Sampling streams of pulses with unknown shapes

Yingsong Zhang*, Pier Luigi Dragotti
Imperial college London
yingsong.zhang@gmail.com, p.dragotti@imperial.ac.uk

***Abstract*—This paper extends the class of continuous-time signals that can be perfectly reconstructed by developing a theory for the sampling and exact reconstruction of streams of short pulses with unknown shapes. The single pulse is modelled as the delayed version of a wavelet-sparse signal, which is normally not band limited. As the delay can be an arbitrary real number, it is hard to develop an exact sampling result for this type of signals. We achieve the exact reconstruction of the pulses by using only the knowledge of the Fourier transform of the signal at specific frequencies. We further introduce a multi-channel acquisition system which uses a new family of compact-support sampling kernels for extracting the Fourier information from the samples. The shape of the kernel is independent of the wavelet basis in which the pulse is sparse and hence the same acquisition system can be used with pulses which are sparse on different wavelet bases. By exploiting the fact that pulses have short duration and that the sampling kernels have compact support, we finally propose a local and sequential algorithm to reconstruct streaming pulses from the samples.**

***Index Terms*—wavelets, sampling, sparsity**

I. Introduction

Sampling theory has experienced a recent revival due in part to the advent of two inter-related theories: Compressed Sensing (CS) [2], [3], and Finite Rate of Innovation (FRI) sampling [4]. Both theories have demonstrated that the prior knowledge that signals can be sparsely described in a proper domain can be used to sample and perfectly reconstruct such signals at a significantly reduced sampling rate.

CS deals with discrete signals (although it is possible to extend it to continuous functions, e.g., [5]–[9]), whereas FRI applies directly to continuous-time functions and streaming signals. Moreover, in FRI, the acquisition set-up is closer to the way traditional A-to-D converters operate and for these reasons, the set-up considered here is similar to the one used in FRI.

FRI has been successfuly used in many signal and image processing applications. In particular, in image super-resolution [10], [11], in neuroscience and healthcare [12], [13] and for channel estimation [14]. Classes of continuous-time signals that can be perfectly reconstructed using the FRI framework include: stream of pulses with known shape and piecewise polynomial signals [4], [15], [16], piecewise sinusoidal signals [17] and classes of 2-D functions [18]–[21].

In this paper we extend the classes of continuous-time signals that can be perfectly reconstructed by developing a theory for the sparse sampling and exact reconstruction of streams of short pulses with unknown shapes. The signals considered can be modelled as follows

$$s(t) = \sum_{l} x_l(t - t_l) \tag{1}$$

where $x_l(t)$ is a short pulse with arbitrary shape and t_l is an arbitrary real number that models the delay of each pulse. Each pulse $x_l(t)$ is different from the others, their shapes are unknown and the only assumption is that they are sparse in a known wavelet basis. An example of the targeted signal is given in Fig. 1. Many signals in real-life applications can be modeled as pulses that are approximately sparse in the wavelet domain. Examples include ECG signals [22] and electro-physiological data that record neuronal activities [23].

We note that, due to the arbitrary delays t_l, these signals do not belong to any union of finite number of shift-invariant sub-spaces. Moreover, while $x_l(t)$ is sparse in a wavelet basis, its delayed version is normally not sparse. This is due to the fact that wavelet bases are shift-variant. These two facts have made it hard in the past to develop an exact sampling result for this type of signals.

We instead achieve an exact reconstruction of $s(t)$ by first showing that each pulse can be reconstructed exactly using only the knowledge of the Fourier transform of $x_l(t)$ at specific frequencies. We then introduce an acquisition system that can extract this Fourier information from the samples. Acquisition is performed using a multi-channel system as shown in Fig. 2. A new family of sampling kernels called MEMS is introduced to allow the estimation of the required Fourier information. MEMS belong to the family of exponential reproducing kernels introduced in [15] in the context of FRI sampling and then extended in [24]–[26].

Multi-channel acquisition systems have been frequently used in sampling, both in the context of sampling signals lying in the union of shift-invariant sub-spaces, e.g., [27]–[29] as well as in the context of sparse sampling, e.g., [16], [30]–[33]. However, an important property of our framework is that the shape of the filters in Fig. 2 is independent of the wavelet basis in which the pulse is sparse. The filter design is only affected by the sparsity parameter (i.e., by the number K of non-zero wavelet coefficients in $x_l(t)$). Consequently the same acquisition system can be used with pulses which are sparse on different wavelet bases as far as the basis is known at the receiver. This universality is an important aspect of the proposed method.

Finally, we propose a sequential reconstruction algorithm to reconstruct streaming pulses where we exploit the fact that pulses have short duration and that the MEMS have compact

This work was partially presented at ICASSP, 2014 [1].

This work is supported by the European Research Council (ERC) starting investigator award Nr. 277800 (RecoSamp).

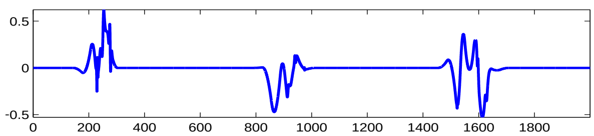


Fig. 1. Example of stream of pulses considered in this paper. This is a section of a long stream. The localized pulses can be different from each other, their shapes are unknown and the pulses are sparse in the wavelet domain.

support.

The paper is organized as follows. In Section II, we first introduce our model of a single pulse and then give an overview of our acquisition and reconstruction processes. Section III and Section IV explain the technical detail of these two processes. In particular, Section III shows how pulses can be reconstructed using only partial Fourier information, while Section IV introduces our new family of acquisition devices. To provide further intuitions, various examples are given in both sections. Section V explains the problem of sampling streams of pulses, where noiseless and noisy examples are also given. Section VIII concludes the paper.

II. Overview

We begin by analysing the perfect reconstruction of a single pulse, and then extend the approach to stream of pulses.

A. Pulse model and notations

Assume that $x(t) \in L_2(\mathbb{R})$ is a pulse that can be represented using a small number of non-zero coefficients in an orthogonal/biorthogonal wavelet basis. Without loss of generality, we assume that the finest scale containing non-zero wavelet coefficients is one, so that the signal can be written as

$$x(t) = \sum_n a_{J,n}\phi_{J,n}(t) + \sum_{m=1}^{J}\sum_n b_{m,n}\psi_{m,n}(t), \qquad (2)$$

where the number of decomposition levels J is known. Functions $\phi(t)$ and $\psi(t)$ are the scaling and wavelet function respectively, and we use the convention that $\phi_{J,n}(t) = 2^{-J/2}\phi(2^{-J}t - n)$ and $\psi_{m,n}(t) = 2^{-m/2}\psi(2^{-m}t - n)$. We assume that $\phi(t)$ and $\psi(t)$ have compact support and that the non-zero wavelet coefficients are all included in $[0,\, 2^S - 1]$. Therefore, $x(t)$ also has compact support. Moreover, the number of wavelet coefficients at scale m is 2^{S-m}. Obviously, the number of non-zero coefficients in each subband is bounded. We assume the bound to be K_m, namely,

$$||\{a_{J,n}\}||_0 \le K_{J+1} \le 2^{S-J}, \; ||\{b_{m,n}\}||_0 \le K_m \le 2^{S-m}, \qquad (3)$$

where the ℓ_0 "norm" is given by counting the number of the non-zero coefficients. We are especially interested in the case where $K_m < 2^{S-m-1}$ $(m = 1, \ldots, J)$ and $K_{J+1} < 2^{S-J-1}$.

Given the description above, we use $\mathcal{M}(\mathbf{K}, 2^S, \phi(t))$ to denote our signal model, where $\mathbf{K} = [K_1, \ldots, K_{J+1}]$. We say that $x(t)$ belongs to $\mathcal{M}(\mathbf{K}, 2^S, \phi(t))$ if $x(t)$ can be represented by (2) and all the non-zero coefficients are supported within the same 2^S interval.

Let $\hat{x}(\omega)$ be the Fourier transform of the signal $x(t)$ at frequency ω:

$$\hat{x}(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt. \qquad (4)$$

We use Ω to denote a set of frequencies, and $\hat{x}(\Omega)$ to denote the set of values of the Fourier transform of $x(t)$ at Ω: $\hat{x}(\Omega) = \{\hat{x}(\omega) : \omega \in \Omega\}$.

This paper is not only interested in the single pulse $x(t)$ that is sparse in the wavelet domain, but also in its delayed version, $x(t - t_0)$ with $0 \le t_0 < 2^J$. For simplicity, throughout the paper, $x(t)$ is always a wavelet-sparse signal, and its delayed version is denoted as $x'(t)$. The Fourier transform of $x(t-t_0)$ is then denoted as $\hat{x}'(\omega)$.

B. Overview of the acquisition and reconstruction process

1) Single pulse case: For a pulse $x(t) \in \mathcal{M}(\mathbf{K}, 2^S, \phi(t))$ that has $J+1$ subbands, the proposed acquisition and reconstruction process is described in Fig. 2.

The acquisition process is carried out by a filter-bank with $J+2$ channels. We show in Section IV how these $J+2$ filters $h_m(t)$ can be designed so that the resulting samples can be exactly mapped into the Fourier transform of $x'(t)$ at selected frequencies $\bar{\Omega}$ (Theorem 3). Specifically, channel m maps the samples $c_m[n]$ to the Fourier transform of $x'(t)$ at frequencies denoted by Ω_m and we denote the combination of all these frequency sets with $\bar{\Omega} = \cup_{m=0}^{J+1}\Omega_m$. Note that the proposed filters are real-valued and of compact support.

Given a proper set of frequencies $\bar{\Omega}$ (e.g. (5)-(9)), we show in Section III (Theorem 1 and Theorem 2) that the wavelet coefficients of $x(t)$ and the delay t_0 can be recovered from $\hat{x}'(\bar{\Omega})$. The pulse $x(t)$ is reconstructed by synthesis filtering the wavelet coefficients and then by delaying the signal by t_0 to obtain $x'(t)$. We note that since $t_0 \in \mathbb{R}$, $x'(t)$ is normally not-sparse even if $x(t)$ is. Therefore, the estimation of t_0 from the samples is of central importance.

Note that pulses that are sparse on different wavelet bases can be recovered using the same set of frequencies $\bar{\Omega}$. Given that the shape of the filters $h_m(t)$ $(m = 0, \ldots, J+1)$ is also unrelated to the wavelet basis, we conclude that the proposed acquisition process is not sensitive to the choice of wavelet basis. While the wavelet basis is not an issue, the sparsity parameter $\mathbf{K}$ and the support of the non-zero coefficients 2^S dictate the choice of frequencies $\bar{\Omega}$. This implies that our acquisition solution is universal for wavelet-sparse pulses whose structure can be described by the parameters $\mathbf{K}$ and 2^S. We also note that we use $J+2$ channels for simplicity and improved stability, however, the acquisition of the Fourier information of the signal can also be achieved using less channels. See Appendix B for more details.

It will also be clear in Section III that $\bar{\Omega}$ only contains a small number of frequencies. Since the acquisition process only needs to provide the required Fourier information at $\bar{\Omega}$, which is not much, it is possible to achieve a low sampling rate even though the pulse itself is not band limited.

2) Sequential reconstruction of streams of pulses: Because our filters have compact supports as well as the pulses of (2), it is possible to sample streams of pulses using the same system shown in Fig. 2. This point will be explored in Section V.

Intuitively, if the pulses are located sufficiently far apart, their samples will be separated from each other's. This is because the sampling filters have compact support and hence a

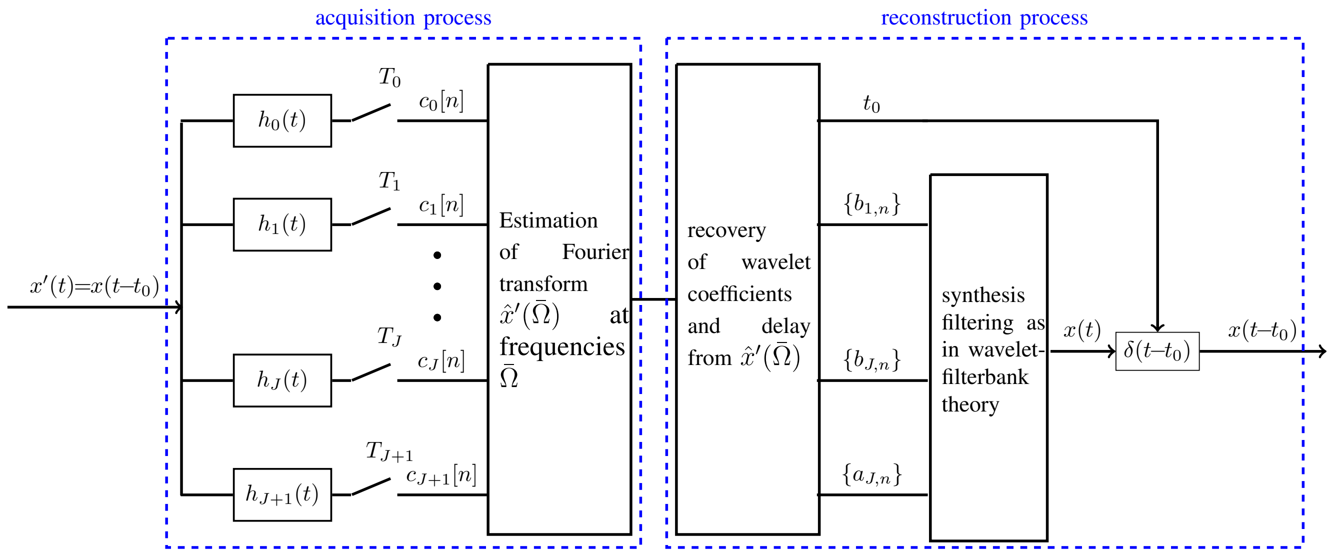


Fig. 2. The proposed acquisition and reconstruction system for single pulse $x(t-t_0)$.

local pulse will only affects samples within a certain interval. If the block of samples corresponding to one specific pulse can be isolated from the rest of the samples, we can use these samples to reproduce the Fourier information of that particular pulse, from which we can then reconstruct the pulse itself.

