

Approximate Strang-Fix: Sampling Infinite Streams of Diracs with any Kernel*

Pier Luigi Dragotti¹, Jon Oñativia¹, Jose Antonio Urigüen¹ and Thierry Blu²

¹ Electrical and Electronic Engineering
Imperial College London, United Kingdom

² Chinese University of Hong Kong

ABSTRACT

In the last few years, several new methods have been developed for the sampling and the exact reconstruction of specific classes of non-bandlimited signals known as signals with finite rate of innovation (FRI). This is achieved by using adequate sampling kernels and reconstruction schemes. An important class of such kernels is the one made of functions able to reproduce exponentials.

In this paper we review a new strategy for sampling these signals which is universal in that it works with any kernel. We do so by noting that meeting the exact exponential reproduction condition is too stringent a constraint, we thus allow for a controlled error in the reproduction formula in order to use the exponential reproduction idea with any kernel and develop a reconstruction method which is more robust to noise.

We also present a novel method that is able to reconstruct infinite streams of Diracs, even in high noise scenarios. We sequentially process the discrete samples and output locations and amplitudes of the Diracs in real-time. In this context we also show that we can achieve a high reconstruction accuracy of 1000 Diracs for SNRs as low as 5dB.

1. INTRODUCTION

Sampling theory provides the bridge between the continuous and discrete-time domains and thus plays a central role in modern signal processing and communications. In the typical sampling setup, the original continuous-time signal $x(t)$ is filtered before being (uniformly) sampled with sampling period T . If we call $y(t) = h(t) * x(t)$ the filtered version of $x(t)$, the samples y_n are given by $y_n = \langle x(t), \varphi(t/T - n) \rangle$ where the sampling kernel $\varphi(t)$ is the scaled and time-reversed version of $h(t)$. This typical sampling set-up is shown in Fig. 1

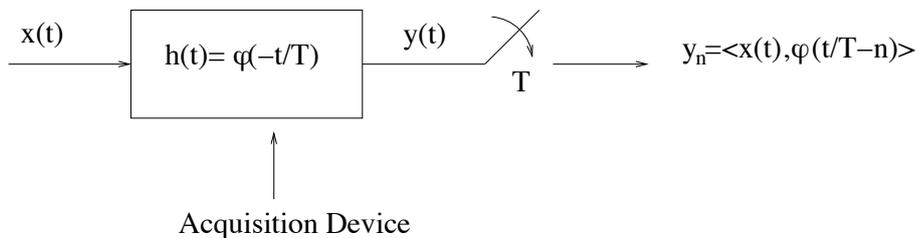


Figure 1. Sampling set-up. Here, $x(t)$ is the continuous-time signal, $h(t)$ the impulse response of the acquisition device and T the sampling period. The measured samples are $y_n = \langle x(t), \varphi(t/T - n) \rangle$.

Recently, it has been shown that it is possible to develop sampling schemes for classes of signals that are neither bandlimited nor belong to a fixed sub-space,²⁴ these signals are completely specified by a finite number of free parameters per unit of time and are called signals with finite rate of innovation (FRI).²⁴

Signals that can be perfectly reconstructed using this framework include: stream of Diracs and piecewise polynomial signals,^{9,24} piecewise sinusoidal signals³ and classes of 2-D signals.^{7,12,14,17,20} The sampling kernels

*The research leading to these results has received funding from the European Research Council under the European Unions Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement Nr. 277800 (RecoSamp).

considered in the original work by Vetterli et al.²⁴ were the Gaussian and the sinc functions. In^{8,9} these were extended to include any polynomial reproducing kernel such as B-splines or exponential reproducing kernels such as E-splines. Multichannel set-ups have been considered in^{1,19} and robust algorithms for the retrieval of FRI signals from noisy samples were presented in.^{4,15}

FRI sampling theory has also had impact in specific applications such as image super-resolution,² for depth sensing,¹³ for calcium transient detection¹⁶ and in compression.^{6,11}

In this paper, we present a new framework for the sampling of FRI signals using arbitrary kernels and also present a local algorithm for the robust reconstruction of infinite streams of Diracs. The paper is organised as follows: in the next section we provide an overview of the theory of sampling signals with FRI, we present the local reconstruction algorithm in Sec. 3 and the approximate FRI framework in Sec. 4. Simulation results are shown in Sec. 5 and we finally conclude in Sec. 6.

