Exact Sampling Results for 1-D and 2-D Signals with Finite Rate of Innovation using Strang-Fix Conditions and Local Reconstruction Algorithms (Invited Paper)*

Pier Luigi Dragotti

[1] Electrical and Electronic Engineering
Imperial College London,
Exhibition Road,
London SW7 2BT, United Kingdom

ABSTRACT

Recently, it was shown that it is possible to sample classes of signals with finite rate of innovation.[18] These sampling schemes, however, use kernels with infinite support and this leads to complex and instable reconstruction algorithms.

In this paper, we show that many signals with finite rate of innovation can be sampled and perfectly reconstructed using kernels of compact support and a local reconstruction algorithm. The class of kernels that we can use is very rich and includes any function satisfying Strang-Fix conditions, Exponential Splines and functions with rational Fourier transforms. Our sampling schemes can be used for either 1-D or 2-D signals with finite rate of innovation.

1. INTRODUCTION

Sampling theory plays a central role in modern signal processing and communications, and has experienced a recent revival thanks, in part, to the recent advances in wavelet theory.[14, 15] 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 $h(t)$.

Recently, it was shown that it is possible to develop sampling schemes for classes of signals that are neither bandlimited nor belong to a fixed sub-space.[18] For instance, it was shown that it is possible to sample streams of Diracs or piecewise polynomial signals using a sinc or a Gaussian kernel. The common feature of such signals is that they have a parametric representation with a finite number of degrees of freedom and are, therefore, called signals with finite rate of innovation (FRI).[18] The reconstruction process is based on the use of a locator or annihilating filter, a tool widely used in spectral estimation[11] and error correction coding.[1]

The fundamental limit of the above method, as well as of the classical Shannon reconstruction scheme, is that they use kernels of infinite support. As a consequence, the reconstruction algorithm is usually physically non-realizable (e.g., realization of an ideal low-pass filter) or, in the case of FRI signals, becomes immediately complex and instable (the complexity is in fact influenced by the global rate of innovation of $x(t)$).

In this paper we show that many signals with a local finite rate of innovation can be sampled and perfectly reconstructed using a wide range of sampling kernels and a local reconstruction algorithm. In particular, we show that the main property the kernel has to satisfy is to be able to reproduce polynomials or exponentials. More precisely, if $\varphi(t)$ is the kernel, we need it to satisfy

$$\sum_n c_{m,n}\varphi(t-n) = t^m \qquad m = 0, 1, ..., N \tag{1}$$

or

$$\sum_n c_{m,n}\varphi(t-n) = e^{\alpha_m t} \qquad \alpha_m = \alpha_0 + m\lambda \text{ and } m = 0, 1, ..., N, \tag{2}$$

*This paper includes research conducted jointly with Martin Vetterli, Thierry Blu and Pancham Shukla.[7, 10] This work is supported in part by the Engineering and Physical Sciences Research Council (EPSRC) of UK under grant GR/S57631/01.

for a proper choice of the coefficients $c_{m,n}$. Interesting enough, the reconstruction algorithm proposed in this paper is also based on the annihilating filter method. It is also possible to show that many kernels with rational transfer functions can be used to sample FRI signals as well. Despite the fact that these kernels have infinite support, the reconstruction algorithm remains local and its complexity still depends on the local, rather than global, rate of innovation of $x(t)$.

These sampling schemes can be used for 2-D signals with finite rate of innovation as well. In particular, sets of 2-D Diracs and polygonal images can be reconstructed exactly using complex moments.

The paper is organized as follows: In the next section we review the notion of signals with finite rate of innovation, and present the families of sampling kernels that are used in our sampling schemes. Section 3 presents our main sampling results for the case of kernels reproducing polynomials. In particular, we show how to sample and perfectly reconstruct streams of Diracs, streams of differentiated Diracs and piecewise polynomial signals. The following section extend the previous results to the case in which the sampling kernel reproduces exponentials. In Section 5 we show how to estimate FRI signals at the output of an RC circuit. Section 6 studies the 2-D case and Section 7 concludes the paper.

2. SIGNALS AND KERNELS

In this section we introduce the notion of signals with finite rate of innovation[18] and present the families of sampling kernels that will be used for the rest of the paper.

2.1. Signals with Finite Rate of Innovation

Consider a signal of the form

$$x(t) = \sum_{n\in\mathbb{Z}} \sum_{k=0}^{K} \lambda_{n,k} \varphi_k \left(\frac{t-t_n}{T} \right). \tag{3}$$

Clearly, if the set of functions $\{\varphi_k(t)\}_{k=0,1,\ldots,K}$ is known, the only free parameters in the signal $x(t)$ are the coefficients $\lambda_{n,k}$ and the time shifts t_n. It is therefore natural to introduce a counting function $C_x(t_a, t_b)$ that counts the number of free parameters in $x(t)$ over an interval $[t_a, t_b]$. The rate of innovation of $x(t)$ is then defined as[18]

$$\rho = \lim_{\tau\to\infty} \frac{1}{\tau} C_x \left(-\frac{\tau}{2}, \frac{\tau}{2} \right). \tag{4}$$

Definition 1 (Vetterli, Marziliano, Blu[18]). *A signal with a finite rate of innovation is a signal whose parametric representation is given in (3) and with a finite ρ as defined in (4).*

It is of interest to note that shift-invariant signals, including bandlimited signals, are included in Definition 3. For instance, if we call f_{max} the maximum non-zero frequency in a bandlimited real signal, then $\rho = 2f_{max}$. Therefore, one possible interpretation is that it is possible to sample bandlimited signals because they have finite rate of innovation and not because they are bandlimited.

