

Super-résolution d'impulsions positives par approximation Toeplitz de rang faible

Laurent CONDAT¹, Akira HIRABAYASHI²

¹GIPSA-lab, Université Grenoble-Alpes, 11 rue des mathématiques, campus BP46, 38402 St Martin d'Hères Cedex

²College of Information Science and Engineering, Ritsumeikan University, Shiga, 525-8577, Japon

laurent.condat@gipsa-lab.grenoble-inp.fr

Résumé – La super-résolution consiste à retrouver les fins détails d'un signal à partir de mesures à basse résolution. Nous considérons ici l'estimation d'une séquence d'impulsions de Dirac, d'amplitudes positives et à des positions arbitraires, à partir d'échantillons filtrés passe-bas bruités. L'estimation au sens du maximum de vraisemblance des paramètres est équivalente à un problème non convexe difficile d'approximation matricielle structurée de rang faible. Pour le résoudre, nous proposons un nouvel algorithme itératif heuristique, qui surpasse l'état de l'art.

Abstract – Super-resolution consists in recovering the fine details of a signal from low-resolution measurements. Here we consider the estimation of Dirac pulses with positive amplitudes at arbitrary locations, from noisy lowpass-filtered samples. Maximum-likelihood estimation of the unknown parameters is equivalent to a difficult nonconvex matrix problem of structured low rank approximation. To solve it, we propose a new heuristic iterative algorithm, which outperforms the state of the art.

1 Formulation du problème

Le cadre de la reconstruction de signaux analogiques à partir de mesures discrètes a récemment été étendu à une large classe de signaux gouvernés par des modèles non linéaires parcimonieux, bien au-delà du paradigme classique ancré dans les travaux de Shannon [1, 2]. Nous considérons ici le problème emblématique, parfois qualifié de *super-résolution* [3], consistant à estimer une séquence d'impulsions de Dirac à partir d'échantillons à bande limitée bruités, ou de manière équivalente à partir de coefficients de Fourier basses-fréquences [1, 3, 4]. Ce problème a un large champ d'applications, incluant la communication en ultra large bande et l'imagerie par ultrasons [5]. Soit un réel $\tau > 0$; on se place sur le tore $\mathbb{T} = \mathbb{R}/\tau\mathbb{Z}$, qui est l'intervalle $[0, \tau[$ avec 0 et τ identifiés. Le nombre d'impulsions, supposé connu, est $K \geq 1$. On introduit la distribution de Dirac $\delta(t)$. Alors, le signal inconnu à estimer est

$$s^*(t) = \sum_{k=1}^K a_k^* \delta(t - t_k^*), \quad \forall t \in \mathbb{T}, \quad (1)$$

où les $(t_k^*)_{k=1}^K$ sont les positions distinctes dans \mathbb{T} et les $(a_k^*)_{k=1}^K$ sont les amplitudes *strictement positives* (l'étoile indiquant une quantité inconnue). Le but est d'estimer ces $2K$ paramètres à partir des données disponibles, qui prennent la forme d'une séquence $(y_n)_{n=0}^{N-1}$ de N mesures uniformes à bande limitée bruitées : $y_n = \int_0^\tau s^*(t) \varphi(\frac{n\tau}{N} - t) dt + \varepsilon_n$, où la fonction d'échantillonnage $\varphi(t) = (\sin(N\pi t/\tau))/(N \sin(\pi t/\tau))$ est la fonction de Dirichlet et les $\varepsilon_n \sim \mathcal{N}(0, \sigma^2)$ sont des réalisations indépendantes de bruit gaussien. On suppose N impair et on pose $M = (N-1)/2$. On suppose aussi que $M \geq K$, une condition

nécessaire et suffisante pour que la reconstruction parfaite de s^* en l'absence de bruit soit possible, au moyen de la méthode de Prony, voir en section 2.

La nouveauté, par rapport au précédent travail des auteurs [4], est l'hypothèse que les amplitudes a_k^* sont *positives*; cette contrainte peut avoir une justification physique, lorsque les amplitudes correspondent à une intensité électrique ou lumineuse. Par conséquent, une méthode d'estimation doit retourner des amplitudes, qui sont positives elles aussi.

