsions de Dirac. Les vraies valeurs sont $(t_1, t_2) = (0.42, 0.52)$ et $(a_1, a_2) = (1, 1)$, avec $\tau = 1$ et $P = M = 5$. (a) En noir, les vraies impulsions. En bleu et rouge, les positions et amplitudes reconstruites par débruitage de Cadzow et l’algorithme proposé, respectivement, pour 500 réalisations de bruit différentes. Le rapport signal-à-bruit (SNR) est de 15dB et le temps de calcul pour chaque reconstruction, avec 50 itérations, est de 14ms. (b) Erreur sur les positions $\min((\tilde{t}_1 - t_1)_\tau^2 + (\tilde{t}_2 - t_2)_\tau^2, (\tilde{t}_1 - t_2)_\tau^2 + (\tilde{t}_2 - t_1)_\tau^2)$, où $(x)_\tau = ((x + \frac{\tau}{2}) \bmod \tau) - \frac{\tau}{2}$, en fonction du SNR, en échelle log-log, et moyennée sur 10 000 réalisations de bruit pour chaque valeur de SNR. Les méthodes comparées sont MUSIC (fonction `pmusic` de Matlab), la méthode de Tufts et Kumaresan [12], la méthode de Prony sans débruitage ($\tilde{\mathbf{T}}_K = \mathbf{T}_K$). Une borne supérieure de l’erreur est donnée par l’estimateur naïf, qui choisit les positions aléatoirement et uniformément dans $[0, \tau[$.

impulsions arbitrairement proches peuvent donner des mesures hautement cohérentes [17]. Ainsi, une méthode heuristique, appelée *débruitage de Cadzow* [18], ou projections alternées, est devenue populaire pour sa simplicité et son efficacité dans la reconstruction d’impulsions [2, 5]. Avant de l’expliciter, introduisons quelques notations.

- Nous nous plaçons dans l’espace de Hilbert réel $\mathcal{H} = \mathbb{C}^{N-P \times P+1}$ des matrices de taille $N - P \times P + 1$ à valeurs complexes, mais ayant une symétrie centro-hermitienne. Le produit scalaire (de Frobenius) entre deux matrices de \mathcal{H} est $\langle \mathbf{X}, \mathbf{X}' \rangle = \sum_{i,j} x_{i,j} x'_{i,j}^* \in \mathbb{R}$.
- Nous notons P_Ω la projection d’un élément vers l’élément le plus proche appartenant à un ensemble Ω .
- Nous définissons \mathcal{T} , le sous-espace linéaire de \mathcal{H} des matrices Toeplitz. $P_{\mathcal{T}}$ consiste simplement à moyenner les éléments de la matrice le long de ses diagonales.
- Nous définissons \mathcal{R}_K , la sous-variété fermée de \mathcal{H} des matrices de rang au plus K . $P_{\mathcal{R}_K}$ correspond à tronquer la décomposition en valeurs singulières (SVD) de la matrice, d’après le théorème de Schmidt-Eckart-Young : si une matrice \mathbf{X} a pour SVD $\mathbf{X} = \mathbf{L}\mathbf{\Sigma}\mathbf{R}^H$, alors $P_{\mathcal{R}_K}(\mathbf{X})$ est obtenue en mettant à zéro les valeurs singulières dans $\mathbf{\Sigma}$, hormis les K plus grandes.

Le débruitage de Cadzow consiste simplement à appliquer alternativement $P_{\mathcal{R}_K}$ et $P_{\mathcal{T}}$ à la matrice $\mathbf{T}_P : \mathbf{T}_P^{(0)} = \mathbf{T}_P$ et $\mathbf{T}_P^{(i+1)} = P_{\mathcal{T}}(P_{\mathcal{R}_K}(\mathbf{T}_P^{(i)}))$, pour tout $i \in \mathbb{N}$. Bien que cette procédure semble toujours converger vers une matrice dans $\mathcal{R}_K \cap \mathcal{T}$,

il n’existe pas de preuve de convergence, contrairement à une idée assez répandue [19]. Quoiqu’il en soit, la matrice obtenue n’est pas un minimiseur local de la fonction coût $\|\cdot - \mathbf{T}_P\|_w^2$ [14, 9]. Ainsi, ce travail vise à proposer une alternative au débruitage de Cadzow, qui fournisse bien une solution locale du problème SLRA (6), donnant ainsi théoriquement une estimation de meilleure qualité.