In some cases it is more convenient to consider a local rate of innovation with respect to a moving window of size τ. The local rate of innovation at time t is thus given by[18]

$$\rho_\tau(t) = \frac{1}{\tau} C_x \left(t - \frac{\tau}{2}, t + \frac{\tau}{2} \right). \tag{5}$$

Clearly $\rho_\tau(t)$ tends to ρ as $\tau \to \infty$.

2.2. Sampling Kernels

As mentioned in the introduction, the signal $x(t)$ is usually filtered before being (uniformly) sampled. The samples y_n of $x(t)$ are given by $y_n = \langle x(t), \varphi(t/T - n)\rangle$, where the sampling kernel $\varphi(t)$ is the time reversed version of filter's impulse response. In some cases, one is free to choose or design the sampling kernel. However, in most realistic situations, this kernel depends on the physical properties of the acquisition device and cannot be modified. It is therefore important to develop sampling schemes that do not require the use of very particular or even physically non-realizable kernels. The classical Shannon sampling theorem is very restrictive in this respect, since it requires an ideal low-pass filter to reconstruct a bandlimited signal from its samples.

In our formulation we can use a wide range of different kernels. For the sake of clarity, we divide them into three different families:

1. Any kernel $\varphi(t)$ that together with its shifted versions can reproduce polynomials:

$$\sum_n c_{m,n}\varphi(t-n) = t^m \qquad m = 0, 1, ..., N \tag{6}$$

2. Any kernel $\varphi(t)$ that together with its shifted versions can reproduce exponentials:

$$\sum_n c_{m,n}\varphi(t-n) = e^{\alpha_m t} \qquad \alpha_m = \alpha_0 + m\lambda \text{ and } m = 0, 1, ..., N, \tag{7}$$

3. Any kernel with rational Fourier transform. That is, any kernel of the form

$$\hat{\varphi}(\omega) = \frac{\prod_i (j\omega - b_i)}{\prod_m (j\omega - \alpha_m)} \qquad a_m = \alpha_0 + m\lambda \text{ and } m = 0, 1, ..., N. \tag{8}$$

   where $\hat{\varphi}(\omega)$ is the Fourier transform of $\varphi(t)$.

In all cases, the choice of N depends on the local rate of innovation of the original signal $x(t)$ as it will become clear later on.

The first family of kernels includes any function satisfying the so called Strang-Fix conditions.[12] Namely, $\varphi(t)$ satisfies Eq. (6) if and only if

$$\hat{\varphi}(0) \neq 0 \text{ and } \hat{\varphi}^{(m)}(2n\pi) = 0 \text{ for } \begin{cases} n \neq 0 \\ m = 0, 1, ..., N \end{cases}$$

where $\hat{\varphi}(\omega)$ is the Fourier transform of $\varphi(t)$. These conditions were originally valid for functions with compact support only, more recently they have been extended to non-compactly supported functions.[2,3,6] It is also of interest to note that this class of kernels contains also any scaling functions that generate wavelets with $N+1$ vanishing moments.[4,8,13,17]

The theory related to the reproduction of exponentials is somewhat more recent and relies on the notion of Exponential Splines (E-Splines).[16] A function $\beta_\alpha(t)$ with Fourier transform

$$\hat{\beta}_\alpha(\omega) = \frac{1 - e^{\alpha - j\omega}}{j\omega - \alpha}$$

is called E-Spline of first order. Notice that α does not have to be real, but can be any complex number. Moreover, notice that $\beta_\alpha(t)$ reduces to the classical zero-order B-Spline when $\alpha = 0$. The function $\beta_\alpha(t)$ satisfies several interesting properties, in particular, it is of compact support and a linear combination of shifted versions of $\beta_\alpha(t)$ reproduces $e^{\alpha t}$. As in the classical case, higher order E-Splines are obtained by successive convolutions of lower-order ones or

$$\hat{\beta}_{\vec{\alpha}}(\omega) = \prod_{n=0}^{N} \frac{1 - e^{\alpha_n - j\omega}}{j\omega - \alpha_n}$$

where $\vec{\alpha} = (\alpha_0, \alpha_1, ..., \alpha_N)$. The higher-order Spline is again of compact support and it is possible to show that it can reproduce any exponential in the subspace spanned by $\{e^{\alpha_0 t}, e^{\alpha_1 t}, ..., e^{\alpha_N t}\}$.[16] Moreover, since the exponential reproduction formula is preserved through convolution,[16] any composite function of the form $\varphi(t) * \beta_{\vec{\alpha}}(t)$ is also able to reproduce exponentials. Therefore, the second group of kernels contains any composite function of the form $\varphi(t) * \beta_{\vec{\alpha}}(t)$ with $\beta_{\vec{\alpha}}(t) = \beta_{\alpha_0}(t) * \beta_{\alpha_1}(t) * ... * \beta_{\alpha_N}(t)$, $\alpha_m = \alpha_0 + m\lambda$ and $m = 0, 1, ..., N$.

Notice that the exponential case reduces to that of reproduction of polynomials when $\alpha_m = 0$ for $m = 0, 1, ..., N$. For this reason we could study our sampling schemes in the exponential case only and then particularize it to the polynomial case. However, we prefer to keep the two cases separated for the sake of simplicity.

