

CRAMÉR-RAO BOUND FOR FINITE STREAMS OF PULSES

Stéphanie Bernhardt, Rémy Boyer, and Sylvie Marcos
 Université Paris-Sud 11
 LSS-Supélec, France
 {bernhardt, rboyer, marcos}@lss.supelec.fr

Yonina C. Eldar
 Technion,
 Department of EE, Israel
 yonina@ee.technion.ac.il

Pascal Larzabal
 ENS-Cachan
 SATIE, France
 pascal.larzabal@satie.ens-cachan.fr

Abstract—Sampling a finite stream of filtered pulses violates the bandlimited assumption of the Nyquist-Shannon sampling theory. However, recent low rate sampling schemes have shown that these sparse signals can be sampled with perfect reconstruction at their rate of innovation. To reach this goal in the presence of noise, an estimation procedure is needed to estimate the time-delay and the amplitudes of each pulse. To assess the quality of any unbiased estimator, it is standard to use the Cramér-Rao Bound (CRB) which provides a lower bound on the Mean Squared Error (MSE) of any unbiased estimator. In this work, analytic expressions of the Cramér-Rao Bound are proposed for an arbitrary number of filtered pulses. Using orthogonality properties on the filtering kernels, an approximate compact expression of the CRB is provided. The choice of the kernel is discussed from the point of view of the estimation accuracy.

I. INTRODUCTION

In classical Nyquist-Shannon sampling theory a *bandlimited* signal can be perfectly reconstructed from its samples, at or above the Nyquist rate. However, in realistic applications, many signals of importance are non-bandlimited and thus the Nyquist-Shannon sampling theory assumption is not met [1], [2]. Finite streams of filtered pulses are an important class of signals since they appear in many applications including bio-imaging, radar, and spread-spectrum communication. Unfortunately, such signals are not bandlimited but fortunately they are sparse in the sense that only a small number of parameters per unit of time are needed to fully describe them. This is the key idea of the finite rate of innovation (FRI) framework introduced in [3]. More precisely, a finite stream of duration N samples of K filtered pulses can be described by $2K$ parameters (a time-delay and an amplitude per pulse) although it couldn't be sampled using Nyquist-Shannon's sampling theory.

In realistic scenarios, the sampling scheme must take into account noise perturbations in analog and digital domains. Analog (*resp.* digital) noise corrupts the signal before (*resp.* after) the uniform sampling [4]. To characterize the estimation performance, it is standard to use the Cramér-Rao Bound (CRB) which provides a lower bound on the Mean Squared Error (MSE) of any unbiased estimator [5]. In [6], the "continuous" CRB is derived for analog noise. The effect of digital noise on the recovery procedure was first analyzed in [7]. In [8], [9], the CRB for deterministic parameters is

derived but no analytic expression is proposed. In [4], [10], the CRB with the same assumptions (*i.e.* for deterministic parameters) is investigated in analytical form but only for a single pulse leading to a rate of innovation of $2/N$. In [11], the authors give an expression of the Fisher information matrix which still has to be numerically inverted.

In this work, we provide analytic expressions of the CRB in the case of digital noise and for an arbitrary number of pulses. We propose a simple approximation for the CRB which allows to easily compare the performance of the possible filtering kernels. We show that the amplitude estimation accuracy does not depend on the choice of the kernel, while the time-delay estimation accuracy depends on the norm of the first-order derivative of the kernel. We apply our results to the Sinc, Gaussian [3] and Sum of Sincs [12] kernels.

II. FINITE STREAMS OF MULTIPLE FILTERED PULSES

Consider a continuous-time signal with a finite number of weighted Diracs:

$$x(t) = \sum_{k=0}^{K-1} a_k \delta(t - \tau_k) \quad (1)$$

where $\boldsymbol{\tau} = [\tau_0, \dots, \tau_{K-1}]^T$ and $\mathbf{a} = [a_0, \dots, a_{K-1}]^T$ are the vectors of the unknown parameters called the time-delays and the amplitudes for each pulse respectively. We consider the problem of estimating the unknown parameters $\boldsymbol{\theta} = [\boldsymbol{\tau}^T \ \mathbf{a}^T]^T$ based on uniform sampling with a sampling interval T_S of a filtered/smoothed version of $x(t)$ according to

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

The real inner product is defined as $\langle g(t), x(t) \rangle = \int_{-\infty}^{\infty} g(t)x(t)dt$ and ϵ_n is a real discrete white Gaussian noise process having mean zero and variance σ^2 (see Fig. 1).