III. The reconstruction process of single pulses

Reconstructing $x(t-t_0)$ from its partial Fourier information is the key to our sampling solution. This section will answer the following questions:

- how to reconstruct non-delayed wavelet-sparse $x(t)$ from its partial Fourier information on selected frequencies (denoted as Ω); and what these frequencies are.
- how to reconstruct $x(t-t_0)$ which requires Fourier information on extra frequencies beyond Ω. The frequencies needed to reconstruct $x(t-t_0)$ are denoted as $\bar{\Omega}$.

We first explicitly give out the Ω and $\bar{\Omega}$ that we use in our system in Fig. 2.

Define

$$\alpha_i = \frac{2\pi}{2^S}(-1)^i\left(\lceil\frac{i}{2}\rceil - \frac{1}{2}\right), \tag{5}$$

then

$$\begin{cases}\Omega_{J+1}^{(N)}=\{\alpha_i : i=1:2N\}, \\ \Omega_m^{(N)}=\{p(\alpha_i,m) : i=1:2N\}, m=1,\ldots,J,\end{cases} \tag{6}$$

where

$$p(\omega,m)=\omega-\mathrm{sign}(\omega)2\pi/2^m. \tag{7}$$

Let

$$K_{\max}=\max\{K_1,\ldots,K_{J+1}\}. \tag{8}$$

Note that K_m is the sparse parameter. Then we have the Ω and $\bar{\Omega}$ defined as follows:

$$\begin{cases}\Omega_0 &= \{\bar{\omega}=p(\omega,0),\forall\omega\in\Omega_1\}, \\ \Omega &= \cup_{m=1}^J\Omega_m^{(K_m)}\cup\Omega_{J+1}^{(K_{\max})}, \\ \bar{\Omega} &= \Omega\cup\Omega_0.\end{cases} \tag{9}$$

Here, the superscript (N) is related to the number of elements in each individual set. In the rest of the paper, we will ignore the superscript when there is no confusion caused. For example, we would just say $\Omega=\cup_{m=1}^{J+1}\Omega_m$ from now on.

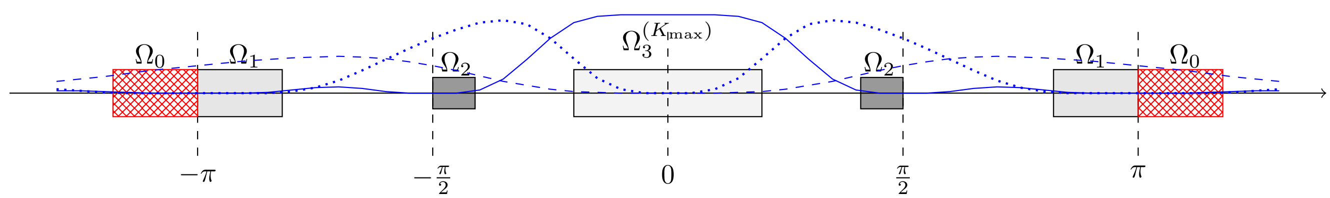
Fig. 3 illustrates two examples of $\bar{\Omega}$ for the case $J=2$ and for two different $\mathbf{K}$. Each $\bar{\Omega}$ is comprised of Ω and Ω_0. The Fourier information on Ω, $\hat{x}(\Omega)$, is what we need to reconstruct $x(t)$.

The construction of Ω is related to the sparsity parameter $\mathbf{K}=[K_1,\ldots,K_{J+1}]$. It is clear that there are $J+1$ elements in the sparsity parameter $\mathbf{K}$ and the frequency set Ω is the combination of $J+1$ subsets. Each of these subsets covers a different range of frequencies. The intuitive motivation for such construction is that the scaling function and wavelets normally have low-pass and band-pass behaviours respectively, and we want to sample the region that correlates most with the specific wavelet/scaling function. In our example, Ω_2 covers the region where energy of the wavelet $\psi_{2,0}(t)$ concentrates most while Ω_3 covers the low-pass region as the energy of the scaling function concentrates there.

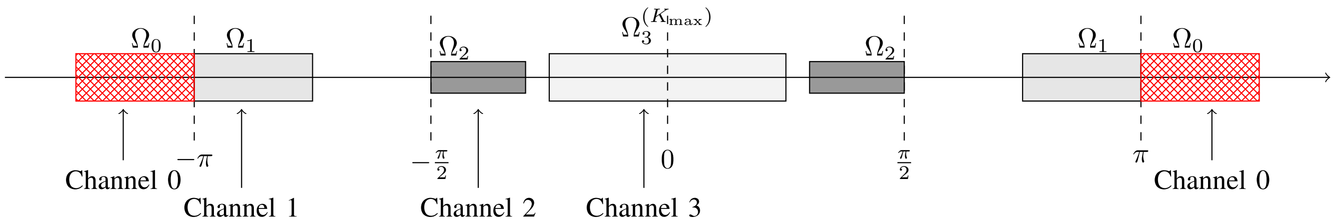
One might have also noticed that Ω_{J+1}, which covers the low frequencies, has the biggest cardinality among all the subsets. This presents some similarities to the variable density/multiscale sampling [34]–[36], and conforms to the empirical study that the low frequencies should be heavily sampled. However, the construction of Ω goes way beyond intuition and empirical study. Our selection is of a deterministic nature. There is no randomness in the choice of location and the cardinality of each subset is directly linked to K_m. This construction guarantees perfect reconstruction in a deterministic manner. We will be able to explain it at the end of Section III.A, after some key information is introduced.

A. *Reconstructing $x(t)$ from partial Fourier information*

This section is mostly based on our previous work [37]. For the sake of clarity, we highlight the most important aspects of the derivation that are useful for the rest of the paper.



(a) $K_1 > K_3 > K_2$



(b) $K_3 > K_1 \geq K_2$

Fig. 3. Illustration of Ω and Ω_0 with $J = 2$ for different $\mathbf{K} = [K_1, K_2, K_3]$. The set $\Omega = \cup_{m=1}^{J+1} \Omega_m$ is only the combination of $3 = J + 1$ subsets, and each of these subsets covers a different range of frequencies. This is partly because the scaling function and wavelets normally have low-pass and band-pass behaviours respectively, as illustrated by blue lines in (a). It is then natural to consider the filter-bank structure of $J + 1 + 1$ channels (the extra channel is for collecting information in Ω_0), of which each channel has a band-pass like filter for obtaining the Fourier information at each of these subsets. Also note that the cardinality of each subset directly links to the sparse parameter K_m. The cardinalities of the subsets are illustrated with shaded areas.

Theorem 1 ([37])**.** *Consider a signal $x(t) \in \mathcal{M}(\mathbf{K}, 2^S, \phi(t))$, and assume*[1]

$$\hat{\phi}(\omega) \neq 0, \forall \omega \in (-\pi, \pi). \tag{10}$$

Then $x(t)$ is uniquely characterized by the knowledge of its Fourier transform at Ω.

The most important aspect of this theorem is that, for any signal $x(t) \in \mathcal{M}(\mathbf{K}, 2^S, \phi(t))$, as long as the scaling function $\phi(t)$ satisfies the mild requirement (10), $x(t)$ can be exactly recovered from $\hat{x}(\Omega)$. The cardinality of Ω is as small as $2K_{\max} + 2\sum_{m=1}^{J} K_m$, which is a small number of samples of the signal's Fourier transform. Also note that, the frequencies in Ω_m are determined by K_m and the support of the non-zero coefficients 2^S, but are not related to the wavelet basis.

Theorem 1 is proved in [37] by giving a reconstruction algorithm (described by Algorithm 1) that perfectly reconstructs $x(t)$ from $\hat{x}(\Omega)$. The key insight of the derivation is that the wavelet representation of (2) can be written recursively as

$$\hat{x}_{m-1}(\omega) = \hat{x}_m(\omega) + \hat{b}_m(\omega)\hat{\psi}_{m,0}(\omega) \tag{11}$$

$$= \hat{a}_m(\omega)\hat{\phi}_{m,0}(\omega) + \hat{b}_m(\omega)\hat{\psi}_{m,0}(\omega) \tag{12}$$

where $x_0(t)$ is $x(t)$. Here $\hat{a}_m(\omega)$ and $\hat{b}_m(\omega)$ denote the DTFT of $b_m[l]$ and $a_m[l]$ which are defined as follows:

$$\begin{aligned} a_m[l] &= \begin{cases} a_{m,n}, & \text{if } l = 2^m n, \\ 0, & \text{otherwise;} \end{cases} \\ b_m[l] &= \begin{cases} b_{m,n}, & \text{if } l = 2^m n, \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{13}$$

Note that $\hat{a}_m(\omega)$ and $\hat{b}_m(\omega)$ are both periodic with period $2\pi/2^m$. Let

$$\omega_m = p(\omega_0, m);$$

because of the periodicity $2\pi/2^m$, we have that

$$\hat{a}_m(\omega_0) = \hat{a}_m(\omega_m) \text{ and } \hat{b}_m(\omega_0) = \hat{b}_m(\omega_m). \tag{14}$$

Writing (12) for frequencies $\omega = \omega_0$ and $\omega = \omega_m$, and expressing it in matrix-vector form yields

$$\underbrace{\begin{bmatrix} \hat{\phi}_{m,0}(\omega_0) & \hat{\psi}_{m,0}(\omega_0) \\ \hat{\phi}_{m,0}(\omega_m) & \hat{\psi}_{m,0}(\omega_m) \end{bmatrix}}_{\mathbf{\Phi}(\omega_0, \omega_m)} \begin{bmatrix} \hat{a}_m(\omega_0) \\ \hat{b}_m(\omega_0) \end{bmatrix} = \begin{bmatrix} \hat{x}_{m-1}(\omega_0) \\ \hat{x}_{m-1}(\omega_m) \end{bmatrix}; \tag{15}$$

where $\mathbf{\Phi}(w_0, w_m)$ is invertible if $\phi(t)$ satisfies (10) [37, Lemma 1]. Therefore $\hat{b}_m(\omega_0)$ can be found from (15). From the knowledge of $\hat{b}_m(\omega)$ at $2K_m$ evenly spaced frequencies, it is then possible to reconstruct the entire sequence $b_m[l]$ using Prony's method.

The reconstruction process (Algorithm 1) is from the finest scale ($m = 1$) to the coarse scale ($m = J$). After recovering $b_m[l]$ from $\hat{x}_{m-1}(\Omega_{J+1}^{(K_m)})$ and $\hat{x}_{m-1}(\Omega_m)$, its Fourier contribution is removed from $\hat{x}_{m-1}(\Omega)$, so as to prepare for the recovery of $b_{m+1}[l]$. At the end, $\hat{x}_J(\omega) = \hat{a}_J(\omega)\hat{\phi}_{J,0}(\omega)$ ($\omega \in \Omega_{J+1}^{(K_1)}$), from which we recover $a_J[l]$ and so $x(t)$[2].

Now with all this information in hand, we are able to explain a bit more on the reconstruction of Ω. The sparsity parameter $\mathbf{K} = [K_1, \ldots, K_{J+1}]$ decides the structure of Ω. There are $J+1$ elements in the sparsity parameter $\mathbf{K}$ and the frequency set Ω is the combination of $J+1$ subsets, where each subset contains at least $2K_m$ frequencies. The frequencies in these

[1]The scaling functions of many orthogonal/biorthogonal bases satisfy (10). See, for example, [38], [39] .

[2]Note that Algorithm 1 is constructive proof of Theorem 1. We use Prony's method in order to get the tight bound on the cardinality of Ω, namely $2K_{\max} + 2\sum_{n=1}^{J} K_n$. In practice, one can use ℓ_1 minimization to replace the whole Algorithm 1 or the stable variants of Prony's method, like matrix pencil algorithm [26], [40], to reconstruct $b_m[l]$ and $a_J[l]$ in step 10 and 14. We note that the ℓ_1 minimization normally requires slightly more frequencies to reconstruct $x(t)$ and it is not deterministically guaranteed to find the correct solution.

Algorithm 1 Recursive algorithm for perfectly reconstructing wavelet-sparse signals

Input the number of non-zeros $K_1, \ldots, K_{J+1}$.
Input $\hat{x}(\omega)$ with $\omega \in \Omega = \cup_{m=1}^{J+1} \Omega_m$.
Set $N_m = \sharp\Omega_m$, the cardinality of Ω_m.
for $m = 1$ to J **do**
  $N = \min(N_m, N_{J+1})/2$;
  for every ω_0 in $\Omega_{J+1}^{(N)}$ **do**
    $\omega_m = p(\omega_0, m) = \omega_0 + \mathrm{sign}(\omega_0)2\pi/2^m$;
    solve the linear equation (15).
  end for
  recover $b_m[l]$ from $\hat{b}_m(\Omega_{J+1}^{(N)})$ with the Prony's method.
  remove $\hat{b}_m(\omega)\hat{\psi}_m(\omega)$ from $\hat{x}_{m-1}(\Omega)$ to obtain $\hat{x}_m(\Omega)$ (11).
end for
for every ω in Ω_{J+1} **do**
  $\hat{a}_J(\omega) = \hat{x}_J(\omega)/\hat{\phi}_J(\omega)$.
end for
recover $a_J[l]$ from $\hat{a}_J(\Omega_{J+1})$
reconstruct $x(t)$ by (2).

subsets have to satisfy

$$\{\omega_0 = p^{-1}(\omega_m, m), \forall \omega_m \in \Omega_m, m = 1, \ldots, J\} \subseteq \Omega_{J+1}, \tag{16}$$

only then the linear equation (15) can be obtained, and it is the key for Algorithm 1 to work. The frequencies in each Ω_m needs to be equally spaced because we use Prony's method and at least $2K$ frequencies are needed to reconstruct a K-sparse signal[3]. This is why we construct Ω_{J+1} with equally spaced frequencies α_i (5).