The last group of kernels includes any *linear differential acquisition device.* That is, any linear device or system for which the input and output are related by linear differential equations. This includes most of the commonly used electrical, mechanical or acoustic systems.

The reason why we can sample signals with finite rate of innovation using such kernels is that we can *convert* a kernel $\varphi(t)$ with rational Fourier transform as in (8) into a kernel that reproduces exponentials. This is achieved by filtering the samples $y_n = \langle x(t), \varphi(t-n) \rangle$ with an FIR filter of form $H(z) = \prod_{m=0}^{N}(1 - e^{\alpha_m} z)$.

For example, assume that $\hat{\varphi}(\omega) = \frac{1}{j\omega - \alpha}$ and $y_n = \langle x(t), \varphi(t-n) \rangle$. Then

$$\begin{aligned} z_n = h_n * y_n &= y_n - e^{\alpha} y_{n+1} = \langle x(t), \varphi(t-n) - e^{\alpha}\varphi(t-n-1) \rangle \\ &= \tfrac{1}{2\pi} \langle X(\omega), e^{-j\omega n} \tfrac{(1-e^{\alpha - j\omega})}{(j\omega - \alpha)} = \langle x(t), \beta_{\alpha}(t-n) \rangle. \end{aligned}$$

Therefore, by filtering the samples y_n with the filter $H(z) = (1 - e^{\alpha} z)$ we get a new set of samples z_n that are equivalent to those that would have been obtained by sampling the original signal $x(t)$ with the E-Spline $\beta_{\alpha}(t)$.

Likewise, when the original kernel has N poles at locations $\vec{\alpha} = (\alpha_0, \alpha_1, ..., \alpha_N)$, by filtering the samples $y_n = \langle x(t), \varphi(t-n) \rangle$ with the filter $H(z) = \prod_{m=0}^{N}(1 - e^{\alpha_m} z)$ we have that

$$z_n = h_n * y_n = \langle x(t), \beta_{\vec{\alpha}}(t-n) \rangle.$$

and the new kernel is of compact support and reproduces the exponentials $\{e^{\alpha_0 t}, e^{\alpha_1 t}, ..., e^{\alpha_N t}\}$.

In Section 5, we use this result to estimate FRI signals at the output of an RC circuit.

3. RECONSTRUCTION OF FRI SIGNALS USING KERNELS THAT REPRODUCE POLYNOMIALS

In this section, we assume that the sampling kernel $\varphi(t)$ satisfies the Strang-Fix conditions,[12] that is, a linear combination of shifted versions of $\varphi(t)$ can reproduce polynomials of maximum degree N (see Equation (6)). We consider the case of streams of Diracs first and derive the other results directly from this case.

3.1. Streams of Diracs

Consider a stream of Diracs $x(t)$. Call y_n the observed samples, that is, $y_n = \langle \varphi(t-n), x(t) \rangle$ where, for simplicity, we have assumed $T = 1$. Assume for now that the signal contains only K Diracs that is, $x(t) = \sum_{k=0}^{K-1} a_k \delta(t - t_k)$, $t \in \mathbb{R}$, and assume that the sampling kernel $\varphi(t)$ is able to reproduce polynomials of maximum degree $N \geq 2K-1$. Under these hypotheses, it is possible to retrieve the locations t_k and the amplitudes a_k of $x(t)$ from its samples y_n. The reconstruction algorithm operates in three steps. First, the first N moments of the signal $x(t)$ are found. Second, the Diracs' locations are retrieved using an annihilating filter. Third, the amplitudes a_k are obtained solving a Vandermonde system.

For a more detailed description of the annihilating filter method we refer to.[11, 18] Here, we only highlight the main steps of our algorithm where the key innovation is in our ability to estimate the moments of $x(t)$ from its samples y_n.

1. Retrieve the first N moments of the signal $x(t)$.

   Call $\tau_m = \sum_n c_{m,n} y_n$, $m = 0, 1, ..., N$ the weighted sum of the observed (non-zero) samples, where the weights $c_{m,n}$ are those in Equation (6). We have that

$$\begin{aligned} \tau_m &= \sum_n c_{m,n} y_n \\ &\overset{(a)}{=} \langle x(t), \sum_n c_{m,n}\varphi(t-n)\rangle \\ &\overset{(b)}{=} \int_{-\infty}^{\infty} x(t) t^m dt \\ &\overset{(c)}{=} \sum_{k=0}^{K-1} a_k t_k^m \qquad m = 0, 1, ..., N \end{aligned} \tag{9}$$

   where (a) follows from the linearity of the inner product, (b) from the polynomial reproduction formula in (6), and (c) from the fact that $x(t) = \sum_{k=0}^{K-1} a_k \delta(t - t_k)$. Hence by opportunely combining the samples y_n, we end-up observing the first N moments τ_m of the signal $x(t)$. From these moments it is then possible to retrieve the amplitudes and locations of the Diracs using the annihilating filter method.