2. OVERVIEW OF FRI SAMPLING THEORY

Given the sampling set-up of Fig. 1, we want to retrieve $x(t)$ from the samples y_n . The acquisition device or sampling kernel plays a central role in this context and a family of kernels that has been successfully used in the past is the family of exponential reproducing functions. A function $\varphi(t)$ is an exponential reproducing function of order P , if together with its shifted versions, it is able to reproduce exponentials:

$$\sum_{n \in \mathbb{Z}} c_{m,n} \varphi(t - n) = e^{\alpha_m t}, \quad (1)$$

for proper coefficients $c_{m,n}$, with $m = 0, \dots, P$ and $\alpha_m \in \mathbb{C}$. It is possible to show that a function satisfies (1) if and only if it meets the generalised Strang-Fix conditions:²³

$$\hat{\varphi}(\alpha_m) \neq 0 \text{ and } \hat{\varphi}(\alpha_m + j2\pi l) = 0 \quad l \in \mathbb{Z} \setminus \{0\}$$

where $\hat{\varphi}(s)$ is the bilateral Laplace transform of $\varphi(t)$.

Exponential reproducing kernels are important because they allow us to map the samples y_n with the Laplace or Fourier transform of $x(t)$ at α_m $m = 0, 1, \dots, P$ and this independently of the property of the incoming signal. For the sake of clarity, assume that the signal $x(t)$ has compact support such that it is characterised by only N non-zero samples. Moreover, assume that $T = 1$. We thus have that the N samples are of the form $y_n = \langle x(t), \varphi(t - n) \rangle$, $n = 0, 1, \dots, N - 1$.

We now linearly combine the samples y_n using the coefficients $c_{n,m}$ of Eq. (1) to obtain:

$$\begin{aligned} s_m &= \sum_{n=0}^{N-1} c_{m,n} y_n \\ &\stackrel{(a)}{=} \langle x(t), \sum_{n=0}^{N-1} c_{m,n} \varphi(t - n) \rangle \\ &\stackrel{(b)}{=} \int_{-\infty}^{\infty} x(t) e^{\alpha_m t} dt, \quad m = 0, 1, \dots, P, \end{aligned}$$

where (a) follows from the linearity of the inner product and (b) is due to Eq. (1) and to the fact that $x(t)$ has compact support.

We note that $\int_{-\infty}^{\infty} x(t) e^{\alpha_m t} dt = \hat{x}(\alpha_m)$ is precisely the bilateral Laplace transform of $x(t)$ evaluated at α_m , $m = 0, 1, \dots, P$. Moreover, when α_m is purely imaginary, i.e. $\alpha_m = j\omega_m$, then the Laplace transform reduces to the Fourier transform. We can write the above equation in matrix/vector form as follows:

$$\mathbf{s} = \mathbf{C}\mathbf{y},$$

where \mathbf{s} is the column vector whose $(P + 1)$ entries correspond to the Fourier or Laplace transform of $x(t)$ at α_m , $m = 0, 1, \dots, P$; \mathbf{y} is the vector with the N samples and \mathbf{C} is the matrix of size $(P + 1) \times N$ whose entry at location (n, m) corresponds to the coefficient $c_{n,m}$ in (1).