B. Reconstructing $x(t - t_0)$

Assume now we only have the Fourier transform of $x(t-t_0)$ ($0 < t_0 < 2^J$), namely $\hat{x}'(\omega)$. Since the wavelet transform is not shift invariant, $x(t - t_0)$ is usually not sparse in the wavelet domain. Therefore, we need to estimate t_0 and correct the shift from the observation $\hat{x}'(\omega)$ to get $\hat{x}(\omega)$ and then reconstruct the wavelet-sparse $x(t)$ using Algorithm 1. The Fourier information $\hat{x}'(\omega)$ on Ω is not sufficient for this task. The estimation of t_0 is only possible when extra information at higher frequencies is available, and this extra information is contained in $\hat{x}(\Omega_0)$.

Theorem 2. *Assume the scaling function satisfies* (10). *If there is at least one frequency $\omega_0 \in \Omega_1$ such that $\hat{x}'(\bar{\omega}_0) \neq 0$ where $\bar{\omega}_0 = p(\omega_0, 0)$, then the non-integer part of the delay can be determined by*

$$r_{t_0} = \mathrm{rem}(t_0, 1) = \log\left(\frac{\hat{x}'(\omega_0)/\hat{\phi}_{0,0}(\omega_0)}{\hat{x}'(\bar{\omega}_0)/\hat{\phi}_{0,0}(\bar{\omega}_0)}\right)/(2\pi j), \tag{17}$$

Proof: First we can write

$$x(t) = \sum_n a_{0,n}\phi_{0,n}(t),$$

since $x(t)$ lies in the subspace spanned by $\{\phi_{0,n}(t)\}_{n\in\mathbb{Z}}$. In Fourier domain, this becomes

$$\hat{x}(\omega) = \hat{a}_0(\omega)\hat{\phi}_{0,0}(\omega). \tag{18}$$

Since $\Omega_0 = p(\Omega_1, 0)$, we can always find the corresponding $\omega_0 \in \Omega_1$ for any $\bar{\omega}_0 \in \Omega_0$ such that $\bar{\omega}_0 = p(\omega_0, 0)$. Without loss of generality, we assume $\omega_0 \leq 0$, and therefore $\bar{\omega}_0 = \omega_0 + 2\pi$. Because the period of $\hat{a}(\omega)$ is 2π, this leads to $\hat{a}_0(\omega_0) = \hat{a}_0(\bar{\omega}_0)$. Moreover, from (18) and assuming $\hat{x}(\bar{\omega}_0) \neq 0$, we have that

$$\hat{a}_0(\omega_0) = \hat{a}_0(\bar{\omega}_0) \neq 0 \text{ and } \hat{\phi}_{0,0}(\bar{\omega}_0) \neq 0.$$

Also since

$$\begin{bmatrix}\hat{x}'(\omega_0)\\ \hat{x}'(\bar{\omega}_0)\end{bmatrix} = \begin{bmatrix}e^{-j\omega_0 t_0}\hat{x}(\omega_0)\\ e^{-j\bar{\omega}_0 t_0}\hat{x}(\bar{\omega}_0)\end{bmatrix} = \begin{bmatrix}e^{-j\omega_0 t_0}\hat{\phi}_{0,0}(\omega_0)\hat{a}_0(\omega_0)\\ e^{-j\bar{\omega}_0 t_0}\hat{\phi}_{0,0}(\bar{\omega}_0)\hat{a}_0(\bar{\omega}_0)\end{bmatrix},$$

we have that

$$1 = \frac{\hat{a}_0(\omega_0)}{\hat{a}_0(\bar{\omega}_0)} = \frac{e^{j\omega_0 t_0}\hat{x}'(\omega_0)/\hat{\phi}_{0,0}(\omega_0)}{e^{j\bar{\omega}_0 t_0}\hat{x}'(\bar{\omega}_0)/\hat{\phi}_{0,0}(\bar{\omega}_0)},$$

which leads to

$$e^{j2\pi t_0} = \frac{\hat{x}'(\omega_0)/\hat{\phi}_{0,0}(\omega_0)}{\hat{x}'(\bar{\omega}_0)/\hat{\phi}_{0,0}(\bar{\omega}_0)}.$$

Note here that $\hat{\phi}_{0,0}(\omega_0) \neq 0$ because $\omega_0 \in \Omega_1 \subset (-\pi, \pi)$ and the scaling function satisfies (10).
Because $e^{j2\pi t}$ is periodic with period 1, we obtain (17). ■

Now let $d = t_0 - rem(t_0, 1)$, then d is an integer and

$$d \in \mathbb{Z} \cap [0, 2^J - 1] = [0, 1, \ldots, 2^J - 1].$$

The delay d can be retrieved by a line search on $\mathbb{Z}\cap[0, 2^J - 1]$. For $k = [0, 1, \ldots, 2^J - 1]$, we apply Algorithm 1 on $e^{-j\omega(k+r)}\hat{x}'(\omega)\big|_{\omega\in\Omega}$. The output of the recursive algorithm is denoted as $x^{(k)}(t)$. If $e^{j\omega(k+r)}\hat{x}^{(k)}(\omega) = \hat{x}'(\omega)$ for all $\omega \in \Omega$, k is the estimate of d.

We summarize the reconstruction of $x(t - t_0)$ in Algorithm 2.

Algorithm 2 Reconstructing wavelet-sparse signals with unknown delay

Input the number of non-zeros $K_1, \ldots, K_{J+1}$.
Input $\hat{x}'(\omega)$ with $\omega \in \bar{\Omega} = \cup_{m=1}^{J+1}\Omega_m \cup \Omega_0$.
Set $N_m = \sharp\Omega_m$, the cardinality of Ω_m.
Pick one $\bar{\omega} \in \Omega_0$ such that $\hat{x}'(\bar{\omega}) \neq 0$;
Find the corresponding ω such that $\bar{\omega} = p(\omega, 0)$.
Compute $\mathrm{rem}(t_0, 1)$ by (17).
Set $\hat{x}'(\omega) = e^{j\omega\mathrm{rem}(t_0,1)}\hat{x}'(\omega),\ \forall\, \omega \in \Omega$.
Set $d = -1$;
repeat
  $d = d + 1$;
  $\hat{x}'(\omega) = e^{j\omega d}\hat{x}'(\omega),\ \forall\, \omega \in \Omega$;
  Apply Algorithm 1 to $\hat{x}'(\Omega)$ to obtain an estimate $x(t)$;
until $\sum_{\omega\in\Omega} |\hat{x}(\omega) - \hat{x}'(\omega)|^2 \leq \Delta$ $\{\Delta$ is a small non-negative number$\}$
delay $x(t)$ by $d + \mathrm{rem}(t_0, 1)$ as an estimate of $x'(t)$.

[3]For an overview of Prony's method, we refer to [41].

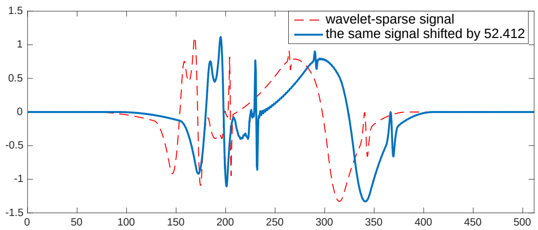


Fig. 4. The wavelet-sparse signal $x(t)$ and its shifted version $x(t-52.412)$. Note that only the non-zero coefficients need to be supported in $[0, 2^S-1]$. It is not necessarily for the wavelets corresponding to non-zero coefficients to be also supported on $[0, 2^S-1]$.

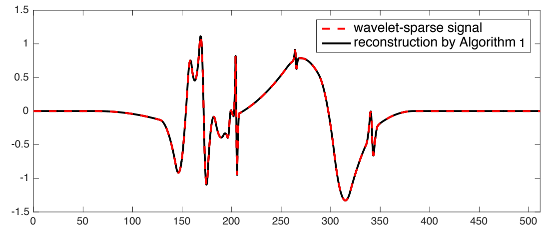
C. Examples and discussion

We now give a few reconstruction examples to demonstrate that $x'(t)$ can be perfectly recovered from $\hat{x}'(\bar{\Omega})$. We highlight here again the fact that the perfect reconstruction is possible because $x(t)$ is sparse in a wavelet basis and that the result is independent of the chosen wavelet.

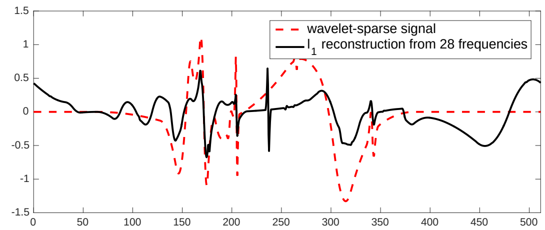
Our example uses the signals in Fig. 4. The signal $x(t)$ is sparse in the cubic-spline wavelet, with the sparse parameter $K_m \leq 2$ and $1 \leq m \leq 7$. The support of the non-zero coefficients is 512.

As shown in Fig. 5, without the shift, both Algorithm 1 and the ℓ_1 minimization can perfectly reconstruct the wavelet-sparse $x(t)$ from its partial Fourier information, even though ℓ_1 requires more frequencies. However, neither Algorithm 1 nor the ℓ_1 minimization can reconstruct the delayed signal $x'(t)$, because $x'(t)$ is not sparse (Fig. 6).

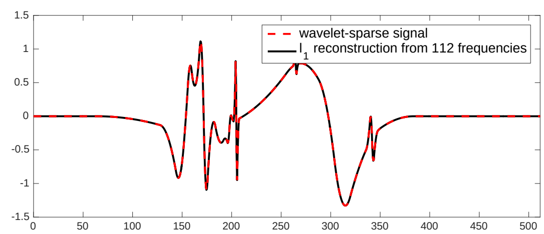
Fig. 7 shows that the delayed signal $x'(t)$ can be perfectly reconstructed by Algorithm 2 with the help of the extra Fourier information at Ω_0. The key to the success of the algorithm is that the shift is estimated and removed in the reconstruction process.



(a) Reconstruction by Algorithm 1

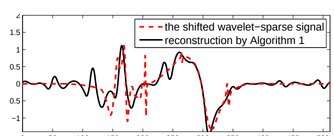


(b) Reconstruction by ℓ_1

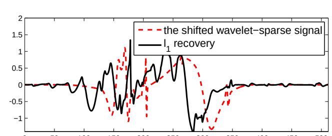


(c) Reconstruction by ℓ_1

Fig. 5. Exact wavelet-sparse signals can be perfectly reconstructed from $\hat{x}(\Omega)$. In this example, $J=6$, $K_m \leq 2$ for $1 \leq m \leq 7$. (a) Algorithm 1 reconstructs $x(t)$ exactly from 28 frequencies ($\sharp\Omega_m = 4$ for all m). (b) The reconstruction by ℓ_1 minimization from the same set of frequencies is not successful. (c) ℓ_1 minimization reconstruct $x(t)$ exactly when the number of frequencies in Ω is increased to 112 ($\sharp\Omega_m = 16$ for all m). While our algorithm uses only 28 frequencies, the ℓ_1 minimization requires more frequencies.



(a) Algorithm 1



(b) ℓ_1 minimization

Fig. 6. Both Algorithm 1 and ℓ_1 minimization fail to reconstruct the delayed wavelet-sparse signal $x'(t)$. The signal $x(t)$ in Fig.5 is delayed by 52.412 to obtain $x'(t)$. There are 112 frequencies in Ω and 16 frequencies in Ω_0. $\bar{\Omega} = \Omega_0 \cup \Omega$.

IV. The acquisition process

When an analog signal is sampled by the ADC unit (Fig. 8), the discrete samples at the output are given by

$$c[n] = \langle x(t), h(-t+nT)\rangle = \langle x(t), \varphi(t/T - n)\rangle. \tag{19}$$

Note that the sampling kernel $\varphi(t) = h(-tT)$ is the scaled and time reversed version of the unit impulse response of the acquisition device.

The acquisition process of Fig. 2 is a multichannel version of the basic ADC structure of Fig. 8. We want the sampling kernel $\varphi_j(t) = h_j(-T_j t)$ ($j = 1, \ldots, J+2$ of Fig. 2 to be able to reproduce exponentials so that we can retrieve the Fourier transform of $x'(t)$ at frequencies $\bar{\Omega}$ from the samples $\{c_m[n]\}_{m=1}^{J+2}$ in order to use Algorithm 1 and Algorithm 2 to reconstruct $x'(t)$. We also want them to be of compact support so that we can apply our reconstruction algorithm sequentially. This point will be more evident in Sec. V. Finally, we want the kernels to be real-valued since normal acquisition devices behave like real-valued filters.

This section will introduce the sampling kernels that satisfy the above requirements. The resulting kernels are called modulated E-Splines with Multiple subbands (MEMS) and belong to the family of exponential reproducing kernels. These are

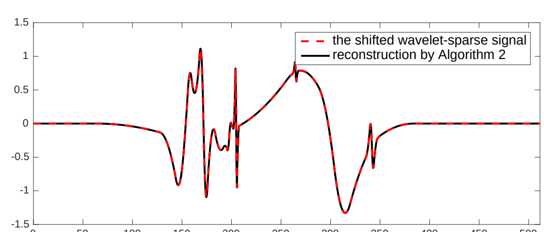


Fig. 7. The delayed wavelet-sparse signal $x'(t)$ can be perfectly reconstructed from $\hat{x}'(\bar{\Omega})$ using Algorithm 2. The signal $x(t)$ in Fig.5 is delayed by 52.412 to obtain $x'(t)$. The frequencies set $\bar{\Omega}$ only contains 32 samples, where $\sharp\Omega_m = 4$ ($m = 0, \ldots, 7$).

