FITTING INSTEAD OF ANNIHILATION: IMPROVED RECOVERY OF NOISY FRI SIGNALS

Christopher Gilliam and Thierry Blu

Department of Electronic Engineering, The Chinese University of Hong Kong
email: {cgilliam, tblu}@ee.cuhk.edu.hk

ABSTRACT

Recently, classical sampling theory has been broadened to include a class of non-bandlimited signals that possess finite rate of innovation (FRI). In this paper we consider the reconstruction of a periodic stream of Diracs from noisy samples. We demonstrate that its noiseless FRI samples can be represented as a ratio of two polynomials. Using this structure as a model, we propose recovering the FRI signal using a model fitting approach rather than an annihilation method. We present an algorithm that fits this model to the noisy samples and demonstrate that it has low computation cost and is more reliable than two state-of-the-art methods.

Index Terms— Finite rate of innovation, sampling theory, noise, recovery of Dirac pulses

1. INTRODUCTION

Signal acquisition and reconstruction relies on the ability to convert a signal between the continuous and discrete-time domains. Unsurprisingly, perfect reconstruction when converting between these domains is highly prized. In 2002, Vetterli et al [1] demonstrated perfect reconstruction for a class of non-bandlimited signals that possess finite rate of innovation (FRI). In other words, they have a finite number of degrees of freedom per unit of time. Specifically, the authors showed how to reconstruct periodic streams of Diracs and piecewise polynomials using the sinc and Gaussian kernels.

Since then the sampling of FRI signals has received wide attention and been extended to broader scenarios [2]. For example, the use of polynomial and exponential reproducing kernels were proposed in [3, 4], and reconstruction of piecewise sinusoidal signals examined in [5]. Recently, recovery from non-uniform samples was examined in [6] and the reconstruction of a long sequence Diracs, 1000 in total, presented in [7]. FRI theory has also been generalised to a spherical coordinate scheme in [8] and higher dimensional signals, such as multi-dimensional Diracs in [9] and curves in [10]. As a result, FRI has found application in the compression of ECG signals [11], the detection of spikes in neurophysiological data [12, 13] and in ultrasound [14]. However, similar to
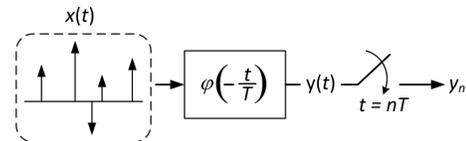
Fig. 1. The FRI acquisition system. The continuous-time input signal $x(t)$, in this case a sequence of K Diracs, is filtered by a sampling kernel $\varphi(-t/T)$ and sampled at a period T.

most applications, perfect FRI reconstruction is only achieved in noiseless conditions. Therefore, a particular topic of interest, examined in [2, 15, 16, 17], is the recovery of FRI signals in noisy conditions.

In this paper, we consider a classic FRI sampling problem; the reconstruction of a periodic stream of Diracs from noisy samples. We demonstrate that, in noiseless conditions, the samples of this periodic FRI signal can be expressed as a ratio of two polynomials. Therefore, the samples of the FRI signal can be represented using only the coefficients of these polynomials. In view of this, we propose a novel method to recover an FRI signal in the presence of noise. The method is based on fitting a model comprising a ratio of two polynomials to the noisy FRI samples. The central concept is that, by minimising the fit between the two, we recover the best estimate of the FRI samples. Accordingly, we present an iterative algorithm that estimates the coefficients of the polynomials by fitting the model to the noisy data. We demonstrate that this algorithm requires less computation time than two state-of-the-art methods - matrix pencil [15] and Cadzow iterative denoising [2] - whilst still achieving the same accuracy.

The paper is organised as follows. In Section 2, we review FRI sampling theory relating to a periodic stream of Diracs in both noiseless and noisy conditions. For a complete review of the state-of-the-art see [18]. In Section 3, we present the framework of our reconstruction algorithm; namely that the samples of a FRI signal can be represented as a ratio of polynomials. Using this framework, we present our reconstruction algorithm in Section 4 and evaluate it using simulations in Section 5. We then conclude in the final section.