2. Find the coefficients h_m $m = 0, 1, ..., K$ of the annihilating filter.

   Call h_m $m = 0, 1, ..., K$ the filter with z-transform

$$H(z) = \sum_{m=0}^{K} h_m z^{-m} = \prod_{k=0}^{K-1} (1 - t_k z^{-1}). \tag{10}$$

   That is, the roots of $H(z)$ correspond to the locations t_k. It clearly follows that

$$h_m * \tau_m = \sum_{i=0}^{K} h_i \tau_{m-i} = 0. \tag{11}$$

   The filter h_m is thus called annihilating filter since it annihilates the observed signal τ_m. The zeros of this filter uniquely define the set of locations t_k since the locations are distinct. The filter coefficients h_m are found from the system of equations in (11). Since $h_0 = 1$, the identity in (11) leads to a Yule-Walker system of equations involving $2K$ consecutive values of τ_m and, in this case, it has a unique solution since h_m is unique for the given signal. Given the filter coefficients h_m, the locations of the Diracs are the roots of polynomial in (10). Notice that, since we need at least $2K$ consecutive values of τ_m to solve the Yule-Walker system, we need the sampling kernel to be able to reproduce polynomial of maximum degree $N \geq 2K - 1$.

3. Find the weight a_k.

   Given the locations $t_0, t_1, ..., t_k$, the weights a_k are obtained by solving, for instance, the first K consecutive equations in (9). These equations form a Vandermonde system which yields a unique solution for the weights a_k given that the t_ks are distinct.

So far, we have not made any assumption about the support of the sampling kernel. However, it is reasonable to assume that $\varphi(t)$ has compact support L. If this is the case, it is clearly possible to reconstruct signals containing more than K Diracs. More precisely, since the kernel is of compact support, only a finite number of samples is influenced by a certain set of Diracs. Thus, if we are sure that the samples generated by different sets of K Diracs do not influence each other, we can still use the above method sequentially. It is easy to see that this happens when there are no more than K Diracs in an interval of size KLT or, using the terminology introduced in the previous section, when the local rate of innovation $\rho_{KLT}(t) \leq 2/LT$.

Thus from the above discussion it follows that:

THEOREM 1. *Given is a sampling kernel $\varphi(t)$ that can reproduce polynomials of maximum degree $N \geq 2K - 1$ and of compact support L. An infinite-length stream of Diracs $x(t) = \sum_{n\in\mathbb{Z}} a_n \delta(t - t_n)$ is uniquely determined from the samples defined by $y_n = \langle \varphi(t/T - n), x(t)\rangle$ if and only if there are at most K Diracs in an interval of length KLT.*

3.2. Stream of Differentiated Diracs

Consider a stream of differentiated Diracs:

$$x(t) = \sum_{k=0}^{K-1} \sum_{m=0}^{M_k-1} a_{k,m} \delta^{(m)}(t - t_k).$$

Note that this signal has K Diracs and $\hat{K} = \sum_{k=0}^{K-1} M_k$ weights. Moreover, recall that the rth derivative of a Dirac is a function that satisfies the property $\int f(t)\delta^{(r)}(t - t_0)dt = (-1)^r f^{(r)}(t_0)$.

Assume that $x(t)$ is sampled with a kernel that can reproduce polynomial of maximum degree $N \geq 2\hat{K} - 1$. As shown previously, we can compute the first N moments of $x(t)$ from its samples y_n. This leads to the following system of polynomial equations

$$\begin{aligned}
&a_{0,0} + a_{1,0} + ... + a_{K-1,0} &&= \tau_0 \\
&a_{0,0}t_0 + ... + a_{K-1,0}t_{K-1} - a_{0,1} - ... - a_{K-1,1} &&= \tau_1 \\
&a_{0,0}t_0^2 + ... + a_{K-1,0}t_{K-1}^2 - 2a_{0,1}t_0 - ... - 2a_{K-1,1}t_{K-1} + 2a_{0,2} + ... + 2a_{K-1,2} &&= \tau_2 \\
&\vdots &&= \vdots \\
&\textstyle\sum_{k=0}^{K-1} \sum_{m=0}^{M_k} (-1)^m a_{k,m} (N)(N-1)\cdots(N-m+1)t_k^{N-m} &&= \tau_N.
\end{aligned} \qquad (12)$$

where we have used the fact that $\int t^n \delta^{(r)}(t - t_0)dt = (-1)^r n(n-1)\cdots(n-r+1)t_0^{n-r}$. Therefore, we can say that what we observe is

$$\tau_n = \sum_{k=0}^{K-1} \sum_{m=0}^{M_k-1} (-1)^m a_{k,m} n(n-1)\cdots(n-m+1)t_k^{n-m}.$$

It can be shown that the filter $(1 - t_k z^{-1})^M$ annihilates the signal $n^r t_k^n$, with $r \leq M - 1$. Therefore the filter h_m with z-transform

$$H(z) = \prod_{k=0}^{K-1} (1 - t_k z^{-1})^{M_k}$$

annihilates τ_n. The $\hat{K}$ unknown coefficients of h_m can be found solving a Yule-Walker system similar to the one in the previous section. We need at least $\hat{K}$ equations to find these coefficients, therefore, we need to know at least $2\hat{K}$ consecutive values of τ_n (this is why $N \geq 2\hat{K} - 1$). From the annihilating filter we obtain the locations $t_0, t_1, ..., t_{K-1}$. We then need to solve the first $\hat{K}$ equations in (12) to obtain the weights $a_{k,m}$. This is a generalized Vandermonde system which has again a unique solution given that the t_ks are distinct.