When $x(t)$ is a specific class of signals with FRI and $\alpha_m = \alpha_0 + m\lambda$, it is then possible to establish a one-to-one mapping between $\hat{x}(\alpha_m)$ and $x(t)$. For example, we may assume that $x(t) = \sum_{k=0}^{K-1} x_k \delta(t - t_k)$ is a stream of K Diracs located at $t_k \in [0, N)$ then we have that

$$\begin{aligned} s_m &= \sum_{n=0}^{N-1} c_{m,n} y_n \\ &= \int_{-\infty}^{\infty} x(t) e^{\alpha_m t} dt, \\ &= \sum_{k=0}^{K-1} x_k e^{\alpha_m t_k} \\ &= \sum_{k=0}^{K-1} \hat{x}_k e^{\lambda m t_k} = \sum_{k=0}^{K-1} \hat{x}_k u_k^m, \quad m = 0, 1, \dots, P, \end{aligned}$$

where we have used the fact that $\alpha_m = \alpha_0 + m\lambda$ and set $\hat{x}_k = x_k e^{\alpha_0 t_k}$, $u_k = e^{\lambda t_k}$.

The quantity

$$s_m = \sum_{k=0}^{K-1} \hat{x}_k u_k^m, \quad m = 0, 1, \dots, P$$

is a sum of exponentials and retrieving the locations u_k and the amplitudes \hat{x}_k from $\{s_m\}_{m=0}^P$ is a classical problem in spectral estimation and was first solved by Gaspard de Prony in 1795.¹⁸ Prony's method allows the exact retrieval of $\{x_k, u_k\}_{k=0}^{K-1}$ using only $2K$ consecutive values s_m . This means that the parameters of a stream of K Diracs can be retrieved when $P \geq 2K - 1$. This condition therefore links the number of free parameters in the FRI signals to the order of the exponential reproducing kernels. Specifically, in this example, the signals is completely determined by $2K$ free parameters -the amplitudes and locations of the Diracs - and this dictates the condition on the order P .

We also note that, while the basic Prony's method requires only $2K$ values s_m , in the case of noisy samples, it is better to use more measurements. This is because when $P > 2K - 1$, it is then possible to use variations of Prony's method which are more resilient to noise. Two of the most successful techniques used for noisy FRI retrieval are Cadzow method⁵ and matrix pencil (e.g.,¹⁰). For an overview of spectral estimation methods we refer to the book.²¹

We also highlight the fact that while in the noiseless settings any exponential reproducing kernel can be used to retrieve the FRI signal, the behaviour of such kernels changes in the presence of noise. In,^{22,23} it was shown that functions reproducing exponentials with purely imaginary exponents (i.e., $\alpha_m = j\omega_m$) are more resilient to noise. Moreover, within this group of functions a subset called exponentials-MOMS (e-MOMS) are the best performing.^{22,23} An example of e-MOMS kernels is shown in Fig. 2.

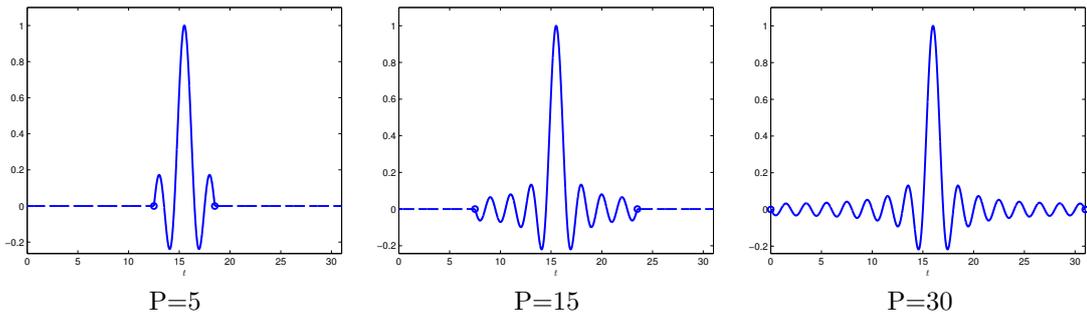


Figure 2. Examples of e-MOMS of different order.