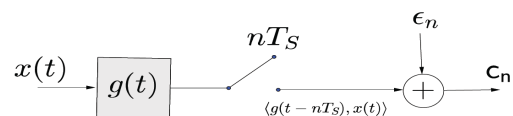


Fig. 1: Uniform sampling scheme corrupted by a digital noise

Considering an analysis duration of N samples, we have

$$\mathbf{c} = [c_0 \ \dots \ c_{N-1}]^T = \boldsymbol{\mu} + \mathbf{e} \quad (3)$$

This work was supported by the European Commission in the framework of the FP7 Network of Excellence in Wireless COMMunications NEW-COM# (Grant agreement no. 318306).

where

$$[\boldsymbol{\mu}]_n = \sum_{k=0}^{K-1} a_k g(\tau_k - nT_S)$$

and $\mathbf{e} = [\epsilon_0 \ \dots \ \epsilon_{N-1}]^T$. Given the measurements \mathbf{c} and the known filter g , it is possible to estimate the amplitude and the time-delay of each component of the signal $x(t)$ under appropriate conditions on $g(t)$. See [12] for details. Thus, without noise, the signal x can be perfectly reconstructed with only a small number of measurements [3], [7], [12]. In this paper we derive the CRB on the amplitude and time-delay estimation of finite rate innovation signals in the presence of digital noise.

III. DETERMINISTIC LOWER BOUND OF THE MSE

The Cramér-Rao Bound (CRB) is a lower bound on the MSE for any unbiased estimator $\hat{\boldsymbol{\theta}}(\mathbf{c})$ of $\boldsymbol{\theta}$ such as

$$\text{MSE} = E\{|\hat{\boldsymbol{\theta}}(\mathbf{c}) - \boldsymbol{\theta}|^2\} \geq \text{CRB} = \text{Tr}(\mathbf{C}(\boldsymbol{\theta})) \quad (4)$$

where $\mathbf{C}(\boldsymbol{\theta})$ is the inverse of the Fisher information matrix of parameters $\boldsymbol{\theta}$ and $\text{Tr}()$ is the trace operator. Observe that the mean of the observation $\boldsymbol{\mu} = E(\mathbf{c}) = \mathbf{G}\mathbf{P}^T\mathbf{P}\mathbf{a}$, with \mathbf{P} a permutation matrix, since $\mathbf{P}^T\mathbf{P} = \mathbf{I}$ where $\mathbf{G} = [\mathbf{g}_0, \dots, \mathbf{g}_{K-1}]$ with $\mathbf{g}_k = [g(\tau_k) \dots g(\tau_k - (N-1)T_S)]^T$ and $\mathbf{a} = [a_0 \ \dots \ a_{K-1}]^T$. For deterministic amplitudes and real Gaussian noise, the inverse of the Fisher information matrix for permuted vector $\boldsymbol{\theta}^{(p)} = (\mathbf{I}_2 \otimes \mathbf{P})\boldsymbol{\theta}$ (where \otimes stands for the Kronecker product and \mathbf{I}_2 is the 2×2 identity matrix), is given by [5]

$$\mathbf{C}(\boldsymbol{\theta}^{(p)}) = \sigma^2 \left(\left(\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}^{(p)}} \right)^T \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}^{(p)}} \right)^{-1} \quad (5)$$

$$= \sigma^2 (\mathbf{I}_2 \otimes \mathbf{P}) (\mathbf{B}^T \mathbf{B})^{-1} (\mathbf{I}_2 \otimes \mathbf{P}^T) \quad (6)$$

where $\mathbf{B} = [\dot{\mathbf{G}}\mathbf{D} \ \mathbf{G}]$ with \mathbf{D} a diagonal matrix containing the amplitudes \mathbf{a} and $\dot{\mathbf{G}} = [\dot{\mathbf{g}}_0, \dots, \dot{\mathbf{g}}_{K-1}]$ with $\dot{\mathbf{g}}_k = [\dot{g}(\tau_k) \dots \dot{g}(\tau_k - (N-1)T_S)]^T$ where we note $\dot{g}(\tau_k - nT_S) = \frac{\partial g(\tau_k - nT_S)}{\partial \tau_k}$. Let \mathbf{e}_k be the vector having "1" at the k -th entry and zero otherwise. By choosing the permutation matrix according to $\mathbf{P}_k = [\mathbf{e}_k \ \mathbf{e}_1 \ \dots \ \mathbf{e}_{k-1} \ \mathbf{e}_{k+1} \ \dots \ \mathbf{e}_K]^T$ and using the inverse of the block-matrix $\mathbf{B}^T \mathbf{B}$, we obtain after straightforward calculus:

