Multichannel Sampling of Signals with Finite Rate of Innovation

Hojjat Akhondi Asl*, Pier Luigi Dragotti and Loic Baboulaz

EDICS: DSP-TFSR, DSP-SAMP.

Abstract

In this paper we present a possible extension of the theory of sampling signals with finite rate of innovation (FRI) to the case of multichannel acquisition systems. The essential issue of a multichannel system is that each channel introduces different unknown delays and gains that need to be estimated for the calibration of the channels. We pose both the synchronization stage and the signal reconstruction stage as a parametric estimation problem and demonstrate that a simultaneous exact synchronization of the channels and reconstruction of the FRI signal is possible. We also consider the case of noisy measurements and evaluate the Cramér-Rao bounds (CRB) of the proposed system. Numerical results as well as the CRB show clearly that multichannel systems are more resilient to noise than the single-channel ones.

I. INTRODUCTION

Sampling theory has experienced a recent revival due in part to a string of results that have extended the classical Shannon sampling framework to classes of non-bandlimited structured signals [7], [6], [18]. In these new sampling schemes, the prior that the signal is sparse in a basis or in a parametric space is taken into account and perfect reconstruction is possible based on a set of suitable measurements. Depending on the set-up used and reconstruction method involved, the above new sampling framework goes under different names like compressed sensing (CS), compressive sampling [7], [6] or sampling signals with finite rate of innovation (FRI) [18], [8].

The set-up considered here is the one in [18], [8], where the acquisition process is modelled as in Figure 1. In [18], [8], it is shown that perfect reconstruction of classes of FRI signals is possible by utilizing the Prony's method, which is also known as the annihilating filter method [15]. Signals that can be sampled with this method include streams of Diracs, piecewise-polynomial and piecewise-sinusoidal signals [18], [8], [3]. Further results on CS and FRI sampling have then appeared providing additional extensions to the original schemes and new insights on performance bounds and reconstruction algorithms. We recommend the recent Signal Processing Magazine issue on the topic for a detailed overview [2]. Most of the papers on
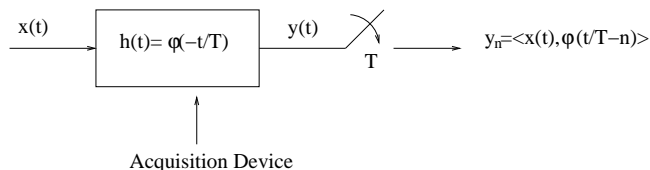

Fig. 1. Sampling set-up. Here, $x(t)$ is the continuous-time signal, $h(t)$ the impulse response of the acquisition device and T the sampling period. The measured samples are $y_n = \langle x(t), \varphi(t/T - n) \rangle$.

sparse sampling focus on a single-channel acquisition model. However, modern and fast Analogue-to-Digital Converters (ADC) use interleaved multichannel converters. This allows a reduction in the complexity of the devices while keeping higher rates of conversion. Given the practical importance of multichannel acquisition devices, it is natural to investigate extensions of sparse sampling theories to the multichannel scenario. The critical issue in multichannel sampling is the precise synchronization of the various channels since different devices introduce different drifts and different gains (due to imperfections of electronic circuits for example) that need to be estimated together with the signal itself. In this paper we therefore consider the multichannel sampling of FRI signals and extend the results in [8] to this new scenario. Multichannel sampling was first proposed by Papoulis in the context of bandlimited signals [13] and extended by Unser et al. [17] for signals lying in shift-invariant subspaces. A further extension related to union of shift-invariant subspaces has been recently considered in [9]. The multichannel sampling of FRI signals has been considered in [12], [10] and [14], but in all cases the drift problem was not addressed.

The paper is organized as follows: In the next section, we review the main results on sampling FRI signals and in particular focus on E-spline sampling kernels. Then, in Section III, we present the extension of this sampling framework to the case of multichannel acquisition devices. In particular, we show that it is possible to estimate simultaneously the channel parameters (i.e., delays and gains) and the signal from the measured samples. This is achieved by operating at a sampling rate proportional to $1/TM$, where M is the number of channels involved. In Section IV we consider the noisy scenario and show with both the Cramér-Rao bounds and numerical simulations that multichannel sampling systems are more resilient to noise than the

single-channel ones. We finally conclude in Section V.