3. SEQUENTIAL LOCAL FRI SAMPLING OF INFINITE STREAMS OF DIRACS

We now consider the case where we have an infinite train of Diracs:

$$x(t) = \sum_{k \in \mathbb{Z}} x_k \delta(t - t_k). \quad (2)$$

The signal $x(t)$ is assumed to have a finite local rate of innovation $2K/\tau$. This means that, if we consider a sliding window of size τ , the number of Diracs that we see inside the window is always at most K . We assume that this signal is sampled using e-MOMS and propose a sequential algorithm that estimates the locations of the Diracs in (2) by using a sliding window that sequentially covers intervals of size τ . The sliding window stepsize corresponds to the sampling period T . For consistency with the previous section we also assume that $\tau = N$ and that $T = 1$.

We thus develop a new strategy based on processing sets of N samples in sequential order. For each window and each group of N samples, we retrieve K Diracs using the algorithm of the previous section. We then store all the locations and amplitudes retrieved in that window. We then slide the window by one sample and repeat the process. When the found locations correspond to real Diracs, they will be consistent among different positions of the sliding window that capture these Diracs. Otherwise, we expect locations that are not correct and are due to noise to be inconsistent amongst windows. For example, in Fig. 3(a) we plot the retrieved locations for different windows. The horizontal axis represents the index of the window corresponding to a retrieved location, and the vertical axis the Dirac location in time. Consistent locations appear as horizontal alignments of dots, overlapping the blue lines. In order to detect consistency, in a second step we compute a histogram of detected locations. Only the peaks of the histogram are assumed to correspond to real Diracs. For a peak in the histogram above a certain threshold, the location of the corresponding Dirac is estimated averaging all the retrieved locations that contributed to that peak. This is illustrated in Fig. 3(b).

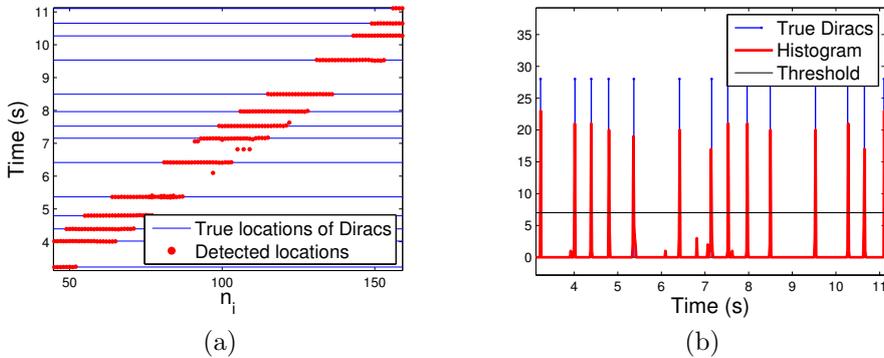


Figure 3. Local Reconstruction algorithm. Part (a) shows as red dots the Diracs retrieved at each window. Note that the x-axis corresponds to the window index. Part (b) shows the histogram of the retrieved locations.

4. APPROXIMATE STRANG-FIX

In the previous sections, we have implicitly assumed that we had a complete control on the design of the acquisition device and assumed that it was possible to have kernels behaving like e-MOMS. In this section we are going to extend the theory of FRI sampling in order to be able to use it with any acquisition device.

4.1 Approximate Reproduction of Exponentials

As highlighted previously, we require the sampling kernel to be able to reproduce exponentials in order to be able to map the samples y_n to the Fourier or Laplace transform of $x(t)$. Consider now an arbitrary kernel $\varphi(t)$, we want to find a linear combination of $\varphi(t)$ with its shifted versions that provides the best approximation to a specific exponential. More precisely, we want to find the coefficients c_n such that:

$$\sum_{n \in \mathbb{Z}} c_n \varphi(t - n) \approx e^{\alpha t}. \quad (3)$$

This approximation is exact only when $\varphi(t)$ satisfies the generalised Strang-Fix conditions, for any other function it is of particular interest to find the coefficients c_n that best fit (3). For the sake of clarity, we assume that