$$\text{CRB}(a_k) = \sigma^2 \left[(\mathbf{P}_k \mathbf{G}^T \mathbf{P}_k^\perp \mathbf{G} \mathbf{P}_k^T)^{-1} \right]_{11} \quad (7)$$

$$\text{CRB}(\tau_k) = \frac{1}{\text{SNR}_k} \left[(\mathbf{P}_k \dot{\mathbf{G}}^T \mathbf{P}_k^\perp \dot{\mathbf{G}} \mathbf{P}_k^T)^{-1} \right]_{11} \quad (8)$$

in which $\text{SNR}_k = a_k^2 / \sigma^2$, $\mathbf{P}_k^\perp = \mathbf{I} - \mathbf{P}_k = \mathbf{I} - \mathbf{G}(\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T$ is the orthogonal projector whose range is $\langle \mathbf{G} \rangle^\perp$ and $\mathbf{P}_k^\perp = \mathbf{I} - \mathbf{P}_k = \mathbf{I} - \dot{\mathbf{G}}(\dot{\mathbf{G}}^T \dot{\mathbf{G}})^{-1} \dot{\mathbf{G}}^T$ is the orthogonal projector whose range is $\langle \dot{\mathbf{G}} \rangle^\perp$.

Let us denote $\dot{\mathbf{G}}^{(k)}$ the matrix extracted from $\dot{\mathbf{G}}$ by removing the k -th column. Using the inverse of a block-matrix, we obtain

$$\left[(\mathbf{P}_k \dot{\mathbf{G}}^T \mathbf{P}_k^\perp \dot{\mathbf{G}} \mathbf{P}_k^T)^{-1} \right]_{11} = \frac{1}{\left\| \mathbf{P}_{[\dot{\mathbf{G}}^{(k)} \ \mathbf{G}]}^\perp \dot{\mathbf{g}}_k \right\|^2} \quad (9)$$

where $\mathbf{P}_{[\dot{\mathbf{G}}^{(k)} \ \mathbf{G}]}^\perp$ is the orthogonal projector whose range is $\langle [\dot{\mathbf{G}}^{(k)} \ \mathbf{G}] \rangle^\perp$. Consequently, the deterministic CRB for the k -th time-delay is given by

$$\text{CRB}(\tau_k) = \frac{1}{\text{SNR}_k} \frac{1}{\left\| \mathbf{P}_{[\dot{\mathbf{G}}^{(k)} \ \mathbf{G}]}^\perp \dot{\mathbf{g}}_k \right\|^2}. \quad (10)$$

The above expression provides a nice geometrical interpretation of the CRB. Indeed, consider the extended subspace $\langle [\dot{\mathbf{G}} \ \mathbf{G}] \rangle$. For the k -th pulse, remove vector $\dot{\mathbf{g}}_k$ from this subspace to obtain a deflated subspace $\langle [\dot{\mathbf{G}}^{(k)} \ \mathbf{G}] \rangle$ of dimension $2K - 1$. The CRB is then proportional to the norm of the projection of vector $\dot{\mathbf{g}}_k$ onto the orthogonal complement of this deflated subspace (a similar geometrical interpretation of the CRB can be found in [13]).

In a similar way, to obtain the CRB for the amplitudes denote $\mathbf{G}^{(k)}$ the matrix extracted from \mathbf{G} by removing the k -th column. Using the inverse of a block-matrix, we obtain

$$\left[(\mathbf{P}_k \mathbf{G}^T \mathbf{P}_k^\perp \mathbf{G} \mathbf{P}_k^T)^{-1} \right]_{11} = \frac{1}{\left\| \mathbf{P}_{[\mathbf{G}^{(k)} \ \dot{\mathbf{G}}]}^\perp \mathbf{g}_k \right\|^2} \quad (11)$$

Finally, the CRB for the k -th amplitude is

$$\text{CRB}(a_k) = \frac{\sigma^2}{\left\| \mathbf{P}_{[\mathbf{G}^{(k)} \ \dot{\mathbf{G}}]}^\perp \mathbf{g}_k \right\|^2} \quad (12)$$

where $\mathbf{P}_{[\mathbf{G}^{(k)} \ \dot{\mathbf{G}}]}^\perp$ is the orthogonal projector whose range is $\langle [\mathbf{G}^{(k)} \ \dot{\mathbf{G}}] \rangle^\perp$. Finally

$$\text{CRB} = \text{Tr}(\mathbf{C}(\boldsymbol{\theta})) = \sum_{k=0}^{K-1} \text{CRB}(\tau_k) + \sum_{k=0}^{K-1} \text{CRB}(a_k) \quad (13)$$

where the CRB for parameters τ_k and a_k is given in expressions (10) and (12). The derived CRB is given for an arbitrary number of pulses and generalizes the derivation given in [4], [8].