4 Algorithme proposé

Nous avons récemment proposé un nouvel algorithme [20] permettant de résoudre, entre autres, les problèmes d’optimisation convexe ayant la forme suivante :

$$\text{trouver } \tilde{x} \in \arg \min_{x' \in \mathcal{H}} F(x') \quad \text{tq } x' \in \Omega_1 \cap \Omega_2, \quad (8)$$

où \mathcal{H} est un espace de Hilbert réel, Ω_1 et Ω_2 sont deux sous-ensembles convexes fermés de \mathcal{H} , et $F : \mathcal{H} \rightarrow \mathbb{R}$ est une fonction différentiable sur \mathcal{H} , de gradient Lipschitz-continu ; c’est-à-dire qu’il existe un réel $\beta > 0$ tel que $\|\nabla F(x') - \nabla F(x)\| \leq \beta \|x - x'\|$, $\forall x, x' \in \mathcal{H}$. L’algorithme proposé dans [20] prend la forme suivante :

Algorithme d’optimisation. Choisir les paramètres $\mu > 0$, $\gamma \in]0, 1[$, tels que $2\gamma > \beta\mu$, et les éléments initiaux $x^{(0)}, s^{(0)} \in \mathcal{H}$. Puis itérer, pour $i \geq 0$,

$$\begin{cases} x^{(i+1)} = P_{\Omega_1}(s^{(i)} + \gamma(x^{(i)} - s^{(i)}) - \mu \nabla F(x^{(i)})) \\ s^{(i+1)} = s^{(i)} - x^{(i+1)} + P_{\Omega_2}(2x^{(i+1)} - s^{(i)}) \end{cases}$$

Moyennant quelques hypothèses faibles supplémentaires, la séquence $(x^{(i)})_{i \in \mathbb{N}}$ converge vers un élément \hat{x} solution de (8) [20]. En l'absence de convexité d'un des ensembles, ce résultat ne s'applique pas. Nous utilisons donc l'algorithme comme un heuristique pour résoudre le problème SLRA (6). En effet, on remarque que celui-ci est une instance particulière de (8), avec, dans les notations de la section précédente, $\Omega_1 = \mathcal{R}_K$ et $\Omega_2 = \mathcal{T}$. La fonction coût est $F(\mathbf{X}) = \frac{1}{2} \|\mathbf{X} - \mathbf{T}_P\|_w^2$, de sorte que $\nabla F(\mathbf{X}) = \mathbf{W} \circ (\mathbf{X} - \mathbf{T}_P)$, où \circ est le produit terme à terme de Hadamard et la matrice \mathbf{W} a pour éléments les $\{w_{i,j}\}$ définissant la norme $\|\cdot\|_w$. La constante de Lipschitz de ∇F est $\beta = \max(\{w_{i,j}\}) = 1$.

On observe empiriquement que l'algorithme converge toujours, pour un certain choix des paramètres μ et γ , vers une matrice Toeplitz et de rang K , qui est un minimiseur local de la fonction coût (voir [11] pour plus de détails). La vitesse de convergence sur les itérées est linéaire (la fonction coût est fortement convexe). De plus, il est possible de vérifier par une recherche combinatoire que l'algorithme fournit bien la solution globale du problème SLRA (6), et donc atteint l'estimation de maximum de vraisemblance, lorsque le niveau de bruit n'est pas trop élevé ($\text{SNR} \geq 12\text{dB}$ dans l'exemple de la figure 1 (b)).

Nous montrons en figure 1 une comparaison de la méthode proposée avec plusieurs autres méthodes. Le gain par rapport au débruitage de Cadzow est de 10% en moyenne pour l'estimation des positions $\{t_k\}$. Ce gain n'est pas très élevé, pour le scénario « orthogonal » considéré ici (impulsions idéales de Dirac et noyau d'échantillonnage de Dirichlet), mais il est à noter que les deux méthodes ont essentiellement la même complexité, dominée par une SVD par itération. Des exemples avec des nombres K d'impulsions plus grands sont donnés dans [11].

