

SAMPLING SIGNALS WITH FINITE RATE OF INNOVATION IN THE PRESENCE OF NOISE

Pier Luigi Dragotti and Felix Homann

Electrical and Electronic Engineering Department
Imperial College London,
Exhibition Road, London SW7-2AZ, UK
e-mail: {p.dragotti,felix.homann04}@imperial.ac.uk,

ABSTRACT

Recently, it has been shown that it is possible to sample non-bandlimited signals that possess a limited number of degrees of freedom and uniquely reconstruct them from a finite number of uniform samples. These signals include, amongst others, streams of Diracs. In this paper, we investigate the problem of estimating the innovation parameters of a stream of Diracs from its noisy samples taken with polynomial or exponential reproducing kernels. For the one-Dirac case, we provide exact expressions for the Cramér-Rao bounds of this estimation problem. Furthermore, we propose methods to reconstruct the location of a single Dirac that reach the optimal performance given by the unbiased Cramér-Rao bounds down to noise levels of 5dB.

Index Terms— Signal sampling, spline functions, spectral analysis, Cramér-Rao Bounds.

1. INTRODUCTION

Recently it was shown that it is possible to sample and perfectly reconstruct parametric non-bandlimited signals [6, 2]. Such signals are called signal with finite rate of innovation (FRI) since they are completely described by a finite number of free parameters.

The acquisition model used in these papers is the one depicted in Fig. 1, where the smoothing function $\varphi(t)$ is called the sampling kernel and normally models the distortion due to the acquisition device. The sampling kernel used in [6] is the sinc function, while the work in [2] uses compactly supported functions like for example polynomial splines (B-splines) [4] or exponential splines (E-splines) [5]. In both works it is shown that perfect reconstruction of classes of FRI signals from the measurements y_n is achievable by using a variation of Prony's method also known as annihilating filter method [3].

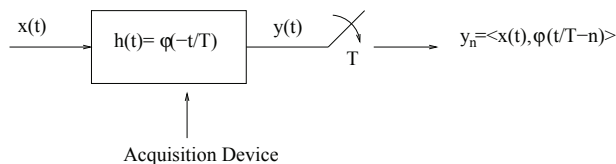


Fig. 1. Sampling setup. Here, $x(t)$ is the continuous-time signal, $h(t)$ the impulse response of the acquisition device and T the sampling period. The measured samples are $y_n = \langle x(t), \varphi(t/T - n) \rangle$.

It is now natural to ask whether these reconstruction algorithms are effective when the measurements are corrupted by noise and

which architecture is the most resilient to this form of distortion. Since the reconstruction of FRI signals is equivalent to a parametric estimation problem, one can compute the Cramér-Rao (CR) bounds to calculate the best possible achievable performance. A first analysis along these lines was done in [1]. That paper only considers the sinc kernel and an interesting conclusion is that the best CR bounds are achieved when the sinc with the maximum possible bandwidth is chosen. Constructive algorithms that reach these bounds are then proposed.

In this paper we extend the results of [1] to the family of kernels introduced in [2]. We compute the CR bounds for these architectures and demonstrate that the kernels with the shortest support have the tightest bounds. We then present a simple reconstruction strategy that achieves these bounds.

The paper is organized as follows: In the next section we review the sampling results of [2] and recall the main properties of the sampling kernels considered. In Sections 3 and 4 we evaluate the CR bounds for the case of a signal made of K Diracs and spline kernels. A comparison between the different architectures is also presented. In the following section, we discuss possible denoising strategies and present simulation results. We finally conclude in Section 6.