Si l'on applique la transformée de Fourier discrète à la séquence $(y_n)_{n=0}^{N-1}$, on obtient N coefficients de Fourier consécutifs de s^* , de la forme [4]

$$\hat{y}_m = \sum_{k=1}^K a_k^* e^{-j2\pi m t_k^*/\tau} + \hat{\varepsilon}_m, \quad \forall m = -M, \dots, M. \quad (2)$$

Ainsi, l'estimation au sens du maximum de vraisemblance (ML) des paramètres revient à résoudre le problème des moindres carrés non linéaires [4] :

$$\underset{\substack{(t_k)_{k=1}^K \in \mathbb{T}^K \\ (a_k)_{k=1}^K \in]0, +\infty[^K}}{\text{minimiser}} \quad \frac{1}{2} \sum_{m=-M}^M \left| \hat{y}_m - \sum_{k=1}^K a_k e^{-j2\pi m t_k/\tau} \right|^2. \quad (3)$$

On a donc affaire à un problème d'estimation spectrale [6]. Le résoudre exactement n'est pas envisageable, car la fonction coût à minimiser est très oscillante, avec un nombre de minima locaux de l'ordre de N^K . Lorsque $N \gg K$ et les positions t_k^* sont suffisamment séparées, les méthodes classiques d'estimation spectrale, comme MUSIC ou ESPRIT, peuvent être utilisées [6]. Elles sont rapides mais sous-optimales : contrairement à l'estimation ML, elles n'atteignent pas les bornes de Cramér-

Rao lorsque le niveau de bruit tend vers zéro [4]. De plus, il n'y a aucune garantie que les amplitudes estimées soient positives.

Dans la section 2, nous reformulons l'estimation ML des paramètres comme un problème d'approximation structurée de rang faible d'une matrice Toeplitz construite à partir des données. L'algorithme proposé est décrit en section 3. Des résultats illustrent son efficacité en section 4.

2 Reformulation comme un problème de débruitage matriciel

Introduisons l'opérateur linéaire Toep qui, à une séquence ou vecteur $(x_m)_{m=-M}^M$ présentant une symétrie hermitienne ($x_{-m} = x_m^*$), associe une matrice Toeplitz hermitienne :

$$\text{Toep} : (x_{-M}, \dots, x_M) \mapsto \begin{pmatrix} x_0 & \cdots & x_{-M} \\ \vdots & \ddots & \vdots \\ x_M & \cdots & x_0 \end{pmatrix}. \quad (4)$$

Nous allons nous appuyer sur le résultat fondamental suivant :

Théorème de Carathéodory [7] : Soit une matrice hermitienne positive \mathbf{T} de taille $(M+1) \times (M+1)$ et de rang K (on a $1 \leq K \leq M$). Soit $\mathbf{x} = (x_m)_{m=-M}^M$ la séquence telle que $\mathbf{T} = \text{Toep}(\mathbf{x})$; autrement dit, les x_m sont les éléments sur les diagonales de \mathbf{T} . Alors il existe des positions distinctes $(t_k)_{k=1}^K \in \mathbb{T}^K$ et des amplitudes $(a_k)_{k=1}^K \in]0, +\infty[^K$, uniques à une permutation sur l'indice k près, telles que, $\forall m = -M \dots, M$, $x_m = \sum_{k=1}^K a_k e^{-j2\pi m t_k / \tau}$.

Ainsi, Toep définit une bijection entre les sommes positives d'exponentielles complexes et les matrices Toeplitz positives de rang faible. De plus, on dispose d'une procédure simple, due à Prony [1], pour trouver les paramètres correspondant à une telle matrice \mathbf{T} :

1) Réarranger les éléments x_m de \mathbf{T} dans une matrice Toeplitz \mathbf{T}_K de taille $(N-K) \times (K+1)$ (on a $N = 2M+1$) et de rang K , telle que

$$\mathbf{T}_K = \begin{pmatrix} x_{-M+K} & \cdots & x_{-M} \\ \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ x_M & \cdots & x_{M-K} \end{pmatrix}. \quad (5)$$

2) Calculer le vecteur singulier à droite $\mathbf{h} = [h_0 \dots h_K]^T$ de \mathbf{T}_K correspondant à la valeur singulière zéro (la plus petite valeur singulière en pratique, ce qui rend la méthode robuste).