5 Conclusion

Nous avons proposé un nouvel algorithme d'optimisation pour l'approximation structurée de rang faible de matrices. Combiné à la méthode de Prony, nous obtenons une méthode efficace et robuste pour l'estimation d'impulsions. L'estimateur du maximum de vraisemblance est atteint, lorsque le bruit ne dépasse pas un certain seuil. De nombreuses questions théoriques liées aux performances de la méthode restent ouvertes. Il serait souhaitable d'avoir des garanties de stabilité du type de celles données pour une relaxation convexe du problème [21, 6]. Une implémentation Matlab est disponible sur la page web de l'auteur.

Références

[1] P.-L. Dragotti, M. Vetterli et T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix", *IEEE Trans. Signal Processing*, vol. 55, n° 5, p. 1741–1757, mai 2007.

[2] T. Blu, P.-L. Dragotti, M. Vetterli, P. Marziliano et L. Coulot, "Sparse sampling of signal innovations", *IEEE Signal Processing Mag.*, vol. 25, n° 2, p. 31–40, mars 2008, Numéro spécial sur le *Compressive Sampling*.

[3] M. Mishali et Y. C. Eldar, "Sub-Nyquist sampling: Bridging theory and practice", *IEEE Signal Processing Mag.*, vol. 28, n° 6, p. 98–124, nov. 2011.

[4] J. Urigüen, Y. C. Eldar, P. L. Dragotti et Z. Ben-Haim, "Sampling at the rate of innovation: Theory and applications", dans *Compressed Sensing: Theory and Applications*, Y. C. Eldar et G. Kutyniok, Eds. Cambridge University Press, 2012.

[5] R. Tur, Y. C. Eldar et Z. Friedman, "Innovation rate sampling of pulse streams with application to ultrasound imaging", *IEEE Trans. Signal Processing*, vol. 59, n° 4, p. 1827–1842, avr. 2011.

[6] E. J. Candès et C. Fernandez-Granda, "Super-resolution from noisy data", preprint arXiv:1211.0290, 2012.

[7] Z. Ben-Haim, T. Michaeli et Y. C. Eldar, "Performance bounds and design criteria for estimating finite rate of innovation signals", *IEEE Trans. Inform. Theory*, vol. 58, n° 8, p. 4993–5015, août 2012.

[8] P. Stoica et R. Moses, *Spectral Analysis of Signals*, Prentice Hall, NJ, 2005.

[9] J. Gillard et A. Zhigljavsky, "Analysis of structured low rank approximation as an optimization problem", *Informatica*, vol. 22, n° 4, p. 489–505, 2011.

[10] T. Blu, "The generalized annihilation property—A tool for solving finite rate of innovation problems", dans *Proc. of Int. Workshop on Sampling Theory and Appl. (SampTA)*, Marseille, France, mai 2009.

[11] L. Condat et A. Hirabayashi, "Cadzow denoising upgraded: A new projection method for the recovery of Dirac pulses from noisy linear measurements", preprint hal-00759253, 2012.

[12] D. W. Tufts et R. Kumaresan, "Frequency estimation of multiple sinusoids: Making linear prediction perform like maximum likelihood", *Proc. IEEE*, vol. 70, n° 9, p. 975–989, sept. 1982.

[13] I. Markovsky, *Low Rank Approximation: Algorithms, Implementation, Applications*, Springer, 2012.

[14] M. T. Chu, R. E. Funderlic et R. J. Plemmons, "Structured low rank approximation", *Linear Algebra Appl.*, vol. 366, p. 157–172, 2003.

[15] M. Schuermans, *Weighted low rank approximation: Algorithms and applications*, Ph.D. thesis, Katholieke Universiteit Leuven, 2006.

[16] I. Markovsky et K. Usevich, "Software for weighted structured low-rank approximation", Tech. Rep. 339974, ECS, Univ. of Southampton, 2012, voir <https://github.com/slra/slra>.

[17] I. Markovsky, "How effective is the nuclear norm heuristic in solving data approximation problems?", dans *Proc. of IFAC Symposium on System Identification (SYSID)*, Brussels, Belgium, juil. 2012.

[18] J. A. Cadzow, "Signal enhancement—A composite property mapping algorithm", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, n° 1, p. 49–62, jan. 1988.

[19] D. R. Luke, "Prox-regularity of rank constraint sets and implications for algorithms", preprint arXiv:1112.0526, 2011.

[20] L. Condat, "A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms", *J. Optim. Theory and Appl.*, 2013, à paraître.

[21] B. N. Bhaskar, G. Tang et B. Recht, "Atomic norm denoising with applications to line spectral estimation", preprint arXiv:1204.0562, 2012.