

Nouveaux noyaux pour l'échantillonnage de signaux impulsions à bande non-limitée

Stéphanie BERNHARDT¹, Rémy BOYER¹, Sylvie MARCOS¹, Yonina C. ELDAR², Pascal LARZABAL³

¹Laboratoire des Signaux et Systèmes (L2S), Université Paris Sud-CNRS-CentraleSupélec, Gif-Sur-Yvette, France

²Technion, Department of EE, Haifa, Israel

³SATIE, ENS-Cachan, Cachan, France

stephanie.bernhardt@lss.supelec.fr, remy.boyer@lss.supelec.fr

Résumé – Nous nous intéressons à l'échantillonnage de signaux constitués de sommes d'impulsions de largeur de bande infinie, qui ne sont donc pas échantillonnables au sens de Shannon. Cependant, on peut montrer que ces signaux sont reconstituables parfaitement grâce au nombre limité de paramètres inconnus par unité de temps nécessaire pour les décrire. Cette reconstruction est possible à l'aide de noyaux judicieusement choisis, dont le noyau somme de sinus cardinal (SoS) sur lequel nous nous focalisons fait partie. Ce noyau a l'avantage d'être paramétrable et donc flexible. Dans cet article, nous obtenons la Borne de Cramer Rao dans le cadre Bayésien (BCRB) relatif aux paramètres d'estimation puis cherchons le noyau SoS maximisant l'information de Fisher. Nous proposons une famille de noyaux pouvant être ajustés par l'utilisateur pour privilégier l'un ou l'autre des paramètres à estimer, et pour rendre le noyau plus robuste au bruit: la famille résultante de noyaux a une variété de forme allant de la sinusoïde au noyau de Dirichlet.

Abstract – We consider the problem of the estimation of finite rate of innovation signals, which is a non-bandlimited class of signals and which thus cannot be sampled in Shannon's sense. However, it has been recently shown that these signals can be sampled perfectly thanks to the limited number of unknown parameters by unit of time needed to describe them. This reconstruction is possible using appropriate kernels, one of those being the sum of sines (SoS) family of kernels on which we chose to focus. Those kernels are of particular interest since they can be user adjusted. In this paper, we derive the Bayesian Cramer Rao Bound (BCRB) on the estimation parameters and derive the optimum SoS kernel in the maximum fisher information sense. We propose a family of optimal kernels that can be user-adjusted to favor either of the parameters. The resulting family has a variety of forms ranging from the sinusoid to well-known Dirichlet kernel.

1 Introduction

Les signaux à sommes d'impulsions se rencontrent dans de nombreuses applications (radar, imagerie médicale...). La forme de l'impulsion étant connue on peut reconstruire le signal en connaissant seulement les retards et les amplitudes des impulsions par unité de temps. C'est ce qui a été montré récemment dans le cadre de l'échantillonnage de signaux FRI introduit en [1] qui utilise un noyau proprement choisi dans ce but [2]. La famille des noyaux somme de sinus cardinaux présentée en [3] est une famille de noyaux paramétrables remplissant les conditions Strang-fix nécessaires à cet échantillonnage (explicitées en [2, 3]).

Dans cet article nous développons les performances de l'estimation des signaux à sommes d'impulsions avec l'utilisation du noyau SoS. Nous calculons pour cela la BCRB [4], qui est une borne inférieure de l'erreur quadratique moyenne d'un estimateur non-biaisé. Notre objectif est de choisir le meilleur paramétrage du noyau SoS au regard de la BCRB, ce qui est obtenu en maximisant l'information de Fisher. Le choix d'un critère Bayésien plutôt que déterministe permet de trouver le noyau optimal indépendamment du tirage des paramètres. Les

noyaux optimaux obtenus sont paramétrables par l'utilisateur afin de choisir l'importance relative de chacun des paramètres à estimer, ainsi que la robustesse vis-à-vis du bruit.