$c_n = c_0 e^{\alpha n}$ and we can show that, for any choice of c_n , the error in approximating $f(t) = e^{\alpha t}$ with the function $s(t) = \sum_{n \in \mathbb{Z}} c_n \varphi(t - n)$ is equal to:

$$\varepsilon(t) = f(t) - e^{\alpha t} = e^{\alpha t} \left[1 - c_0 \sum_{l \in \mathbb{Z}} \hat{\varphi}(\alpha + j2\pi l) e^{j2\pi l t} \right]. \quad (4)$$

Note that, if $\varphi(t)$ satisfies the generalised Strang-Fix conditions, then the error is zero when $c_0 = \hat{\varphi}(\alpha)^{-1}$. If $\varphi(t)$ does not satisfy the conditions but its Laplace transform decays sufficiently quickly, very few terms of the Fourier series expansion are needed to have an accurate bound for the error. Moreover, the error won't be zero but might be very small. We denote such kernels as approximate Strang-Fix kernels.

A natural choice of the coefficients $c_n = c_0 e^{\alpha n}$ is the one leading to the *least-squares approximation* which is obtained by computing the orthogonal projection of $f(t)$ onto the subspace spanned by $\varphi(t - n)$.

The least-squares approximation has the disadvantage that it requires exact knowledge of $\varphi(t)$. However, as we stated before, if the Laplace transform of $\varphi(t)$ decays sufficiently quickly, we can assume the terms $\hat{\varphi}(\alpha + j2\pi l)$ are close to zero for $l \in \mathbb{Z} \setminus \{0\}$. In this case we have that the error in (4) is easily minimised by choosing $c_0 = \hat{\varphi}(\alpha)^{-1}$. We denote this second type of approximation *constant least-squares*. Besides its simplicity, a second advantage of choosing $c_n = \hat{\varphi}(\alpha)^{-1} e^{\alpha n}$ is that it requires only the knowledge of the Laplace transform of $\varphi(t)$ at α .

In Fig. 4, we show an example where a linear spline is used to approximate four different exponentials. We note that linear splines are able to reproduce exactly polynomials up to degree one as shown in Fig. 4 (a)(b). The figures highlight the fact that the approximate reproduction of low-frequency exponential is very accurate and the accuracy reduces when trying to approximate higher-frequency exponentials. We also note that, in theory, the number of exponentials we reproduce using the approximate framework is arbitrary, this is in contrast with the exact reproduction framework where the number of exponentials is fixed and depend on the order P of the kernel.

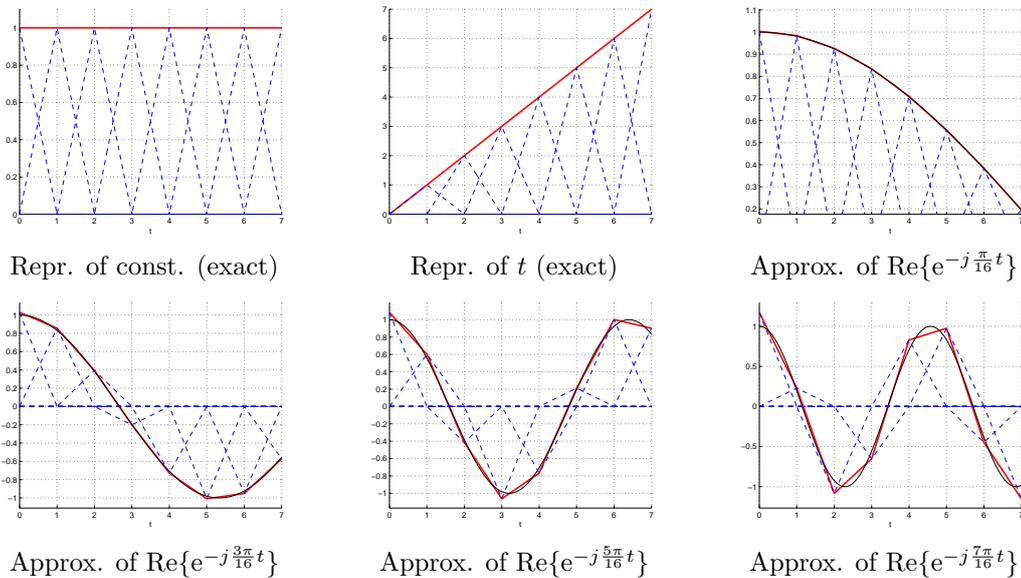


Figure 4. Example of approximate reproduction of exponentials using the linear spline.