2. SAMPLING SIGNALS WITH FINITE RATE OF INNOVATION

For the sake of clarity we restrict our analysis to the case where the acquired signal $x(t)$ is a τ -periodic stream of K Diracs with amplitudes a_k located at distinct instants $t_k \in [0, \tau[$:

$$x(t) = \sum_{l \in \mathbb{Z}} \sum_{k=0}^{K-1} a_k \delta(t - t_k - l\tau). \quad (1)$$

Furthermore, we assume the sampling period is $T = \tau/N$. Consequently, the measurements are

$$\begin{aligned} y_n &= \langle x(t), \varphi(t/T - n) \rangle \\ &= \sum_{l \in \mathbb{Z}} \sum_{k=0}^{K-1} a_k \varphi(t_k/T - n - lN), \quad n = 0, 1, \dots, N-1. \end{aligned}$$

In [2], it was shown that with a proper choice of the acquisition kernel, it is possible to reconstruct $x(t)$ from the samples y_n exactly. The kernels used in [2] include:

- *Polynomial reproducing kernels:* Any kernel that satisfies

$$\sum_{n \in \mathbb{Z}} c_{m,n} \varphi(t - n) = t^m \quad m = 0, 1, \dots, P \quad (2)$$

for a proper choice of coefficients $c_{m,n}$.

- *Exponential reproducing kernels*: Any kernel that satisfies

$$\sum_{n \in \mathbb{Z}} c_{m,n} \varphi(t - n) = e^{\alpha_m t} \quad \begin{array}{l} \text{with } \alpha_m = \alpha_0 + m\lambda \\ \text{and } m = 0, 1, \dots, P \end{array} \quad (3)$$

for a proper choice of coefficients $c_{m,n}$.

B-splines belong to the family of kernels that can reproduce polynomials. A B-spline of order P is a function of compact support $L = P + 1$ and can reproduce polynomials up to degree P . It is obtained by the $(P + 1)$ -fold convolution of the zero order B-spline and has the following Fourier transform

$$\hat{\beta}_P(\omega) = \left(\frac{1 - e^{j\omega}}{j\omega} \right)^{P+1}.$$

The family of E-splines represents an extension of the polynomial splines and the Fourier transform of the P -th order E-spline is:

$$\hat{\beta}_{\bar{\alpha}}(\omega) = \prod_{m=0}^P \left(\frac{1 - e^{\alpha_m - j\omega}}{j\omega - \alpha_m} \right). \quad (4)$$

The above spline is able to reproduce the exponentials $e^{\alpha_m t}$, $m = 0, 1, \dots, P$. Notice that the exponent α_m in Eq. (4) can be complex which indicates that E-splines are usually complex functions. However, this can be avoided by choosing complex conjugate exponents. In Section 4.2, we use purely imaginary exponents equally spaced around the origin leading to the so called trigonometric E-splines. For example, for P even, we have $\alpha_m = -j\omega_0 P/2 + jm\omega_0$, $m = 0, 1, \dots, P$.

The reconstruction scheme of [2] operates as follows: First the samples are linearly combined with the coefficients $c_{m,n}$ of (2),(3) to obtain the new measurements

$$s_m = \sum_{n=0}^N c_{m,n} y_n \quad m = 0, 1, \dots, P. \quad (5)$$

Then, if the original signal is a stream of Diracs as the one in (1), one can show that $s_m = \sum_{k=0}^{K-1} a_k u_k^m$, where $u_k = t_k/T$ when polynomial splines are used and $u_k = e^{\lambda t_k/T}$ when exponential splines are involved¹. In either cases, the pairs of unknowns $\{a_k, u_k\}$ can be retrieved from the power series $s_m = \sum_{k=0}^{K-1} a_k u_k^m$ using the classical Prony's method. The key ingredient of this method is the annihilating filter. Call h_m , $m = 0, 1, \dots, K$ the filter with z -transform $H(z) = \sum_{m=0}^K h_m z^{-m} = \prod_{k=0}^{K-1} (1 - u_k z^{-1})$. That is, the roots of $H(z)$ correspond to the locations u_k . It clearly follows that

$$h_m * s_m = \sum_{i=0}^K h_i s_{m-i} = \sum_{i=0}^K \sum_{k=0}^{K-1} a_k h_i u_k^{m-i} = 0. \quad (6)$$

The filter h_m is thus called annihilating filter since it annihilates the observed series s_m . Moreover, the zeros of this filter uniquely define the set of locations u_k since the locations are distinct. The identity in (6) can be written in matrix/vector form as follows: $SH = 0$ which reveals that the Toeplitz matrix S is rank deficient. By solving the above system, we retrieve the u_k 's and therefore the locations t_k . Given the locations, the weights a_k are then obtained by solving a system of linear equations. Notice that the problem can be solved only when $P \geq 2K - 1$.

