ON THE EXPONENTIAL REPRODUCING KERNELS FOR SAMPLING SIGNALS WITH FINITE RATE OF INNOVATION

Jose Antonio Urigüen[†], *Pier Luigi Dragotti*[†] *and Thierry Blu*[*]

[†]Imperial College of London
[*]The Chinese University of Hong Kong

jose.uriguen08@imperial.ac.uk, p.dragotti@imperial.ac.uk, thierry.blu@m4x.org

ABSTRACT

The theory of Finite Rate of Innovation (FRI) broadened the traditional sampling paradigm to certain classes of parametric signals. In the presence of noise, the original procedures are not as stable, and a different treatment is needed.

In this paper we review the ideal FRI sampling scheme and some of the existing techniques to combat noise. We then present alternative denoising methods for the case of exponential reproducing kernels. We first vary existing subspace-based approaches. We also discuss how to design exponential reproducing kernels that are most robust to noise.

Keywords— FRI, Sampling, Noise, Subspace, SVD

1. INTRODUCTION

Recently, Vetterli et al. demonstrated how certain classes of non-bandlimited signals can be sampled and perfectly reconstructed using the sinc and the Gaussian kernels [1]. These signals are completely determined by a finite number of degrees of freedom and are called signals with Finite Rate of Innovation (FRI). In [2], these results were extended to the case of sampling kernels with compact support and, in particular, to exponential reproducing kernels such as E-Splines [3]. In the presence of noise, however, these approaches become unstable. In [4] and [5] improved alternatives to the original methods were presented, focused on a subspace perspective for signal retrieval.

This paper focuses on the use of exponential reproducing kernels in the noisy scenario. Our contribution is twofold. First, we discuss variations of the algorithms considered in [5] when exponential reproducing kernels are involved. Second, we present a methodology to design exponential reproducing kernels that are most robust against noise.

The outline of the paper is as follows. In Section 2 we review the noiseless scenario of [2]. Then, in Section 3 we give an overview of the denoising techniques of [5]. We also introduce our modified procedures. In Section 4 we connect the Sum of Sincs kernel of [6] with the exponential reproducing kernels. Finally, in Section 5 we show simulation results, to then conclude in Section 6.

2. SAMPLING SIGNALS WITH FRI

For the sake of clarity we consider that $x(t)$ is a stream of K Diracs with amplitudes $\{a_k\}_{k=0}^{K-1}$ located at instants of time $\{t_k\}_{k=0}^{K-1} \in [0, \tau[$

$$x(t) = \sum_{k=0}^{K-1} a_k \delta(t - t_k). \tag{1}$$

We assume the sampling period is $T = \frac{\tau}{N}$. The measurements obtained sampling with $\varphi\left(-\frac{t}{T}\right)$, for $n = 0, 1, \ldots, N-1$, are

$$y_n = \left\langle x(t), \varphi\left(\frac{t}{T} - n\right)\right\rangle = \sum_{k=0}^{K-1} a_k \varphi\left(\frac{t_k}{T} - n\right). \tag{2}$$

In [1, 2] it was shown that, with a proper choice of the acquisition kernel, a perfect reconstruction of $x(t)$ from the samples y_n is possible. In this paper we concentrate on a specific class of kernels, used in [2], that are able to reproduce exponentials. An *exponential reproducing kernel* is any function $\varphi(t)$ that satisfies

$$\sum_{n\in\mathbb{Z}} c_{m,n}\varphi(t-n) = \mathrm{e}^{\alpha_m t} \quad \text{with } \alpha_m \in \mathbb{C}, \tag{3}$$

for a proper choice of coefficients $c_{m,n}$. The coefficients $c_{m,n}$ in the above equation are given by

$$c_{m,n} = \int_{-\infty}^{\infty} \mathrm{e}^{\alpha_m t}\tilde{\varphi}(t-n)\,\mathrm{d}t, \tag{4}$$

where $\tilde{\varphi}(t)$ is chosen to form with $\varphi(t)$ a quasi-biorthonormal set [2]. Note that the coefficients $c_{m,n}$ are discrete-time exponentials. This can be shown by making a change of variable in equation (4), which yields:

$$c_{m,n} = \int_{-\infty}^{\infty} \mathrm{e}^{\alpha_m x} e^{\alpha_m n}\tilde{\varphi}(x)\mathrm{d}x = \mathrm{e}^{\alpha_m n} c_{m,0}. \tag{5}$$