4.2 Approximate FRI Recovery

We now go back to the problem of reconstructing a stream of K Diracs $x(t) = \sum_{k=0}^{K-1} x_k \delta(t - t_k)$ from samples which are now taken with an arbitrary kernel $\varphi(t)$. We thus observe: $y_n = \langle x(t), \varphi(t - n) \rangle$, $n = 0, 1, \dots, N - 1$, but now we make no assumption on the sampling kernel. We find proper coefficients for $\varphi(t)$ to approximate the exponentials $e^{\alpha_m t}$, where $m = 0, \dots, P$, $\alpha_m = \alpha_0 + m\lambda$, and $\alpha_0, \lambda \in \mathbb{C}$. From the previous section we know

that a good approximation is achieved if we choose $c_{m,n} = c_{m,0}e^{\alpha_m n}$ with $c_{m,0} = \hat{\varphi}(\alpha_m)^{-1}$. We thus only need to know the Laplace transform of $\varphi(t)$ at α_m , $m = 0, \dots, P$. Also, note that P can be chosen arbitrarily.

We then operate as if we were in the exact exponential reproduction framework. We therefore linearly combine the samples y_n using the new coefficients $c_{m,n}$ to obtain:

$$\begin{aligned} s_m &= \sum_{n=0}^{N-1} c_{m,n} y_n = \left\langle x(t), \underbrace{\sum_{n=0}^{N-1} c_{m,n} \varphi(t-n)}_{e^{\alpha_m t - \varepsilon_m(t)} \right\rangle \\ &= \sum_{k=0}^{K-1} x_k u_k^m - \underbrace{\sum_{k=0}^{K-1} x_k \varepsilon_m(t_k)}_{\zeta_m} \end{aligned} \quad (5)$$

where $\hat{x}_k = x_k e^{\alpha_0 t_k}$ and $u_k = e^{\lambda t_k}$.

There is a model mismatch due to the approximation error $\varepsilon_m(t)$ of (4). We treat it as noise and retrieve the parameters of the signal using matrix pencil. The model mismatch depends on the quality of the approximation, dictated by the coefficients $c_{m,n}$, the parameters α_m and P . The estimation of the Diracs can be refined using the iterative algorithm shown in the box Algorithm 1. The basic idea of the algorithm is that, given an estimate of the locations of the Diracs, we can compute an approximation of ζ_m and use it to refine the computation of the values s_m . In noisy scenarios, if ζ_m is negligible when compared to other forms of noise then the procedure is sufficiently good.

Algorithm 1 Recovery of a train of K Diracs using approximation of exponentials

- 1: Calculate the approximation coefficients $c_{m,n}$.
- 2: Compute the moments $s_m^0 = \sum_n c_{m,n} y_n$, from the original data y_n , $n = 0, \dots, N-1$ and set $s_m^i = s_m^0$, $m = 0, \dots, P$. The iteration is $i = 1$.
- 3: Obtain the values $\{u_k^i, x_k^i\}_{k=0}^{K-1}$ by applying either total least-squares and Cadzow or matrix pencil to the sequence s_m^i . Determine the locations t_k^i and amplitudes a_k^i , for the i th iteration.
- 4: Recalculate the moments for the next iteration $i+1$ by removing the model mismatch from s_m^0 . This can be done by using (5) as follows:

$$s_m^{i+1} = s_m^0 + \sum_{k=0}^{K-1} a_k^i \varepsilon_m \left(\frac{t_k^i}{T} \right), \quad (6)$$

for $m = 0, \dots, P$ and where $\varepsilon_m(t)$ is the error of the approximation (4).