IV. APPROXIMATED CRB EXPRESSIONS

A. Sampling kernels

In this work, we study the sinc function, Gaussian [3] and the sum of sincs (SoS) kernels [12] which is defined as

$$g_{\text{SoS}}(t) = \text{rect}\left(\frac{t}{NT_S}\right) \sum_{l=-p}^p b_l e^{j\frac{2\pi l t}{NT_S}} \quad (14)$$

where $\text{rect}(t)$ denotes the rectangular function.

Equation (14) represents a class of kernels determined by the parameters $\{b_l\}_{l \in [-p, \dots, p]}$. We will name "SoS" the sum of sincs filters where the b_l 's form a rectangular window and "SoS Hamming" the filter where the b_l 's form a symmetric Hamming window (See (26) in [12]). To derive the CRB we need to express $\dot{g}(t) = \partial g(t) / \partial t$. Unfortunately, the first-order derivative of the SoS kernel does not exist. To circumvent this problem, we approximate $\text{rect}(t)$ using the Generalized Gaussian [14] function with a large shape parameter.

B. Orthogonality properties of the kernel and its first-order derivative

Let us assume that the kernel $g(t)$ and its first-order derivative $\dot{g}(t)$ verify the following properties:

$$\mathbf{G}^T \mathbf{G} \approx \mathbf{F} \quad (15)$$

$$\dot{\mathbf{G}}^T \dot{\mathbf{G}} \approx \mathbf{E} \quad (16)$$

$$\dot{\mathbf{G}}^T \mathbf{G} = \mathbf{G}^T \dot{\mathbf{G}} \approx \mathbf{0}_K \quad (17)$$

where $\mathbf{0}_K$ is the $K \times K$ null matrix, $[\mathbf{E}]_{kk'} = \gamma(\tau_k) = \|\dot{\mathbf{g}}_k\|^2$ for $k = k'$ and 0 otherwise and $[\mathbf{F}]_{kk'} = \tilde{\gamma}(\tau_k) = \|\mathbf{g}_k\|^2$ for $k = k'$ and 0 otherwise. At an intuitive level, property (15) is fulfilled for disjoint pulse supports. This means that the time-delays are assumed to be not too closely spaced. An interpretation of properties (16) and (17) seems hard to provide in a general context because they are closely related to the behavior of the first-order derivative of each kernel. But these properties will be validated in section B.2) for the considered kernels.

Let us denote

$$\gamma_k(\tau_1, \dots, \tau_K) = \left\| \mathbf{P}_{[\dot{\mathbf{G}}^{(k)} \mathbf{G}]}^\perp \dot{\mathbf{g}}_k \right\|^2$$

and

$$\tilde{\gamma}_k(\tau_1, \dots, \tau_K) = \left\| \mathbf{P}_{[\mathbf{G}^{(k)} \dot{\mathbf{G}}]}^\perp \mathbf{g}_k \right\|^2.$$

Using the above properties, we have

$$\mathbf{P}_{[\dot{\mathbf{G}}^{(k)} \mathbf{G}]}^\perp \approx \mathbf{I}_N - \dot{\mathbf{G}}^{(k)} \mathbf{E}^{(k)-1} \dot{\mathbf{G}}^{(k)T} - \mathbf{G} \mathbf{F}^{-1} \mathbf{G}^T,$$

$$\mathbf{P}_{[\mathbf{G}^{(k)} \dot{\mathbf{G}}]}^\perp \approx \mathbf{I}_N - \mathbf{G}^{(k)} \mathbf{F}^{(k)-1} \mathbf{G}^{(k)T} - \dot{\mathbf{G}} \mathbf{E}^{-1} \dot{\mathbf{G}}^T$$

where $\mathbf{E}^{(k)}$ and $\mathbf{F}^{(k)}$ are the $(K-1) \times (K-1)$ matrices extracted from \mathbf{E} and \mathbf{F} respectively by removing the k -th column and row. Using the above properties, we have