3) Calculer les racines $(z_k)_{k=1}^K$ du polynôme $\sum_{k=0}^K h_k z^k$. On en déduit les positions $(t_k)_{k=1}^K$ par : $t_k = \frac{\tau}{2\pi} \arg_{[0, 2\pi[}(z_k)$.

4) Étant données les positions, le vecteur $\mathbf{a} = [a_1 \dots a_K]^T$ des amplitudes est obtenu en résolvant le système linéaire des moindres carrés

$$\mathbf{U}^H \mathbf{U} \mathbf{a} = \mathbf{U}^H \mathbf{x}, \quad (6)$$

où \cdot^H est la transposée hermitienne, $\mathbf{x} = [x_{-M} \dots x_M]^T$ et

$$\mathbf{U} = \begin{pmatrix} e^{j2\pi M t_1 / \tau} & \cdots & e^{j2\pi M t_K / \tau} \\ \vdots & \ddots & \vdots \\ e^{-j2\pi M t_1 / \tau} & \cdots & e^{-j2\pi M t_K / \tau} \end{pmatrix}. \quad (7)$$

Notons qu'en l'absence de bruit, cette procédure appliquée à $\mathbf{x} = \hat{\mathbf{y}}$ fournit de manière exacte les paramètres $(t_k^*)_{k=1}^K$ et $(a_k^*)_{k=1}^K$. De plus, en présence de bruit, étant données des estimées des positions, les estimées ML des amplitudes sont obtenues en résolvant (6), avec \mathbf{x} remplacé par $\hat{\mathbf{y}}$.

Ainsi, en vertu du théorème de Carathéodory, le problème d'estimation ML (3) peut être reformulé comme

$$\begin{aligned} & \underset{\mathbf{x} \in \mathbb{C}^N}{\text{minimiser}} \frac{1}{2} \|\mathbf{x} - \hat{\mathbf{y}}\|_2^2 \quad \text{sujet à : } \text{rang}(\text{Toep}(\mathbf{x})) \leq K \text{ et} \\ & \text{Toep}(\mathbf{x}) \succcurlyeq 0, \end{aligned} \quad (8)$$

où $\succcurlyeq 0$ indique la positivité. En utilisant l'injectivité de Toep, on peut réécrire (8) comme un problème d'approximation structurée de rang faible (SLRA) :

$$\begin{aligned} & \underset{\mathbf{T} \in \mathcal{H}}{\text{minimiser}} \frac{1}{2} \|\mathbf{T} - \text{Toep}(\hat{\mathbf{y}})\|_w^2 \quad \text{sujet à : } \mathbf{T} \text{ est Toeplitz} \\ & \text{et } \text{rang}(\mathbf{T}) \leq K \text{ et } \mathbf{T} \succcurlyeq 0, \end{aligned} \quad (9)$$

où \mathcal{H} est l'espace de Hilbert réel des matrices hermitiennes de taille $(M+1) \times (M+1)$ et la norme pondérée $\|\cdot\|_w$ est telle que $\|\mathbf{x} - \hat{\mathbf{y}}\|_2^2 = \|\text{Toep}(\mathbf{x}) - \text{Toep}(\hat{\mathbf{y}})\|_w^2$; c.-à-d. $\|\mathbf{A}\|_w^2 = \sum_{i=1}^{M+1} \sum_{j=1}^{M+1} w_{i,j} |a_{i,j}|^2$, où $w_{i,j} = 1/(M+1-|i-j|)$. Ainsi, (9) peut être vu comme un problème de *débruitage de matrice* : la matrice $\text{Toep}(\hat{\mathbf{y}})$, ou de manière équivalente les données $\hat{\mathbf{y}}$, est débruitée en trouvant la matrice la plus proche consistante avec le modèle, à partir de laquelle les paramètres estimés sont extraits par la procédure de Prony. La formulation SLRA (9) a plusieurs avantages par rapport à la forme paramétrique (3); par exemple, la question de l'initialisation disparaît, puisque la matrice bruitée $\text{Toep}(\hat{\mathbf{y}})$ est l'estimée initiale naturelle de la solution.