- 5: Repeat steps 3 and 4 until convergence of the values $\{t_k^i, a_k^i\}_{k=0}^{K-1}$.
-

5. SIMULATION RESULTS

In Fig. 5, we show the reconstruction of an infinite stream of Diracs using the local reconstruction algorithm of Sec. 3. Almost all Diracs can be retrieved up to noise levels of 5dBs.

The example in Fig. 6 shows the merits of the approximate Strang-Fix framework. In this example, $K = 4$ Diracs are sampled with a B-spline of order $P = 5$. This means that in the exact framework at most $K = 3$ Diracs can be reconstructed as shown in part (a). The approximate framework, instead, allows us to choose P arbitrarily and consequently allows us to reconstruct the $K = 4$ Diracs almost perfectly.

6. CONCLUSIONS

In this paper we have shown how to sample FRI signals with arbitrary kernels and that a novel local reconstruction algorithm can be fast and very resilient to noise.

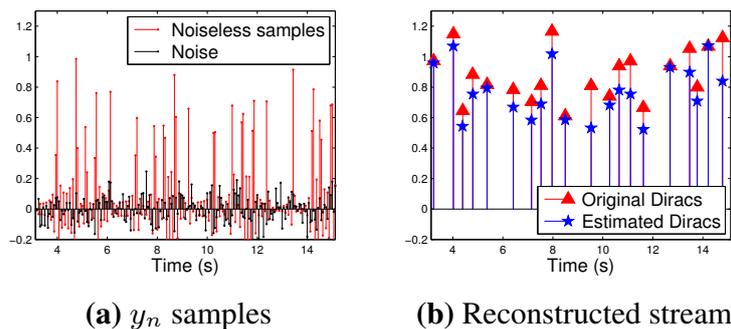


Figure 5. Reconstruction of an infinite stream of Diracs using the local reconstruction algorithm of Sec. 5.

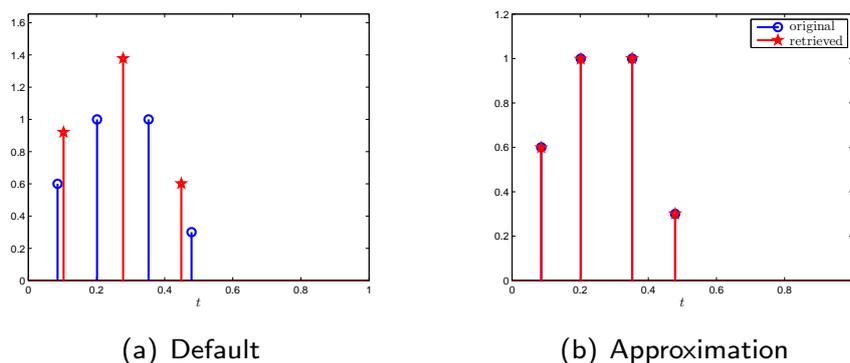


Figure 6. Reconstruction of $K = 4$ Diracs using the default strategy, part (a), and the approximate framework, part (b).