Les notations suivantes sont utilisées dans cet article : \mathbf{e}_m est le vecteur avec un 1 au m^{eme} élément et des zéros ailleurs. $\mathbb{1}_L$ est une matrice carrée $L \times L$ uniquement composée de 1. $\text{tr}(\cdot)$ est la trace d'une matrice. Le produit scalaire est défini par $\langle g(t), x(t) \rangle = \int_{-\infty}^{\infty} g(t)x(t)dt$.

2 Présentation du modèle

2.1 Signaux sommes d'impulsions

Le signal qui nous intéresse est un signal continu composé d'un nombre fini de L impulsions, pondérées et décalées en temps :

$$x(t) = \sum_{l=0}^{L-1} a_l h(t - \tau_l) \quad (1)$$

où $\boldsymbol{\tau} = [\tau_0, \dots, \tau_{L-1}]^T$ et $\mathbf{a} = [a_0, \dots, a_{L-1}]^T$ sont tous les paramètres à estimer pour pouvoir reconstruire parfaitement ce signal, ils correspondent à l'ensemble des retards et des amplitudes de chaque impulsion. $h(t)$ est une fonction symétrique définissant la forme de l'impulsion. $x(t)$ fait partie de la classe

Ce travail est soutenu par les projets NEWCOM# (Grant agreement no. 318306), ANR MAGELLAN et MI-CNRS Titan.

de signaux FRI (pour *Finite Rate of Innovation*) car il est totalement déterminé par seulement un nombre fini de paramètres par intervalle temporel, ici $2L$ sur le temps de mesure fini considéré.

Nous étudions le problème de l'estimation du vecteur de paramètres inconnus $\theta = [\boldsymbol{\tau}^T \mathbf{a}^T]^T$ basé sur un nombre fini d'échantillons N , échantillonnés uniformément avec un intervalle d'échantillonnage T_S . Nous nous intéressons au cas Bayésien où les paramètres d'amplitudes a_l sont des variables aléatoires indépendantes et uniformément distribuées suivant une distribution centrée $p(a_l)$ de variance σ_a^2 , tandis que les paramètres de retards τ_l sont indépendamment et uniformément distribués suivant une distribution de gaussienne généralisée (notée GN) de paramètre de position nul, de paramètre d'échelle $\alpha > 0$ et de paramètre de forme $\beta > 0$ notée $\tau_l \sim \mathcal{GN}(0, \alpha, \beta)$, on note σ_τ^2 la variance de $p(\tau_l)$. La distribution GN est une distribution intéressante car elle est générale, et englobe les distributions courantes laplaciennes ($\beta = 1$), gaussiennes ($\beta = 2$) et uniformes ($\beta \rightarrow \infty$) [5].

L'échantillonnage de $x(t)$ se fait après une étape de filtrage par un noyau $g(t)$ (voir le schéma Figure 2.1) :

$$y_n = \langle g(t - nT_S), x(t) \rangle + w_n = c_n + w_n \quad (2)$$

où w_n est un bruit blanc gaussien de moyenne nulle et de variance σ^2 .

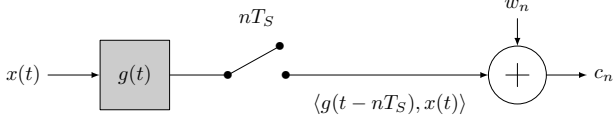


FIGURE 1 – Schéma de l'échantillonnage du signal de somme d'impulsions

2.2 Matrice d'information Bayésienne de Fisher

Les mesures (2) peuvent se mettre sous une forme vectorielle avec $\mathbf{y} = [y_1, \dots, y_N] = \mathbf{c} + \mathbf{w}$ où $\mathbf{c} = [c_1, \dots, c_N]$ et $\mathbf{w} = [w_1, \dots, w_N]$.