2. SAMPLING FRI SIGNALS

The generic FRI sampling problem presented in [1] involves the recovery of a continuous-time FRI signal, $x(t)$, from a set

of N samples, $\{y_n\}_{n=0}^{N-1}$. These samples are obtained from an analogue-to-digital acquisition system; the continuous-time signal $x(t)$ is filtered using a kernel, with impulse response $\varphi(-t/T)$, and then uniformly sampled in time. Assuming a sampling period T, the samples we obtain are

$$y_n = \int_{-\infty}^{\infty} x(t)\varphi\left(\frac{t}{T} - n\right)\mathrm{d}t = \left\langle x(t), \varphi\left(\frac{t}{T} - n\right)\right\rangle, \quad (1)$$

where $n = 0, 1, \ldots, N-1$. Figure 1 illustrates this acquisition system using a stream of Diracs (a standard FRI signal).

In this paper, we consider the specific case presented in [1, 2]: the signal $x(t)$ is a τ-periodic stream of K Diracs and the sampling kerenel is a sinc window with bandwidth $B = 1/T$. This type of FRI signal has a rate of innovation $2K/\tau$ and is defined as

$$x(t) = \sum_{k=1}^{K}\sum_{l\in\mathbb{Z}} x_k\delta\left(t - t_k - l\tau\right), \quad (2)$$

where $\{x_k, t_k\}_{k=1}^{K}$ are respectively the amplitudes and locations of the K Diracs. Note that the locations are restricted such that $t_k \in [0, \tau[$. Therefore, using the definition of the Dirichlet kernel (or τ-periodic sinc function), the samples of (2) we obtain using the sinc kernel are

$$y_n = \sum_{k=1}^{K} x_k \frac{\sin(\pi B(nT - t_k))}{B\tau\sin(\pi(nT - t_k)/\tau)}. \quad (3)$$

Note that $T = \tau/N$ in this framework thus $N = B\tau$ as $B = 1/T$. Also, without loss of generalisation, we shall assume $B\tau$ is an odd integer.

Now, as demonstrated in [2], the signal $x(t)$ can be perfectly reconstructed from a set of N samples provided $N \geq 2K + 1$. The authors' reconstruction scheme involves determining the locations $\{t_k\}_{k=1}^{K}$ via the annihilation filter method (also known as Prony's methods) and then determining the amplitudes $\{x_k\}_{k=1}^{K}$ via least mean squares.

In brief, the annihilation filter method involves determining a filter A whose coefficients $\mathbf{a} = [a_0, a_1, \ldots, a_K]^{\mathrm{T}}$ satisfy $a * \hat{y}_n = 0$, where $*$ represents convolution and $\hat{y}_n$ is the discrete Fourier transform (DFT) of the samples. Using matrix notation, this convolution can be written as

$$\underbrace{\begin{bmatrix} \hat{y}_{-M+L} & \hat{y}_{-M+L-1} & \cdots & \hat{y}_{-M} \\ \hat{y}_{-M+L+1} & \hat{y}_{-M+L} & \cdots & \hat{y}_{-M+1} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_M & \hat{y}_{M-1} & \cdots & \hat{y}_{M-L} \end{bmatrix}}_{\mathbf{Y}_{L=K}=\mathbf{Y}_K} \mathbf{a} = \mathbf{0} \quad (4)$$

where $M = \lfloor N/2 \rfloor$, $L = K$ and $\mathbf{Y}_K$ is a Toeplitz matrix of size $(N - K) \times (K + 1)$. The locations $\{t_k\}_{k=1}^{K}$ are then determined from the roots of filter A. Since the samples $\hat{y}_n$ are annihilated in (4) the filter A is called the annihilation filter. For further details of the annihilation method see [2].