The above analysis can be summarized in the following theorem:

THEOREM 2. *Given is a sampling kernel $\varphi(t)$ that can reproduce polynomials of maximum degree $N \geq 2\hat{K} - 1$ and of compact support L. An infinite-length stream of differentiated Diracs $x(t) = \sum_{n=\mathbb{Z}} \sum_{m=0}^{M_n-1} a_{n,m}\delta^{(m)}(t - t_n)$ is uniquely determined by the samples $y_n = \langle \varphi(t/T - n), x(t) \rangle$ if and only if there are at most K differentiated Diracs with $\hat{K}$ weights in an interval of length KLT.*

3.3. Piecewise polynomial signals

A signal $x(t)$ is piecewise polynomial with pieces of maximum degree M if and only if its $(M+1)$ derivative is a stream of differentiated Diracs or $x(t)^{(M+1)}(t) = \sum_{n\in\mathbb{Z}} \sum_{m=0}^{M} a_{n,m}\delta^{(m)}(t - t_n)$. This means that if we are able to relate the samples of $x(t)$ to those of $x^{(M+1)}(t)$, we can use Theorem 2 to reconstruct $x(t)$. This is indeed possible by recalling the link existing between discrete differentiation and derivation in continuous domain.

Consider the samples $y_n = \langle x(t), \varphi(t-n)\rangle$ where $\varphi(t)$ is a generic sampling kernel. Let $z_n^{(1)}$ denote the finite difference $y_{n+1} - y_n$. It follows that

$$\begin{aligned} z_n^{(1)} &= \langle x(t), \varphi(t-n-1) - \varphi(t-n)\rangle \\ &= \tfrac{1}{2\pi}\langle X(\omega), \hat{\varphi}(\omega)e^{-j\omega n}(e^{-j\omega}-1)\rangle \\ &= \tfrac{1}{2\pi}\langle X(\omega), -j\omega\hat{\varphi}(\omega)e^{-j\omega n}\left(\tfrac{1-e^{-j\omega}}{j\omega}\right)\rangle \\ &= \langle \tfrac{dx(t)}{dt}, \varphi(t-n) * \beta_0(t-n)\rangle \end{aligned}$$

where $\beta_0(t)$ is the B-Spline of order zero. This means that the coefficients $z_n^{(1)}$ represent the samples given by the inner products of the derivative of $x(t)$ with the new kernel $\varphi(t) * \beta_0(t)$. In the same way, it is straight-forward to show that the $(M+1)$th finite differences $z_n^{(M+1)}$ represent the samples obtained by sampling $x^{(M+1)}(t)$ with the kernel $\varphi(t) * \beta_M(t)$, where $\beta_M(t)$ is the B-Spline of degree M.

Now, assume that $\varphi(t)$ is of compact support L and that it can reproduce polynomials of maximum degree N. Then $\varphi(t) * \beta_M(t)$ has support $L + M + 1$ and can reproduce polynomials of maximum degree $N + M + 1$. Thus, if the new kernel satisfies the hypotheses of Theorem 2, the samples $z_n^{(M+1)}$ are a sufficient representation of $x^{(M+1)}(t)$ and, therefore, of $x(t)$. This leads to the following theorem

THEOREM 3. *Given is a sampling kernel $\varphi(t)$ of compact support L and that can reproduce polynomials of maximum degree N. An infinite-length piecewise polynomial signal with pieces of maximum degree $M - 1$ is uniquely defined by the samples $y_n = \langle\varphi(t/T - n), x(t)\rangle$ if and only if there are at most $K + 1$ polynomials in an interval of size $(L + M)KT$ and $2KM - 1 \leq (M + N)$.*

Proof: Assume again $T = 1$. Given the samples y_n, compute the Mth finite difference $z_n^{(M)}$. As shown before, $z_n^{(M)} = \langle x^{(M)}(t), \varphi(t-n) * \beta_{M-1}(t-n)\rangle$ and $x(t)^{(M)}(t) = \sum_{n\in\mathbb{Z}}\sum_{m=0}^{M-1} a_{n,m}\delta^{(m)}(t-t_n)$. The new kernel $\varphi(t) * \beta_{M-1}(t)$ has support $L + M$ and can reproduce polynomials of maximum degree $N + M$. Since for hypothesis $x(t)$ has at most $K + 1$ polynomial in an interval of size $(L + M)K$, $x^{(M)}(t)$ has at most K Diracs in that interval with a total number of weights $\hat{K} = KM$. Since we are assuming $2KM - 1 \leq N + M$, the hypotheses of Theorem 2 are satisfied, thus, the samples $z_n^{(M)}$ are sufficient to reconstruct $x^{(M)}(t)$ and therefore $x(t)$.[†]

□

4. THE EXPONENTIAL CASE

In the previous section we have used the property that $\varphi(t)$ can reproduce polynomials to reduce our sampling problem to that of finding the coefficients a_k and t_k of the discrete signal $\tau_m = \sum_{k=0}^{K-1} a_k t_k^m$, $m = 0, 1, ..., N$ and this is achieved using the annihilating filter method. The interesting point is that the annihilating filter method can be used also in the case the observed signal is of the form $s_m = \sum_{k=0}^{K-1} a_k e^{\lambda m t}$ and $\alpha_m = \alpha_0 + m\lambda$. For this reason, FRI signals can be sampled and reconstructed using kernels that reproduce exponentials. The reconstruction scheme is the same as in the polynomial case. First, the signal $s_m = \sum_{k=0}^{K-1} a_k e^{\lambda m t}$ is estimated from the samples y_n, then locations and amplitudes of the Diracs are retrieved from s_m.