We thus conclude that perfect reconstruction of a stream of Diracs is possible with either B-splines or E-splines. The reconstruction

¹For the sake of simplicity, we do not discuss the border effects here.

procedure is the same, the only difference is in the choice of the coefficients $c_{n,m}$, which depends on the properties of the chosen kernel. We now want to investigate whether these differences lead to different performances when the samples are corrupted by noise.

3. DERIVATION OF THE CRAMÉR-RAO BOUNDS

Any practical acquisition device introduces noise during the acquisition process. We assume the noise is introduced after sampling, consequently, the new measurements are

$$\hat{y}_n = \langle x(t), \varphi(t/T - n) \rangle + \epsilon_n, \quad n = 0, 1, \dots, N - 1,$$

where ϵ_n is assumed to be i.i.d. additive Gaussian noise with zero mean and variance σ^2 . We denote with

$$\Theta = (a_0, a_1, \dots, a_K, t_0, t_1, \dots, t_K)^T$$

the vector of the unknown parameters of $x(t)$ and our aim is to estimate Θ as precisely as possible from the noisy measurements. The performance of any unbiased estimator $\hat{\Theta}$ is lower bounded by the Cramér-Rao bound: $\text{var}(\hat{\Theta}) \geq I^{-1}(\Theta)$, where $I(\Theta)$ is the Fisher Information Matrix (FIM) defined as

$$I(\Theta) = E \left(\nabla l(\Theta) \nabla l(\Theta)^T \right)$$

and $l(\Theta)$ is the log-likelihood function. It is therefore of primary importance to evaluate this bound in order to analyze the performance of any reconstruction algorithm and to compare different architectures.

For simplicity we denote \hat{y}_n as follows:

$$\hat{y}_n = f(\Theta, n) + \epsilon_n, \quad n = 0, 1, \dots, N - 1.$$

Now notice that

$$p_{\hat{y}}(\hat{y}_n | \Theta) = p_{\epsilon}(\hat{y}_n - f(\Theta, n)).$$

Hence, using independency of the noise samples we have:

$$\begin{aligned} l(\theta) &= \ln P(\hat{y}_0, \hat{y}_1, \dots, \hat{y}_{N-1} | \Theta) \\ &= \sum_{n=0}^{N-1} \ln p_{\epsilon}(\hat{y}_n - f(\Theta, n)). \end{aligned}$$

Next, we compute the partial derivatives of the log-likelihood function with respect to the parameters θ_i . It is easy to show that

$$\frac{\partial l(\Theta)}{\partial \theta_i} = \frac{1}{\sigma^2} \sum_{n=0}^{N-1} \epsilon_n \frac{\partial f(\Theta, n)}{\partial \theta_i}. \quad (7)$$

Hence $\nabla l(\Theta) = \frac{1}{\sigma^2} \sum_{n=0}^{N-1} \epsilon_n \nabla f(\Theta, n)$ and the Fisher information matrix is given by:

$$\begin{aligned} I(\Theta) &= E \left(\nabla l(\Theta) \nabla l(\Theta)^T \right) \\ &= E \left(\frac{1}{\sigma^4} \sum_n \sum_m \epsilon_n \epsilon_m \nabla f(\Theta, n) \nabla f(\Theta, m)^T \right) \\ &= \frac{1}{\sigma^2} \sum_{n=0}^{N-1} \nabla f(\Theta, n) \nabla f(\Theta, n)^T, \end{aligned} \quad (8)$$

where we have used the linearity property of the expectation and the fact the noise is uncorrelated (independent). The Cramér-Rao bound is thus given by:

$$\text{CRB}(\Theta) = \sigma^2 \left(\sum_{n=0}^{N-1} \nabla f(\Theta, n) \nabla f(\Theta, n)^T \right)^{-1}.$$

4. EVALUATION OF THE CR BOUNDS

In order to evaluate the CR bounds we assume that $K = 1$. In this case, the parametric space is reduced to $\Theta = (a_0, t_0)^T$ and the noisy samples are given by:

$$\begin{aligned}\hat{y}_n &= \langle x(t), \varphi(t/T - n) \rangle + \epsilon_n \\ &= \sum_{l \in \mathbb{Z}} a_0 \varphi(t_0/T - n - lN) + \epsilon_n \quad n = 0, 1, \dots, N-1.\end{aligned}$$

We now evaluate and analyze the CR bounds for the case of B-splines and E-splines.