Les problèmes SLRA ont de nombreuses applications, mais à cause de la non-convexité du rang, ils sont très difficiles [8,9]. Plusieurs méthodes ont été proposées pour obtenir une solution locale d'un problème SLRA, voir la bibliographie dans [8] et [4]. À notre connaissance, la seule bibliothèque publiquement accessible et efficace est celle en cours de développement par Ivan Markovsky [10]. Cependant, elle ne traite que des données à valeurs réelles, alors que les matrices dans (9) sont complexes. Dans la section suivante, nous proposons un algorithme simple à implémenter et efficace. Dans le reste de cette section, nous présentons deux alternatives fournissant une solution approchée du problème.

Une alternative consiste à "convexifier" le problème en remplaçant le comptage des impulsions par la norme atomique, qui est ici la somme des amplitudes [11]. Le problème devient :

$$\underset{\mathbf{x} \in \mathbb{C}^N}{\text{minimiser}} \frac{1}{2} \|\mathbf{x} - \hat{\mathbf{y}}\|_2^2 + \lambda x_0 \quad \text{sujet à } \text{Toep}(\mathbf{x}) \succcurlyeq 0, \quad (10)$$

pour un certain paramètre $\lambda > 0$. En effet, la décomposition de la solution $x_m = \sum_{k=1}^{K'} a_k e^{-j2\pi m t_k / \tau}$ donne simplement $x_0 = \sum_{k=1}^{K'} a_k$. Notons qu'il est difficile de choisir λ et qu'il n'y a aucune garantie que $K' = K$. Nous laissons pour un travail futur l'évaluation des performances de cette approche, ainsi que l'étude des liens entre les différents paramètres, les bornes de Cramér-Rao et le SNR [12–15].

Mentionnons finalement la méthode heuristique simple des projections alternées, aussi appelée débruitage de Cadzow [16],

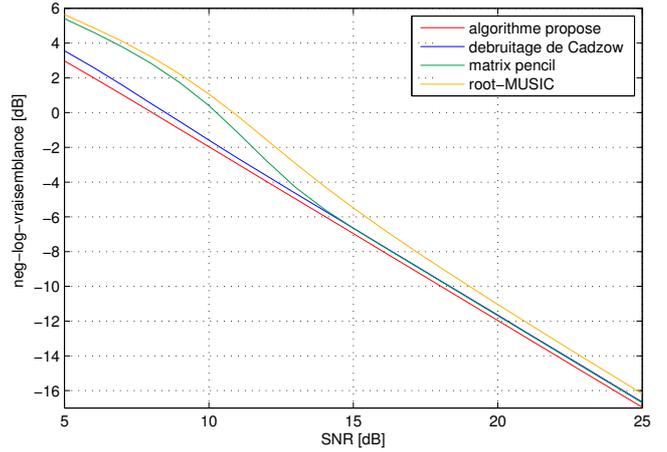
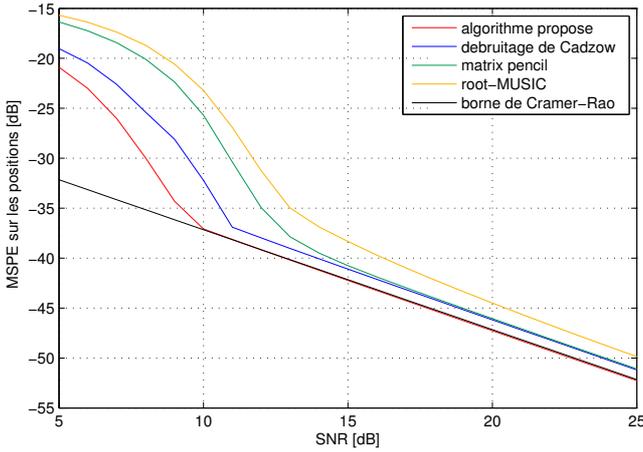


FIGURE 1 – Tracé en échelle log-log de la MSPE (à gauche) et de la neg-log-vraisemblance (à droite) en fonction du SNR, pour l’expérience 1 : le signal s^* consiste en $K = 2$ pics, aux positions $(0.42, 0.52)$ ($\tau = 1$), d’amplitudes $(1, 1)$. Les données consistent en $N = 11$ échantillons basses-fréquences. Les valeurs sont moyennées sur 10^4 réalisations de bruit.