2.1. Model Mismatch

Unfortunately, in practice, no acquisition system is perfect thus the samples we obtain are corrupted either by noise or more generally mode mismatch. We denote these noise corrupted samples as $\tilde{y}_n$. The presence of noise means that the annihilation equation in (4) is no longer valid. In other words, if $\tilde{\mathbf{Y}}_K$ is the Toeplitz matrix formed from the noisy samples then $\tilde{\mathbf{Y}}_K\mathbf{a} \neq \mathbf{0}$. One simple solution is to assume $\tilde{\mathbf{Y}}_K\mathbf{a} \approx \mathbf{0}$ and estimate $\mathbf{a}$ using the total least squares (TLS) method. However, when the noise level increases this approach becomes unreliable [18].

More sophisticated approaches involve extending the annihilation equation in (4) such that $L = \lfloor N/2 \rfloor$, hence we obtain a new noisy matrix $\tilde{\mathbf{Y}}_L$ that is $(L + 1)$-square in size. In [2, 19], the authors' presented iterative algorithms that exploit the K-rank Toeplitz structure of $\mathbf{Y}_L$ in noiseless conditions. Blu et al [2] used Cadzow's iterative denoising (CID) algorithm [20] whereas [19] used structured low rank approximation [21]. Once the approximate matrix is obtained it is reshaped such that $L = K$ and the locations determined by TLS. A non-iterative approach using $\tilde{\mathbf{T}}_L$ was proposed in [15] and subsequently used in [16, 4]. This approach involves applying the matrix pencil method [22] to $\tilde{\mathbf{T}}_L$ to obtain an estimate of t_k directly. The performance achieved using the matrix pencil method is similar to that achieved using the CID algorithm [4]. Finally, a stochastic algorithm was proposed in [23] to reconstruct FRI signals in the presence of noise.

A disadvantage of both the iterative and non-iterative algorithms is that they require singular valued decomposition (SVD) to be performed on $\tilde{\mathbf{Y}}_L$. Thus, their computational cost rapidly increases as the number of samples N increases (remember that $L = \lfloor N/2 \rfloor$). Note that the iterative algorithms are the worst offenders as this SVD is performed at each iteration. In this paper, we propose a new approach to dealing with noise that does not require us to compute large SVD.

2.2. Assessment of FRI Signal Recovery

Often, the accuracy of an FRI recover algorithms is based on how accurate the parameters $\{x_k, t_k\}_{k=1}^{K}$ have been estimated in comparison to the originals. This may however be unreliable as it depends on the order of the reconstructed locations. Instead, we propose assessing recovery based on the mean squared error (MSE) between the reconstructed FRI samples, y_n', and $\tilde{y}_n$, which we term $\mathrm{MSE_R}$. As a result, we can then construct a criteria to decide if the FRI recovery has been successful. This criteria is

$$\mathrm{MSE_R} < \mathrm{MSE_{IN}}, \quad (5)$$

where $\mathrm{MSE_{IN}}$ is the input MSE between the $\tilde{y}_n$ and y_n. Note that it is reasonable to assume prior knowledge of $\mathrm{MSE_{IN}}$ via noise statistics so (5) can be used directly when recovering FRI signals.

3. FRI SAMPLES AS A RATIO OF POLYNOMIALS

A fundamental concept in this paper is that, in the noiseless scenario, the samples of a FRI signal can be expressed as a ratio of two polynomials: a numerator polynomial of order $K-1$ and a denominator of order K. To demonstrate this relationship, consider again the samples of the τ-periodic stream of K Diracs defined in (3). Using the identities $N = B\tau$ and $T = \tau/N$, these samples are

$$y_n = \sum_{k=1}^{K} x_k \frac{\sin(\pi\,(n - \bar{t}_k))}{N \sin(\pi\,(n - \bar{t}_k)\,/N)}, \qquad (6)$$

where $\bar{t}_k = t_k/T$. Now, if we use Euler's formula, the samples are expressed as a ratio of complex exponentials. In particular, the denominator is dependent on a $j2\pi n/N$ term. Therefore, we can manipulate the complex exponentials such that we obtain:

$$y_n = e^{j2\pi nM/N} \frac{P_{K-1}(e^{j2\pi n/N})}{Q_K(e^{j2\pi n/N})}, \qquad (7)$$

where $P_{K-1}(e^{j2\pi n/N})$ is a polynomial of order $K - 1$ and $Q_K(e^{j2\pi n/N})$ is a polynomial of order K. In fact, the polynomial Q_K is the annihilation filter mentioned in the previous section, hence we shall refer to it now as A.