4.1. B-splines

Here $\varphi(t)$ is a B-spline of order $P > 0$ whose analytical expression is given by:

$$\varphi(t) = \frac{1}{P!} \sum_{l=0}^{P+1} \binom{P+1}{l} (-1)^l \left(t - l + \frac{P+1}{2} \right)_+^P, \quad (9)$$

where $(t)_+^P = t^P$ when $t \geq 0$ and 0 otherwise.

The expression of the 2×2 FIM can be obtained by replacing Eq. (9) in (8). Unfortunately, the resulting equations are not easy to simplify especially for large values of P . We therefore first evaluate the bounds numerically. We assume that t_0 is uniformly distributed over the period and, for a fair comparison, adjust the elements of the FIM so that the kernels have unit norm for any choice of P and T . The results shown in Fig. 2 reveal that it is better to use short splines to minimize the uncertainty on the location. On the other hand, one can show that higher-order splines lead to a marginally more accurate evaluation of the amplitude of the Dirac². In many

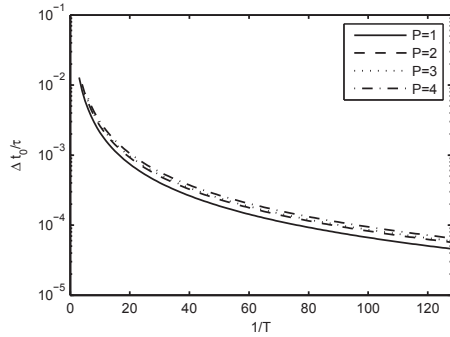


Fig. 2. Uncertainty on the location when sampling a single Dirac with B-splines of different orders. Recall that here $T = \tau/N$ and $\tau = 1$, $PSNR = a_0^2/\sigma^2 = 100$.

situations, it is more important to evaluate the location of the Dirac precisely. Therefore, we focus on the linear-spline case, i.e., $P = 1$, which leads to the best uncertainty for the location. Without loss of generality, we assume that $t_0 \in [0, T[$. The FIM reduces to:

$$I(\Theta) = \frac{1}{\sigma^2} \frac{1}{\|\varphi(t)\|^2 T} \begin{pmatrix} \frac{2a_0^2}{T^2} & \frac{a_0}{T} \left(\frac{2t_0}{T} - 1 \right) \\ \frac{a_0}{T} \left(\frac{2t_0}{T} - 1 \right) & \frac{2t_0^2}{T^2} - \frac{2t_0}{T} + 1 \end{pmatrix}. \quad (10)$$

²The figure is omitted due to the lack of space.

The Cramér-Rao bound is obtained by inverting the FIM and then by taking the expected value of t_0 over $[0, T[$. This leads to the following expression:

$$CRB(\Theta) = \sigma^2 \|\varphi(t)\|^2 T \begin{pmatrix} \frac{2}{3} \frac{T^2}{a_0^2} & 0 \\ 0 & 2 \end{pmatrix}.$$

The above bound yields the following uncertainty relationships:

$$\frac{\Delta t_0}{\tau} \geq \frac{2}{3N} \sqrt{\frac{\tau}{N}} \cdot PSNR^{-1/2}, \quad \frac{\Delta a_0}{|a_0|} \geq \frac{2}{\sqrt{3}} \sqrt{\frac{\tau}{N}} \cdot PSNR^{-1/2}, \quad (11)$$

where $PSNR = a_0^2/\sigma^2$.