souvent utilisée pour résoudre des problèmes d’approximation avec deux contraintes, en projetant alternativement sur celles-ci. Elle est utilisée dans [1, 5] pour la reconstruction d’impulsions. Appliquée au problème (9), et dans les notations de la section suivante, on a $\mathbf{T}^{(l+1)} = P_{\Omega_2}(P_{\Omega_1}(\mathbf{T}^{(l)}))$, où l est le compteur d’itérations et $\mathbf{T}^{(0)} = \text{Toep}(\hat{\mathbf{y}})$. On observe que cette méthode converge, mais pas vers une solution locale, car la fonction coût à minimiser n’est pas prise en compte.

3 Un nouvel algorithme pour l’approximation structurée de rang faible

Considérons le problème d’optimisation générique :

$$\underset{z \in \mathcal{Z}}{\text{minimiser}} F(z) \quad \text{sujet à } z \in \Omega_1 \cap \Omega_2, \quad (11)$$

où \mathcal{Z} est un espace de Hilbert réel, Ω_1 et Ω_2 sont deux ensembles fermés de \mathcal{Z} et $F : \mathcal{Z} \rightarrow \mathbb{R}$ est une fonction convexe et différentiable de gradient β -Lipschitz continu, pour un certain $\beta > 0$: $\|\nabla F(z) - \nabla F(z')\| \leq \beta \|z - z'\|$, $\forall (z, z') \in \mathcal{Z}^2$. L’algorithme proposé par le premier auteur pour résoudre ce problème est [17] :

Algorithme d’optimisation proposé. Choisir les paramètres $\mu > 0$, $\gamma \in]0, 1[$, tels que $2\gamma > \beta\mu$, et les estimées initiales $z^{(0)}, v^{(0)} \in \mathcal{Z}$. Puis itérer, pour tout $l \geq 0$,

$$\begin{cases} z^{(l+1)} = P_{\Omega_1}(v^{(l)} + \gamma(z^{(l)} - v^{(l)}) - \mu \nabla F(z^{(l)})), \\ v^{(l+1)} = v^{(l)} - z^{(l+1)} + P_{\Omega_2}(2z^{(l+1)} - v^{(l)}). \end{cases}$$

Si les ensembles Ω_1 et Ω_2 sont convexes, l’algorithme converge vers une solution, sous certaines hypothèses faibles [17].

On reconnaît alors que le problème SLRA (9) est une instance de (11), avec $\mathcal{Z} = \mathcal{H}$ et les correspondances suivantes :

- Ω_1 est le sous-ensemble de \mathcal{H} des matrices positives de rang $\leq K$. La projection $P_{\Omega_1}(\mathbf{A}) = \mathbf{Q} \text{diag}(\mathbf{d}') \mathbf{Q}^H$ sur cet ensemble, pour la norme de Frobenius, d’une matrice

$\mathbf{A} = \mathbf{Q} \text{diag}(\mathbf{d}) \mathbf{Q}^H$, est obtenue en tronquant sa décomposition en éléments propres : les valeurs propres dans \mathbf{d}' sont obtenues en mettant tous les éléments de \mathbf{d} à zéro, sauf les (au plus) K plus grands éléments positifs.

- Ω_2 est le sous-espace vectoriel de \mathcal{H} des matrices Toeplitz. La projection P_{Ω_2} consiste à remplacer les éléments d’une même diagonale par leur moyenne.
- La fonction coût est $F(\mathbf{A}) = \frac{1}{2} \|\mathbf{A} - \text{Toep}(\hat{\mathbf{y}})\|_w^2$, donc $\nabla F(\mathbf{A}) = \mathbf{W} \circ (\mathbf{A} - \text{Toep}(\hat{\mathbf{y}}))$, où \circ est le produit terme à terme et \mathbf{W} est la matrice ayant pour éléments les $w_{i,j}$. Sa constante de Lipschitz est $\beta = \max(\{w_{i,j}\}) = 1$.