II. SAMPLING SIGNALS WITH FINITE RATE OF INNOVATION

For the sake of clarity we assume that $x(t)$ is a stream of K Diracs with amplitudes a_k located at distinct instants $t_k \in [0, \tau[$ (other 1-D FRI signals could also be considered):

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

Furthermore, we assume the sampling period is $T = \tau/N$ where N is the number of samples. Consequently, the measurements are:

$$
\begin{aligned}
y_n &= \langle x(t), \varphi(t/T - n) \rangle \\
&= \sum_{k=0}^{K-1} a_k \varphi(t_k/T - n), \quad n, = 0, 1, ...N - 1.
\end{aligned}
$$

In [18], [8], it was shown that with a proper choice of the acquisition kernel, a perfect reconstruction of $x(t)$ from the samples y_n is possible. The kernels used in [18] are the sinc and the Gaussian functions. In this paper, we will focus on a specific class of kernels, used in [8], that are able to reproduce real or complex exponentials. A kernel $\varphi(t)$ is able to reproduce exponentials up to order P if there exists coefficients $c_{m,n} \in \mathbb{C}$ such that:

$$\sum_{n \in \mathbb{Z}} c_{m,n} \varphi(t - n) = e^{\alpha_m t} \quad \text{with } \alpha_m \in \mathbb{C} \quad \text{and } m = 0, 1, ..., P. \tag{2}$$

The coefficients $c_{m,n}$ in the above equation are given by the following equation:

$$c_{m,n} = \frac{1}{T} \int_{-\infty}^{\infty} e^{\alpha_m t} \tilde{\varphi}(t/T - n) dt,$$

where $\tilde{\varphi}(t)$ is chosen to form with $\varphi(t)$ a quasi-biorthonormal set [5]. This includes the particular case where $\tilde{\varphi}(t)$ is the dual of $\varphi(t)$, that is, $\langle \tilde{\varphi}(t - n), \varphi(t - k) \rangle = \delta_{n,k}$. In the sampling set-up of [8] the choice of the exponents in (2) is restricted to $\alpha_m = \alpha_0 + m\lambda$ with $\alpha_0, \lambda \in \mathbb{C}$ and $m = 0, 1, ..., P$. This is done to allow the use of the annihilating filter method at the reconstruction stage. This will be more evident when the reconstruction method is described later on.

Exponential splines (E-splines) [16] are central to the exponential reproduction property. The Fourier transform of the $(P+1)$-th order E-spline is:

$$\hat{\beta}_{\vec{\alpha}_P}(\omega) = \prod_{m=0}^{P} \left(\frac{1 - e^{\alpha_m - j\omega}}{j\omega - \alpha_m} \right). \tag{3}$$

The above E-spline is able to reproduce the exponentials $e^{\alpha_m t}$ with $m = 0, 1, ..., P$. It is interesting to point out that by choosing complex conjugate exponents for the E-spline sampling kernel, real-valued functions will be obtained.

The time domain expression of a spline of order two is given by:

$$\beta_{\vec{\alpha}_P}(t) = \begin{cases} \frac{e^{\alpha_0 t} - e^{\alpha_1 t}}{\alpha_0 - \alpha_1} & 0 \le t < 1 \\ \frac{e^{\alpha_1 t + \alpha_0 - \alpha_1} - e^{\alpha_0 t + \alpha_1 - \alpha_0}}{\alpha_0 - \alpha_1} & 1 \le t < 2. \end{cases}$$

Moreover, since the exponential reproduction formula is preserved through convolution [16], any composite function of the form $\varphi(t) * \beta_{\vec{\alpha}_P}(t)$ is also able to reproduce exponentials. Finally, in the case of E-splines, it can be shown that the exponential reproduction coefficients can be computed using the following expression:

$$c_{m,n} = \begin{cases} \frac{e^{am}}{\sum_{k=-P+1}^{1} e^{\alpha_m k} \beta_{\vec{\alpha}_P}(1-k)} & n = 0 \\ e^{\alpha_m n} c_{0,m} & n \ne 0. \end{cases}$$