Assume that our kernel is of compact support and that it is able to reproduce exponential of the form $e^{\alpha_m t}$ with $\alpha_m = \alpha_0 + m\lambda$ and $m = 0, 1, ..., N$. For instance, $\varphi(t)$ might be an E-Spline $\beta_{\vec{\alpha}}(t)$ with $\vec{\alpha} = (\alpha_0, \alpha_1, ..., \alpha_N)$ and $\alpha_m = \alpha_0 + m\lambda$ or a composite function $\varphi(t) * \beta_{\vec{\alpha}}(t)$. Consider again a stream of K Diracs $x(t) = \sum_{k=0}^{K-1} a_k\delta(t - t_k)$. The samples y_n are then given by $y_n = \langle x(t), \varphi(t-n)\rangle$ and, using Eq. (7), it follows that

$$\begin{aligned} s_m &= \textstyle\sum_n c_{m,n} y_n = \int_{-\infty}^{\infty} x(t) e^{\alpha_0 + m\lambda t} dt \\ &= \textstyle\sum_{k=0}^{K-1} a_k e^{\alpha_0 + m\lambda t_k} \qquad m = 0, 1, ..., N. \end{aligned}$$

[†]Note that the mean of $x(t)$ is obtained directly.

This means that, as in the polynomial case, by opportunely combining the samples y_n we end-up observing a signal s_m of the form $s_m = \sum_{k=0}^{K-1} a_k e^{\alpha_0 + m\lambda t_k}$. The amplitudes and locations of the Diracs are then retrieved from s_m using the annihilating filter method. Again, this reconstruction algorithm can be applied to any stream of Diracs with local rate of innovation $\rho_{KLT}(t) \leq 2/LT$. Thus, we can summarize the above analysis as follows

THEOREM 4. *Given is a sampling kernel $\varphi(t)$ of compact support L and that can reproduce exponentials $e^{\alpha_0 + m\lambda t}$ with $m = 0, 1, ..., N$ and $N \geq 2K - 1$. An infinite-length stream of Diracs $x(t) = \sum_{n \in \mathbb{Z}} a_n \delta(t - t_n)$ is uniquely determined from the samples defined by $y_n = \langle x(t), \varphi(t/T - n)\rangle$ if and only if there are at most K Diracs in an interval of length KLT.*

5. RECONSTRUCTION OF FRI SIGNALS AT THE OUTPUT OF AN ELECTRIC CIRCUIT

The sampling schemes of the previous section are very important in practice since kernels with rational Fourier transform can be turned into kernels that reproduce exponentials. As already mentioned, most of the commonly used electric circuits fall into this category and thus can be used in our sampling context. As an illustrative example, we show how to estimate a piecewise constant signal at the output of an RC circuit.

Consider the classical RC circuit shown in Figure 1 and call $H(\omega) = \alpha/(\alpha + j\omega)$ with $\alpha = 1/RC$ its transfer function. Assume that the input voltage is $x(t) = Au(t - t_0)$. The output $y(t) = h(t) * x(t)$ is clearly given by $y(t) = Au(t-t_0) - Ae^{-\alpha(t-t_0)}u(t-t_0)$. The output voltage is then uniformly sampled with sampling period $T = 1$ leading to the samples $y_n = Au(n-t_0) - Ae^{-\alpha(n-t_0)}u(n-t_0)$. Alternatively, we can say that $y_n = \langle x(t), \varphi(t-n)\rangle$ with $\varphi(t) = h(-t)$.[‡] Our aim is to retrieve $x(t)$ from the samples y_n.

First, compute the following difference

$$\begin{aligned} z_n &= e^{\alpha} y_{n+1} - y_n = \langle x(t), e^{\alpha}\varphi(t - n - 1) - \varphi(t - n)\rangle \\ &= \tfrac{1}{2\pi}\langle X(\omega), \alpha e^{-j\omega n}\tfrac{(1-e^{\alpha - j\omega})}{(j\omega - \alpha)} = \langle x(t), \alpha\beta_\alpha(t - n)\rangle. \end{aligned}$$

Then compute the first order difference

$$z_n^{(1)} = z_{n+1} - z_n = \langle \frac{dx(t)}{dt}, \alpha\beta_\alpha(t-n) * \beta_0(t-n)\rangle.$$

Now, the signal $dx(t)/dt$ is a Dirac centered at t_0 and with amplitude A. The new kernel $\varphi_\alpha(t) = \alpha\beta_\alpha(t) * \beta_0(t)$ is of compact support $L = 2$, and can reproduce a constant function or the exponential $e^{\alpha t}$. More precisely

$$\frac{1}{e^{\alpha} - 1}\sum_n \varphi_\alpha(t - n) = 1$$

and

$$\frac{1}{1 - e^{-\alpha}}\sum_n e^{\alpha n}\varphi_\alpha(t - n) = e^{\alpha t}.$$

This means that

$$\frac{1}{e^{\alpha} - 1}\sum_n z_n^{(1)} = A$$

and that

$$\frac{1}{1 - e^{-\alpha}}\sum_n e^{\alpha n} z_n^{(1)} = Ae^{\alpha t_0}.$$

Thus, we retrieve the amplitude A from the first sum and the location t_0 from the second one.

[‡]Recall that, since $h(t)$ is real, $\varphi(t) = h(-t)$ implies that $\hat{\varphi}(\omega) = H(-\omega)$.