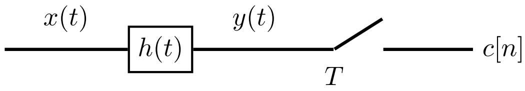


Fig. 8. The structure of ADC sampling where $h(t) = \varphi(-t/T)$. The function $h(t)$ is the unit impulse response of the acquisition device, and $\varphi(t)$ is the sampling kernel.

developed using the E-Spline, which is the elementary member of the family of exponential reproducing functions [42].

A. Overview of E-Splines

This section provides a brief overview of E-Splines. For more information on the topic, we refer to [15], [26], [42] and references therein.

An E-Spline $\beta_{\mathbf{a}}(t)$ is a function with Fourier transform

$$\hat{\beta}_{\mathbf{a}}(\omega) = \prod_{m=0}^{P} \frac{1 - e^{a_m - j\omega}}{j\omega - a_m}, \tag{20}$$

and has compact support $P+1$. Here $\mathbf{a} = [a_0, a_1, \ldots, a_P]^T$ is the vector containing the generating parameters and a_m can be a real or complex number. Moreover, when the parameter α_m are in complex conjugate pairs, the E-Spline is real valued. In this paper, we only consider the special case where a_m is purely imaginary, i.e. $a_m = j\omega_m$ with ω_m being real. Therefore, for the rest of the paper we use

$$\mathbf{a} = j[\omega_0, \omega_1, \ldots, \omega_n]^T = j\boldsymbol{\omega}. \tag{21}$$

E-Spline $\beta_{j\boldsymbol{\omega}}(t)$ can reproduce exponential $e^{j\omega_m t}$. This is to say that there exist weights $\gamma_{m,n}$ such that

$$\sum\nolimits_{k\in\mathbb{Z}} \gamma_{m,n}\beta_{j\boldsymbol{\omega}}(t-n) = e^{j\omega_m t}. \tag{22}$$

The weights $\gamma_{m,n}$ are given by [15], [26]:

$$\gamma_{m,n} = \gamma_{m,0} e^{j\omega_m n} \text{ with } \gamma_{m,0} = \frac{1}{\hat{\beta}_{j\boldsymbol{\omega}}(\omega_m)}. \tag{23}$$

Using purely imaginary generating parameters allows us to estimate the Fourier transform of $x(t)$. Indeed, once we have sampled the input signal $x(t)$ with the exponential reproducing kernel $h(t) = \beta_{j\boldsymbol{\omega}}(-t/T)$, we obtain $\hat{x}(-\omega_m/T)$ from

$$\begin{aligned} s_m =& \sum_{k\in\mathbb{Z}} \gamma_{m,n} c[n] = \sum_{k\in\mathbb{Z}} \gamma_{m,n} \langle x(t), \beta_{j\boldsymbol{\omega}}(t/T - k)\rangle \\ =& \langle x(t), \sum_{k\in\mathbb{Z}} \gamma_{m,n} \beta_{j\boldsymbol{\omega}}(t/T - k)\rangle \\ =& \langle x(t), e^{j\frac{\omega_m}{T}t}\rangle = \hat{x}(-\frac{\omega_m}{T}). \end{aligned} \tag{24}$$

This convenient relationship between the discrete spatial samples and the Fourier transform is at the heart of our sampling method.

We note that E-Splines are sometimes not stable [24], [25], especially when the generating parameters are located in band-pass regions. Take the E-spline shown in Fig. 9 as an example. This E-spline is unstable because the ratio $\frac{\min_{w_n} |\hat{\beta}_\omega(w_n)|}{\max_{w_n} |\hat{\beta}_\omega(w_n)|}$ (about 10^{-2}) is small. Therefore we need to introduce a new construction of exponential reproducing kernels to overcome this limitation. MEMS is stable because the ratio $\frac{\min_{w_n} |\hat{\beta}_\omega(w_n)|}{\max_{w_n} |\hat{\beta}_\omega(w_n)|} \approx 1$ by design.

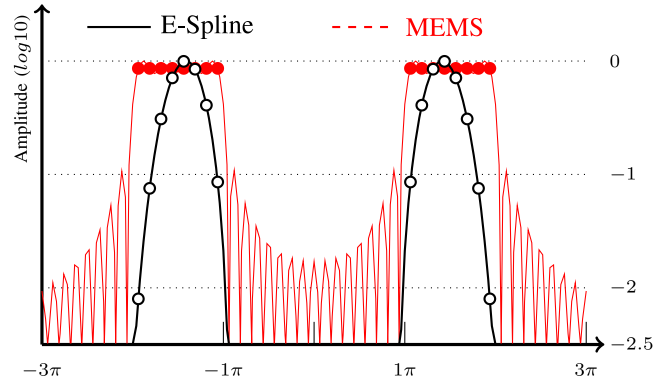


Fig. 9. Illustration of the instability problem of the E-spline. The conventional E-spline $\beta_{\mathbf{w}}(t)$ reproduces exponential e^{jwt} with $w \in \boldsymbol{w} = \{\frac{\pm k}{16} | k = 17 : 2 : 31\}$ and the frequency amplitudes at these frequencies are shown by black circles. Meanwhile, the MEMS reproducing the same frequencies is shown in thin red lines and the amplitudes at the reproduced frequency are marked by red solid dots. The kernels are scaled so that the maximum frequency amplitude is 1.

B. Modulated real E-Spline with multiple subbands (MEMS)

We define a MEMS as follows:

$$\varphi_{j\boldsymbol{\omega}}^{M,N}(t) = \beta_{j\boldsymbol{\omega},2M}(t) \left(\sum_{i=1}^{N} 2b_i \cos\left(\frac{2k_i+1}{2M}\pi t \right) \right), \tag{25}$$

where

$$\hat{\beta}_{j\boldsymbol{\omega},2M}(w) = \prod_{m=0}^{P} \frac{1 - e^{2Mj(\omega_m - \omega)}}{2Mj(\omega - \omega_m)}, \tag{26}$$

b_i is real and non-zero and $N \leq M$. Note that, as the bandwidth of the sampling device is fixed to be 2π and each term in (26) occupies $2\pi/M$, we can have at most M and we set $N = M$ for the most efficient use of the bandwidth. For the the rest of the paper, we use $\hat{\varphi}_{j\boldsymbol{\omega}}^{M}(t)$ to denote $\hat{\varphi}_{j\boldsymbol{\omega}}^{M,M}(t)$.

An important property of MEMS is that it has compact support $2M(P+1)$. This is because $\beta_{j\boldsymbol{\omega},2M}(t)$ has by construction compact support $2M(P+1)$.

Theorem 3. *Let $r_i = rem(k_i, 2M)$. Assume*

$$\begin{cases} \omega_0 \leq \omega_1 \leq \cdots \leq \omega_P, |\omega_m| < \dfrac{\pi}{2M}, \\ \omega_m = -\omega_{P-m}, \end{cases} \tag{27}$$

then MEMS $\varphi_{j\boldsymbol{\omega}}^{M}(t)$ can reproduce exponential $e^{\pm j\omega_{m,i} t}$ if

$$r_i + r_l \neq 2M - 1 \text{ and } \min_{i \neq l} |r_i - r_l| \neq 0. \tag{28}$$

Here $\omega_{m,i} = \omega_m + \frac{2k_i+1}{2M}\pi$.

The proof of Theorem 3 is given in Appendix A and it is based upon the properties of E-Splines.

Once we set the parameters $M, \boldsymbol{\omega}$ and $\{k_i\}$, we can design the MEMS and calculate the weights $\gamma_{m,i,n}$ for generating $e^{j\omega_{m,i} t}$ by replacing $\hat{\beta}_{j\boldsymbol{\omega}}(\omega_m)$ with $\hat{\varphi}_{j\boldsymbol{\omega}}^{M}(\omega_{m,i})$ in (23), namely,

$$\gamma_{m,i,n} = c_{m,i,0} e^{j\omega_{m,i} n} \text{ with } \gamma_{m,i,0} = \frac{1}{\hat{\varphi}_{j\boldsymbol{\omega}}^{M}(\omega_{m,i})}. \tag{29}$$

C. Setting the acquisition filters of Fig. 2

For a pulse $x(t)$ from $\mathcal{M}(\mathbf{K}, 2^S, \phi(t))$, we need $\hat{x}'(\bar{\Omega})$ in order to perfectly reconstruct $x'(t)$. This section focuses on the problem of designing the appropriate MEMS as the sampling kernel, so that the samples $c_m[n]$ can be used to reproduce the appropriate Fourier information at $\bar{\Omega}$. For example, if the sampling kernel in Channel 1 of Fig. 2 is $h_1(t) = \varphi^M_{j\boldsymbol{\omega}}(-t/T_1)$, where $\varphi^M_{j\boldsymbol{\omega}}(t)$ is MEMS, then the Fourier information that $c_1[n]$ can reproduce is $\hat{x}(\pm\omega_{n,i}/T_1)$, and we need

$$\Omega_1 \subseteq \{\pm\omega_{n,i}/T\}_{n=0,\dots,P,\, i=1,\dots,M}, \tag{30}$$

which is equivalent to

$$\Omega_1 T \subseteq \{\pm\omega_{n,i}\}_{n=0,\dots,P,\, i=1,\dots,M}. \tag{31}$$

To achieve this, we need to set the parameters $\{\boldsymbol{\omega}, T, M, k_1, \dots, k_M\}$ correctly.

In Theorem 3, there are a number of constraints that need to be met to design the right MEMS. While this could be achieved in several ways, here we only consider the following two more intuitive cases,

1) $M = 1$ and $P + 1 = 2^{\lceil \log_2(\sharp\Omega_m)\rceil - 1}$;
2) $P = 1$ and $2M = 2^{\lceil \log_2(\sharp\Omega_m)\rceil - 1}$.

Here $\sharp\Omega_m$ denotes the cardinality of Ω_m. In both cases, the other parameters are given by

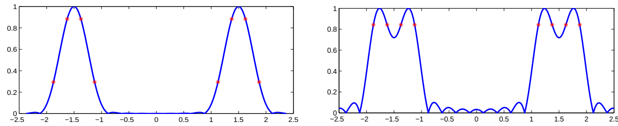
$$\begin{aligned}
\boldsymbol{\omega} &= \{\pm\frac{2n-1}{P+1}\frac{\pi}{2M}, n = 1, \dots, P\} \\
K &= \begin{cases} \frac{2^{S+1-\max(m,1)}}{2M(P+1)} - (m \geq 1), & \text{if } m {\leq} J{+}1 \\ 0, & \text{if } m{=}J{+}1 \end{cases} \\
k_i &= KM + i - 1, i = 1, \dots, M; \\
T_m &= \frac{2^S}{2M(P+1)}.
\end{aligned} \tag{32}$$

These calculated parameters satisfy Theorem 3, so the generated MEMS can reproduce exponentials $e^{\pm j\omega_{m,i}t}$ with $\omega_{m,i} = \omega_m + \frac{2k_i+1}{2M}\pi$. The resulting filter behaves like low-pass/band-pass filters. An example of band-pass MEMS is illustrated in Fig. 10 and further intuition on this construction is provided in Appendix B. Usually, the setting $P = 1$ and $2M = 2^{\lceil \log_2(\sharp\Omega_m)\rceil - 1}$ results in filters that are stabler and hence are preferred. Using these settings, the overall sampling rate of the system is

$$\sum_{m=0}^{J+1} \frac{1}{T_m} = \frac{1}{2^S} \sum_{m=0}^{J+1} 2^{\lceil \log_2(\sharp\Omega_m)\rceil}.$$

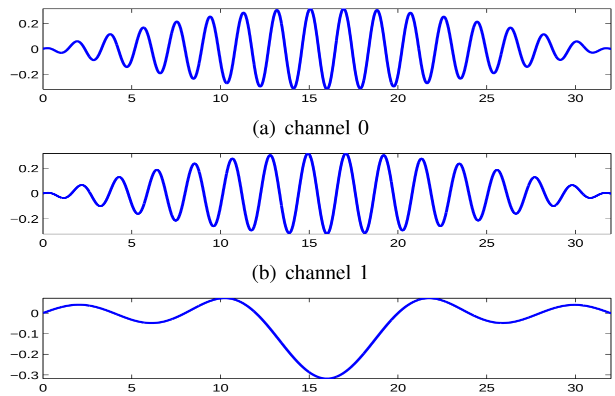
Because $\sharp\Omega_m$ is related to the sparse parameter $\mathbf{K}$, $2^{\lceil \log_2(\sharp\Omega_m)\rceil}/2^S$ can be very small if $K_m/2^S$ is small, i.e. the signal is sparse. It means the overall sampling rate can be significantly smaller than 1. An example will be given in Section IV.D to explain this further.

Note that (32) corresponds to one of the simplest stable MEMS, but there are alternative MEMS that can be used in our formulation. In fact, MEMS and the multichannel system can even be used to obtain Fourier information at pseudo random frequencies. (See Appendix B.)



(a) $M = 1, P = 3$ (b) $P = 1, M = 2$

Fig. 10. Example of band-pass MEMS. Here the red $*$ corresponds to $|\hat{\varphi}^M_{j\boldsymbol{\omega}}(\omega_{n,i})|$. Both band-pass MEMS are constructed using the two simplified cases, 1) $M = 1, P = 3$, 2) $M = 2, P = 1$. Both functions can reproduce the same exponentials, but the function in (b) is more stable in reproducing exponentials. A filter is more stable when $\frac{\min_{\omega_n} \hat{\varphi}^M_{j\boldsymbol{\omega}}(\omega)}{\max_{\omega_n} \hat{\varphi}^M_{j\boldsymbol{\omega}}(\omega)}$ is close to 1. One unit on the x-axis corresponds to π.