L’ensemble Ω_1 n’étant pas convexe, l’algorithme devient heuristique, sans garantie de convergence. Toutefois, avec $\mu = 1.3$, $\gamma = 0.51\mu$, $z^{(0)} = v^{(0)} = \text{Toep}(\hat{\mathbf{y}})$, on observe qu’il converge presque toujours vers une solution locale de (9) (le critère d’arrêt est $\|z^{(l+1)} - P_{\Omega_2}(z^{(l+1)})\|_F < 1e-12$ et l’algorithme fournit $P_{\Omega_2}(z^{(l+1)})$). Dans de très rare cas, quand le niveau de bruit est très élevé, il entre dans un cycle au lieu de converger; mais il suffit alors de diminuer μ et γ pour obtenir la convergence. Ce comportement a déjà été observé dans [4], mais ici l’absence de convergence est encore plus rare : la contrainte $\mathbf{T} \succcurlyeq 0$ semble rendre l’algorithme encore plus robuste.

4 Résultats expérimentaux

L’algorithme proposé est comparé au débruitage de Cadzow, à la fonction `rootmusic` de Matlab, implémentant une variante de MUSIC, et à la méthode “matrix pencil” [18]. Comme ces deux dernières ne fournissent pas toujours des impulsions positives, une fois obtenues les positions, nous obtenons les amplitudes en remplaçant la solution des moindres carrés (6) par la solution des moindres carrés positifs, obtenue par la commande `lsqnonneg` de Matlab; mais comme certaines amplitudes peuvent être nulles, l’approche peut échouer à délivrer K impulsions. Pour comparer les méthodes, nous défi-

nissons deux critères d’erreur : le premier, lorsque les amplitudes a_k^* sont égales, correspond à l’erreur sur les positions, pour la meilleure affectation entre les vrais pics et leurs estimés. Ainsi, dans l’expérience 1 ci-dessous, on calcule l’erreur quadratique périodique moyenne (MSPE) $\min((t_1 - t_1^*)^2 + (t_2 - t_2^*)^2, (t_1 - t_2^*)^2 + (t_2 - t_1^*)^2)/2$, où $(x)_\tau = ((x + \frac{\tau}{2}) \bmod \tau) - \frac{\tau}{2}$. Ce critère a l’avantage de pouvoir être comparé aux bornes de Cramér-Rao, calculées comme dans [1]. Dans le cas général, puisqu’on vise l’estimation ML, on peut comparer les valeurs, obtenues par différentes méthodes, de la neg-log-vraisemblance, qui est la fonction coût en (3), (8) et (9).

L’expérience 1 consiste à estimer $K = 2$ pics de même amplitude. Les résultats sont illustrés en figure 1. La méthode proposée est la seule à atteindre la borne de Cramér-Rao, pour un rapport signal à bruit (SNR) (égal à $10 \log_{10}(\|\hat{y} - \hat{e}\|_2^2 / \|\hat{e}\|_2^2)$) supérieur à 10dB, ce qui montre qu’elle fournit bien l’estimateur ML dans ce régime, autrement dit la solution globale du problème.

L’expérience 2 est illustrée en figure 2. La neg-log-vraisemblance n’est pas affichée, mais le bilan est similaire à celui de la figure 1 à droite : l’algorithme proposé surpasse les autres.

5 Conclusion

Nous avons proposé une formulation non convexe, basée sur l’approximation structurée de rang faible, de l’estimation, au sens du maximum de vraisemblance, d’une séquence d’impulsions de Dirac d’amplitudes positives, à partir d’échantillons basses-fréquences bruités. Nous avons aussi proposé un algorithme heuristique, simple à implémenter et efficace : il converge vers la solution globale du problème, lorsque le niveau de bruit n’est pas trop élevé. Le code Matlab implémentant les différentes méthodes comparées est disponible sur la page web du premier auteur.