Exponential splines (E-Splines) [3] are central to the exponential reproduction property. The Fourier transform of the P-th order E-Spline is:

$$\hat{\beta}_{\vec{\alpha}_P}(\omega) = \prod_{m=0}^{P}\left(\frac{1-\mathrm{e}^{\alpha_m - j\omega}}{j\omega - \alpha_m}\right). \tag{6}$$

The above function is able to reproduce the exponentials $\mathrm{e}^{\alpha_m t}$, $m = 0, 1, \ldots, P$. Moreover, since the exponential reproduction formula is preserved through convolution [3], any composite function of the form $\gamma(t) * \beta_{\vec{\alpha}_P}(t)$ is also able to reproduce exponentials.

In the reconstruction scheme of [2] the samples y_n are first combined linearly with the coefficients $c_{m,n}$ to obtain the new measurements

$$s_m = \sum_{n=0}^{N-1} c_{m,n} y_n = \hat{a}_k u_k^m \tag{7}$$

for $m = 0, \ldots, P$, and where $\hat{a}_k = a_k \mathrm{e}^{\alpha_0 \frac{t_k}{T}}$ and $u_k = \mathrm{e}^{\lambda\frac{t_k}{T}}$. Here we have used the fact that the original signal is a stream of Diracs, we have combined equations (2) and (4), and used $\alpha_m = \alpha_0 + m\lambda$. The values s_m represent the exact exponential moments of the continuous-time signal $x(t)$.

Jose Antonio Urigüen is sponsored by the non-profit organisation "*Fundación Caja Madrid*"

Then, the new pairs of unknowns $\{\hat{a}_k, u_k\}_{k=0}^{K-1}$ can be retrieved from the power series in (7) using the classical Prony's method. The key ingredient of this method is the annihilating filter. Call $\{h_i\}_{i=0}^{K}$ the filter with z-transform $\hat{h}(z) = \sum_{i=0}^{K} h_i z^{-i} = \prod_{k=0}^{K-1}(1 - u_k z^{-1})$. That is, the roots of $\hat{h}(z)$ correspond to the locations u_k. It clearly follows that $h_m * s_m = 0$ [2]. 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 last identity can be written in matrix-vector form as follows:

$$\mathbf{S}\mathbf{h} = \mathbf{0} \tag{8}$$

which reveals that the Toeplitz matrix $\mathbf{S}$ is rank deficient. By solving the above system, we find the filter coefficients h_m and then retrieve the u_k's by computing the roots of $\hat{h}(z)$. Given the u_k's we can find the locations t_k. Finally, we obtain the weights a_k by solving, for instance, the first K consecutive equations in (7). Notice that the problem can be solved only when $P \geqslant 2K - 1$.

We thus conclude that perfect reconstruction of a stream of K Diracs is possible with any kernel able to reproduce exponentials, namely, with any kernel $\varphi(t)$ of the form $\varphi(t) = \gamma(t) * \beta_{\vec{\alpha}_P}(t)$.

3. SAMPLING AND RECONSTRUCTION IN THE PRESENCE OF NOISE

When noise is present in the acquisition process we do not have access to the ideal measurements. In contrast, we get the following set of samples:

$$\hat{y}_n = y_n + \epsilon_n. \tag{9}$$

The natural question we want to address is what is the best method to handle the noise effectively. As we explain in the following subsections, one way to control how noise affects the measurements is by studying the rank deficiency property of $\mathbf{S}$, shown by equation (8). An alternative to this involves designing a specific kernel in the family of exponential reproducing functions that is more resilient to noise. For both approaches we assume for simplicity that ϵ_n is a set of i.i.d. additive Gaussian measurements, with zero mean and variance σ^2.

3.1. Denoising: A subspace approach

Before discussing in detail other alternatives, we review briefly the denoising strategy that has been used successfully in [5].

First of all, because of noise, equation (8) is not satisfied any more. The reason is that, now, matrix $\mathbf{S}$ is perturbed and it becomes:

$$\hat{\mathbf{S}} = \mathbf{S} + \mathbf{B}. \tag{10}$$

However, it is reasonable to look for a solution that minimises $\|\hat{\mathbf{S}}\mathbf{h}\|^2$ under the constrain that $\|\mathbf{h}\|^2 = 1$. This is a classical total-least-square (TLS) problem that can be solved using Singular Value Decomposition (SVD).