(a) channel 0

(b) channel 1

(c) channel 2

Fig. 11. Examples of acquisition filters $h_0(t)$, $h_1(t)$ and $h_2(t)$.

D. Examples

We now give an example of constructing the right MEMS to sample pulse $x'(t)$. Here $x'(t)$ is the delayed version of $x(t) \in \mathcal{M}([4, 2], 32, \phi(t))$. We consider $2^S = 32, \sharp\Omega_1 = 4$ and $\sharp\Omega_2 = 8$. By fixing $P = 1$, one can easily calculate the parameters for the MEMS at each channel using (32), and then calculate the reproduced frequencies of these MEMS using Theorem 3. All these parameters are shown in Table I.

With the three sets of parameters, we can construct the corresponding MEMS using (25), and the resulting kernels are denoted as $\varphi^M_{j\boldsymbol{\omega}_0}(t)$, $\varphi^M_{j\boldsymbol{\omega}_1}(t)$ and $\varphi^M_{j\boldsymbol{\omega}_2}(t)$. We set the filters at Channel 0, 1 and 2 as

$$h_m(t) = \varphi^M_{j\boldsymbol{\omega}_m}(-t/T_m),\ m = 0,\ 1,\ 2$$

and these filters are shown in Fig. 11 (a)-(c).

Fig. 12 shows an example where the samples taken with these filters lead to an exact reproduction of the signal's Fourier information at Ω_0, Ω_1 and Ω_2. The wavelet-sparse pulse $x(t) \in \mathcal{M}([4, 2], 32, \phi(t))$, where $\phi(t)$ is a cubic-spline, is shown in Fig. 12 (a) together with $x'(t)$. The samples, $c_0[n]$, $c_1[n]$ and $c_2[n]$, are shown in (b)-(d). The exact reproduction of the Fourier transform of $x(t)$ and $x'(t)$ at Ω_0, Ω_1 and Ω_2 is shown in Fig. 12 (e) together with the complete Fourier transform. From this reproduced Fourier information, we can exactly reconstruct $x(t)$ or $x'(t)$ using Algorithm 1 and Algorithm 2.

Now we show that our method has an overall sampling rate smaller than 1. For the single pulse $x(t)$ without any delay, we only need Channel 1-2 and the overall sampling rate is

TABLE I

PARAMETERS OF MEMS FOR SAMPLING WAVELET-SPARSE SIGNALS WITH PARAMETERS $\mathbf{K} = [4,\ 2]$ AND $2^S = 32$

Channel	$\Omega_n (n = 0,\ 1\ 2)$	Parameters for MEMS (fixing $P = 1$)	Reproduced frequencies
0	$\Omega_0 = \{\pm\frac{16.5}{16}\pi, \pm\frac{17.5}{16}\pi\}$	$M = 1,\ \boldsymbol{\omega} = [-\frac{1}{2}, \frac{1}{2}]\frac{\pi}{2},\ K = 8,\ k_1 = 8,\ T_0 = 8$	$\frac{\pm(\boldsymbol{\omega}+\frac{2k_1+1}{2M})}{T_0} = \pm\frac{16.5}{2T_1}\pi, \pm\frac{17.5}{2T_0}\pi$
1	$\Omega_1 = \{\pm\frac{15.5}{16}\pi, \pm\frac{14.5}{16}\pi\}$	$M = 1,\ \boldsymbol{\omega} = [-\frac{1}{2}, \frac{1}{2}]\frac{\pi}{2},\ K = 7,\ k_1 = 7,\ T_1 = 8$	$\frac{\pm(\boldsymbol{\omega}+\frac{2k_1+1}{2M})}{T_1} = \pm\frac{14.5}{2T_1}\pi, \pm\frac{15.5}{2T_1}\pi$
2	$\Omega_2 = \{\pm\frac{0.5}{16}\pi, \pm\frac{1.5}{16}\pi, \pm\frac{2.5}{16}\pi, \pm\frac{3.5}{16}\pi\}$	$M = 2,\ \boldsymbol{\omega} = [-\frac{1}{2}, \frac{1}{2}]\frac{\pi}{4},\ K = 0,\ k_1 = 0,\ k_2 = 1,$ $T_2 = 4$	$\frac{\pm(\boldsymbol{\omega}+\frac{2k_1+1}{2M})}{T_2} = \pm\frac{0.5}{4T_2}\pi, \pm\frac{1.5}{4T_2}\pi$ $\frac{\pm(\boldsymbol{\omega}+\frac{2k_2+1}{2M})}{T_2} = \pm\frac{2.5}{4T_2}\pi, \pm\frac{3.5}{4T_2}\pi$

obtained by adding together the sampling rate of each channel:

$$\sum_{m=1}^{2} \frac{1}{T_m} = \frac{12}{32}, \tag{33}$$

where $T_1 = 8$ and $T_2 = 4$ as calculated before in Table I. For sampling $x'(t)$, our system needs information on Ω_0 in order to recover the delay. The overall sampling rate is

$$12/32 + 1/T_0 = 16/32.$$

Here $T_0 = 8$ as in Table I.

When the signal is sparser, the system can reach an even lower sampling rate. For example, to sample a signal in $\mathcal{M}([4, 2], 512, \phi(t))$, the overall sampling rate could be as low as $16/512$. Note that the conventional wavelet-filterbank always has an overall sampling rate of 1; moreover it can not reconstruct $x'(t)$ exactly, because $x'(t)$ is not in the subspace spanned by the scaling function at scale zero.

V. ACQUISITION AND RECONSTRUCTION OF STREAMS OF PULSES

Using the findings of the previous sections, we are now able to sample and reconstruct streams of pulses using the acquisition system of Fig. 2. Our approach takes advantage of the fact that the filters in Fig. 2 are MEMS which are by construction of compact support. Moreover, we assume that the pulses have compact support 2^S.

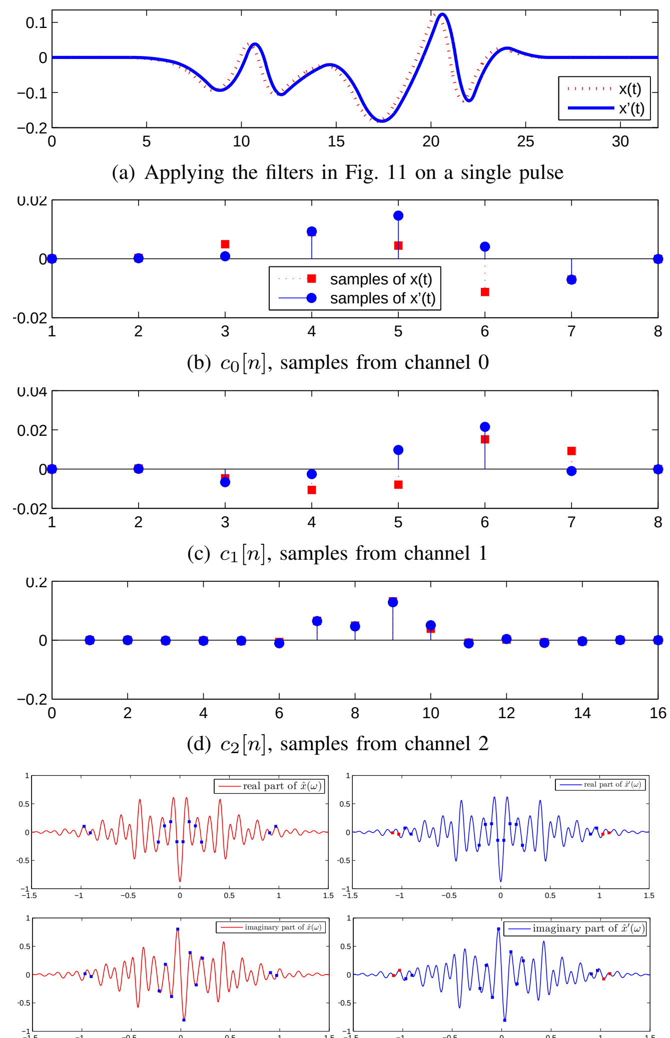
Specifically, we assume that the signal is made of unknown pulses $x_l(t - t_l)$ with $x_l(t) \in \mathcal{M}(\mathbf{K}, 2^S, \phi(t))$. We also assume that the distance between any two neighbouring pulses satisfies

$$\min_l(|t_l - t_{l+1}| - 2^S) > L_{\max} + T_{\max}, \tag{34}$$

where

$$T_{\max} = \max_m T_m \text{ and } L_{\max} = \max_m L_m. \tag{35}$$

Here L_m is the support of the filter $h_m(t)$. The former requirement means that the pulses have the same sparsity pattern. The minimum-distance requirement (34) ensures that the filtered pulses will not overlap with each other in any channel, and that the blocks of non-zero samples from any two neighbouring pulses are separated by at least one zero sample. This is illustrated in Fig. 13. By locating these zeros, we can separate the two blocks. The requirement (34) is the necessary condition for a successful separation. However, we



(a) Applying the filters in Fig. 11 on a single pulse

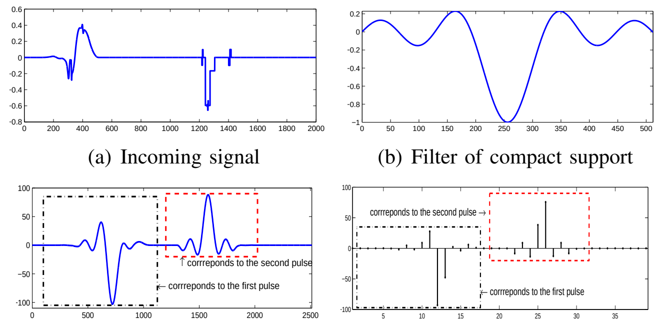
(b) $c_0[n]$, samples from channel 0

(c) $c_1[n]$, samples from channel 1

(d) $c_2[n]$, samples from channel 2

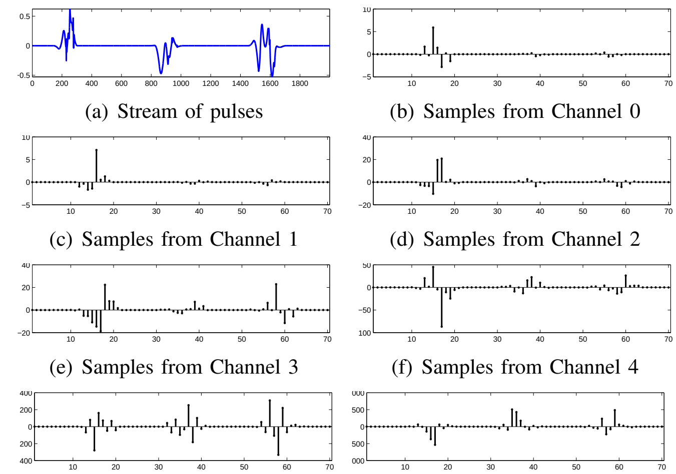
(e) The reproduced Fourier information at $\bar{\Omega}_0$ (red dots), Ω_1 (blue dots) and Ω_2 (blue dots) is exact. The plots on the right is for signal $x(t)$, while the left is for $x'(t)$.

Fig. 12. Applying the proposed sampling setting in sampling single pulses, $x(t)$ and its delayed version $x'(t)$. Here $x(t)$ is sparse on the cubic-spline wavelet basis. Sampling $x(t)$ only requires Channel 1 and 2, while sampling $x'(t)$ requires the extra Channel 0. The reproduced Fourier information is exact, and hence the reconstruction of the pulses is exact.



(a) Incoming signal (b) Filter of compact support (c) Filtered analogue signal (d) Samples of (c)

Fig. 13. Illustration of the sampling of streaming pulses. If the localized pulses are sufficiently far apart and are sampled with a compact-support filter, their samples will be separated. If the samples corresponding to one specific pulse can be isolated from the rest of the samples, they can be used to reproduce the Fourier information of that pulse, which will be later utilized to reconstruct the pulse.



(a) Stream of pulses (b) Samples from Channel 0 (c) Samples from Channel 1 (d) Samples from Channel 2 (e) Samples from Channel 3 (f) Samples from Channel 4 (g) Samples from Channel 5 (h) Samples from Channel 6

Fig. 14. Example of sampling streaming pulses. The reconstruction is carried out locally and sequentially. If the pulses are far enough from each other, the block of non-zero samples from one pulse can be separated, and this block of non-zero samples can be used to reconstruct the corresponding pulse. Once the pulse has been reconstructed, the algorithm moves from there, and look for the next block of non-zero samples.

usually need the distance between neighbouring pulses to be bigger than the necessary condition, so as to make it easier to detect the blocks of non-zero samples.

Under these two assumptions, we use the following steps to reconstruct the pulses from our samples. The whole process is carried out locally and sequentially:

a) separate the blocks of samples from different pulses.
b) reproduce the Fourier information of the pulse under consideration at $\bar{\Omega}$
c) reconstruct $x_l(t-t_l)$ from the Fourier information using Algorithm 2.

As long as the filter-bank system is properly designed, the Fourier information reproduced in Step b) is exact, and hence the reconstruction in Step c) is exact.

We now apply our method to acquire and reconstruct the stream of pulses shown at the beginning of this paper (Fig. 1). Every single pulse is sparse on the cubic-spline wavelet basis, with $J = 5$ and the number of non-zero coefficient at each subband is no bigger than 4. The pulse and all its non-zero coefficients are included in an interval of 512s. With this information, we know from Section III that any single pulse from this model can be recovered exactly from $\bar{\Omega}$ that is constructed according to (9) with $\sharp\Omega_m = 8$. From $\bar{\Omega}$ and (32), we can work out the parameters of MEMS for the $J+2 = 7$ channels in the multi-channel system in Fig. 2, and generate the MEMS $\varphi^m_{j\boldsymbol{\omega}_m}(t)$ and set the filters as

$$h_m(t) = \varphi^m_{j\boldsymbol{\omega}_m}(-t/T_m),$$

just like in the example of Section IV.