$$\begin{aligned} \mathbf{P}_{[\dot{\mathbf{G}}^{(k)} \mathbf{G}]}^\perp \dot{\mathbf{g}}_k &\approx \dot{\mathbf{g}}_k - \dot{\mathbf{G}}^{(k)} \mathbf{E}^{(k)-1} \underbrace{\dot{\mathbf{G}}^{(k)T} \dot{\mathbf{g}}_k}_{=0} \\ &\quad - \mathbf{G} \mathbf{F}^{(k)-1} \underbrace{\mathbf{G}^H \dot{\mathbf{g}}_k}_{=0} \approx \dot{\mathbf{g}}_k, \\ \mathbf{P}_{[\mathbf{G}^{(k)} \dot{\mathbf{G}}]}^\perp \mathbf{g}_k &\approx \mathbf{g}_k - \mathbf{G}^{(k)} \mathbf{F}^{(k)-1} \underbrace{\mathbf{G}^{(k)T} \mathbf{g}_k}_{=0} \\ &\quad - \dot{\mathbf{G}} \mathbf{E}^{-1} \underbrace{\dot{\mathbf{G}}^T \mathbf{g}_k}_{=0} \approx \mathbf{g}_k. \end{aligned}$$

Therefore, the norms of the above terms can be approximated in the following manner:

$$\gamma_k(\tau_1, \dots, \tau_K) \approx \gamma(\tau_k) = \|\dot{\mathbf{g}}_k\|^2, \quad (18)$$

$$\tilde{\gamma}_k(\tau_1, \dots, \tau_K) \approx \tilde{\gamma}(\tau_k) = \|\mathbf{g}_k\|^2. \quad (19)$$

1) CRB expressions:

Result 1. Given a kernel $g(t)$ under conditions (15), (16) and (17), the CRB can be approximated by the following expression:

$$\text{CRB} \approx \sum_{k=0}^{K-1} \frac{1}{\text{SNR}_k \|\dot{\mathbf{g}}_k\|^2} + \sigma^2 \sum_{k=0}^{K-1} \frac{1}{\|\mathbf{g}_k\|^2}. \quad (20)$$

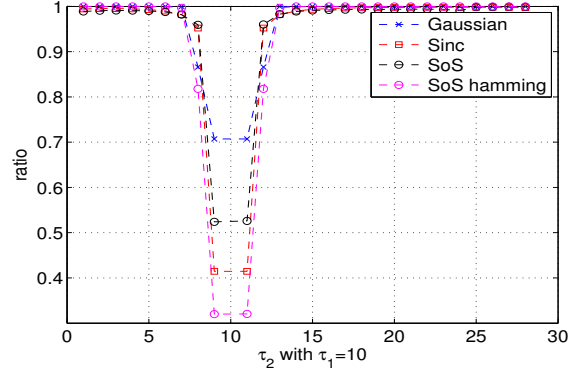


Fig. 2: Ratio $\gamma_1(\tau_1, \tau_2)/\gamma(\tau_1)$ for normalized kernels.

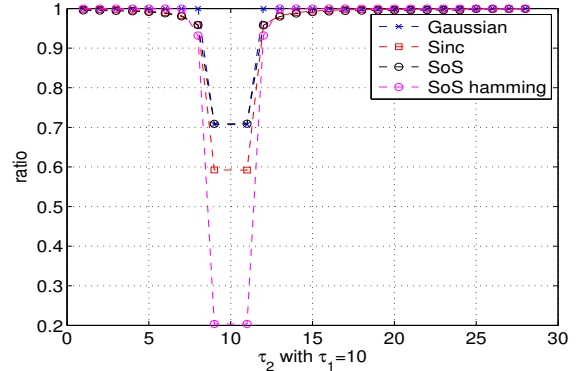


Fig. 3: Ratio $\tilde{\gamma}_1(\tau_1, \tau_2)/\tilde{\gamma}(\tau_1)$ for normalized kernels.

2) *Validity of the orthogonality properties:* To compare the kernels, we fix a same bandwidth of $B = 1/T_S$ for all the filters. The variance $\sigma_s^2 = \ln(2)/(\pi B)^2$ is chosen such that the 3dB bandwidth of the Gaussian kernel is equal to B . The bandwidth of the sum of sines kernel is directly related to the parameter p by $p = \lceil \frac{B\tau-1}{2} \rceil$.