Références

- [1] T. Blu, P.-L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, “Sparse sampling of signal innovations,” *IEEE Signal Processing Mag.*, vol. 25, no. 2, pp. 31–40, Mar. 2008, Special issue on Compressive Sampling.
- [2] J. Urigüen, Y. C. Eldar, P. L. Dragotti, and Z. Ben-Haim, “Sampling at the rate of innovation: Theory and applications,” in *Compressed Sensing: Theory and Applications*, Y. C. Eldar and G. Kutyniok, Eds., pp. 148–209. Cambridge University Press, 2012.
- [3] E. J. Candès and C. Fernandez-Granda, “Super-resolution from noisy data,” *Journal of Fourier Analysis and Applications*, vol. 13, no. 6, pp. 1229–1254, Dec. 2013.
- [4] L. Condat and A. Hirabayashi, “Cadzow denoising upgraded: A new projection method for the recovery of Dirac pulses from noisy linear measurements,” *Sampling Theory in Signal and Image Proc.*, vol. 14, no. 1, pp. 17–47, 2015.
- [5] R. Tur, Y. C. Eldar, and Z. Friedman, “Innovation rate sampling of pulse streams with application to ultrasound imaging,” *IEEE Trans. Signal Processing*, vol. 59, no. 4, pp. 1827–1842, Apr. 2011.
- [6] P. Stoica and R. Moses, *Spectral Analysis of Signals*, Prentice Hall, NJ, 2005.

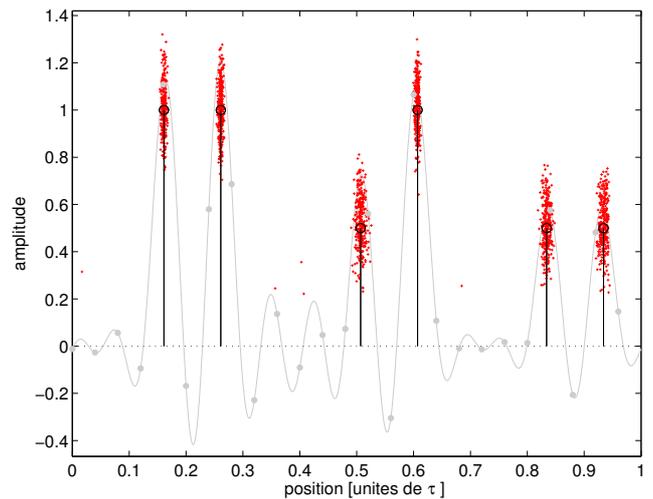


FIGURE 2 – Expérience 2 : s^* consiste en $K = 6$ pics, en noir. On a $N = 25$ échantillons basses-fréquences (points gris) corrompus par du bruit de SNR 12dB ($\tau = 1$). Les points rouges montrent les pics estimés par la méthode proposée, pour 300 réalisations de bruit.

- [7] C. Carathéodory, “Über den Variabilitätsbereich der Fourier-schen Konstanten von positiven harmonischen Funktionen,” *Rendiconti del Circolo Matematico di Palermo*, vol. 3, pp. 193–217, 1911.
- [8] I. Markovsky, *Low Rank Approximation: Algorithms, Implementation, Applications*, Springer, 2012.
- [9] J. W. Gillard and A. A. Zhigljavsky, “Optimization challenges in the structured low rank approximation problem,” *Journal of Global Optimization*, vol. 53, pp. 733–751, 2013.
- [10] I. Markovsky and K. Usevich, “Software for weighted structured low-rank approximation,” *J. Comput. Appl. Math.*, vol. 256, pp. 278–292, 2014.
- [11] B. N. Bhaskar, G. Tang, and B. Recht, “Atomic norm denoising with applications to line spectral estimation,” *IEEE Trans. Sig. Proc.*, vol. 61, no. 23, pp. 5987–5999, Dec. 2013.
- [12] D. L. Donoho, “Superresolution via sparsity constraints,” *SIAM J. Math. Anal.*, vol. 23, pp. 1309–1331, Sept. 1992.
- [13] L. Demanet and N. Nguyen, “The recoverability limit for super-resolution via sparsity,” preprint arXiv:1502.01385, 2015.
- [14] V. I. Morgenshtern and E. J. Candès, “Super-resolution of positive sources: the discrete setup,” preprint arXiv:1504.00717, 2015.
- [15] X. Wei and P. L. Dragotti, “Guaranteed performance in the FRI setting,” *IEEE Signal Processing Letters*, 2015, to appear.
- [16] J. A. Cadzow, “Signal enhancement—A composite property mapping algorithm,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 1, pp. 49–62, Jan. 1988.
- [17] L. Condat, “A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms,” *J. Optim. Theory Appl.*, vol. 158, no. 2, pp. 460–479, 2013.
- [18] Y. Hua and T. K. Sarkar, “Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 5, pp. 814–824, May 1990.