The important aspects of this relationship are as follows. First, the samples of the FRI signal are defined by the coefficients of the two polynomials. In other words, if we know these coefficients we can construct the FRI samples and in turn recover the FRI signal. Second, as the order of the polynomials is related to the number of Diracs, this representation of the FRI samples is independent of the number of samples. Finally, this is a sparse representation of the samples in the time domain.

Now, in noisy conditions, the expression in (7) does not hold. However, we can use this ratio structure as a model and performing model-fitting on the noisy samples. The principle is that by fitting the model to the noisy samples we obtain a sparse representation of the samples that fits the FRI framework. Therefore, if we minimise this fitting procedure, we obtain the best sparse representation of the samples and a FRI signal. As a result, we present a new approach to the recovery of FRI signals in the presence of noise; Rather than annihilating the noisy samples, we propose fitting a model comprising a ratio of two polynomials to them. This fitting procedure is summarised as

$$\min_{A,P} \sum_{n=0}^{N-1} \left| \tilde{v}_n - \frac{P_{K-1}(e^{j\omega_n})}{A_K(e^{j\omega_n})} \right|^2, \qquad (8)$$

where $\tilde{v}_n = y_n e^{-j2\pi nM/N}$ and $\omega_n = 2\pi n/N$.

Notice that this nicely coincides with the criteria defined in (5). By minimising (8) we should obtain a sparse representation of the noisy samples that minimises the MSE between the reconstructed samples and the noise samples.

4. MODEL FITTING ALGORITHM

4.1. Algorithm

Unfortunately, the minimisation as stated in (8) is non-linear in nature. Therefore, its direct computation would require non-linear based methods, such as the Gauss-Newton method or the Levenberg-Marquard algorithm. Instead, we propose to estimate the solution in a linear manner using the following iterative minimisation

$$\min_{A^i, P} \sum_{n=0}^{N-1} \left| \frac{A_K^i(e^{j\omega_n})\tilde{v}_n - P_{K-1}(e^{j\omega_n})}{A_K^{i-1}(e^{j\omega_n})} \right|^2, \qquad (9)$$

where i represents the iteration number. In the following section we shall detail our implementation of (9). Note, however, that this type of algorithm is similar to the Steiglitz-McBride algorithm [24] for system identification and Sanathanan-Koerner algorithm [25] for transfer function synthesis. Accordingly, alternative minimisation procedures are possible, for example see [26].

4.2. Implementation

We start by rewriting (9) in terms of the product of Fourier matrices. First, by defining $\mathbf{W}_{N,M}$ as a $N \times M$ inverse DFT matrix, we have the following

$$\left\{ P_K(e^{j\omega_n}) \right\}_{n=0}^{N-1} = \mathbf{W}_{N,K}\mathbf{p}$$
$$\left\{ A_{K+1}(e^{j\omega_n}) \right\}_{n=0}^{N-1} = \mathbf{W}_{N,(K+1)}\mathbf{a},$$

where $\mathbf{p}$ is the K coefficients of P. Now, we define $\tilde{\mathbf{v}} = [\tilde{v}_0, \tilde{v}_1, \ldots, \tilde{v}_{N-1}]^{\mathrm{T}}$ as the vector relating to $\tilde{v}_n$ and $\tilde{\mathbf{V}}$ as the diagonal matrix constructed from $\tilde{\mathbf{v}}$. Using this new notation, (9) becomes

$$\min_{\mathbf{a}_i,\mathbf{p}} \|\mathbf{A}_{i-1}\mathbf{a}_i - \mathbf{B}_{i-1}\mathbf{p}\|^2, \qquad (10)$$

where the matrices $\mathbf{A}_{i-1}$ and $\mathbf{B}_{i-1}$ are defined as

$$\mathbf{A}_{i-1} = \tilde{\mathbf{V}}\mathbf{R}_{i-1}\mathbf{W}_{N,(K+1)}, \quad \text{and} \quad \mathbf{B}_{i-1} = \mathbf{R}_{i-1}\mathbf{W}_{N,K},$$

where $\mathbf{R}_{i-1} = \left[\mathrm{diag}\left\{ \mathbf{W}_{N,(K+1)}\mathbf{a}_{i-1} \right\} \right]^{-1}$.