The multi-channel system has $J+2$ channels, and the samples from each channels are shown in Fig. 14. The reconstruction is carried out locally and sequentially. The blocks of non-zero samples are first identified and then used to recover single pulses. The reconstruction of the pulse streams is exact.

VI. Noise resilience and model mismatch

Our algorithm can perfectly reconstruct single pulses and streams of pulses in noiseless settings. It is then natural to wonder how it would perform when noise is present or when there is model mismatch. For both cases, we need more Fourier samples. This is to say $\sharp\Omega_m$ should be bigger than $2K_m$, where K_m is the sparse parameter in Theorem 1.

A. Noise in the Fourier information

1) Shift estimation: Retrieving the non-integer part of the shift using (17) with only one pair of frequencies leads to estimates with high variance. To bring down the variance, we can simply average various estimates, i.e.

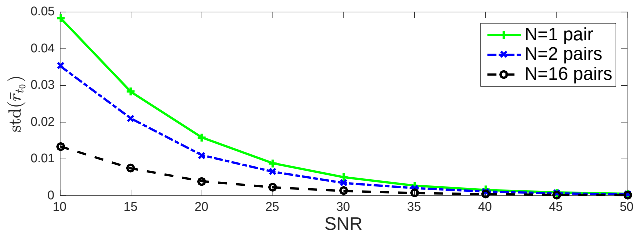
$$\bar{r}_{t_0} = \frac{1}{N}\sum_{n=1}^{N} r^{(n)}_{t_0}$$

where $r^{(n)}_{t_0}$ is obtained by applying (17) to different pairs of frequencies. Fig. 15 (a) shows that using more frequencies effectively brings down the variance for a given shift r_{t_0}. Fig. 15 (b) shows that the variance is mostly affected by noise but not the value of the shift. As a result, the estimates are more reliable when more frequencies are used, and this independently of the value of the shift.

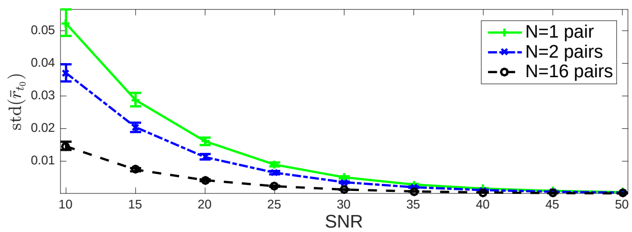
2) Reconstruction of the pulse: Using more frequencies also helps the reconstruction of pulses. It is well known that the original Prony's method is not numerically stable and hence it should be replaced by numerically stabler algorithms. Here we consider two alternatives. One is to replace the Prony's method with a stable variant known as the matrix pencil algorithm [26], [40]. The other is to replace the whole Algorithm 1 with ℓ_1 minimization. We compare these two different alternatives in Fig. 16 on different noise level. The performance is measured by the mean relative reconstruction error (rerr), where

$$\mathrm{rerr} = \frac{||\tilde{x}(t) - x(t)||}{||x(t)||}$$

and $\tilde{x}(t)$ is the estimated signal. We see that the reconstruction errors of both methods are comparable. Therefore to better compare the two approaches, we also analyse their success



(a) $r_{t_0} = 0.512$



(b) Errobar of 100 random r_{t_0}

Fig. 15. The standard deviation of the estimates of non-integer shift in noisy situations, using the signal shown in Fig. 4. The shift is estimated by averaging the results obtained by applying (17) to 1 pairs (green), 2 pairs (blue) and (16) pairs (black) of frequencies. We then obtained (a) the standard deviation $std(\bar{r}_{t_0})$ from 100 independent random realizations of noise added to the Fourier samples, and finally (b) the errorbar by averaging $std(\bar{r}_{t_0})$ from 100 non-integer shifts r_{t_0} randomly picked from $[0, 1)$.

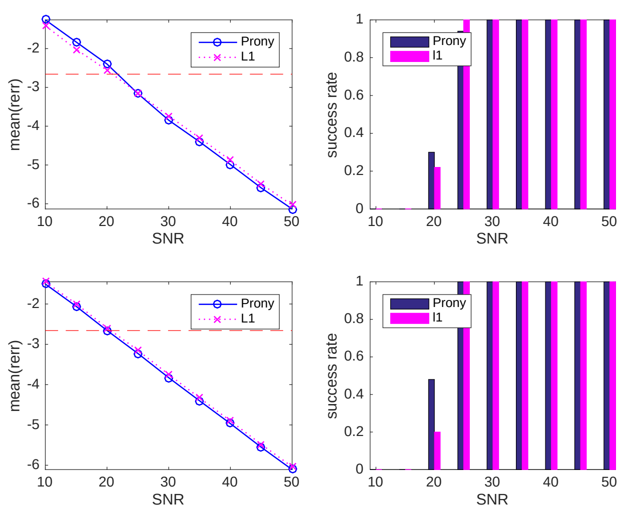
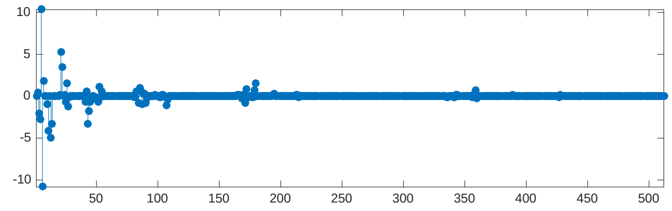
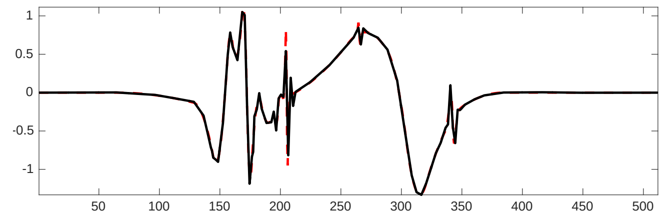


Fig. 16. Comparison of applying Algorithm 2 with Algorithm 1 and ℓ_1 minimization serving in Step 12. In the experiments, Prony's method in Algorithm 1 is replaced by its stabler variant, the matrix pencil. The number of frequencies used are, from top to bottom, 192+32, 320+64. The extra frequencies are used for recovering the shift. The left column shows the mean relative reconstruction error (rerr, shown in log scale) is obtained by 50 random realization of noise added to the Fourier samples of the shifted signal of Fig. 4. In the right column, we count the recovery as a successful reconstruction if rerr is smaller than 0.07. The success rate on different noise level with the same number of frequencies are shown in the right column. The red dashed line mark the threshold of 0.07 in rerr.

rate. We deem a reconstruction is successful when the relative reconstruction error (rerr) is smaller than 0.07. The comparison is shown by a bar plot in the right column of Fig. 16, from which we can observe that using more frequencies can improve the success rate of both algorithms. At low SNR situation (20dB), our algorithm outperforms the ℓ_1 minimization in terms of success rate.



(a) Wavelet coefficients



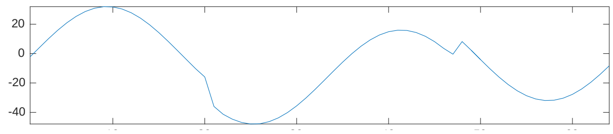
(b) Reconstruction from 336 frequencies using Algorithm 1

Fig. 17. Reconstruction of the wavelet-sparse signal in a slightly different wavelet basis. The signal is exactly sparse in cubic-spline wavelet basis ($\phi_1(t)$), but it is only approximately sparse in the quadratic spline wavelet basis ($\phi_2(t)$). The sparse parameter needs to be re-calibrated on the new basis, which is achieved by looking at the rank of the matrices built when using Prony's method. By doing so we reach a faith reconstruction despite the model mismatch.

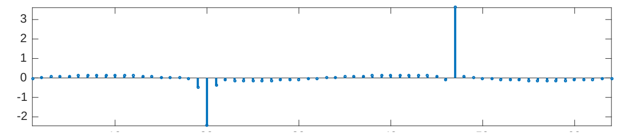
B. Model mismatch

We now consider the model-mismatch problem, namely, the case where we assume the signal to be sparse in a wavelet basis and instead the signal is either sparse in a different basis (example in Fig. 17) or is not sparse at all (example in Fig. 18 and 19). In the first case, the signal is exactly sparse in the biorthogonal cubic-spline wavelet with the sparse parameter $\mathbf{K} = [2, \ldots, 2]$, but its wavelet representation in the biorthogonal quadratic spline wavelet basis is not. In fact, there are 340 non-zero coefficients in scale 1-7 in this case (Fig. 17 (a)). In much the same way as for the noisy scenario, we see that if we allow redundancy we can achieve a very faithful reconstruction (i.e., Fig. 17(b)) even in the case of model mismatch. In this experiment, K_m is estimated by looking at the rank of the matrices built when using Prony's method.

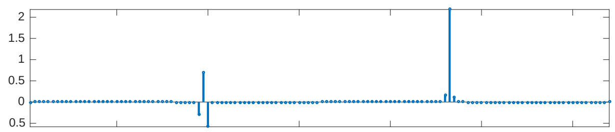
However, a lot of signals can be compressible in a couple of different wavelet bases, and the reconstruction of the signal on different bases for small number of Fourier coefficients are similarly good. We use the Heavisine signal as the example. The signal length is 4096. The signal is zero outside the support range. We show its wavelet coefficients on Daubechies' wavelets (db2 and db3) in Fig. 18 and 19. The Heavisine signal is not exactly sparse in any of the two wavelet bases, but it is compressible on both bases. The sub-captions in Fig. 18 and 19 are the estimation of the upper bound of the sparse parameters K_m, which is obtained by counting the number of coefficients with amplitude greater than 2% of the maximum amplitude in the scale. We truncate the representation at scale 1 to ignore all the non-zero coefficients at scale finer than scale 1. Therefore, $\tilde{K}_0 = 0$. We oversample the Fourier domain approximately by 2.5, and that corresponds to about $2.5 \times (2 \sum \tilde{K}_m)$ frequencies. We then apply Algorithm 2 to reconstruct the signal directly



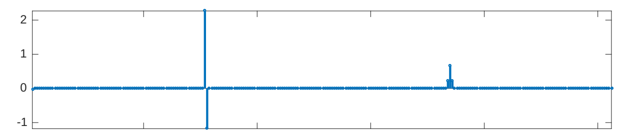
(a) $\tilde{K_6} = 64$



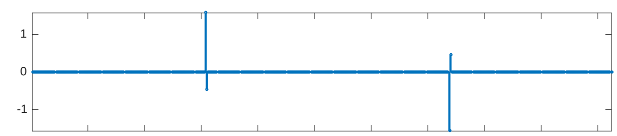
(b) $\tilde{K_5} = 44$



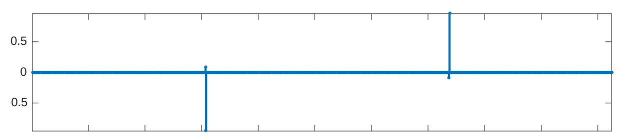
(c) $\tilde{K_4} = 6$



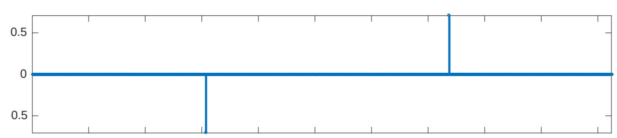
(d) $\tilde{K_3} = 5$



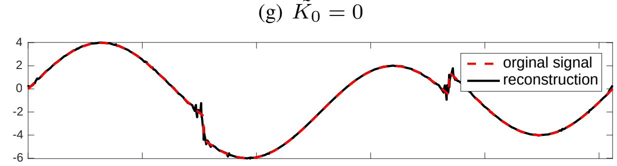
(e) $\tilde{K_2} = 4$



(f) $\tilde{K_1} = 4$

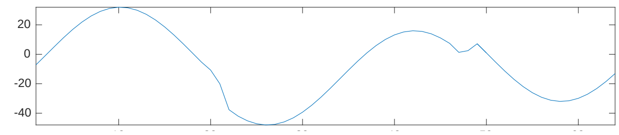


(g) $\tilde{K_0} = 0$

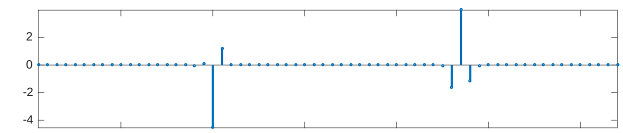


(h) Reconstruction

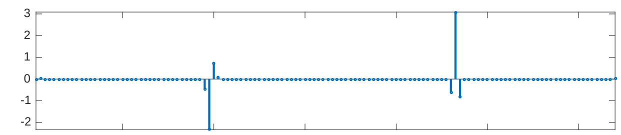
Fig. 18. The wavelet coefficients of the Heavisine signals on Daubechies' 4-tag wavelet (db2) and the reconstruction from its partial Fourier transform of 242+24 frequencies.The length of the signal is 4096. Heavisine signal is only approximately sparse on db2. The upper bound estimation $\tilde{K}_m$ can only approximately model the signal. The sub-caption of figure (a) - (g)are the estimated upper bounds of K_m.