La matrice de Fisher Bayésienne (BIM) est définie comme l'espérance sur les paramètres d'intérêts de la matrice d'information de Fisher. Les échantillons conditionnés à un tirage donné suivent une distribution gaussienne $\mathbf{y}|\theta \sim \mathcal{N}(\mathbf{c}, \sigma^2 \mathbf{I}_N)$. D'après la formule de Slepian-Bang, la BIM est [4] :

$$\mathbf{J}(\theta) = \frac{1}{\sigma^2} E_\theta \left\{ \frac{\partial \mathbf{c}^T}{\partial \theta} \frac{\partial \mathbf{c}}{\partial \theta^T} \right\} - E_\theta \left\{ \frac{\partial^2 \ln p(\theta)}{\partial \theta \partial \theta^T} \right\}. \quad (3)$$

Quand la variance du bruit σ^2 est faible, le second terme est négligeable par rapport au premier [6].

3 Performances du noyau SoS

3.1 Calcul de la borne de Cramer Rao

Nous nous intéressons au noyau SoS défini dans [3] par

$$g(t) = \text{rect} \left(\frac{t}{NT_S} \right) \sum_{k=-K/2}^{K/2} b_k e^{\frac{j2\pi kt}{NT_S}}. \quad (4)$$

Ce noyau est particulièrement intéressant car on peut le moduler à l'aide du vecteur de poids $\mathbf{b} = [b_{-K/2}, \dots, b_0, \dots, b_{K/2}]^T$. Il est donc possible de choisir les poids optimisant une fonction de coût spécifique.

En mettant $g(t)$ sous forme matricielle, on peut réécrire les échantillons $\mathbf{c} = [c_1, \dots, c_N]$ comme

$$\mathbf{c} = \mathbf{V}(-\mathbf{t}_S) \mathbf{B} \mathbf{x}, \quad (5)$$

où $\mathbf{t}_S = \{nT_S : 0 \leq n \leq N-1\}$ est le vecteur des temps d'échantillonnage, $\mathbf{V}(-\mathbf{t}_S)$ est une matrice $N \times (K+1)$ dont le $(m, k)^{eme}$ élément est $e^{-j2\pi kn/N}$ et \mathbf{B} est une matrice diagonale de dimensions $K \times K$ ayant $\{b_k\}$ comme éléments diagonaux. \mathbf{x} est le vecteur des $K+1$ coefficients de la transformée de Fourier continue de $x(t)$. On peut l'écrire

$$\mathbf{x} = \mathbf{H} \mathbf{V}(\boldsymbol{\tau}) \mathbf{a}, \quad (6)$$

où $\mathbf{V}(\boldsymbol{\tau})$ est la matrice $(K+1) \times L$ dont le $(k, l)^{eme}$ élément est $e^{-j2\pi k\tau_l/(NT_S)}$, \mathbf{H} est la matrice diagonale de dimensions $(K+1) \times (K+1)$ dont le k^{eme} élément est $h_k = \frac{1}{NT_S} H \left(\frac{2\pi k}{NT_S} \right)$ où $H(w)$ est la transformée de Fourier continue de l'impulsion $h(t)$.

Finalement on a $\mathbf{c} = \mathbf{V}(-\mathbf{t}_S) \mathbf{B} \mathbf{H} \mathbf{V}(\boldsymbol{\tau}) \mathbf{a} \triangleq \boldsymbol{\Delta} \mathbf{V}(\boldsymbol{\tau}) \mathbf{a}$, dont les dérivées partielles relatives au paramètre sont $\frac{\partial \mathbf{c}}{\partial \tau_l} = a_l \boldsymbol{\Delta} \dot{\mathbf{g}}(\tau_l)$ et $\frac{\partial \mathbf{c}}{\partial a_l} = \boldsymbol{\Delta} \mathbf{g}(\tau_l)$ avec le vecteur $\mathbf{g}(\tau_l) = e^{\frac{j2\pi k\tau_l}{NT_S}}$ où $\mathbf{k} = [-\frac{K}{2}, \dots, \frac{K}{2}]^T$ et le vecteur $\dot{\mathbf{g}}(\tau_l) = \frac{\partial \mathbf{g}(\tau_l)}{\partial \tau_l}$. $\boldsymbol{\Delta}^H \boldsymbol{\Delta}$ est une matrice diagonale dont le k^{eme} élément est $N b_k^2 h_k^2$. En utilisant ce résultat et le fait que $p(a_l)$ suit une distribution centrée en zéro, les termes croisés s'annulent. De plus on a $[\mathbf{J}_\tau]_{ll'} = 0$ pour l différent de l' . Les termes restants de la FIM sont donc :

