FINDING THE MINIMUM RATE OF INNOVATION IN THE PRESENCE OF NOISE

Christopher Gilliam and Thierry Blu

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

ABSTRACT

Recently, 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 problem of determining the minimum rate of innovation (RI) in a noisy setting. First, we adapt a recent model-fitting algorithm for FRI recovery and demonstrate that it achieves the Cramér-Rao bounds. Using this algorithm, we then present a framework to estimate the minimum RI based on fitting the sparsest model to the noisy samples whilst satisfying a mean squared error (MSE) criterion - a signal is recovered if the output MSE is less than the input MSE. Specifically, given a RI, we use the MSE criterion to judge whether our model-fitting has been a success or a failure. Using this output, we present a Dichotomic algorithm that performs a binary search for the minimum RI and demonstrate that it obtains a sparser RI estimate than an existing information criterion approach.

***Index Terms*—** Finite rate of innovation, model order, model-fitting, sampling theory, recovery of Dirac pulses

1. INTRODUCTION

A crucial element in the acquisition of all real world signals is the ability to convert a signal between the continuous and discrete-time domains. Unsurprisingly, perfect reconstruction when converting between these domains is highly prized. Recently, 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 time. Specifically, the authors showed that a periodic stream of Diracs and a piecewise polynomial could be perfectly reconstructed using a sinc or Gaussian sampling kernel.

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 sampling kernels were proposed in [3], and reconstruction of piecewise sinusoidal signals examined in [4]. More recently, sampling and reconstruction of FRI signals using arbitrary kernels was presented in [5] and recovery from non-uniform samples was examined in [6, 7]. FRI theory has also been generalised to spherical coordinate schemes in [8, 9] and higher dimensional signals, such as multi-dimensional Diracs in [10] and curves in [11]. As a result, FRI has found application in noisy channel detection in ECG [12], reconstruction of MRI data [13], the detection of spikes in neurophysiological data [14] and in ultrasound imaging in [15]. A key requirement in all of the work outlined so far is knowledge of the rate of innovation (RI) of the signal. In practice, however, such knowledge is likely to be unknown thus an important topic in FRI sampling is the estimation of this rate.

This work was supported by a grant #CUHK14600615 from the Hong Kong Research Grants Council.

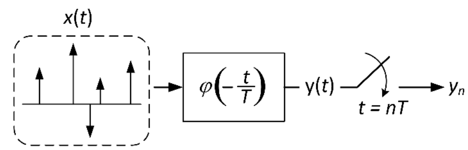


Fig. 1. The FRI acquisition system. The continuous-time input signal $x(t)$, in this case a sequence of K Diracs, is filtered by a sampling kernel $\varphi(-t/T)$ and sampled at a period T.

In this paper, we present a novel framework to determine the minimum rate of innovation of a noisy FRI signal. Instead of testing all possible RI values and choosing one that achieves the minimum MSE [16, 17], our framework is based on the following mean squared error (MSE) criterion: an FRI signal has been successfully recovered from its noisy samples if the MSE between the recovered signal and the noisy samples is less than the input MSE (i.e. original noise level). Given this MSE criterion, we adapt a recent model-fitting algorithm [18] so that it reliably achieves the criterion when the RI is correct. Consequently, we can use the algorithm to judge a given RI value: if the criterion is met then a lower RI may exist, whereas if the reverse is true then the RI needs to be increased. Accordingly, we propose a Dichotomic algorithm that uses the model-fitting method to perform an efficient binary search for the minimum RI of a noisy FRI signal. We demonstrate that this algorithm estimates a sparser RI than the standard Bayesian Information Criterion [19] and that the model-fitting algorithm used reaches the Cramér-Rao bounds.

The paper is organised as follows. In Section 2, we review the 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 [20]. In Section 3, we examine the concept of using the mean squared error (MSE) as a criterion for assessing the recovery of an FRI signal. Using this MSE criterion, we adapt and analyse a robust model-fitting method for FRI recovery in Section 4. Next, using this model-fitting method, we present a novel algorithm to determine the minimum RI in Section 5 and evaluate its performance in Section 6. 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

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

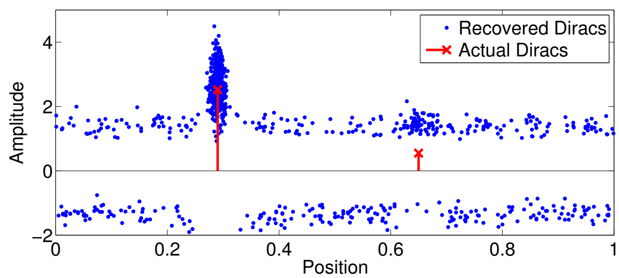


Fig. 2. Example of recovering 2 Diracs in heavy noise (SNR = 0 dB). The blue dots indicate the recovered Diracs using maximum likelihood over 500 realisations. Note that $N = 21$ samples.