Having gone through the sampling process, we will now explain the reconstruction process [8]. The reconstruction scheme operates as follows: First the samples are linearly combined with the coefficients $c_{m,n}$ of (2) to obtain the surrogate measurements:

$$s_m = \sum_{n=0}^{N-1} c_{m,n} y_n \quad m = 0, 1, ..., P. \tag{4}$$

Then, assuming that the original signal is a stream of Diracs as the one in (1), we have that:

$$s_m = \langle x(t), \sum_n c_{m,n} \varphi(t/T - n) \rangle = \int_{-\infty}^{\infty} x(t) e^{\alpha_m t/T} dt = \sum_{k=0}^{K-1} \hat{a}_k u_k^m \tag{5}$$

where $m = 0, 1, ..., P$, $\hat{a}_k = a_k e^{\alpha_0 t_k/T}$ and $u_k = e^{\lambda t_k/T}$. Here we have used the fact that $\alpha_m = \alpha_0 + m\lambda$. Notice also that s_m represent the exact exponential moments of the continuous-time signal $x(t)$. The new pairs of unknowns $\{\hat{a}_k, u_k\}$ can then be retrieved from the power series $s_m = \sum_{k=0}^{K-1} \hat{a}_k u_k^m$ using the classical Prony's method. The key ingredient of this method is the annihilating filter. Call h_m, $m = 0, 1, ..., K$ the filter with z-transform:

$$H(z) = \sum_{m=0}^{K} h_m z^{-m} = \prod_{k=0}^{K-1} (1 - u_k z^{-1})$$

where the roots of $H(z)$ correspond to the locations u_k. It clearly follows that:

$$h_m * s_m = \sum_{i=0}^{K} h_i s_{m-i} = \sum_{k=0}^{K-1} \hat{a}_k u_k^m \underbrace{\sum_{i=0}^{K} h_i u_k^{-i}}_{H(u_k)} = 0. \tag{6}$$

The filter h_m is thus called annihilating filter since it annihilates the observed series s_m. Moreover, the zeros of this filter uniquely define the set of locations u_k as the locations are distinct. The identity in (6) can be written in matrix/vector form as follows:

$$SH = 0 \tag{7}$$

which reveals that the Toeplitz matrix S is rank deficient. By solving the above system, we find the filter h_m and then retrieve the u_k's by computing the roots of $H(z)$. Given the u_k's we retrieve the locations t_k
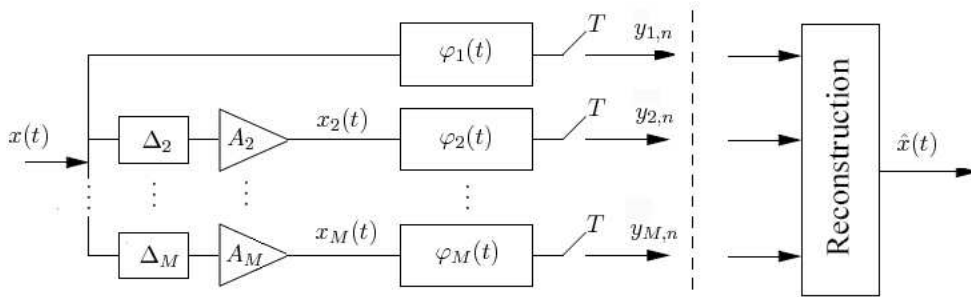
Fig. 2. Multichannel sampling set-up. Here, the continuous-time signal $x(t)$ is received by multiple channels with multiple acquisition devices. The samples $y_{i,n}$ from each channel are utilized jointly for the reconstruction process.

and finally the weights a_k are obtained by solving, for instance, the first K consecutive equations in (5). Notice that the problem can be solved only when $P \geq 2K - 1$.

We thus conclude that perfect reconstruction of a stream of K Diracs is possible with any kernel able to reproduce exponentials. If $x(t)$ has more than K Diracs or possibly an infinite number of Diracs, we cannot use the above method directly. However, since the kernels considered have compact support, the above scheme can be applied sequentially. More precisely, it was shown in [8] that if there are no more than K Diracs in an interval of size $\tau = 2KLT$, then we are guaranteed that two groups of K consecutive Diracs are sufficiently distant and that they are separated by some zero samples. By locating these zeros, one can separate the two groups and apply the above reconstruction method on each group independently. Here L indicates the support of the sampling kernel. If $x(t)$ has more than K Diracs in an interval of size $2KLT$ then the only way to sample it is by increasing the sampling rate. In the next section we show that this can be avoided by using a multichannel acquisition system.