4.2. E-splines

We now consider the case where the sampling kernel is an E-spline. In order to have fair comparisons, however, we restrict our attention to real-valued E-splines which are obtained when the complex exponents are chosen to be complex conjugate. In our analysis we further assume that the exponents are purely imaginary and equally spaced around the origin. Fig. 3 shows the uncertainty on the location with the trigonometric E-splines of order one to four. As in the B-spline case, we observe that shorter E-splines lead to better uncertainties on the location of the Dirac. It is also of interest to observe that performance improves when ω_0 increases. However, ω_0 cannot be chosen arbitrarily. There are two major constraints that have to be considered. First, the difference between any two distinct exponential coefficients cannot be a multiple of $j2\pi$, otherwise an E-spline stops being a basis. Second, to avoid ambiguities in the estimation of the location of the Dirac, $\omega_0 \leq 2\pi T/\tau$ for P even, $\omega_0 \leq \pi T/\tau$ for P odd and the difference between two complex parameters should not exceed $j2\pi$.

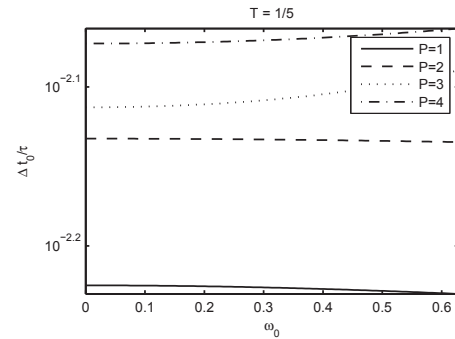


Fig. 3. Uncertainty on the location of a single Dirac when sampling with trigonometric E-splines of different orders.

Given that we are more interested in the estimation of the location of the Dirac, we now concentrate on the shortest real-valued E-spline. Namely a spline of order $P = 1$ with exponents $\alpha_1 = j\omega_0$ and $\alpha_{-1} = -j\omega_0$. Its analytical expression is given by:

$$\beta_{j\omega_0}(t) = \begin{cases} \frac{\sin \omega_0 t}{\omega_0} & 0 \leq t < 1, \\ -\frac{\sin \omega_0 (t-2)}{\omega_0} & 1 \leq t < 2. \end{cases}$$

By replacing the above expression in Eq. (8), we obtain the corresponding FIM. The CR bound is obtained by inverting the FIM and

by taking the expected value of t_0 over $[0, \tau]$. In this way, we obtain the following uncertainty relationships:

$$\frac{\Delta t_0}{\tau} \geq \frac{\omega_0 - \sin \omega_0 \cos \omega_0}{\omega_0^2 \sin \omega_0} \frac{1}{N} \sqrt{\frac{\tau}{N}} \cdot \text{PSNR}^{-1/2},$$

$$\frac{\Delta a_0}{|a|} \geq \frac{\sqrt{\omega_0^2 - \cos^2 \omega_0 + \cos^4 \omega_0}}{\omega_0 \sin \omega_0} \sqrt{\frac{\tau}{N}} \cdot \text{PSNR}^{-1/2}.$$

There are several elements in these uncertainty relationships that are worth mentioning. First, one can easily see that both uncertainties converge to those in Eq. (11) when $\omega_0 = 0$. This is expected since the E-spline converges to the corresponding linear B-Spline when $\omega_0 = 0$. Second, both uncertainties obtain a minimum at $\omega_0 = \pi/2$ and, in the interval $[0, \pi/2]$, both factors decrease monotonically with ω_0 . This is consistent with the plots shown in Fig. 3 where the CRB had been calculated numerically.

Since as discussed before $\omega_0 \leq \pi T = \pi \tau / N$, the exponential coefficient of the spline tends to zero when the sampling period decreases. In Fig. 4, we show the best bounds for the two different sampling kernels considered, together with the bound for the sinc kernel derived in [1]. All bounds scale in the same way with regard to the sampling period. However, as anticipated, the E-spline bounds converge to the ones computed for the B-splines as T decreases, since ω_0 tends to zero.

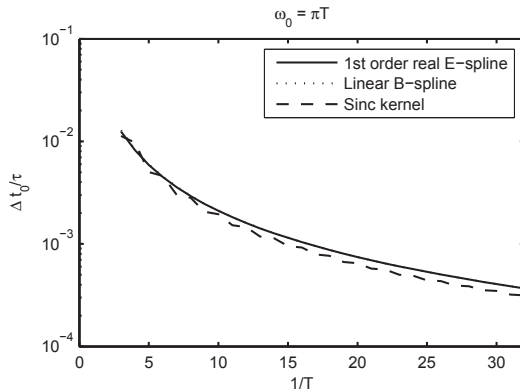


Fig. 4. The performances of the different sampling kernels are very similar and decay in the same way with N . Recall that $T = \tau/N$.