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 that are characterised by a set of locations $\{t_k\}_{k=1}^{K}$ and a set of amplitudes $\{x_k\}_{k=1}^{K}$. This type of FRI signal has a rate of innovation of $2K/\tau$ and is defined as

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

Note that the locations are restricted such that $t_k \in [0, \tau[$. This signal is then filtered with a sinc kernel with bandwidth $B = 1/T$. Therefore, using the definition of the Dirichlet kernel (or τ-periodic sinc function), the samples of (2) we obtain are

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

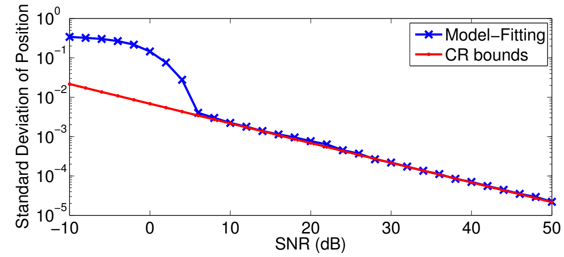
Note that $T = \tau/N$ in this framework thus $N = B\tau$ as $B = 1/T$. Also, without loss of generality, we shall assume $B\tau$ is an odd integer.

Now, from [1, 2, 3], the standard framework for recovering the signal $x(t)$ in noiseless conditions is as follows. The first element is to map the FRI samples y_n in such a way that the resulting sample moments s_m have a power sum form. The mapping in question is dependent upon the sampling kernel, e.g. an exponential reproducing mapping is used for arbitrary sampling kernel in [5]. In the case of a sinc kernel, however, it is simply the discrete Fourier transform of the samples:

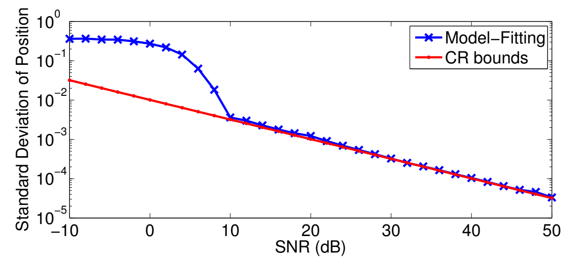
$$s_m = \sum_{n=0}^{N-1} y_n \mathrm{e}^{-j2\pi mn/N} = \sum_{k=1}^{K} x_k \mathrm{e}^{-j2\pi m t_k/\tau}, \tag{4}$$

for $m = -M, \ldots, M$, where $M = \lfloor N/2 \rfloor$ and $N = B\tau$.

Given this power sum form, the locations $\{t_k\}_{k=1}^{K}$ are determined using the non-linear annihilating filter method (also known as Prony's method). In brief, this method involves determining a filter H whose coefficients $\mathbf{h} = [h_0, h_1, \ldots, h_K]^{\mathrm{T}}$ satisfy $h * s_m = 0$, where $*$ represents convolution. The locations of the Diracs are then determined from the roots of the annihilating filter H. For further details of the method see [2]. Finally, the amplitudes $\{x_k\}_{k=1}^{K}$ are determined via least mean squares. Note that $N \geq 2K + 1$ samples are required for perfect recovery.



(a) $\{t_1, x_1\} = \{0.39, 2.52\}$



(b) $\{t_2, x_2\} = \{0.65, 1.7\}$

Fig. 3. Comparing the performance of the model-fitting approach to the Cramér-Rao (CR) bounds. The graphs show the standard deviation of the position estimate as the noise level increases. Note that $K = 2$ Diracs and $N = 21$ samples.

2.1. Model Mismatch

Unfortunately, in practice, the samples we obtain are corrupted by noise or more generally model mismatch. We denote the noisy samples as $\tilde{y}_n$ and the noisy moments as $\tilde{s}_m$. The presence of this noise means that the annihilation equation, $h * s_m = 0$, is no longer valid.