Now, similar to [2, 15], we shall constraint the minimisation in (10) such that the Euclidean norm of $\mathbf{a}_i$ is equal to one (i.e. $\|\mathbf{a}_i\|^2 = 1$). Using this constraint, we decouple the minimisation in (10) to obtain the following in just $\mathbf{a}_i$:

$$\min_{\mathbf{a}_i} \left\| \left(\mathbf{I} - \mathbf{B}_{i-1}\mathbf{B}_{i-1}^{\dagger} \right) \mathbf{A}_{i-1}\mathbf{a}_i \right\|^2 \text{ s.t. } \|\mathbf{a}_i\|^2 = 1, \quad (11)$$

where $(\mathbf{B}_{i-1})^{\dagger} = \left[\mathbf{B}_{i-1}^{\mathrm{H}}\mathbf{B}_{i-1} \right]^{-1} \mathbf{B}_{i-1}^{\mathrm{H}}$. We estimate the next iteration $\mathbf{a}_i$ by solving (11) using total least squares, thus we require SVD. However, unlike [2, 15], the SVD is only performed on a square matrix of size $(K + 1)$. Therefore, it scales only with the number of Diracs not the number of samples hence this method can be applied to large sets of samples.

Final, once a suitable $\mathbf{a}_i$ is obtained, we determine $\mathbf{p}$ by solving

$$\min_{\mathbf{p}} \left\| \tilde{\mathbf{v}} - \mathbf{A}_i \mathbf{W}_{N,K} \mathbf{p} \right\|^2 . \qquad (12)$$

Two important aspect of this algorithm are of course deciding $\mathbf{a}_0$ and deciding when to stop. For the first, we perform TLS on the noisy matrix $\tilde{\mathbf{Y}}_K$ and determine $\mathbf{a}_0$. For the second aspect, we follow our criteria set in (5), thus we iterate until the (5) is satisfied. Alternatively, we could iterate for a fixed number of iterations. The complete procedure is defined in Algorithm 1.

Algorithm 1 Model-fitting method for recovery of a FRI signal in the presence of noise.

---

1. Initialisation: Calculate $\mathbf{a}_0$ as the Total Least Squares solution of $\tilde{\mathbf{Y}}_K \mathbf{a} \approx \mathbf{0}$ subject to $\|\mathbf{a}\|^2 = 1$.
2. Set $i = i + 1$.
3. Solve the decoupled minimisation in (10) to obtain $\mathbf{a}_i$. Solution involves SVD of $(K+1)$-square matrix.
4. Using $\mathbf{a}_i$, solve the minimisation defined in (12).
5. Reconstruct samples y_n' and check against the criteria (5). If it is not satisfied, and i less than a threshold, return to Step 2.
6. If required, determine $\{x_k, t_k\}_{k=1}^{K}$ from y_n'.

---

5. SIMULATIONS

We now compare the performance of our algorithm against the CID algorithm presented in [2] and the matrix pencil method presented in [15].

5.1. Comparison of Computation Time

To perform this comparison, we use a FRI signal comprising $K = 60$ Diracs and calculate the computation time required as a function of the sample number, N. In more detail, we start with 120 samples of the FRI signal, i.e. the signal is critically sampled, and then gradually increase until we reach 2401 samples, which equates to oversampling by a factor of 20. Note that the noise level is fixed such that the SNR is $5\,\mathrm{dB}$. The resulting computation times of each algorithm are shown in Figure 2.