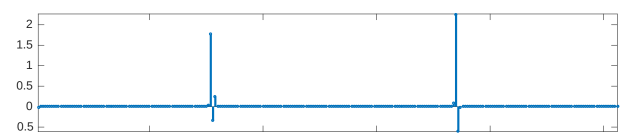
(a) $\tilde{K_6} = 64$



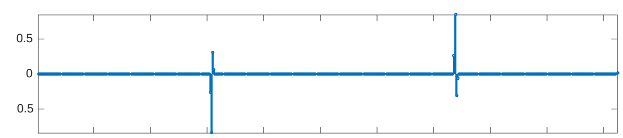
(b) $\tilde{K_5} = 6$



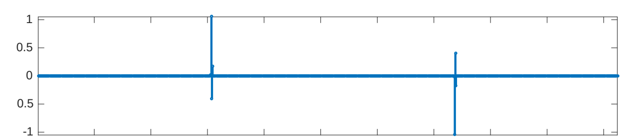
(c) $\tilde{K_4} = 7$



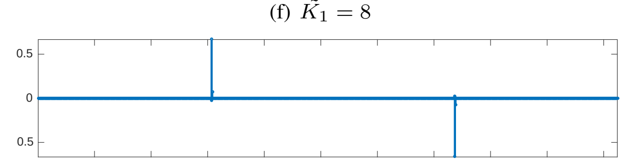
(d) $\tilde{K_3} = 6$



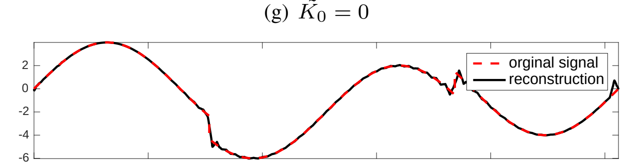
(e) $\tilde{K_2} = 8$



(f) $\tilde{K_1} = 8$



(g) $\tilde{K_0} = 0$



(h) Reconstruction

Fig. 19. The wavelet coefficients of the Heavisine signals on Daubechies' 6-tag wavelet (db3) and the reconstruction from its partial Fourier transform of 278+36 frequencies. The length of the signal is 4096. Heavisine signal is only approximately sparse on db3. The upper bound estimation $\tilde{K}_m$ can only approximately model the signal.

from the partial Fourier information. As shown in the figures, the reconstructions are very faithful to the original signals despite the fact that we only use less that 8% of the Fourier information for both bases.

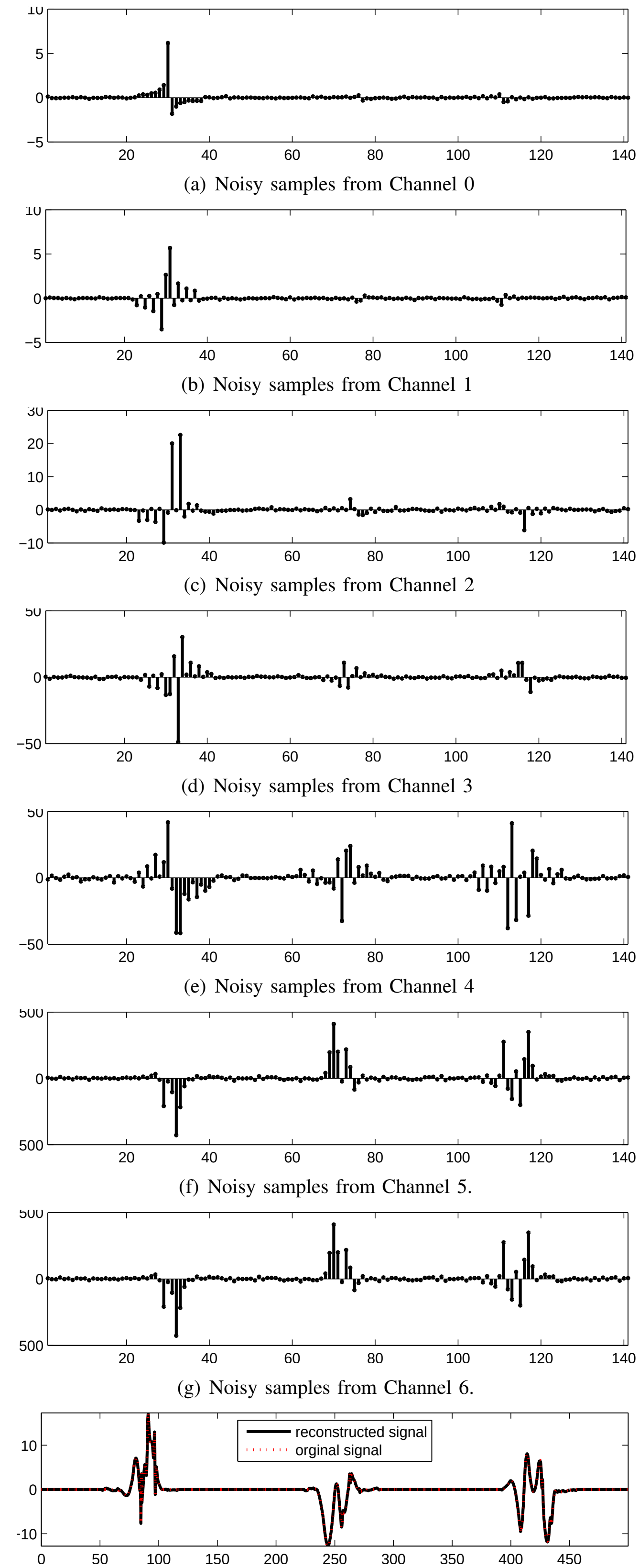
C. Noise in the discrete samples

We now consider a more difficult situations, sampling the stream of pulses in noisy conditions where the input continuous-time signal $s(t)$ is noiseless and the noise is introduced by the acquisition device after sampling. Thus, the samples are

$$\bar{c}[n] = c[n] + e_n$$

where e_n is the additive Gaussian white noise. We use the same stream of pulses of Fig. 14 as the example, which is also shown at the beginning of this paper (Fig. 1) . As said before, we need Fourier information at more frequencies to work robustly with noise. Therefore, $\bar{\Omega}$ is constructed according to (9) with $\sharp\Omega_m = 32$, which means we used 192+32 frequencies in total.

We also do oversampling (oversampling rate is 2) in the time domain to improve the SNR of Fourier information [15]. The samples and the reconstruction are shown in Fig. 20. The reconstruction is done by our algorithm using matrix pencil. In this example, the reconstruction is very accurate with an SNR of 24.87 dB. We can even reduce the SNR to 10 dB at the cost of using more Fourier samples and time-domain discrete samples. The reconstruction is with an SNR of $15.02dB$ (see Fig. 21).



(a) Noisy samples from Channel 0

(b) Noisy samples from Channel 1

(c) Noisy samples from Channel 2

(d) Noisy samples from Channel 3

(e) Noisy samples from Channel 4

(f) Noisy samples from Channel 5.

(g) Noisy samples from Channel 6.

(h) Reconstruction

Fig. 20. Example of sampling streaming pulses in noisy condition ($SNR = 20dB$). White noise is added to the samples. In order to counteract the noise, we acquire Fourier information on more frequencies and do oversampling in the time domain as well. The oversampling rate is 2 and $\bar{\Omega}$ is constructed with $\sharp\Omega_m = 32$. The reconstruction from the samples is very accurate with $SNR = 24.87dB$.

VII. Discussion

The critical role of the sparsity parameter $\mathbf{K}$ is in the design of the sampling pattern in the frequency domain, namely Ω. We prove the one-to-one mapping relationship between the Fourier information in Ω and the wavelet-sparse signal with sparse parameter $\mathbf{K}$.

In the reconstruction process, our algorithm needs to make use of the sparse parameter $\mathbf{K}$, while ℓ_1 minimization does not. In this repect, the ℓ_1 minimization is more universal. However, at the same time, the knowledge of $\mathbf{K}$ in our method allows to reduce substantially the number of the samples needed. This paper uses Prony's method in the noiseless setting to get a tight bound on the cardinality of Ω. Prony's method only requires $2k$ frequencies to recover a k-sparse signal, which is far less than what ℓ_1 requires. As shown in Fig. 5, our algorithm can perfectly reconstruct the signal from merely 28 Fourier samples, while ℓ_1 minimizations requires 112.

When $\mathbf{K}$ is unknown, one needs to use the upper bound of $\mathbf{K}$ which results in redundancy in the sampling pattern. Both our algorithm and ℓ_1 minimization perform very well if there is enough redundancy in the sampling pattern Ω.

In the reconstruction of streams of pulses, we requires the distribution of the pulses to satisfy the minimum separation condition (34). This is because we assumes the shapes of the pulses to be unknown and they can be different from each other. This is very different from the separation conditions discussed in [43] and a few more recent papers in the field of super-resolution [44]–[46] where they need separation even though they assume that the shape of the pulses is identical and known.

VIII. Conclusion

We have proved that the wavelet-sparse signal with sparse parameter $\mathbf{K}$ can be exactly reconstructed from the partial

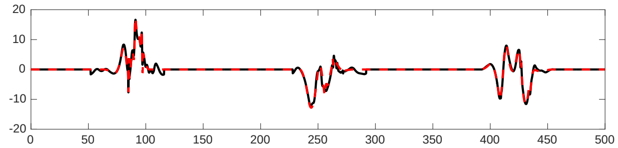


Fig. 21. Example of sampling streaming pulses in noisy condition ($SNR = 10dB$). The relative reconstruction error from the samples is $SNR = 15.02dB$

Fourier information in Ω, where the cardinality of Ω can be as small as

$$2\max\{K_1, \ldots, K_{J+1}\} + 2\sum_{m=1}^{J} K_m,$$

which is, as far as we know, the tightest bound for an exact result. We have also presented a new scheme to collect discrete samples that leads to exact Fourier informations, and apply the new scheme to sample and perfectly reconstruct 1-D stream of pulses with unknown shapes at a significantly reduced rate. Our sampling scheme has the following important features:

1) the acquisition set-up is close to the conventional A-to-D converters;
2) the shape of the filters are not related to the wavelet bases that the pulses are sparse in;
3) the filters are of compact support and can stably reproduce exponentials;
4) the reconstruction of streaming pulses is carried out locally and sequentially.

Our sampling scheme may have an impact in practical engineering applications in the future because of these features. One possible application of our theory is to record electrophysiological data for spike sorting [47], where different neuron tends to fire spikes of different shapes.

The streams of pulses concerned in this paper are mono-dimension signals, and our scheme can now only apply to mono-dimensional signals.

APPENDIX A
PROOF OF THEOREM 3

To prove Theorem 3, we need the following two lemmas.

Lemma A1 (the generalized Strang-Fix conditions [48])**.** *The sufficient and necessary conditions for a function $f(t)$ to be able to reproduce exponentials $e^{j\omega_0 t}$ are that*[4]

$$\hat{f}(\omega_0) \neq 0 \text{ and } \hat{f}(\omega_0 + 2\pi l) = 0 \ (\forall l \neq 0). \tag{36}$$

Lemma A2. *Assume $\boldsymbol{\omega}$ satisfies* (27)*. Besides the Strang-Fix conditions* (36)*, function $\beta_{j\boldsymbol{\omega},2M}(t)$* (26) *also satisfies the following strengthen Strang-Fix conditions:*

$$\hat{\beta}_{j\boldsymbol{\omega},2M}(t)(\omega_n) \neq 0 \text{ and } \hat{\beta}_{j\boldsymbol{\omega},2M}(t)(\omega_n + \frac{k\pi}{M}) \neq 0. \tag{37}$$

The proof of Lemma A2 follows directly from the expression in (26). Now we prove Theorem 3 by verifying that the proposed MEMS satisfies the Strang-Fix conditions. Without loss of generality, we assume $b_i = 2$ and have

$$\begin{aligned} &2\beta_{j\boldsymbol{\omega},2M}(t)\cos\left(\frac{k\pi}{2M}t\right) \\ =&\beta_{j\boldsymbol{\omega},2M}(t)\left(\exp(j\frac{k\pi}{2M}t) + \exp(-j\frac{k\pi}{2M}t)\right) \\ =&\beta_{j\boldsymbol{\omega}+j\frac{k\pi}{2M},2M}(t) + \beta_{j\boldsymbol{\omega}-j\frac{k\pi}{2M},2M}(t) \\ =&\beta_{j\boldsymbol{\omega}+j\frac{k\pi}{2M},2M}(t) + \beta_{-j\boldsymbol{\omega}-j\frac{k\pi}{2M},2M}(t). \end{aligned} \tag{38}$$

We denote $\boldsymbol{\omega}_i = \boldsymbol{\omega} + \frac{2k_i+1}{2M}$ and this leads to

$$\hat{\varphi}^M_{j\boldsymbol{\omega}}(\omega) = \sum_{i=1}^{M}\left(\hat{\beta}_{j\boldsymbol{\omega}_i,2M}(\omega) + \hat{\beta}_{-j\boldsymbol{\omega}_i,2M}(\omega)\right). \tag{39}$$

Because of (27) and (28), we have, for $l \neq i$,

$$\begin{aligned} &\omega_{m,i} = -\omega_{P-m,i} + (k_i+1)\frac{\pi}{M} \\ &\omega_{m,i} + 2k\pi = \omega_{m,l} + (2kM + k_i - k_l)\frac{\pi}{M} = C_1\frac{\pi}{M} \\ &\omega_{m,i} + 2k\pi = -\omega_{P-m,l} + (2kM + k_i + k_l + 1)\frac{\pi}{M} \\ &= C_2\frac{\pi}{M}. \end{aligned} \tag{40}$$

Since $rem(k_i - k_l, 2M) \neq 0$ and $rem(r_n + r_m + K_0, 2M) \neq 0$, neither C_1 nor C_2 can be zero.