FRI algorithms designed to overcome this issue with noise can be split into four categories. The first category exploits the observation: a Toeplitz matrix formed from the moments s_m will have a rank of K. Accordingly, denoising is achieved by enforcing, in an iterative manner, the K-rank Toeplitz structure onto the noisy moments $\tilde{s}_m$. This operation was performed using Cadzow denoising [21] in [2], and using structured low rank approximation [22] in [23]. The second category uses subspace methods, e.g. the matrix pencil [24], to directly estimate the locations $\{t_k\}_{k=1}^{K}$. This type of approach was first proposed for FRI in [25] and has been subsequently used in [5, 26]. The third category covers stochastic methods for FRI recovery such as Gibbs sampling in [27] and a genetic algorithm in [28].

The final category is based on model-fitting [18, 16]. Instead of trying to solve the annihilation equation, the central concept is to fit an FRI model to the noisy samples (or moments) and thus recover an estimate of the FRI samples. In this paper, we adapt the model-fitting algorithm in [18] in order to determine the minimum rate of innovation of a signal.

3. MSE CRITERION FOR FRI RECOVERY

Often, the performance of FRI recovery algorithms is based on how accurately the parameters $\{t_k, x_k\}_{k=1}^{K}$ have been estimated in comparison to the Cramér-Rao (CR) bounds [2, 5, 20]. However, this accuracy may be unreliable, e.g. FRI algorithms meet the CR bounds only up to a certain breakdown Signal-to-Noise-Ratio (SNR) [29]. Also, in practice, we do not have access to the original parameters.

Instead, we follow the concept introduced in [18] - assessing

FRI recovery based on a mean squared error (MSE) between the reconstructed FRI samples, $\hat{y}_n$, and the noisy samples, $\tilde{y}_n$, which is termed MSE_R. In more detail, rather than just trying to minimise this value, e.g. maximum likelihood estimation [30, 31], the authors constructed the following criterion based on the input MSE, MSE_IN, between y_n and $\tilde{y}_n$:

$$\text{Criterion:} \qquad \text{MSE}_\text{R} < \text{MSE}_\text{IN}. \tag{5}$$

Thus, the aim when recovering an FRI signal is to minimise MSE_R until the criterion above is satisfied.

In this paper, we want to use this criterion to determine the unknown rate of innovation for an FRI signal. However, as (5) can be satisfied with any RI that is larger than the original, we evoke the concept of parsimony and aim to estimate the minimum RI required to satisfy the criterion. A consequence of this approach is that depending on the value of MSE_IN we may estimate a RI lower than the original, i.e. we lose Diracs. To understand how this could happen, consider estimating the two Diracs shown in Figure 2. The figure shows that under heavy noise corruption, $\text{SNR} = 0\,\text{dB}$, the estimate of the smaller Dirac is unstable even when performing maximum likelihood estimation. In other words, the smaller Dirac is indistinguishable from the noise level MSE_IN thus a sparser model is more appropriate to approximate the FRI signal.

4. MODEL-FITTING USING A RATIO OF POLYNOMIALS

The central concept of the model-fitting approach presented in [18] is that the noiseless samples y_n can be expressed as a ratio of two polynomials: a numerator P of order $K-1$, with coefficients $\mathbf{p}$, and a denominator H (the annihilation filter) of order K, with coefficients $\mathbf{h}$. Therefore, in the presence of noise, the authors propose minimising, subject to the MSE criterion, the fit between this model and noisy samples $\tilde{y}_n$:

$$\min_{H,P} \sum_{n=0}^{N-1} \left| \tilde{v}_n - \frac{P\big(\mathrm{e}^{j\omega_n}\big)}{H\big(\mathrm{e}^{j\omega_n}\big)} \right|^2, \tag{6}$$

where $\tilde{v}_n = \tilde{y}_n \mathrm{e}^{-j2\pi n M/N}$ and $\omega_n = 2\pi n/N$. The advantage of this approach is that the FRI samples are completely defined by the coefficients of the respective polynomials. To overcome the non-linear nature of the problem, [18] used a iterative linear minimisation strategy, which is similar to the Steiglitz-McBride algorithm [32] and Sanathanan and Koerner algorithm [33].