$$[\mathbf{J}_\tau]_{ll} = N \sigma_a^2 \sum_{k=-K/2}^{K/2} k^2 b_k^2 h_k^2, \quad [\mathbf{J}_\mathbf{a}]_{ll} = N \sum_{k=-K/2}^{K/2} b_k^2 h_k^2$$

$$[\mathbf{J}_\mathbf{a}]_{ll'} = N \sum_{k=-K/2}^{K/2} b_k^2 h_k^2 E_\tau \left\{ e^{\frac{j2\pi k\tau}{NT_S}} \right\} \cdot E_\tau \left\{ e^{-\frac{j2\pi k'\tau}{NT_S}} \right\}.$$

Si on prend une distribution approximativement uniforme entre $-N/2$ et $N/2$ pour les τ_l ($\beta \rightarrow \infty$), ce qui est le choix le plus naturel, on obtient :

$$\mathbf{J}_\tau \approx \left(\frac{\sigma_a^2 4\pi^2}{\sigma^2 NT_S^2} \sum_{k=-K/2}^{K/2} k^2 h_k^2 b_k^2 \right) \mathbf{I}_L$$

$$= \left(\frac{4\sigma_a^2 \pi^2}{\sigma^2 NT_S^2} \mathbf{b}^T \mathbf{D} \mathbf{b} \right) \mathbf{I}_L \quad (7)$$

$$\begin{aligned} \mathbf{J}_a &\approx \left(\frac{N^2}{\sigma^2} \sum_{k=-K/2}^{K/2} h_k^2 b_k^2 \right) \mathbf{I}_L + \frac{N}{\sigma^2} (h_0^2 b_0^2 (\mathbf{1}_L - \mathbf{I}_L)) \\ &= \frac{N^2}{\sigma^2} ((\mathbf{b}^T \bar{\mathbf{D}} \mathbf{b}) \mathbf{I}_L + h_0^2 b_0^2 (\mathbf{1}_L - \mathbf{I}_L)) \end{aligned} \quad (8)$$

où \mathbf{D} et $\bar{\mathbf{D}}$ sont des matrices diagonales avec $[\mathbf{D}]_{kk} = k^2 h_k^2$ et $[\bar{\mathbf{D}}]_{kk} = h_k^2$.

La BCRB est la trace de l'inverse de la BIM [6]. Avec la formule d'inversion de matrices de Sherman-Morison on obtient que la BCRB à haut RSB est

$$\text{BCRB}(\tau_l) \approx \frac{\sigma^2 N T_S^2}{4\pi^2 \sigma_a^2} \frac{1}{\sum_{k=-K/2}^{K/2} k^2 h_k^2 b_k^2} \quad (9)$$

$$\begin{aligned} \text{BCRB}(a_l) &\approx \frac{\sigma^2}{N \left(\sum_{\substack{k=-K/2 \\ k \neq 0}}^{K/2} h_k^2 b_k^2 \right)} \\ &\cdot \left(\frac{h_0^2 b_0^2 + (K-1) \sum_{\substack{k=-K/2 \\ k \neq 0}}^{K/2} h_k^2 b_k^2}{h_0^2 b_0^2 + K \sum_{\substack{k=-K/2 \\ k \neq 0}}^{K/2} h_k^2 b_k^2} \right). \end{aligned} \quad (10)$$

3.2 Optimisation du noyau SoS

La somme des BCRB (9) et (10) n'est pas une fonction convexe, c'est pourquoi au lieu de chercher à minimiser la BCRB nous allons maximiser l'information disponible relative à chaque paramètre. Cette information est mesurée par $\text{tr}(\mathbf{J}_\tau)$ et $\text{tr}(\mathbf{J}_a)$ respectivement. De plus pour pouvoir comparer les paramètres tirés de distributions différentes, nous normalisons la BIM : pour normaliser la BCRB on divise par la variance de chaque paramètre et par conséquent normaliser la BIM revient à multiplier \mathbf{J}_τ par σ_τ^2 et \mathbf{J}_a par σ_a^2 .