The algorithm may be further improved by denoising $\hat{\mathbf{S}}$ before applying TLS. This is done by using the Cadzow iterative algorithm [7]. Cadzow algorithm is based on the fact that, in the absence of any perturbation, the matrix $\mathbf{S}$ is Toeplitz and rank deficient (i.e., it has rank K). When noise is present $\hat{\mathbf{S}}$ becomes full rank. So, in the first step of the Cadzow iteration an SVD of $\hat{\mathbf{S}}$ is performed leading to $\hat{\mathbf{S}} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}$, where $\mathbf{\Lambda}$ is a diagonal matrix. Then, only the first K diagonal elements of $\mathbf{\Lambda}$ are kept and $\hat{\mathbf{S}}$ is reconstructed. The new matrix $\hat{\mathbf{S}}$ is now rank deficient by construction, but is not Toeplitz anymore. This condition is then imposed by averaging the diagonal elements of $\hat{\mathbf{S}}$. The procedure is then iterated.

3.2. Modified Total Least Squares and Cadzow algorithm

The minimisation problem described so far was derived assuming that the perturbation in $\hat{\mathbf{S}}$ is due to additive white Gaussian noise. This is indeed the case in [5], because the properties of the noise added to the samples are preserved when their discrete Fourier transform is computed. This causes that matrix $\mathbf{B}$ in (10) to have a covariance matrix multiple of the identity matrix.

However, for exponential reproducing kernels, this assumption does not hold any more. More specifically, when the samples are corrupted by noise, equation (7) becomes

$$\hat{s}_m = \sum_{n=0}^{N-1} c_{m,n}\hat{y}_n = s_m + \sum_{n=0}^{N-1} c_{m,n}\epsilon_n \quad m = 0, \dots, P. \tag{11}$$

We can rewrite the above equation as follows:

$$\hat{\mathbf{s}} = \mathbf{s} + \hat{\mathbf{e}} = \mathbf{C}\mathbf{y} + \mathbf{C}\mathbf{e} \tag{12}$$

where $\mathbf{C}$ is the $(P+1) \times N$ matrix with the exponential reproducing coefficient $c_{m,n}$ at location (m, n), and $\mathbf{s}$ is a vector of length $P+1$.

As a consequence, the entries of the noise matrix $\mathbf{B}$ in (10), formed from the filtered noise $\hat{\mathbf{e}}$, no longer have the same variance. In order for SVD to provide a reliable separation of the signal and noise subspaces it is necessary to "pre-whiten" the noise. This is a well known approach proposed by various authors in the spectral estimation community (for instance by De Moor in [8]). It was also successfully used in the context of FRI in [4].

The key idea is that we need to know the covariance matrix of the noise $\mathbf{B}$ up to a constant factor λ [8]. In this situation, and assuming $\mathbf{R} = \lambda\mathbf{B}^*\mathbf{B}$ is positive definite, we can factor it as $\mathbf{R} = \mathbf{Q}^T\mathbf{Q}$ (using Cholesky decomposition). Then, we can still recover the appropriate subspaces by considering the SVD of $\hat{\mathbf{S}}' = \hat{\mathbf{S}}\mathbf{Q}^{-1}$. It is easy to see that the entries of the weighted noisy matrix $\hat{\mathbf{S}}'$ have now the same variance, which is the reason why the subspace division is again successful. This is not the only way to make the noise white (see for example [4]). In any case, once we modify the singular values of $\hat{\mathbf{S}}'$ to denoise the matrix, we need to revert the effect, reconstructing $\tilde{\mathbf{S}} = \hat{\mathbf{S}}'\mathbf{Q}$. The explicit use of $\mathbf{Q}^{-1}$ may result in inaccurate data calculations [8]. This can be avoided by using the quotient singular value decomposition (QSVD) of the pair $(\hat{\mathbf{S}}, \mathbf{Q})$.