Now, in this paper, we introduce two novel elements to the iterative minimisation to improve the robustness of the model-fitting approach. First, we use a new solving constraint proposed in [16]: $\mathbf{h}_0^{\text{H}}\mathbf{h}_i = 1$. Thus, the minimisation we wish to solve at each iteration is

$$\min_{H_i,P} \sum_{n=0}^{N-1} \left| \frac{H_i\big(\mathrm{e}^{j\omega_n}\big)\tilde{v}_n - P\big(\mathrm{e}^{j\omega_n}\big)}{H_{i-1}\big(\mathrm{e}^{j\omega_n}\big)} \right|^2, \ \text{s.t.} \ \mathbf{h}_0^{\text{H}}\mathbf{h}_i = 1, \tag{7}$$

where i represents the iteration number and $\mathbf{h}_0$ the initial value of the coefficients. The benefit of this constraint is that the solution to (7) is equivalent to solving a small linear system of equations (i.e. very efficient and fast). The second element we introduce is the idea of using a sequence of random initialisations when trying to solve (7). In other words, if, after a finite number of iterations, the minimisation in (7) fails to satisfy the MSE criterion then a new initialisation is chosen and the process is repeated. Although this may seem costly, as we shall now demonstrate, few initialisations are needed in the vast majority of cases.

Table 1. Analysis of the number of failure cases (out of 500,000 tests) of the Model-Fitting Algorithm as the number of random initialisations varies.

	SNR (dB)											
# Rand Init	20	18	16	14	12	10	8	6	2	4	0	-2
100	0	0	1	2	4	3	3	6	5	2	0	0
50	2	2	4	6	15	18	23	15	17	7	0	0
30	18	25	22	32	56	69	75	58	42	9	1	0
$\text{Q}_{99.9}$	11	12	13	14	15	15	15	11	11	6	3	2

* FRI signal setting: $K = 6$ Diracs and $N = 51$ samples.
** $\text{Q}_{99.9}$ = 99.9th Quantile of the number of initialisations required

Table 2. Estimating the number of Diracs for an FRI signal.

	SNR (dB)					
	30	25	20	15	10	5
Dichotomic Algorithm	**12**	**11**	**10**	**10**	**8**	**7**
BIC [19] w. Cadzow denoising [2]	13	13	11	11	11	8

* FRI signal setting: $K = 12$ Diracs and $N = 97$ samples.
** Bold values indicate the minimum number of Diracs

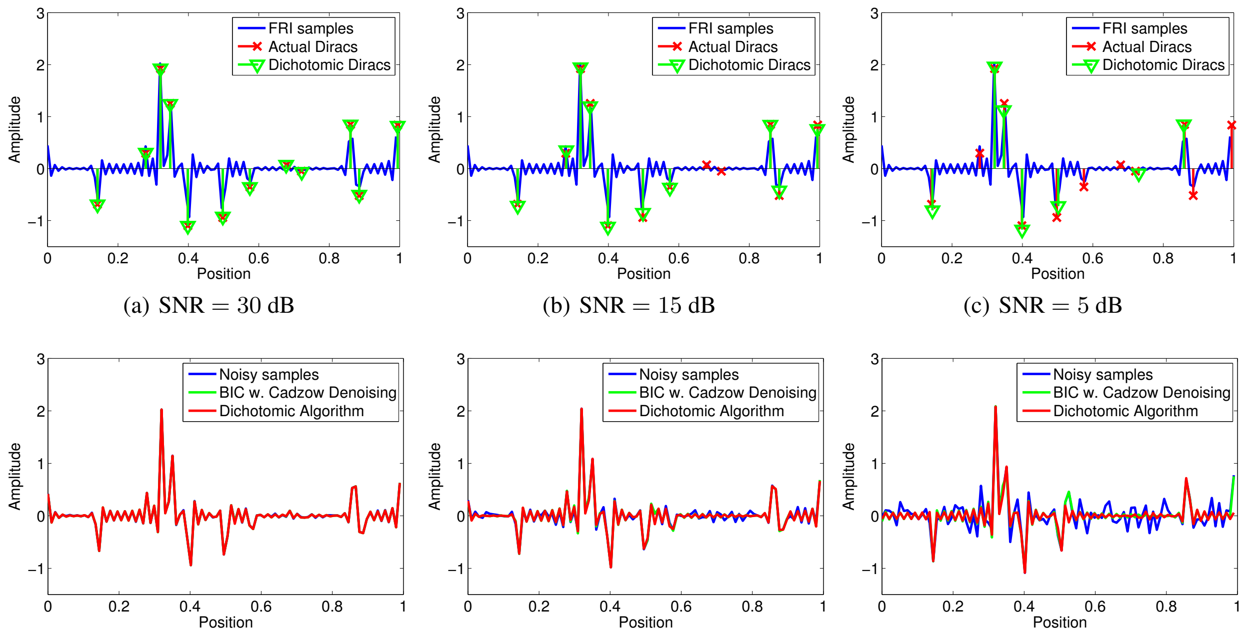
4.1. Validating the Robustness of the Model-Fitting