REFERENCES

- [1] H. Akondi Asl, P. L. Dragotti, and L. Baboulaz. Multichannel sampling of signals with finite rate of innovation. *Signal Processing Letter*, 17(8):762–765, August 2010.
- [2] L. Baboulaz and P. L. Dragotti. Exact feature extraction using finite rate of innovation principles with an application to image super-resolution. *IEEE Trans. on Image Processing*, 18(2):281–298, February 2009.
- [3] J. Berent, P. L. Dragotti, and T. Blu. Sampling piecewise sinusoidal signals with finite rate of innovation methods. *IEEE Trans. on Signal Processing*, 58(2):613–625, February 2010.
- [4] 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.
- [5] J.A. Cadzow. Signal enhancement: A composite property mapping algorithm. *IEEE Trans. on Acoustic, Speech and Signal Processing*, 36(1):149–62, January 1988.
- [6] V. Chaisinthop and P.L. Dragotti. Centralized and distributed semi-parametric compression of piecewise smooth functions. *IEEE Trans. on Signal Processing*, 59(7):3071–3085, July 2011.
- [7] C. Chen, P. Marziliano, and A. C. Kot. 2d finite rate of innovation reconstruction method for step edge and polygon signals in the presence of noise. *IEEE Trans. Signal Processing*, 60(6):2851–2859, June 2012.
- [8] P.L. Dragotti, M. Vetterli, and T. Blu. Exact sampling results for signals with finite rate of innovation using Strang-Fix conditions and local kernels. In *Proc. of IEEE Int. Conf. on acoustic, speech and signal processing (ICASSP)*, Philadelphia (PA), March 2005.
- [9] 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.
- [10] M. Elad, P. Milanfar, and G. H. Golub. Shapes from moments- an estimation theory perspective. *IEEE Trans. Signal Processing*, 52(7):1814–1829, July 2004.

- [11] N. Gehrig and P.L. Dragotti. Distributed compression in camera sensor network. In *Proc. of IEEE Workshop on Multimedia Signal Processing (MMSP)*, Siena, Italy, September 2004.
- [12] A. Hirabayashi and P.L. Dragotti. E-spline sampling for precise and robust line-edge extraction. In *Proc. of IEEE Int. Conf. on Image Processing (ICIP)*, Kong Kong, October 2010.
- [13] A. Kirmani, A. Colaco, F. N. C. Wong, and V. K. Goyal. Exploiting sparsity in time-of-flight range acquisition using a single time-resolved sensor. *Optics Express*, 19(22):21485–21507, October 2011.
- [14] I. Maravic and M. Vetterli. Exact sampling results for some classes of parametric nonbandlimited 2-d signals. *IEEE Trans. on Signal Processing*, 52(1):175–189, January 2004.
- [15] 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*, 53(8):2788–2805, August 2005.
- [16] J. Onativia, S. R. Schultz, and P.L. Dragotti. A finite rate of innovation algorithm for fast and accurate spike detection from two-photon calcium imaging. *Journal of Neural Engineering*, 2013.
- [17] H. Pan, P.L. Dragotti, and T. Blu. Sampling curves with finite rate of innovation. In *Proc. of IEEE International Conference on Sampling Theory and Applications (SampTA)*, Singapore, May 2011.
- [18] G. C. F. M. R. Prony. Essai expérimental et analytique sur les lois de la dilatabilité des fluides élastiques et sur celles de la force expansive de la vapeur de l' eau et de la vapeur de l' alkool, à différentes températures. *J. de l' École Polytechnique*, 1:24–76, 1795.
- [19] C.S. Seelamantula and M. Unser. A generalized sampling method for finite-rate-of-innovation-signal reconstruction. *Signal Processing Letter*, 15:813–816, t 2008.
- [20] P. Shukla and P.L. Dragotti. Sampling schemes for multidimensional signals with finite rate of innovation. *IEEE Trans. on Signal Processing*, 55(7):3670–3686, July 2007.
- [21] P. Stoica and R. Moses. *Spectral Analysis of Signals*. Englewood Cliffs,NJ, Prentice-Hall, 2005.
- [22] J.A. Uriguen, T. Blu, and P.L. Dragotti. On the exponential reproducing kernels for sampling signals with finite rate of innovation. In *Proc. of IEEE International Conference on Sampling Theory and Applications (SampTA)*, Singapore, May 2011.
- [23] J.A. Uriguen, T. Blu, and P.L. Dragotti. Fri sampling with arbitrary kernels. *IEEE Trans. Signal Processing*, submitted December 2012.
- [24] M. Vetterli, P. Marziliano, and T. Blu. Sampling signals with finite rate of innovation. *IEEE Trans. Signal Processing*, 50(6):1417–1428, June 2002.