We set $N = 30$ and $T_S = 1$ s and normalize the kernels such that:

$$\tilde{\gamma}(\tau_k) = \sum_{n=-\frac{N}{2}+1}^{\frac{N}{2}} g(nT_S)^2 = 1. \quad (21)$$

In Fig. 2 and 3, we have drawn ratios $\gamma_1(\tau_1, \tau_2)/\gamma(\tau_1)$ and $\tilde{\gamma}_1(\tau_1, \tau_2)/\tilde{\gamma}(\tau_1)$ for $\tau_1 = 10T_S$ and a varying $\tau_2 = nT_S$. Those ratios measure the relevance of the approximation, a ratio of one meaning that the approximate CRB equals the CRB. Both figures show that the ratio is always close to one for a wide range of time-delay values. So, we conclude that if the delays are spaced enough, the CRB of multiple pulses reduces to the CRB for one pulse and the orthogonality approximation can be made for all kernels (a similar observation has been made in [15]).

C. Independence on the time delay

In this section, we show that $\gamma(\tau_k)$ and $\tilde{\gamma}(\tau_k)$ can be approximated as functions independent of the time-delay τ_k . To reach this goal, we derive the first-order derivative of

$\gamma(\tau_k)$ which is given by

$$\left. \frac{\partial \gamma(\tau)}{\partial \tau} \right|_{\tau=\tau_k} = \sum_{n=0}^{N-1} f(\tau_k - nT_S)$$

where $f(t) = 2\dot{g}(t)\ddot{g}(t)$ in which $\dot{g}(t)$ is the derivative of $g(t)$ according to t .

We consider that the function $f(t)$ has the following properties:

- i) $f(t)$ is an odd function i.e. $f(t) = -f(-t)$ (in particular, note that $f(0) = 0$),
- ii) $f(t)$ has a finite support of length $(2N_S + 1)T_S$ i.e. $f(t) = 0$ for $|t| \geq N_S T_S$.

Considering $\tau_k = n_k T_S$, we have

$$\begin{aligned} \left. \frac{\partial \gamma(\tau)}{\partial \tau} \right|_{\tau=n_k T_S} &= \sum_{n=0}^{N-1} f((n_k - n)T_S) \\ &= f(0) + \sum_{n=0}^{n_k-1} f((n_k - n)T_S) + \sum_{n=n_k+1}^{N-1} f((n_k - n)T_S) \end{aligned}$$

Using property ii) and $N_S < n_k < N - N_S - 1$ and $f(0) = 0$, the sum becomes

$$\begin{aligned} &= \sum_{n=n_k-N_S}^{n_k-1} f((n_k - n)T_S) + \sum_{n=n_k+1}^{n_k+N_S} f((n_k - n)T_S) \\ &= \sum_{n=1}^{N_S} \underbrace{f(nT_S) + f(-nT_S)}_{=0 \text{ using property i}} = 0. \end{aligned}$$

Depending on the studied kernel, property ii) is only an approximation, however the norm of the residual defined as

$$\sum_{n=0}^{n_k-N_S-1} f((n_k - n)T_S)^2 + \sum_{n=n_S+1}^{N-1} f((n_k - n)T_S)^2$$

is assumed to be negligible. As the derivative of $\gamma(\tau_k)$ is approximately zero, we can conclude that $\gamma(\tau_k)$ depends weakly on τ_k under the condition that τ_k is away enough from the borders (i.e. $N_S < n_k < N - N_S - 1$). In this case, we will denote $\gamma(\tau_k) = \gamma_{\text{kernel}}$.

To illustrate our assumptions, we have plotted $f(t)$ in Fig. 4. We can check the validity of properties i) and ii) for the considered kernels.

The same methodology can be made concerning $\tilde{\gamma}(\tau_k)$ if we consider a function $f'(t) = 2g(t)\dot{g}(t)$ that shares the same properties as $f(t)$. Fig. 5 shows function $f'(t)$ and confirms that the assumptions are verified. Therefore, for time-delays away enough from the border, $\tilde{\gamma}(\tau_k) = \tilde{\gamma}_{\text{kernel}}$ does not depend on the time delay.

D. Expression of the approximate CRB

Under the above conditions, the CRB for the k -th parameters can be written:

$$\text{CRB}(\tau_k) \approx \frac{1}{\gamma_{\text{kernel}} \text{SNR}_k} \quad (22)$$

$$\text{CRB}(a_k) \approx \frac{\sigma^2}{\tilde{\gamma}_{\text{kernel}}} = \frac{a_k^2}{\tilde{\gamma}_{\text{kernel}} \text{SNR}_k} \quad (23)$$

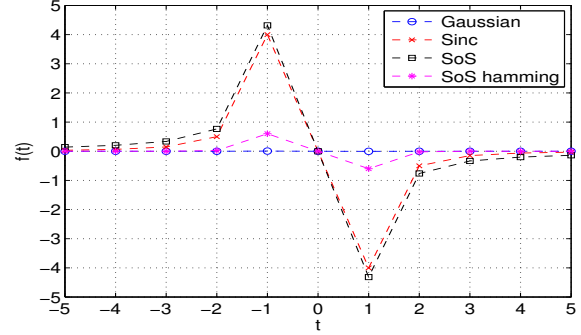


Fig. 4: $f(t)$ for the considered kernels.