To conclude, note that the modified TLS approach can be combined with a further enhancement, in a similar fashion as explained in the previous section. It is possible to do an SVD on $\hat{\mathbf{S}}'$, find the low rank approximation $\hat{\mathbf{S}}'_{\mathrm{LR}}$ by keeping the K largest singular values, and then compute the closest Toeplitz matrix to $\hat{\mathbf{S}}'_{\mathrm{LR}}\mathbf{Q}$. The process can then be iterated. We will call this the modified Cadzow procedure.

3.3. Modified E-Splines

As we mentioned before, it would be desirable for the noise term $\hat{\mathbf{e}}$ of (12) to be i.i.d. This is equivalent to saying that we want $\mathbf{R}_{\hat{\mathbf{e}},\hat{\mathbf{e}}} = \sigma^2\mathbf{I}$, and this can be achieved by making the rows of $\mathbf{C}$ orthonormal. The issue then is to find the exponential reproducing kernels that lead to the desired $\mathbf{C}$. Recall from (5) that for any $m = 0, 1, \dots, P$, we have that

$$c_{m,n} = \mathrm{e}^{\alpha_m n} c_{m,0}, \quad n = 0, 1, \dots, N-1. \tag{13}$$

Thus, orthogonality of the rows of $\mathbf{C}$ is achieved by choosing $\alpha_m = j\frac{2\pi m}{N}$. The norm of each row is then equal to $|c_{m,0}|$ and, by using (5), we have that

$$|c_{m,0}| = \left|\int_{-\infty}^{\infty} \tilde{\varphi}(t)\, \mathrm{e}^{j\frac{2\pi m}{N}t} \mathrm{d}t\right| = \left|\hat{\tilde{\varphi}}\left(\frac{2\pi m}{N}\right)\right|, \tag{14}$$

for $m = 0, 1, \ldots, P$, where $\hat{\tilde{\varphi}}(\omega)$ is the Fourier transform of the dual of $\varphi(t)$. Therefore, orthonormality is achieved when

$$\left|\hat{\tilde{\varphi}}\left(\frac{2\pi m}{N}\right)\right| = 1, \quad m = 0, 1, \ldots, P. \tag{15}$$

Since any exponential reproducing kernel $\varphi(t)$ can be written as $\varphi(t) = \gamma(t) * \beta_{\vec{\alpha}_P}(t)$, we have that $\hat{\varphi}(\omega) = \hat{\gamma}(\omega)\hat{\beta}_{\vec{\alpha}_P}(\omega)$. For the particular case that we are considering, the Fourier transform of $\beta_{\vec{\alpha}_P}(t)$ is given by

$$\hat{\beta}_{\vec{\alpha}_P}(\omega) = \prod_{m=0}^{P} \mathrm{e}^{-j\frac{\omega-\omega_m}{2}} \operatorname{sinc}\left(\frac{\omega-\omega_m}{2}\right), \tag{16}$$

where the E-Spline parameters satisfy $\alpha_m = j\omega_m = j\frac{2\pi m}{N}$. Moreover, it is well known (e.g. [9]) that the Fourier transforms of $\varphi(t)$ and of its dual $\tilde{\varphi}(t)$ are related as follows:

$$\hat{\tilde{\varphi}}(\omega) = \frac{\hat{\varphi}(\omega)}{\sum_{k\in\mathbb{Z}} |\hat{\varphi}(\omega + 2\pi k)|^2}, \tag{17}$$

which reveals that

$$\hat{\tilde{\varphi}}(\omega_m) = \frac{\hat{\varphi}(\omega_m)}{|\hat{\varphi}(\omega_m)|^2}. \tag{18}$$

Here we have used the fact that $\hat{\beta}_{\vec{\alpha}_P}(\omega_m + 2\pi k) = 0$ for $k \in \mathbb{Z}$ and $k \neq 0$ (see equation (16)). We thus conclude that it is possible to solve the original problem in (15) by just imposing

$$|\hat{\varphi}(\omega_m)| = |\hat{\gamma}(\omega_m)\hat{\beta}_{\vec{\alpha}_P}(\omega_m)| = 1$$
$$\Leftrightarrow \quad |\hat{\gamma}(\omega_m)| = |\hat{\beta}_{\vec{\alpha}_P}(\omega_m)|^{-1}. \tag{19}$$