The figure illustrates two points: first, for all values of N, our algorithm requires the least computation time; and second, the difference in computation time increases with N. These results are not surprising as both the other two algorithms require the SVD of a L-square matrix, where $L = \lfloor N/2 \rfloor$, which has a high computational cost. In contrast, we only require an SVD of a $(K+1)$-square matrix. Therefore, our algorithm is more suited to recovering FRI signals from a large number of samples.

5.2. Success Percentage of the Algorithm

Having examined the computation time in the previous section, we now analyse the average performance of the algo-
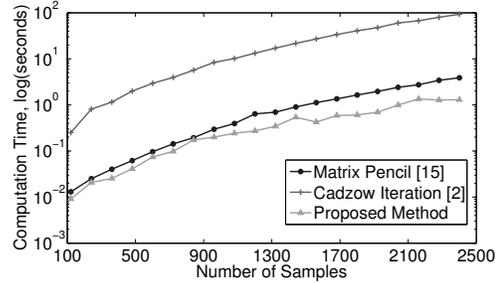

Fig. 2. Graphs showing the computations time, in $\log(\text{seconds})$, for three FRI recovery algorithms as the number of samples increases. The algorithms are: the matrix pencil [15], Cadzow iterative algorithm [2] and our proposed method.
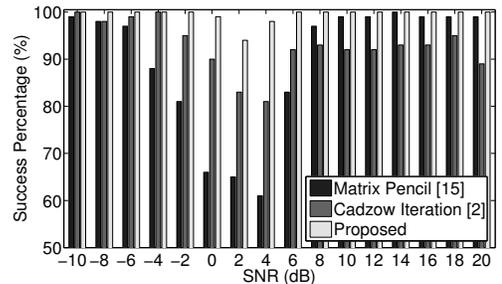

Fig. 3. Graph comparing the success percentage of three FRI recovery algorithms for a periodic stream of 6 Diracs in the presence of noise. The algorithms are: the matrix pencil [15], Cadzow iterative algorithm [2] and our proposed method. A success is decided using the criteria in (5).

rithms. In particular, we are interested in how often the algorithms achieve the criteria defined in (5). We define this characteristic as the success percentage. To compare the success percentage of the algorithms, we use a FRI signal comprising $K = 6$ Diracs and sampled using 51 samples, an oversampling factor of 4. Using 100 realisations of a noise level, we calculate the success percentage for the algorithms as the noise level increases from SNR $= 20\,\mathrm{dB}$ to $-10\,\mathrm{dB}$. The results are shown in Figure 3. The bar graph in Figure 3 demonstrates that our algorithm is the most consistent of the three and the matrix pencil method is the most inconsistent, in particular when the noise level is between $4\,\mathrm{dB}$ and $-2\,\mathrm{dB}$.

6. CONCLUSIONS

In this paper, we demonstrated that the samples of a periodic stream of finite Diracs can be expressed as a ratio of two polynomials. From this, we proposed a novel method to recover the FRI signal in the presence of noise. The method is based on fitting a model comprising a ratio of two polynomials to the noisy FRI samples. Accordingly, we presented an iterative algorithm that estimates the coefficients of the polynomials by minimising Euclidean distance between the model and the noisy data. Finally, we showed that our algorithm has lower computational cost and is more reliable than two state-of-the-art methods.

7. REFERENCES

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

[2] T. Blu, P. L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse sampling of signal innovations," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 31 –40, march 2008.

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

[4] J.A. Urigüen, T. Blu, and P.L. Dragotti, "FRI sampling with arbitrary kernels," *IEEE Trans. Signal Process.*, vol. 61, no. 21, pp. 5310–5323, 2013.

[5] J. Berent, P. L. Dragotti, and T. Blu, "Sampling piecewise sinusoidal signals with finite rate of innovation methods," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 613 –625, feb. 2010.

[6] X. Wei, T. Blu, and P. L. Dragotti, "Finite rate of innovation with non-uniform samples," in *Proc. Int. Conf. Signal Process., Commun. and Comput., Hong Kong*, August 2012.

[7] J. Onativia, J. Uriguen, and P. L. Dragotti, "Sequential local FRI sampling of infinite streams of Diracs," in *Proc. Int. Acoust., Speech and Signal Process. (ICASSP), Vancouver, Canada*, 2013, pp. 5440–5444.