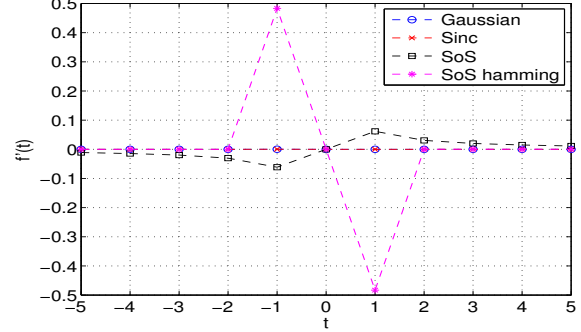


Fig. 5: $f'(t)$ for all the considered kernels

where γ_{kernel} and $\tilde{\gamma}_{\text{kernel}}$ only depend on the kernel function $g(t)$.

Result 2. For normalized kernels (i.e. following (21)), the CRB for the K amplitudes is linear in the noise variance, i.e., $\text{CRB}(a_k) = K\sigma^2$.

For the time-delay estimation, the larger the norm of the derivative of the kernel, the better the performance.

Result 3. Given a kernel following the orthogonality properties (16), (17) and (15) and supposing that the time delays are well spaced and away from the borders, the CRB can then be approximated by

$$\text{CRB} \approx \frac{K\sigma^2}{\tilde{\gamma}_{\text{kernel}}} + \frac{1}{\gamma_{\text{kernel}}} \sum_{k=0}^{K-1} \frac{1}{\text{SNR}_k}, \quad (24)$$

$$= \sum_{k=0}^{K-1} \frac{1}{\text{SNR}_k} \left(\frac{1}{\gamma_{\text{kernel}}} + \frac{a_k^2}{\tilde{\gamma}_{\text{kernel}}} \right). \quad (25)$$

V. NUMERICAL ILLUSTRATIONS

For all the simulations we use $N = 100$ samples and $T_S = 1$ and kernels presented in section IV-A. The signal $x(t)$ is composed of 3 weighted Diracs with $\tau = [20T_S, 60T_S, 85T_S]^T$ and $\mathbf{a} = [1, 2, 1]^T$. Fig. 6 plots the CRB for the Gaussian, sinc and SoS filters with first a rectangular window and second a hamming window. The abscissa shows the global SNR (which is defined by $\text{SNR} = \frac{\mathbf{a}^T \mathbf{G}^T \mathbf{G} \mathbf{a}}{\sigma^2}$) in dB. We first note that the CRB from (13) is exactly the same

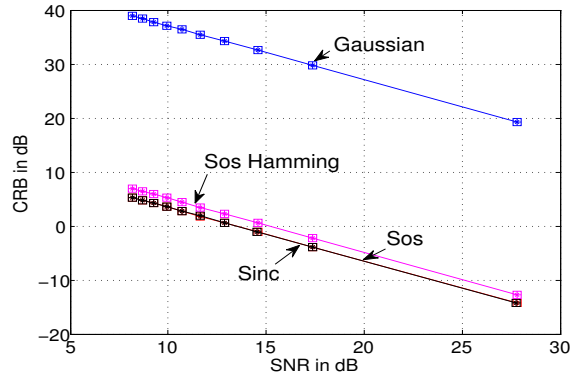


Fig. 6: CRB with well-spaced time-delays for the different kernels, the continuous line is the numerical CRB using (5), the star corresponds to the CRB using expression (13) and the square is obtained using (24)

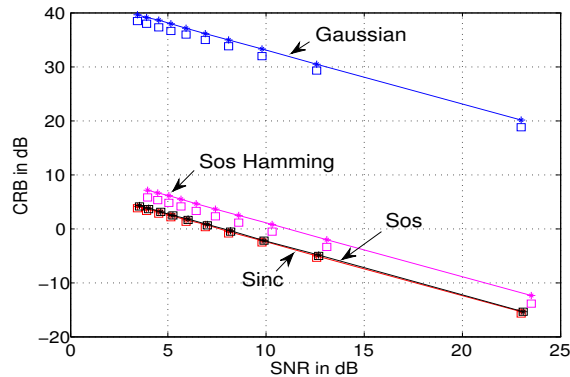


Fig. 7: CRB with closely spaced time-delays for the different kernels, the continuous line is the numerical CRB using (5), the star corresponds to the CRB using expression (13) and the square is obtained using (24)

as the CRB computed by inverting directly the FIM. The approximated CRB is very close to the numerical CRB which confirms that the different approximations are valid even for a large variety of kernels, and for a N relatively small.