Among all the admissible kernels satisfying (19), we are interested in the kernel with the shortest support. We therefore consider the kernels given by a linear combination of various derivatives of the original E-Spline $\beta_{\vec{\alpha}_P}(t)$, i.e.:

$$\varphi(t) = \sum_{i=0}^{P-1} d_i \beta^{(i)}_{\vec{\alpha}_P}(t), \tag{20}$$

where $\beta^{(i)}_{\vec{\alpha}_P}(t)$ is the ith derivative of $\beta_{\vec{\alpha}_P}(t)$, $\beta^{(0)}_{\vec{\alpha}_P}(t) = \beta_{\vec{\alpha}_P}(t)$, and d_i is a set of coefficients. This is like saying that $\gamma(t)$ is a distribution. These kernels are clearly still able to reproduce the exponentials and are a variation of the maximal-order minimal-support kernels introduced in [10, 11]. The advantage of this formulation is twofold: first the modified kernel $\varphi(t)$ is of minimum support $P+1$, the same as that of $\beta_{\vec{\alpha}_P}(t)$; second we only need to find the coefficients d_i to satisfy (19).

Using the Fourier transform of (20), which is given by:

$$\hat{\varphi}(\omega) = \hat{\beta}_{\vec{\alpha}_P}(\omega) \sum_{i=0}^{P-1} d_i (j\omega)^i, \tag{21}$$

we realise that we can satisfy (19) by choosing the coefficients d_i so that the resulting polynomial $\hat{\gamma}(\omega) = \sum_i d_i (j\omega)^i$ interpolates the set of points $(\omega_m, |\hat{\beta}_{\vec{\alpha}_P}(\omega_m)|^{-1})$ for $m = 0, 1, \ldots, P$.

One last consideration is in order here. When $j\omega_m = j\frac{2\pi m}{N}$, $m = 0, 1, \ldots, P$, the resulting E-Spline and modified kernels are complex-valued functions. This can be avoided by choosing

$$j\omega_m = \begin{cases} j\frac{2\pi}{N}(2m - P) & \text{when } P \text{ is odd} \\ j\frac{2\pi}{N}\frac{2m-P}{2}, & \text{when } P \text{ is even} \end{cases} \tag{22}$$

for $m = 0, 1, \ldots, P$. The conditions derived throughout the section are still valid for the new choices of parameters α_m. To conclude, in Figure 1 we present some of the kernels obtained by the above procedure.

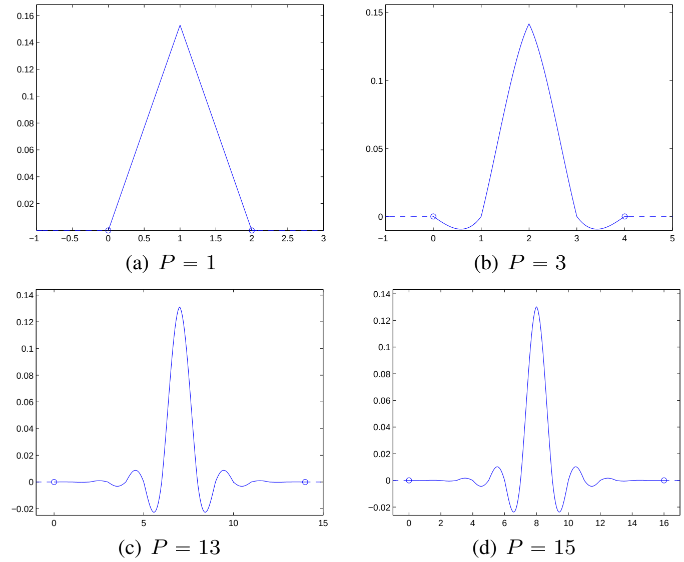


Fig. 1. *Examples of modified E-Splines*. These are four of the eight possible functions with P odd, support $P+1$ (indicated -o-) and $N = 32$ samples. The first one is identical to the original E-Spline of order $P = 1$. The last one is exactly half period of the Dirichlet kernel of period $2(P+1) = N$.

4. THE SOS: A MODIFIED E-SPLINE

Following the E-Spline modification presented in the previous section, we are now going to show that the family of Sum of Sincs (SoS) kernels introduced in [6] is a particular instance of exponential reproducing kernels. We restrict our analysis to P being even and the number of samples equal to $N = P+1$. A similar development applies when P is odd.