III. THE MULTICHANNEL SCENARIO

In this section, we consider the multichannel sampling set-up shown in Figure 2. Here $x(t)$ is a stream of K Diracs located at distinct instants $t_k \in [0, \tau[$ and is acquired with a system made of M channels. Each channel introduces a different unknown delay and gain. The sampling kernel in each channel is able to reproduce exponentials, in fact, for simplicity we assume that each kernel is an E-spline of order $P + 1$. The samples of the i-th channel are thus given by:

$$y_{i,n} = \langle A_i x(t - \Delta_i), \varphi_i(t/T - n) \rangle,$$

with $A_1 = 1$ and $\Delta_1 = 0$. Our goal is to have a reconstruction system that can perfectly retrieve both the

input signal and the unknown delays and gains. Let us first assume that $M = 2$ and that:

$$\hat{\varphi}_1(\omega) = \prod_{m=0}^{P} \left(\frac{1 - e^{\alpha_m - j\omega}}{j\omega - \alpha_m} \right) \text{ and } \hat{\varphi}_2(\omega) = \prod_{m=P-1}^{2P-1} \left(\frac{1 - e^{\alpha_m - j\omega}}{j\omega - \alpha_m} \right).$$

Specifically, both kernels can reproduce the exponentials $e^{\alpha_{P-1}t}$ and $e^{\alpha_P t}$. We now combine the samples with the proper coefficients $c_{m,n}^{(i)}$, $i = 1, 2$ to obtain:

$$s_m^{(1)} = \sum_{n=0}^{N} c_{m,n}^{(1)} y_{1,n} = \int_{-\infty}^{\infty} x(t) e^{\alpha_m t} dt, \quad m = 0, 1, \ldots, P$$

and

$$s_m^{(2)} = \sum_{n=0}^{N} c_{m,n}^{(2)} y_{2,n} = \int_{-\infty}^{\infty} A_2 x(t - \Delta_2) e^{\alpha_m t} dt = A_2 s_m^{(1)} e^{\alpha_m \Delta_2},$$

$m = P - 1, P, \ldots, 2P - 1$. As shown in the equations above, in our set-up we have two common parameters between the exponents of the two channels, that is at $m = P - 1$ and $m = P$. The gain factor and the delay factor can therefore be calculated from the following two equations:

$$\Delta_2 = \frac{1}{\alpha_{P-1} - \alpha_P} \ln \left(\frac{s_P^{(1)} s_{P-1}^{(2)}}{s_{P-1}^{(1)} s_P^{(2)}} \right) \tag{8}$$

and

$$A_2 = \frac{s_P^{(2)}}{s_P^{(1)}} e^{-\alpha_m \Delta_2}. \tag{9}$$

This reveals that, independently of $x(t)$, it is possible to synchronize the two channels exactly from the samples $y_{i,n}$. However, a few considerations need to be addressed here. In the above analysis we have implicitly assumed that $s_m^{(i)} \neq 0$ for $m = P - 1, P$ and $i = 1, 2$ and this is not always true. Thus, for a guaranteed synchronization of the channels, some constraints need to be imposed on the signal. In our context we are interested in streams of Diracs and in this case the simple assumption that, given the K locations of the Diracs, the amplitudes are drawn from a non-singular distribution over $\mathbb{R}^K$ guarantees that the event $s_m^{(i)} = 0$ has probability zero. This is clearly a fairly mild hypothesis and similar type of conditions can be imposed on any other FRI signal. This type of recovery is also known as almost sure recovery [11].

Let us now return to the original problem of reconstructing the stream of Diracs $x(t)$. Given the exact gain and delay of channel two, we can now estimate the moments $s_m^{(1)}$, with $m = P + 1, P + 2, ..., 2P - 1$ from $s_m^{(2)}$ as follows:

$$s_m^{(1)} = \frac{s_m^{(2)}}{A_2} e^{-\alpha_m \Delta_2}, \quad m = P + 1, P + 2, ..., 2P - 1.$$

It then follows that if $2P - 1 \geq 2K - 1$ or, more simply, $P \geq K$ then a perfect recovery of $x(t)$ is possible from the moments $s_m^{(1)}$, $m = 0, 1, ...2P - 1$ by using the annihilating filter method discussed in the previous section. The advantage of the new set-up is that we now require splines of lower order (i.e., $P \geq K$ rather than $P \geq 2K - 1$) and this leads to shorter kernels.