According to Lemma A2, we have that, $\forall k \in \mathbb{Z}$,

$$\begin{aligned} &\hat{\beta}_{-j\omega_i,2M}(\omega_{m,i} + 2k\pi) = 0 \\ &\hat{\beta}_{j\omega_l,2M}(\omega_{m,i} + 2k\pi)) = 0,\ l \neq i \\ &\hat{\beta}_{-j\omega_l,2M}(\omega_{m,i} + 2k\pi) = 0,\ l \neq i. \end{aligned}$$

Since

$$\hat{\varphi}^M_{j\boldsymbol{\omega}}(\omega_{m,i} + 2k\pi) = \hat{\beta}_{j\omega_i,2M}(\omega_{n,i} + 2k\pi), \tag{41}$$

using again Lemma A2, we have that

$$\hat{\varphi}^M_{j\boldsymbol{\omega}}(\omega_{m,i}) \neq 0 \text{ and } \hat{\varphi}^M_{j\boldsymbol{\omega}}(\omega_{m,i} + 2k\pi) = 0, \forall k \in \mathbb{Z} \setminus 0.$$

Therefore, $\varphi^M_{j\boldsymbol{\omega}}(t)$ can reproduce the exponential $e^{j\omega_{m,i}t}$. Similarly, we can shown that $\varphi^M_{j\boldsymbol{\omega}}(t)$ can reproduce the exponential $e^{-j\omega_{m,i}t}$ ■

APPENDIX B
MORE ABOUT MEMS AND THE MULTICHANNEL SYSTEM

Our multichannel system uses $J+2$ channels. The intuition behind this setting is illustrated in Fig. 3. The set $\bar{\Omega}$ is the combination of $J+2$ subsets in different ranges of frequencies. It is then natural to consider the structure of $J+2$ channels. This setting also leads to easier design of the right MEMS, as explained in Section IV-C. If the number of channels L is smaller than $J+2$, we will need to find the combination of $k_{i,m}$, M_m and T_m such that

$$\cup_{m=0}^{L-1}\tilde{\Omega}_m T_m \supseteq \bar{\Omega} \tag{42}$$

where $P = 1$, $\tilde{\Omega}_m = \frac{\pi}{4M_m} + \frac{2k_{i,m}+1}{2M_m}\pi$ and $k_{i,m}$ must satisfy Theorem 3. The combinatorial search when using a smaller L is not as straightforward as (32).

[4]This is the special case when the exponent is purely imaginary; otherwise, the Laplace transform of $f(t)$ should replace the Fourier transform of $f(t)$ in (36).

The estimate of the Fourier information of $x(t)$ is also more reliable if the energy of the signal at the selected frequencies is similar in a given channel. Since the signal energy often varies a lot over a wide range of frequencies, we don't want the selected frequencies in one channel to distribute too widely and hence adopt a $J+2$ low-pass/band-pass setting. For example, the signal $x'(t)$ in Fig. 12 has its energy mostly concentrated at low-frequencies. If we construct a stable filter that reproduces Fourier information at hight frequencies $\{\pm\frac{16.5}{16}\pi, \pm\frac{17.5}{16}\pi\}$ as well as at low frequencies $\{\pm\frac{2.5}{16}\pi, \pm\frac{3.5}{16}\pi\}$ to sample $x'(t)$, the reproduced Fourier information at high frequencies is more prone to noise interference. In simulations when using this filter with $20dB$ additive white noise, we note that the SNR of the estimates are about $25dB$ at low frequencies and about $1dB$ at the high frequencies.

We note here that the multichannel system can also be used to stably obtain Fourier information at pseudo random frequencies. This is very useful in other contexts, especially when CS analysis is applied [2]. It is achieved by generating a group of random k_i that satisfy Theorem 3 and using them for constructing MEMS. Every channel can choose the parameters of their MEMS independently. Therefore, the multichannel system with more channels has more freedom in selecting frequencies at which the whole system can reproduce Fourier information.

References

[1] Y. Zhang and P. L. Dragotti, "the Modulated E-spline with Multiple subbands and its application to sampling wavelet-sparse signals," *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, pp. 2371–2375, May 2014.

[2] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *Information Theory, IEEE Transactions on*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[3] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[4] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Signal Processing*, vol. 50, no. 6, pp. 1417–1428, Jun. 2002.

[5] M. Mishali and Y. C. Eldar, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE Journal Select. Topics Sig. Proc.*, vol. 4(2), no. 2, pp. 375–391, Apr. 2010.

[6] J. Tropp, J. Laska, M. Duarte, J. Romberg, and R. Baraniuk, "Beyond Nyquist: Efficient sampling of sparse bandlimited signals," *IEEE Trans. Information Theory*, vol. 56(1), pp. 520–544, Jan. 2010.

[7] M. A. Lexa, M. E. Davies, and J. Thompson, "Reconciling Compressive Sampling Systems for Spectrally-sparse Continuous-time Signals," *IEEE Trans. on Sig. Proc.*, vol. 60(1), pp. 155–171, 2012.

[8] B. Adcock and A. C. Hansen, "Generalized sampling and infinite-dimensional compressed sensing," *Preprint*, 2011.

[9] Y. C. Eldar, "Compressed sensing of analog signals in shift-invariant spaces," *Signal Processing, IEEE Transactions on*, vol. 57, no. 8, pp. 2986–2997, 2009.

[10] L. Baboulaz and P. L. Dragotti, "Exact feature extraction using finite rate of innovation principles with an application to image super-resolution," *Image Processing, IEEE Transactions on*, vol. 18, no. 2, pp. 281–298, 2009.

[11] X. Wei and P. L. Dragotti, "FRESH - FRI-based single-image super-resolution algorithm," *Image Processing, IEEE Transactions on*, 2016, to appear.

[12] 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*, vol. 10, no. 4, p. 46017, 2013.

[13] Y. Hao, P. Marziliano, M. Vetterli, and T. Blu, "Compression of ecg as a signal with finite rate of innovation," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*. IEEE, 2006, pp. 7564–7567.

[14] J. Kusuma, I. Maravić, and M. Vetterli, "Sampling with finite rate of innovation: channel and timing estimation for uwb and gps," in *Communications, 2003. ICC'03. IEEE International Conference on*, vol. 5. IEEE, 2003, pp. 3540–3544.

[15] P. L. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang–Fix," *Signal Processing, IEEE Transactions on*, vol. 55, no. 5, pp. 1741–1757, 2007.

[16] E. Matusiak and Y. C. Eldar, "Sub-Nyquist sampling of short pulses," *Signal Processing, IEEE Transactions on*, vol. 60, no. 3, pp. 1134–1148, 2012.

[17] J. Berent, P. L. Dragotti, and T. Blu, "Sampling Piecewise Sinusoidal Signals with Finite Rate of Innovation Methods," *IEEE Trans. on Signal Processing*, vol. 58(2), pp. 613–625, Feb. 2010.

[18] P. Shukla and P. L. Dragotti, "Sampling Schemes for Multidimensional Signals with Finite Rate of Innovation," *IEEE Trans. on Signal Processing*, vol. 55(7), pp. 3670–3686, Jul. 2007.

[19] I. Maravic and M. Vetterli, "Exact Sampling Results for Some Classes of Parametric Nonbandlimited 2-D Signals," *IEEE Trans. on Signal Processing*, vol. 52(1), pp. 175–189, Jan. 2004.

[20] H. Pan, T. Blu, and P. L. Dragotti, "Sampling Curves with Finite Rate of Innovation," *IEEE Trans. Signal Processing*, vol. 62(2), pp. 458–471, Jan. 2014.

[21] 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*, vol. 60(6), pp. 2851–2859, Jun. 2012.

[22] B. Rajoub *et al.*, "An efficient coding algorithm for the compression of ecg signals using the wavelet transform," *Biomedical Engineering, IEEE Transactions on*, vol. 49, no. 4, pp. 355–362, 2002.

[23] A. B. Wiltschko, G. J. Gage, and J. D. Berke, "Wavelet filtering before spike detection preserves waveform shape and enhances single-unit discrimination," *Journal of neuroscience methods*, vol. 173, no. 1, pp. 34–40, 2008.

[24] J. A. Uriguen, T. Blu, and P. L. Dragotti, "On the Exponential Reproducing Kernels for Sampling Signals with Finite Rate of Innovation," in *Proceedings of the Sampling Theory and Application Conference, Singapore*, Singapare, May 2011.

[25] R. Tur, Y. C. Eldar, and Z. Friedman, "Innovation rate sampling of pulse streams with application to ultrasound imaging," *Signal Processing, IEEE Transactions on*, vol. 59, no. 4, pp. 1827–1842, 2011.

[26] J. A. Uriguen, T. Blu, and P. L. Dragotti, "FRI Sampling With Arbitrary Kernels," *Signal Processing, IEEE Transactions on*, vol. 61, no. 21, pp. 5310–5323, Nov. 2013.

[27] A. Papoulis, "Generalized sampling expansion," *IEEE Transactions on Circuits and Systems*, vol. 24, no. 11, pp. 652–654, Nov. 1977.

[28] M. Unser and J. Zerubia, "A generalized sampling theory without band-limiting constraints," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 45, pp. 959–969, 1998.

[29] Y. Bresler, "Spectrum-blind sampling and compressive sensing for continuous-index signals," in *2008 Information Theory and Applications Workshop - Conference Proceedings, ITA*, 2008, pp. 547–554.

[30] H. Akhondi Asl, P. L. Dragotti, and L. Baboulaz, "Multichannel sampling of signals with finite rate of innovation," *Signal Processing Letters, IEEE*, vol. 17, no. 8, pp. 762–765, 2010.

[31] C. S. Seelamantula and M. Unser, "A generalized sampling method for finite-rate-of-innovation-signal reconstruction," *Signal Processing Letter*, vol. 15, pp. 813–816, 2008.

[32] K. Gedalyahu, R. Tur, and Y. C. Eldar, "Multichannel sampling of pulse streams at the rate of innovation," *Signal Processing, IEEE Transactions on*, vol. 59, no. 4, pp. 1491–1504, 2011.

[33] M. Mishali, Y. C. Eldar, and A. J. Elron, "Xampling: Signal Acquisition and Processing in Union of Subspaces," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4719–4734, Oct. 2011.

[34] G. Puy, P. Vandergheynst, and Y. Wiaux, "On variable density compressive sampling," *Signal Processing Letters, IEEE*, vol. 18, no. 10, pp. 595–598, 2011.

[35] F. Krahmer and R. Ward, "Stable and robust sampling strategies for compressive imaging." *IEEE transactions on image processing*, vol. 23, no. 2, pp. 612–622, Nov. 2013.

[36] B. Adcock, A. C. Hansen, C. Poon, and B. Roman, "Breaking the coherence barrier: A new theory for compressed sensing," *arXiv preprint arXiv:1302.0561*, 2013.

[37] Y. Zhang and P. L. Dragotti, "On the Reconstruction of Wavelet-Sparse Signals From Partial Fourier Information," *Signal Processing Letters, IEEE*, vol. 22, no. 9, pp. 1234–1238, Sep. 2015.

[38] S. Mallat, *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. Academic Press, Dec. 2008.
[39] I. Daubechies, *Ten lectures on wavelets*. Philadelphia: Society for Industrial Mathematics and applied mathematics, 1992.
[40] G. H. Golub, P. Milanfar, and J. Varah, "A stable numerical method for inverting shape from moments," *SIAM Journal on Scientific Computing*, vol. 21, no. 4, pp. 1222–1243, 1999.
[41] P. Stoica and R. Moses, *Introduction to Spectral Analysis*. Englewood Cliffs, NJ, Prentice-Hall, 2000.
[42] M. Unser and T. Blu, "Cardinal exponential splines: part I - theory and filtering algorithms," *IEEE Transactions on Signal Processing*, vol. 53, no. 4, pp. 1425–1438, Apr. 2005.
[43] T. Bendory, "Robust recovery of positive stream of pulses," *arXiv preprint arXiv:1503.08782*, 2015.
[44] E. J. Candès and C. Fernandez-Granda, "Towards a Mathematical Theory of Super-resolution," *Communications on Pure and Applied Mathematics*, 2013.
[45] G. Schiebinger, E. Robeva, and B. Recht, "Superresolution without separation," *arXiv preprint arXiv:1506.03144*, 2015.
[46] V. Duval and G. Peyré, "Exact support recovery for sparse spikes deconvolution," *Foundations of Computational Mathematics*, vol. 15, no. 5, pp. 1315–1355, 2015.
[47] R. Q. Quiroga, Z. Nadasdy, and Y. Ben-Shaul, "Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering," *Neural computation*, vol. 16, no. 8, pp. 1661–1687, 2004.
[48] I. Khalidov, T. Blu, and M. Unser, "Generalized L-spline wavelet bases," in *Optics & Photonics 2005*. International Society for Optics and Photonics, 2005, pp. 59 140F–59 140F.

Zhang, Yingsong is currently a data scientist. She received the B.Sc. degree from Mathematics Department, East China Normal University in 2003, the M.Eng. degree from Automation Department, Tsinghua University, China, in 2007, received her Ph.D. degree (from Cambridge University, UK, in 2012. She was then a research associate at the Electrical and Electronic Engineering Department at Imperial College London, till 2015. Her research interests include sampling theory, wavelet theory and its applications, image image enhancement techniques (denoising,deconvolution and super-resolution).

Pier Luigi Dragotti is Professor of Signal Processing with the Electrical and Electronic Engineering Department at Imperial College London. He received the Laurea degree (summa cum laude) in electrical and electronic engineering from the University of Naples Federico II, Naples, Italy, in 1997; the Masters degree in communications systems from the Swiss Federal Institute of Technology of Lausanne (EPFL), Lausanne, Switzerland, in 1998, and the Ph.D. degree from EPFL in April 2002. He has held several visiting positions at different universities and research centers. He was a Visiting Student with Stanford University, Stanford, CA, USA, in 1996, a Researcher with the Mathematics of Communications Department, Bell Labs, Lucent Technologies, Murray Hill, NJ, USA, in 2000 and a Visiting Scientist with the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2011. He was Technical Co-Chair for the European Signal Processing Conference in 2012, an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING from 2006 to 2009 and an Elected Member of the IEEE Image, Video and Multidimensional Signal Processing Technical Committee (2008–2013). He is currently a member of the IEEE Signal Processing Theory and Methods technical committee and a recipient of the ERC Starting Investigator Award for the project RecoSamp. His research interests include sampling theory, wavelet theory and its applications, image and video compression, image-based rendering, and image super-resolution