The sinc kernel is the best kernel for a weighted Dirac signal. The SoS kernel with a rectangular window being an approximation of the sinc kernel has very close performances. The Gaussian kernel has bad results compared to the others for this type of input signal. This can be understood since the derivative of the Gaussian has short support and small amplitude, and therefore a small norm. The SoS kernel is very interesting since its parameters can be optimized to fit different type of signals. Choosing an appropriate window gives the same performances as the optimal kernel for a specific input signal (here the sinc kernel for a weighted sum of diracs).

Fig. 7 plots the CRB for time-delays which do not meet the orthogonality assumptions. We choose two closely spaced time-delays: $\tau = [20T_S, 22T_S]^T$ and $\mathbf{a} = [1, 1]^T$. In this case we see on Fig. 2 and Fig. 3 that the ratios are between 0.8 and 0.98 depending on the kernels. Consequently, the

approximated CRB is no longer valid but remains close to the CRB without the orthogonal approximations.

VI. CONCLUSION

In this work, we have proposed analytical expressions of the Cramér-Rao Bound (CRB) of unknown time-delays and amplitudes for finite streams of an arbitrary number of filtered pulses. Our analytical expressions are sufficiently general to encompass the important and difficult case of multiple pulses. This is a major difference with the existing contributions where only the single pulse case is derived in closed-form. We obtain several new results: We show how the CRB can be approximated by a very compact expression by exploiting the orthogonal properties of the kernels. A first interesting result is that for normalized kernels, the amplitude estimation accuracy does not depend on the kernel. A second result is that for the time-delays, the kernels with large first-order derivative norm lead to the best performance in the sense of the CRB.

REFERENCES

- [1] Y. Eldar and T. Michaeli, "Beyond bandlimited sampling," *Signal Processing Magazine, IEEE*, vol. 26, no. 3, pp. 48–68, May 2009.
- [2] M. Mishali and Y. Eldar, "Sub-nyquist sampling," *Signal Processing Magazine, IEEE*, vol. 28, no. 6, pp. 98–124, Nov 2011.
- [3] 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.
- [4] Y. C. Eldar and G. Kutyniok, *Compressed sensing: theory and applications*. Cambridge University Press, 2012.
- [5] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Upper Saddle River, New Jersey: Prentice-Hall, 1993.
- [6] Z. Ben-Haim, T. Michaeli, and Y. C. Eldar, "Performance bounds and design criteria for estimating finite rate of innovation signals," *Information Theory, IEEE Transactions on*, vol. 58, no. 8, pp. 4993–5015, 2012.
- [7] I. Maravic and M. Vetterli, "Sampling and reconstruction of signals with finite rate of innovation in the presence of noise," *Signal Processing, IEEE Transactions on*, vol. 53, no. 8, pp. 2788–2805, 2005.
- [8] T. Blu, P.-L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse sampling of signal innovations," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 31–40, 2008.
- [9] X. Wei, T. Blu, and P.-L. Dragotti, "Finite rate of innovation with non-uniform samples," in *Signal Processing, Communication and Computing (ICSPCC), 2012*. IEEE, 2012, pp. 369–372.
- [10] C. R. Comsa and A. M. Haimovich, "Performance bound for time delay and amplitude estimation from low rate samples of pulse trains," in *Signal Processing Conference (EUSIPCO), 2012*. IEEE, 2012, pp. 455–459.
- [11] A. Angierski and V. Kuehn, "Cramer-rao bound for sampling and reconstruction of fri signals," in *Acoustics, Speech and Signal Processing (ICASSP), 2014*. IEEE, 2014.
- [12] 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.
- [13] L. L. Scharf and L. McWhorter, "Geometry of the cramer-rao bound," *Signal Processing*, vol. 31, no. 3, pp. 301 – 311, 1993.
- [14] S. Nadarajah, "A generalized normal distribution," *Journal of Applied Statistics*, vol. 32, no. 7, pp. 685–694, 2005.
- [15] Y. Barbotin, A. Hormati, S. Rangan, and M. Vetterli, "Estimation of sparse mimo channels with common support," *Communications, IEEE Transactions on*, vol. 60, no. 12, pp. 3705–3716, December 2012.