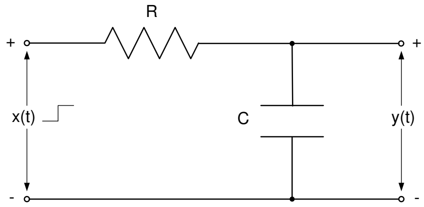


Figure 1. A typical RC circuit. In our example the input voltage is $x(t) = Au(t - t_0)$ and the output $y(t)$ is uniformly sampled. The samples y_n are sufficient to reconstruct $x(t)$ exactly.

Let us verify the above analysis for our specific example. Recall that in our case $y_n = Au(n - t_0) - Ae^{-\alpha(n-t_0)}u(n - t_0)$ and assume for simplicity that $t_0 \in [0, 1]$, then

$$z_n = e^{\alpha} y_{n+1} - y_n = \begin{cases} 0 & \text{for } n < 0 \\ A(e^{\alpha} - e^{\alpha t_0}) & \text{for } n = 0 \\ A(e^{\alpha} - 1) & \text{for } n > 0 \end{cases}$$

and

$$z_n^{(1)} = z_{n+1} - z_n = \begin{cases} A(e^{\alpha} - e^{\alpha t_0}) & \text{for } n = -1 \\ A(e^{\alpha t_0} - 1) & \text{for } n = 0 \\ 0 & \text{otherwise} \end{cases}$$

Therefore, it is clearly true that $1/(e^{\alpha} - 1)\sum_n z_n^{(1)} = A$ and that $1/(1 - e^{-\alpha})\sum_n e^{\alpha n} z_n^{(1)} = Ae^{\alpha t_0}$.

Notice that with this RC circuit we can sample any piecewise constant signal that has at most one discontinuity in an interval of length $2T$. To sample signals with higher local rate of innovation, we need an electrical circuit with more than one pole. For instance, we can use a properly designed cascade of RC circuits.

6. SAMPLING SCHEMES FOR 2-D SIGNALS WITH FINITE RATE OF INNOVATION

In this section we concentrate only on kernels that reproduce polynomials. In particular, we assume that the 2-D sampling kernel $\varphi_{xy}(x, y)$ is given by the tensor product of a 1-D function $\varphi(x)$ that reproduces polynomials. That is, $\varphi_{x,y}(xy) = \varphi(x)\varphi(y)$ and $\varphi(x)$ satisfies Eq (6).

The sampling schemes of Section 3 are based on the fact that a stream of K Diracs is uniquely determined by its first $2K$ moments. Since it is possible to retrieve these moments from the samples y_n, it is possible to reconstruct the original signal. The situation in 2-D is very similar, but complex rather than real moments are needed in this context.

Consider first a set of K 2-D Diracs. That is, $f(x, y) = \sum_{k=0}^{K} a_k \delta(x - x_k, y - y_k)$. The samples are $y_{n,m} = \langle f(x, y), \varphi_{xy}(x - n, y - m)\rangle$ and, by construction, the kernel $\varphi_{xy}(x, y)$ is able to reproduce polynomials of the form $x^n y^l$, $n = 0, 1, ..., N$, $l = 0, 1, ..., N$. It is easy to show that with the right linear combination of the samples $y_{n,m}$, we can estimate the complex moments of $f(x, y)$ in much the same way as we estimated the real moments in the 1-D case. Thus, we end-up observing

$$\tau_m = \int\int f(x, y)(x + jy)^m dxdy \qquad m = 0, 1, ..., N.$$

Since $f(x, y)$ is a set of K Diracs, the complex moments of $f(x, y)$ have the following form

$$\tau_m = \sum_{k=0}^{K-1} a_k z_k^m$$

where the z_ks represent the locations of the K Diracs in complex form: $z_k = x_k + jy_k$. As in the 1-D case, the complex locations of the Diracs and their amplitudes are found using the annihilating filter method. Therefore, as in the 1-D case, the reconstruction algorithm in 2-D operates in three steps:

1. Estimate the first $N \geq (2K-1)$ complex moments τ_m of $f(x, y)$ from the samples $y_{n,m}$.
2. Find the filter h_m that annihilates τ_m. The roots of the filter represents the locations of the Diracs in complex form.
3. Estimate the amplitudes of the Diracs by solving a Vandermonde system.

If the kernel has compact support, it is possible to sample sets of Diracs with more than K Diracs. We just need group of at most K Diracs to be separated enough so that they can be reconstructed independently.

Bi-level polygonal images are also uniquely determined by their complex moments.[5,9] Consider a simply connected convex polygon with K vertices, it is possible to show that[5,9]

$$\hat{\tau}_m = m(m-1) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) z^{m-2} dx dy = \sum_{k=0}^{K-1} \rho_k z_k^m,$$

where the z_ks represent the locations of the vertices of the polygon in complex coordinates. Therefore, as in the previous case, by estimating the complex moments from the samples $y_{n,m}$ and by using the annihilating filter, we can retrieve the locations of the vertices of the polygons and, therefore, the original signal. An example of this sampling scheme is shown in Figure 2.