Un paramètre λ est introduit afin de régler l'importance relative attribuée à chaque paramètre. Pour $\lambda = 0.5$ chaque BIM normalisée a la même importance. De toutes ces considérations découle la fonction objective suivante :

$$f_\lambda(\mathbf{b}) = \lambda \sigma_\tau^2 \text{tr}(\mathbf{J}_\tau) + (1 - \lambda) \sigma_a^2 \text{tr}(\mathbf{J}_a) \quad (11)$$

comme les éléments diagonaux sont tous égaux on a :

$$f_\lambda(\mathbf{b}) = \mathbf{b}^T \Delta_\lambda \mathbf{b}. \quad (12)$$

avec

$$\Delta_\lambda = \text{RSBL} \left(\lambda \frac{\sigma_\tau^2 4\pi^2}{N T_S^2} \mathbf{D} + (1 - \lambda) N^2 \bar{\mathbf{D}} \right) \quad (13)$$

où $\text{RSB} = \sigma_a^2 / \sigma^2$.

Remarquons que les matrices \mathbf{D} et $\bar{\mathbf{D}}$ sont des matrices positives définies, par conséquent cette fonction à optimiser est convexe.

Pour que le noyau SoS soit valable, il faut que les b_k soient supérieurs à zéro. Etant donné que seul le carré des coefficients est pris en compte dans la fonction objectif, cette contrainte est équivalente à la contrainte que $b_k \geq \epsilon > 0$ où le paramètre ϵ est

réglé par l'utilisateur. Il faut également contraindre la puissance disponible, car une solution simple consisterait à augmenter l'amplitude des coefficients à l'infini, c'est pourquoi nous introduisons également la contrainte que $\|\mathbf{b}\|^2 = K + 1$. Finalement, le problème d'optimisation que nous résolvons est :

$$\max_{\mathbf{b}} \mathbf{b}^T \Delta_\lambda \mathbf{b} \text{ s.t. } \begin{cases} \|\mathbf{b}\|^2 = K + 1 \\ \forall k \in [-K/2 \dots K/2] : b_k \geq \epsilon \end{cases} \quad (14)$$

Pour avoir une idée générale du type de solution que l'on obtient, nous considérons d'abord le problème simplifié pour $\epsilon = 0$. La fonction de Lagrange obtenue dans ce cas là est :

$$\forall \lambda \in [0, 1], \mathcal{L}_\lambda(\mathbf{b}, \alpha) = \mathbf{b}^T \Delta_\lambda \mathbf{b} - \alpha (\|\mathbf{b}\|^2 - (K + 1)). \quad (15)$$

Le vecteur \mathbf{b}^* optimal est obtenu en dérivant $\mathcal{L}_\lambda(\mathbf{b}, \alpha)$ par rapport à \mathbf{b} , la solution est le vecteur propre associé à la plus grande valeur propre de la matrice Δ_λ . Comme Δ_λ est une matrice diagonale et que $[\Delta_\lambda]_{-m} = [\Delta_\lambda]_m$, le vecteur optimal dépendant de λ \mathbf{b}_λ^* est donné par :

$$\mathbf{b}_\lambda^* = \sqrt{\frac{K+1}{2}} (\mathbf{e}_{m_\lambda^*} + \mathbf{e}_{-m_\lambda^*}) \quad (16)$$

car la valeur propre maximale de Δ_λ est de multiplicité 2. La contrainte $b_{-m} = b_m$ est donc respectée, et on a bien $\|\mathbf{b}_\lambda^*\|^2 = K + 1$.

Remarquons que si l'on considère que $h(t)$ est une impulsion de Dirac, les coefficients fréquentiels h_k sont tous égaux, ce qui a pour conséquence que \mathbf{J}_a ne dépend pas de \mathbf{b} . C'est pourquoi quand seules les amplitudes sont considérées ($\lambda = 0$) tous les noyaux possibles sont des solutions, dont en particulier le noyau de Dirichlet (où tous les b_k sont égaux).