Let us consider the kernel $\varphi'(t) = \varphi\left(t + \frac{P+1}{2}\right)$, centred in zero, with $\varphi(t) = \gamma(t) * \beta_{\vec{\alpha}_P}(t)$, where $\beta_{\vec{\alpha}_P}(t)$ is defined to have Fourier transform (16). We relax condition (19) and now we allow

$$|\hat{\varphi}(\omega_m)| = |\hat{\gamma}(\omega_m)\hat{\beta}_{\vec{\alpha}_P}(\omega_m)| = b_m. \tag{23}$$

The next step is to use the periodic extension of $\varphi'(t)$, which can be written as follows:

$$b(t) = \sum_{l\in\mathbb{Z}} \varphi'(t + lN) = \frac{1}{P+1} \sum_{k\in\mathbb{Z}} \hat{\varphi}'\left(\frac{2\pi k}{P+1}\right) \mathrm{e}^{\frac{2\pi k}{P+1}t}, \tag{24}$$

where the last term follows from the application of Poisson summation formula, and $N = P+1$.

Now, note that the Fourier transform of the shifted kernel can be written using (16) as:

$$\hat{\varphi}'(\omega) = \gamma(\omega) \prod_{m=0}^{P} \operatorname{sinc}\left(\frac{\omega-\omega_m}{2}\right). \tag{25}$$

In (24) the Fourier transform $\hat{\varphi}'(\omega)$ is evaluated at $\omega_k = \frac{2\pi k}{P+1}$. If we use (22) for P even then we have $\omega_k = \omega_m$ for $\mathcal{K} = \{k : k = \frac{2m-P}{2}, m = 0, \ldots, P\}$. We have designed the filter $\gamma(\omega)$ so that (23) holds. As a consequence, whenever $k \in \mathcal{K}$, it follows that $\hat{\varphi}'(\omega_k) = b_k$. In contrast, for any $k \notin \mathcal{K}$, we have that $\hat{\varphi}'(\omega_k) = 0$ because we can find a term in the product of (25) equal to $\operatorname{sinc}(\ell\pi)$ with $\ell \in \mathbb{Z}$. Therefore, equation (24) can be reduced to:

$$b(t) = \frac{1}{P+1} \sum_{k=-\frac{P}{2}}^{\frac{P}{2}} b_k \mathrm{e}^{\frac{2\pi k}{P+1}t}. \tag{26}$$

If, now, we consider just one period of (26) and we use $t = \frac{x}{T}$, we get precisely the time domain definition of the SoS kernel:

$$b\left(\frac{x}{T}\right) = g(x) = \mathrm{rect}\left(\frac{x}{\tau}\right) \frac{1}{N} \sum_{k \in \mathcal{K}} b_k \mathrm{e}^{\frac{2\pi k}{\tau} x}. \qquad (27)$$

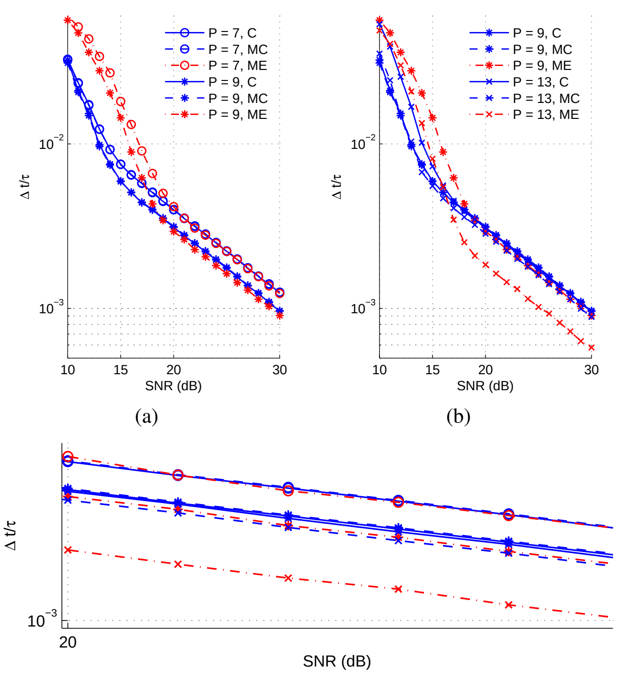
To conclude, note that when the values $b_k = 1, \forall k$, then equation (26) reduces to the Dirichlet kernel of period $N = P + 1$. This is an example of the modified E-Spline kernels with orthonormal rows of coefficients, and is therefore the SoS kernel most robust against noise.