7. CONCLUSIONS

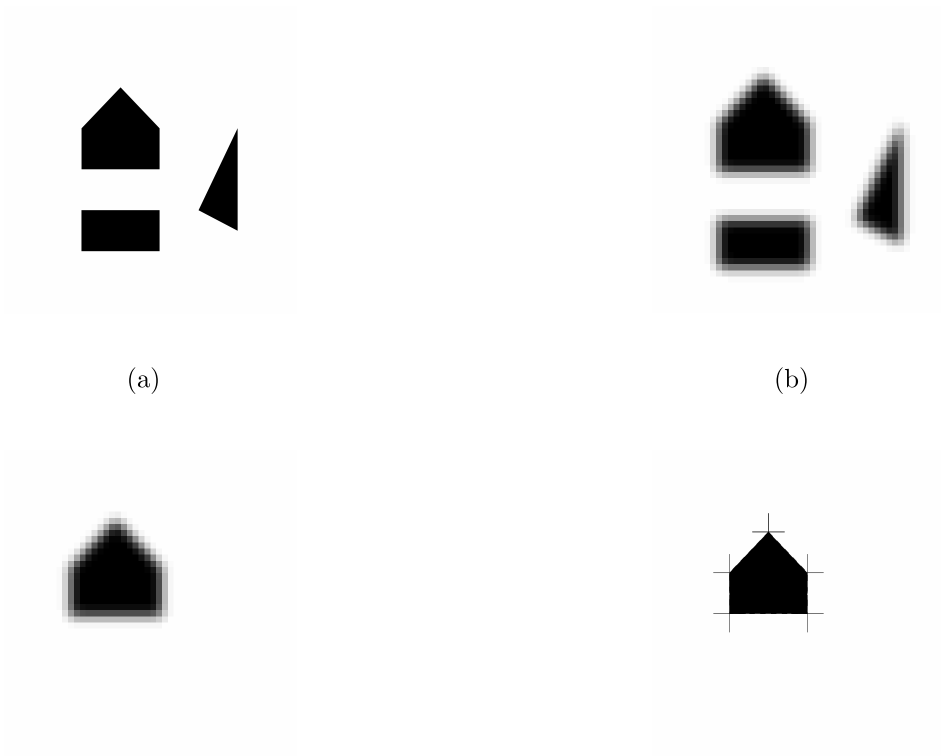
In this paper we have shown that it is possible to sample FRI signals with kernels that can reproduce polynomials or exponentials and with a local reconstruction algorithm. Applications of these sampling theorems can potentially be found in image super-resolution, communication systems and biological systems.

8. ACKNOWLEDGEMENT

The author would like to thank Prof. M. Unser for several stimulating and fruitful discussions.

REFERENCES

1. R.E. Blahut. *Theory and Practice of Error Control Codes.* Addison-Wesley, 1983.
2. T. Blu and M. Unser. Approximation errors for quasiinterpolators and (multi-) wavelet expansions. *Applied and Computational Harmonic Analysis*, 6 (2):219–251, March 1999.
3. E.W. Cheney and W.A. Light. Quasiinterpolation with basis functions having noncompact support. *Constr. Approx.*, 8:35–48, 1992.
4. I. Daubechies. *Ten Lectures on Wavelets.* Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.
5. P. J. Davis. Triangle formulas in the complex plane. *Math. Comput.*, 18:569–577, 1964.
6. C. de Boor, R. A. Devore, and A. Ron. Approximation from shift-invariant subspaces of $L_2(\mathbb{R}^d)$. *Trans. Amer. Math. Society*, 341(2):787–806, February 1994.
7. 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.



(a) (b)

(c) (d)

Figure 2. (a) An original image $g(x, y)$ of size 3767×3767 pixels consists of three bilevel polygons: triangle, rectangle, and pentagon. (b) The set of 50×50 samples produced by the inner products of $g(x, y)$ with a B-Spline sampling kernel $\varphi_{xy}(x, y) = \beta^9_{xy}(x, y)$ with support 631×631 pixels that can reproduce polynomials up to degree nine. The original image is reconstructed from this samples exactly. (c) Sampled version of the pentagon. (d) Original pentagon and the five reconstructed corner points (with +).

8. S. Mallat. *A Wavelet Tour of Signal Processing.* Academic Press, 1998.
9. P. Milanfar, G. C. Verghese, W.C. Karl, and A. S. Willsky. Reconstructing polygons from moments with connections to array processing. *IEEE Trans. Signal Processing*, 43(2):432–443, February 1995.
10. P. Shukla and P.L. Dragotti. Sampling schemes for 2-d signals with finite rate of innovation using kernels that reproduce polynomials. In *Proc. of IEEE Int. Conf. on Image Processing(ICIP)*, Genova (Italy), September 2005.
11. P. Stoica and R Moses. *Introduction to Spectral Analysis.* Englewood Cliffs,NJ, Prentice-Hall, 2000.
12. G. Strang and Fix. G. A Fourier analysis of the finite element variational method. In *Constructive Aspect of Functional Analysis*, pages 796–830, Rome, Italy, 1971.
13. G. Strang and T. Nguyen. *Wavelets and Filter Banks.* Wellesley-Cambridge Press, Boston, 1996.
14. M. Unser. Sampling-50 years after Shannon. *Proc. IEEE*, 88:569–587, April 2000.
15. M. Unser and A. Aldroubi. A general sampling theory for nonideal acquisition devices. *IEEE Trans. Signal Processing*, 42(11):2915–2925, November 1994.
16. 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.
17. M. Vetterli and J. Kovačević. *Wavelets and Subband Coding.* Prentice-Hall, 1995.
18. M. Vetterli, P. Marziliano, and T. Blu. Sampling signals with finite rate of innovation. *IEEE Trans. Signal Processing*, 50(6):1417–1428, June 2002.