Pour résoudre le problème d'optimisation (14) on utilise le package CVX [7], avec le changement de variable $\beta_k = b_k^2$. Le résultat des simulations montre que le noyau obtenu est similaire à (16) pour $\epsilon = 0$, mais avec un palier de niveau ϵ (voir Figure 2). Le pic obtenu est égal à $u = \sqrt{\frac{(1-\epsilon^2)(K+1)}{2}}$ pour $\epsilon \neq 1$ de telle sorte que $\sum_{-K/2}^{K/2} \beta_k = K + 1$.

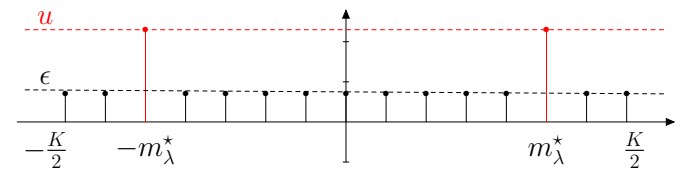


FIGURE 2 – Coefficients optimaux obtenus

Le noyau obtenu dans le domaine temporel (voir un exemple Figure 3) est :

$$g^*(t) = 2(u - \epsilon) \cos\left(\frac{2\pi m_\lambda^* t}{N T_S}\right) + \epsilon D_{K/2}\left(\frac{2\pi t}{N T_S}\right) \quad (17)$$

où $D_p(x) = \sum_{k=-p}^p e^{jkx}$ est la fonction de Dirichlet. Quand $\epsilon = 0$ on a une simple sinusoïde tandis que pour $\epsilon = 1$ on a le noyau de Dirichlet. En optimisant à partir de la BCRB, on minimise l'erreur théorique dans le régime asymptotique, le noyau optimal dans ce cas est obtenu pour $\epsilon \rightarrow 0$. Cependant,

augmenter ϵ permet d'augmenter la valeur du pic par rapport au niveau de la sinusoïde. Ce pic est nécessaire pour pouvoir distinguer le retard de la source, et le prendre infiniment petit rendrait la source indétectable en présence de bruit. Le choix d' ϵ va donc être naturellement dépendant du niveau de bruit considéré, plus ϵ est grand, plus le noyau sera robuste.

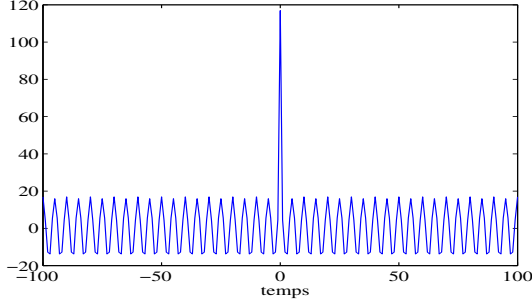


FIGURE 3 – Noyau optimal pour $N = K = 200$ et $\epsilon = 0.5$

4 Simulation

Pour les simulations nous prenons $N = 400$ échantillons, et une largeur de noyau de $K = 300$ avec un RSB de 50dB. En premier lieu nous nous intéressons au cas de l'échantillonnage de Diracs (avec $\epsilon = 0.5$) pour lesquels tous les coefficients en fréquence sont égaux. Pour cette raison la BCRB sur l'amplitude est constante et égale à 3.32×10^{-8} quel que soit le noyau choisi (Dirichlet ou optimisé indépendamment du λ . Quand $\lambda = 0$ et qu'on ne s'occupe pas des retards, le résultat de l'optimisation choisit un noyau parmi les autres possibles, qui donne une BCRB sur les retards similaire à celle du Dirichlet (5.35×10^{-10}). Dès que $\lambda > 0$, la BIM sur les retards décide du résultat de l'optimisation : comme le poids le plus important est associé à la fréquence la plus haute, le résultat est naturellement que $m^* = K$, noyau pour lequel on a $BCRB_\tau = 2.15 \times 10^{-10}$.