More precisely, for E-splines of order P we know that the support is $L = P+1$. So, in the single channel case, the sampling of an infinite stream of Diracs requires that there are no more than K Diracs in an interval of size $2KLT = 2K(P+1)T \leq 4K^2T$, where we have used the fact that in the single channel set-up $P \geq 2K - 1$. In the case of the two-channel acquisition system, the interval is reduced to $(2K^2 + 2K)T$ since $P \geq K$. This indicates that in the new set-up we can either sample signals with a higher concentration of Diracs or alternatively for the same signal we can almost halve the sampling rate.

The extension to the case of M channels is now straightforward. By designing each sampling kernel so that pairs of channels have two moments in common, it is possible to synchronize the channels using equations (8),(9) and then reconstruct $x(t)$. For M channels and K Diracs the requirement is now that $MP - (M-1) \geq 2K - 1$ which implies $P \geq \lceil (2K-2)/M \rceil + 1$. This indicates that by using an M-channel system we can either sample the same FRI signals with a reduced sampling rate proportional to $\sim 1/TM$ or sample signals with a much higher density of Diracs.

In the next section, we will assess the resilience to noise of the proposed multichannel system.

IV. NOISY SCENARIO

We saw in the previous section that, in the noiseless case, a multichannel acquisition system achieves perfect reconstruction of FRI signals with a sampling rate proportional to $1/TM$. Thus, perfect reconstruction is achieved at lower sampling rates. We now show that, if we do not reduce the sampling rate and leave it fixed at T, a multichannel system is more resilient to noise than a single-channel one.

Consider the multichannel sampling set-up of Figure 2 and let us now assume that, due to the noise, the samples we measure are given by:

$$\hat{y}_{i,n} = \langle A_i x(t - \Delta_i), \varphi_i(t/T - n) \rangle + \epsilon_{i,n},$$

where $n = 0, 1, ...N - 1$, $i = 1, 2, ..., M$, $\epsilon_{i,n}$ is i.i.d. additive Gaussian noise with zero mean and variance σ^2 and $T = \tau/N$, with $\tau = 1$ seconds. For the sake of simplicity, we assume that the gains introduced by the channels are all equal and known a-priori, that is, $A_i = 1$, $i = 1, 2, ..., M$, therefore only the delays need to be estimated. Moreover, we assume that the input signal $x(t)$ has $K = 3$ Diracs with known fixed amplitudes. Our aim is to compare the performance of the single channel set-up (i.e., $M = 1$) against the two-channel and three-channel systems.