5. SIMULATIONS

We have implemented the denoising approaches explained in Section 3 to test the retrieval of $K = 2$ Diracs. The samples provided by an exponential reproducing kernel are corrupted by additive white Gaussian noise (AWGN). We have added the noisy samples to the ideal measurements, calculating the variance of the noise according to the desired signal-to-noise ratio (SNR) we wanted to test.

For the original E-Spline, we have used $\alpha_m = j\frac{2\pi}{50}(2m - P)$, and for the modified E-Spline $\alpha_m = j\frac{2\pi}{N}(2m - P)$, where $N = 32$, and $m = 0, \ldots, P$. The kernels are, thus, real. We use the original and modified Cadzow algorithms for the E-Spline, and Cadzow for the modified E-Spline, iterating the routine 30 times.

The results, shown in Figure 2, are the average of 10000 realisations. They reveal, as expected, that the modified E-Spline kernels have a better performance than the original E-Splines, and this improves as the order P increases. On the other hand, the modified Cadzow algorithm, for low and intermediate orders, performs almost the same as the original subspace method. For the highest order ($P = 13$) it beats the best performance of the original algorithm, even though more marginally than the modified kernels.



(a) (b)

(c) Zoom, SNR=20–25dB

Fig. 2. *Retrieval of $K = 2$ Diracs in the presence of noise*. The continuous line, dashed line and dash-dotted line are the performances of the original E-Spline kernel with Cadzow algorithm (C), with the new Cadzow (MC) and of the modified E-Spline (ME) respectively.

6. CONCLUSIONS AND FUTURE WORK

In this paper we have reviewed the exponential reproducing kernels used to sample signals with Finite Rate of Innovation (FRI). We have also considered the noisy scenario for the same type of kernels. Our contribution is that we have adapted the denoising methods of [5] to the case of exponential reproducing kernels. In addition, we have presented a methodology to design exponential reproducing kernels that are the most resilient to noise. We have also connected the family of Sum of Sincs (SoS) kernels presented in [6] with that of exponential reproducing kernels.

Future work will consider the subspace perspective of the denoising algorithms presented in this paper more in depth. There exist alternatives to the Total Least Squares solution and Cadzow's iterative algorithm, as well as other approaches, which are relevant for further development.

7. REFERENCES

[1] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 50, 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," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 1741–1757, 2007.

[3] M. Unser and T. Blu, "Cardinal Exponential Splines: Part I—Theory and Filtering Algorithms," *IEEE Transactions on Signal Processing*, vol. 53, pp. 1425–1438, April 2005.

[4] I. Maravic and M. Vetterli, "Sampling and reconstruction of signals with finite rate of innovation in the presence of noise," *IEEE Transactions on Signal Processing*, vol. 53, pp. 2788–2805, 2005.

[5] T. Blu, P.-L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse Sampling of Signal Innovations," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 31–40, 2008.

[6] R. Tur and Y. C. Eldar, "Low Rate Sampling of Pulse Streams with Application to Ultrasound Imaging," *Submitted to IEEE Transactions on Signal Processing*, 2010.

[7] J. A. Cadzow, "Signal Enhancement – A Composite Property Mapping Algorithm," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 36, pp. 49–62, January 1988.

[8] B. De Moor, "The Singular Value Decomposition and Long and Short Spaces of Noisy Matrices," *Signal Processing, IEEE Transactions on*, vol. 41, pp. 2826–2838, September 1993.

[9] M. Unser, "Sampling-50 years after Shannon," in *Proceedings of the IEEE*, pp. 569–587, 2000.

[10] T. Blu, P. Thevenaz, and M. Unser, "MOMS: maximal-order interpolation of minimal support," *Image Processing, IEEE Transactions on*, vol. 10, pp. 1069–1080, July 2001.

[11] A. Ron, "Factorization theorems for univariate splines on regular grids," *Israel J. Math*, vol. 70, no. 1, pp. 48–68, 1990.