[8] S. Deslauriers-Gauthier and P. Marziliano, "Sampling signals with a finite rate of innovation on the sphere," *IEEE Trans. Signal Process.*, vol. 61, no. 18, pp. 4552–4561, 2013.

[9] P. Shukla and P. L. Dragotti, "Sampling schemes for multidimensional signals with finite rate of innovation," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3670–3686, 2007.

[10] H. Pan, T. Blu, and P. L. Dragotti, "Sampling curves with finite rate of innovation," in *Proc Int. Workshop Samp. Theory Appl. (SampTA2011), Singapore*, 2011.

[11] G. Baechler, N. Freris, R.F. Quick, and R.E. Crochiere, "Finite rate of innovation based modeling and compression of ECG signals," in *Proc. Int. Acoust., Speech and Signal Process. (ICASSP), Vancouver, Canada*, 2013, pp. 1252–1256.

[12] J. Caballero, J. A. Urigüen, S. R. Schultz, and P. L. Dragotti, "Spike sorting at sub-nyquist rates," in *Proc. Int. Conf. Acoustic, Speech and Signal Process. (ICASSP), Kyoto, Japan*, Kyoto, Japan, 2012.

[13] J. Onativia, S. Schultz, and P. L. Dragotti, "A finite rate of innovation algorithm for fast and accurate spike detection from two-photon calcium imaging," *J. Neural Eng.*, vol. 10, no. 4, pp. 046017, 2013.

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

[15] I. Maravic and M. Vetterli, "Sampling and reconstruction of signals with finite rate of innovation in the presence of noise," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2788–2805, aug. 2005.

[16] A. Erdozain and P.M. Crespo, "Reconstruction of aperiodic FRI signals and estimation of the rate of innovation based on the state space method," *Signal Process.*, vol. 91, no. 8, pp. 1709 – 1718, 2011.

[17] L. Condat and A Hirabayashi, "Cadzow denoising upgraded: A new projection method for the recovery of Dirac pulses from noisy linear measurements," JSPS long-term fellowship, Mar. 2013.

[18] J.A. Urigüen, P. L. Dragotti, Y. C. Eldar, and Z. Ben-Haim, "Sampling at the rate of innovation: Theory and applications," in *Compressed Sensing: Theory and Applications*, Y. C. Eldar and G. Kutyniok, Eds., pp. 148–209. Cambridge University Press, 2012.

[19] L. Condat, A. Hirabayashi, and Y. Hironaga, "Recovery of nonuniform Dirac pulses from noisy linear measurements," in *Proc. Int. Conf. Acoust., Speech and Signal Processing (ICASSP), Vancouver, Canada*, 2013, pp. 6014–6018.

[20] J. Cadzow, "Signal enhancement - A composite property mapping algorithm," *IEEE Trans. Acoust., Speech and Signal Process.*, vol. 36, no. 1, pp. 49–62, 1988.

[21] Ivan Markovsky, "Structured low-rank approximation and its applications," *Automatica*, vol. 44, no. 4, pp. 891 – 909, 2008.

[22] Y. Hua and T.K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoust., Speech and Signal Process.*, vol. 38, no. 5, pp. 814–824, 1990.

[23] V.Y.F. Tan and V.K. Goyal, "Estimating signals with finite rate of innovation from noisy samples: A stochastic algorithm," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5135–5146, 2008.

[24] K. Steiglitz and L. E. McBride, "A technique for the identification of linear systems," *IEEE Trans. Automatic Control*, vol. 10, no. 4, pp. 461–464, 1965.

[25] C. K. Sanathanan and J. Koerner, "Transfer function synthesis as a ratio of two complex polynomials," *Automatic Control, IEEE Transactions on*, vol. 8, no. 1, pp. 56–58, 1963.

[26] A.K. Shaw, "Optimal design of digital IIR filters by model-fitting frequency response data," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 42, no. 11, pp. 702–710, 1995.