Since for all cases we have a standard parametric estimation problem, we use Cramér-Rao bounds to compare the minimum bounds of the different set-ups and probe whether the multichannel systems are in theory more resilient to noise. In the single-channel scenario, the reconstruction of $x(t)$ involves the estimation of three unknown parameters (i.e., the three locations of the Diracs). In the multichannel set-up instead, one has to retrieve the unknown delays for the synchronization process and the unknown locations
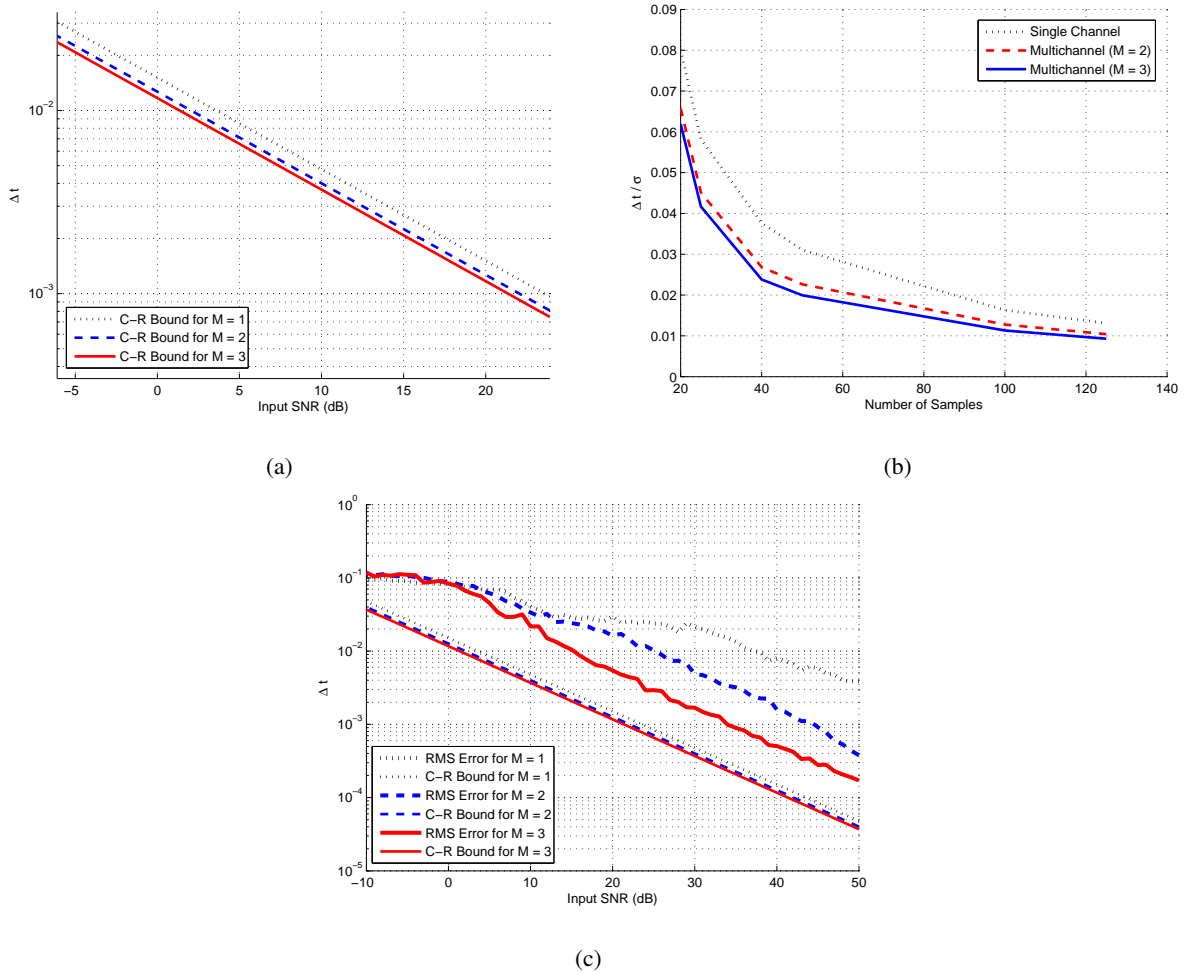
(a)

(b)

(c)

Fig. 3. (a) CRB for single channel and multichannel sampling systems. The input SNR is calculated as $10log_{10}\frac{||y||^2}{\sigma^2}$ where σ^2 is the noise variance and Δt is the uncertainty on the estimated locations. (b) Theoretical uncertainties on the estimated locations with varying sampling rates. (c) Numerical results with single and multichannel sampling. Dirac locations are set at 0.5, 0.6 and 0.7 set for all cases. The delays Δ_2 and Δ_3 are fixed at $\frac{T}{2}$ and T respectively.

of the Diracs. Therefore, there are four parameters to be estimated when $M = 2$ and five parameters when $M = 3$. As the large number of unknown parameters leads to a fairly large Fisher information matrices, it is simpler to evaluate the CRB numerically for all cases. The sampling kernel used in our analysis is a complex-valued E-spline function with equally spaced, purely imaginary exponents with $\alpha_0 = 0$ and $\lambda = 0.18j$. The order of E-splines for each channel is fixed and set to $P = 9$.

The CRB for the estimation of the Diracs for the three cases discussed above, assuming a fixed number of samples $N = 20$, are shown in Figure 3(a). Interestingly, the results reveal that the CRB improves with the number of channels. More precisely, the CRB improvement when going from single-channel to two channels is approximately 0.86dB, while the improvement when going from single channel to three channel system is approximately 1.1dB. It is interesting to see that, despite the fact that the unknown delays need

to be estimated in order to synchronize the channels, there is still a noticeable gain by using multichannel sampling set-up when compared to the single channel sampling set-up. Furthermore, in Figure 3(b) we show the CRB of each sampling system at varying sampling rates. We can see that at a given uncertainty of the estimated locations, there is a reduction in the number of samples needed when going from single-channel to multichannel sampling systems. For example, at the reconstruction quality of $\frac{\Delta t}{\sigma} = 0.04$, the number of samples could be reduced from 38 samples to 27 samples when going from the single channel to the three channel set-up.