Dans la Figure 4 nous étudions le comportement de l'optimisation pour une impulsion gaussienne, dans le cas où $\epsilon = 0.1$ et $\epsilon = 0.5$. Le paramètre λ règle le compromis entre l'influence du terme h_k^2 qui a son pic en zéro et tend donc à rapprocher $m^* = K$ de zéro, et le terme $k^2 h_k^2$ qui augmente m^* . Plus λ est faible, plus les amplitudes sont prédominantes et plus proche sera m^* de zéro, dans ce cas la BCRB de l'amplitude optimisée est bien meilleure que celle du noyau Dirichlet tandis que celle des retards est moins bonne. Quand λ augmente, la BCRB sur les amplitudes augmente mais reste meilleure que celle du noyau Dirichlet, tandis que la borne sur les retards diminue. Un palier est atteint quand le choix de m^* ne change plus. Pour cette valeur, la BCRB sur les retards optimisée est meilleure que celle du noyau Dirichlet. Notons également que plus ϵ est faible, plus la borne est faible également puisque $\epsilon = 0$ est la valeur minimisant la borne, mais l'ordre de grandeur et l'allure de la courbe restent les mêmes.

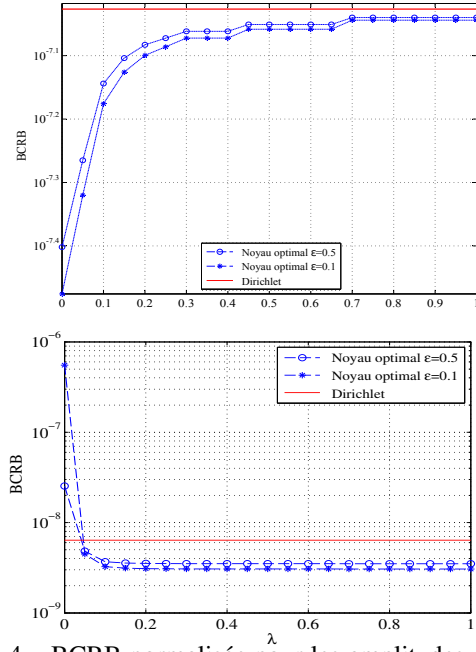


FIGURE 4 – BCRB normalisée pour les amplitudes en haut et les retards en bas

5 Conclusion

Dans cet article nous avons donné une expression analytique de la BCRB pour l'estimation des paramètres nécessaires à l'échantillonnage des signaux à somme d'impulsions. En nous basant sur la matrice d'information de Fisher, nous avons proposé une fonction objectif permettant de choisir un noyau SoS optimal laissant à l'utilisateur le choix de la relative importance des paramètres ainsi que la hauteur du pic permettant de distinguer le signal du bruit. Les simulations montrent que le noyau obtenu a de meilleures performances en terme de BCRB que le noyau de Dirichlet couramment choisi, pour des impulsions de Dirac et gaussiennes.

Références

- [1] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *Signal Processing, IEEE Transactions on*, vol. 50, no. 6, pp. 1417–1428, 2002.
- [2] P.L. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation : Shannon meets strang fix," *Signal Processing, IEEE Transactions on*, vol. 55, no. 5, pp. 1741–1757, May 2007.
- [3] R. Tur, Y. C. Eldar, and Z. Friedman, "Innovation rate sampling of pulse streams with application to ultrasound imaging," *Signal Processing, IEEE Transactions on*, vol. 59, no. 4, pp. 1827–1842, 2011.
- [4] S.M. Kay, *Fundamentals of Statistical Signal Processing : Estimation theory*, Fundamentals of Statistical Signal Processing. Prentice-Hall PTR, 1993.
- [5] S. Nadarajah, "A generalized normal distribution," *Journal of Applied Statistics*, vol. 32, no. 7, pp. 685–694, 2005.
- [6] H. L. Van Trees, *Detection, estimation, and modulation theory*, John Wiley & Sons, 2004.
- [7] M. Grant and S. Boyd, "CVX : Matlab software for disciplined convex programming, version 2.1," <http://cvxr.com/cvx>, Mar. 2014.