As a consequence of the fact that splines have compact support L , the above analysis can be extended to the case $K > 1$ if we assume that the minimum distance between the two closest Diracs is larger than LT . In fact in this case we are guaranteed that each Dirac influences a different set of samples and for this reason the uncertainties derived for a single Dirac hold exactly. The more complicated case of closely spaced Diracs is under investigation.

5. SIMULATION RESULTS

If the sampling kernel is a B-spline or an E-spline, we can use the fact that the kernel has compact support to reduce the noise in the signal. As mentioned before, if we sample a single Dirac, we get exactly L non-zero samples where L is the support of the kernel. Our hard thresholding variant exploits this fact to remove part of the noise in the sense that we first detect the indices of the non-zero signal samples and then set to zero all other non-zero samples as they are assumed to be entirely due to additive noise. More precisely, given

the noisy samples we compute the moving sum of L consecutive samples wrapping around to the beginning of the array if necessary. We compute the absolute value of the elements of the resulting array and determine its maximum. Since the kernel is non-negative, the maximum corresponds to the element that is the sum of all L nonzero signal samples for moderate noise levels. Once the location of the maximum is found we keep the $L - 1$ samples in the immediate neighborhood of the maximum and set to zero the others.

When $P > 1$ we also use Cadzow's method [1] to further de-noise the moments s_m . In Fig. 5 we show the results for the case of a real-valued E-spline and $P = 1, 2$. As revealed by the plots the case $P = 1$ leads to the best performance and achieves the CRB.

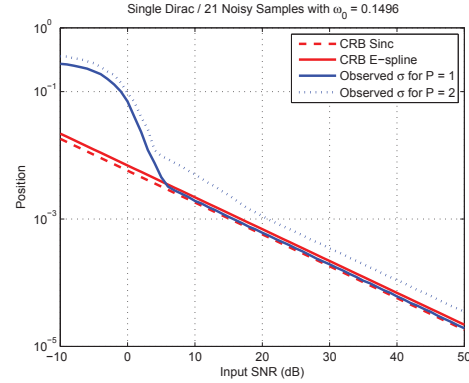


Fig. 5. Sampling a single Dirac with E-splines of order one and two. Hard-thresholding the samples obtained with a first order E-spline leads to optimal performance up to 5dB of SNR.

6. CONCLUSIONS

In this paper, we have analyzed the sampling of FRI signals in the presence of noise. For the case of a single Dirac, we have derived exact expressions for the CR bounds and showed that kernels with the shortest support lead to the best uncertainties. Then we have showed that a simple algorithm based on hard thresholding achieves the bound for a wide range of SNRs. In future, we aim to perform a precise analysis of the CR bounds when more than one Dirac is present in the signal.

7. REFERENCES

- [1] T. Blu, P.L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot. Sparse sampling of signal innovations: Theory, algorithms and performance bounds. *IEEE Signal Processing Magazine*, 25(2):31–40, March 2008.
- [2] P.L. Dragotti, M. Vetterli, and T. Blu. Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix. *IEEE Trans. on Signal Processing*, 55(5):1741–1757, May 2007.
- [3] P. Stoica and R. Moses. *Introduction to Spectral Analysis*. Englewood Cliffs, NJ, Prentice-Hall, 2000.
- [4] M. Unser. Splines: a perfect fit for signal and image processing. *IEEE Signal Processing Magazine*, 16(6):22–38, November 1999.
- [5] M. Unser and T. Blu. Cardinal Exponential Splines: Part I-theory and filtering algorithms. *IEEE Trans. on Signal Processing*, 53(4):1425–1438, April 2005.
- [6] M. Vetterli, P. Marziliano, and T. Blu. Sampling signals with finite rate of innovation. *IEEE Trans. Signal Processing*, 50(6):1417–1428, June 2002.