To analyze the performance of the reconstruction algorithm, Figure 3(c) presents some actual numerical results on the uncertainty of the estimated locations which are also compared against the theoretical bounds from Figure 3(a). The delays are estimated using Eq. (8), while the locations of the Diracs are obtained using a variation of the annihilating filter method, known as the matrix-pencil method [15] and also the Cadzow's algorithm [4] to further denoise the surrogate measurements s_m. While none of the algorithms achieve the CRB, the obtained results show that the gain in performance with multichannel sampling over single channel sampling can be significant. For instance, at input SNR= 15dB, the gain in performance from single channel to three channels is approximately 4.4dB.

V. CONCLUSIONS

In this paper we have extended the theory of sampling FRI signals to the case of multichannel acquisition systems. We illustrated that by synchronizing the different channels of the proposed multichannel sampling set-up, one can estimate the unknown delays and gains introduced within the channels, regardless of the input FRI signal. For the case of noisy measurements, we showed by evaluating the CRB that, despite the fact that the unknown delays are required to be estimated for synchronization, the number of samples required can be reduced by the multichannel set-up. Moreover, our numerical results with the multichannel sampling set-up revealed that the performance of the reconstruction algorithm improves over the single channel case.

REFERENCES

[1] L. Baboulaz and P.L. Dragotti. Distributed acquisition and image super-resolution based on continuous moments from samples. In *Proc. of IEEE Int. Conf. on Image Processing(ICIP)*, Atlanta (GA), October 2006.

[2] R. G. Baraniuk, E. Candès, R. Nowak, and M. Vetterli. Compressive sampling. *IEEE Signal Processing Magazine*, 25(2):12–13, March 2008.

[3] J. Berent, P.L. Dragotti, and T. Blu. Sampling piecewise sinusoidal signals with finite rate of innovation methods. *IEEE Trans. on Signal Processing*, 58(2):613–625, February 2010.

[4] T. Blu, P.L. Dragotti, M. Vetterli, Marziliano P., and Coulot L. Sparse sampling of signal innovations: Theory, algorithms and performance bounds. *IEEE Signal Processing Magazine*, 25(2):31–40, March 2008.

[5] T. Blu and M. Unser. Approximation errors for quasiinterpolators and (multi-) wavelet expansions. *Applied and Computational Harmonic Analysis*, 6 (2):219–251, March 1999.

[6] E. Candès, J. Romberg, and T. Tao. Robust uncertainty principle: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Information Theory*, 52(2):489–509, February 2006.

[7] D.L. Donoho. Compressed sensing. *IEEE Trans. Information Theory*, 52(4):1289–1306, April 2006.

[8] P.L. Dragotti, M. Vetterli, and T. Blu. Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix. *IEEE Trans. on Signal Processing*, 55(5):1741–1757, May 2007.

[9] Y. Eldar. Compressed sensing of analog signals in shift-invariant spaces. *IEEE Trans. Signal Processing*, 57(8):2986–2997, August 2009.

[10] Olkkonen H. and Olkkonen J.T. Measurement and reconstruction of impulse train by parallel exponential filter. *IEEE Signal Processing Letter*, 15:241–244, February 2008.

[11] A. Hormati, O. Roy, Lu. Y.M., and M. Vetterli. Distributed sampling of correlated signals linked by sparse filtering: Theory and applications. *IEEE Trans. on Signal Processing*, 58(3):1095–1109, March 2010.

[12] J. Kusuma and V.K. Goyal. On the accuracy and resolution of powersum-based sampling methods. *IEEE Trans. Signal Processing*, 57(1):182–193, January 2009.

[13] A Papoulis. Generalized sampling expansions. *IEEE Trans. on Circuits Syst.*, 24:652–654, 1977.

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

[15] P. Stoica and R. Moses. *Spectral Analysis of Signals*. Prentice Hall, April 2005.

[16] M. Unser and T. Blu. Cardinal Exponential Splines: Part I-theory and filtering algorithms. *IEEE Trans. on Signal Processing*, 53(4):1425–1438, April 2005.

[17] M. Unser and J. Zerubia. A generalized sampling theory without bandlimiting constraints. *IEEE Trans. on Circuits Syst. II*, 45(8):959–969, August 1998.

[18] M. Vetterli, P. Marziliano, and T. Blu. Sampling signals with finite rate of innovation. *IEEE Trans. Signal Processing*, 50(6):1417–1428, June 2002.