We start this validation by comparing the performance of the model-fitting algorithm to CR bounds when estimating an FRI signal comprising $K = 2$ and $N = 21$ samples. The resulting standard deviation of the location estimate for each Dirac is shown in Figure 3. Importantly, the figure demonstrates that the MSE criterion still allows the algorithm to reach the CR bounds for high SNR values. In the second validation, we examine the relationship between the number of random initialisations and the number of times the model-fitting algorithm fails to satisfy the MSE criterion. For this validation, we use an FRI signal comprising $K = 6$ Diracs and $N = 51$ samples, and use a limit of 50 iterations for the fitting algorithm. Using 500,000 realisations of each noise level, the number of failure cases for a varying number of random initialisations and SNR levels are detailed in Table 1. The table shows that the number of failure cases quickly decreases as the number of initialisations increases; in particular, the maximum 99.9% quantile of the number of initialisations required never exceeds 15.

5. DETERMINING THE RATE OF INNOVATION

Given the model-fitting method described above and an accuracy MSE_IN, we now present our algorithm to determine the minimum RI for an FRI signal. The central element of the algorithm is that our model-fitting approach has a binary outcome - it either succeeds or fails when trying to met the MSE criterion for a certain RI - and, as demonstrated in the previous section, this outcome is very reliable. Accordingly, we formulate a dichotomic algorithm to determine the minimum RI using the success/failure of the model-fitting. More precisely, we perform a binary search for the minimum K, i.e. the minimum number of Diracs in the signal. Given that a set of FRI samples can reconstruct at most $L = \lfloor N/2 \rfloor$ Diracs, then this search spans the following range $K \in [1, L]$. The full details of this Dichotomic algorithm are given in Alg. 1. The advantage of using this search method is that it requires at most $\mathcal{I}$ calls of the model-fitting approach, where

$$\mathcal{I} = \lceil \log_2(L) \rceil . \tag{8}$$



(a) SNR = 30 dB (b) SNR = 15 dB (c) SNR = 5 dB

(d) SNR = 30 dB (e) SNR = 15 dB (f) SNR = 5 dB

Fig. 4. Results of the Dichotomic algorithm when recovering an FRI signal, with unknown rate of innovation, under three different noise levels: SNR $= 30$ dB, 15 dB and 5 dB. Graphs (a), (b) and (c) compare the noiseless FRI samples and the original Diracs to the recovered Diracs. Graphs (d), (e) and (f) compare the noisy FRI samples to the reconstructions achieved by Dichotomic and BIC with Cadzow denoising. Note that $K = 12$ Diracs and $N = 97$ samples.

Consequently, the dichotomic algorithm can very efficiently determine the minimum rate of innovation of a signal.

Algorithm 1 Dichotomic method to estimate the rate of innovation of an FRI signal.

Inputs: Noisy samples $\tilde{y}_n$ and MSE_{IN}

1: Set $K_{\max} = \lfloor N/2 \rfloor + 1$, $K_{\min} = 0$, $K_{\text{test}} = (K_{\max} + K_{\min})/2$ and $K_{opt} = \lfloor N/2 \rfloor$.
2: Using K_{test} and $\tilde{y}_n$, run the model-fitting algorithm described in Section 4 to obtain the reconstructed samples $\hat{y}_n$. Calculate MSE_{R}
3: Check the criterion in (5). If true, set $K_{\max} = K_{\text{test}}$ and $K_{\text{opt}} = K_{\text{test}}$, else $K_{\min} = K_{\text{test}}$.
4: Repeat steps 2 and 3 until $K_{\max} - K_{\min} = 0$.
5: Minimum rate of innovation is K_{opt}.

6. SIMULATIONS

We now analyse the performance of our Dichotomic algorithm against the standard Bayesian Information Criterion (BIC) outlined in [19]. To compute the BIC, we use Cadzow denoising from [2] to obtain a proxy of the maximum likelihood estimation at each value of K. Also, note that the BIC algorithm is applied to the noisy FRI moments $\tilde{s}_m$ hence the required signal model is a sum of sinusoids.

To perform this analysis, we use an FRI signal comprising $K = 12$ Diracs and sampled using $N = 97$ samples. The corresponding FRI samples are then subjected to a noise level varying from SNR $= 30$ dB to 5 dB. Note that the model-fitting algorithm is set to use 50 random initialisations and has a limit of 50 iterations. The resulting estimates of the number of Diracs for both algorithms are shown in Table 2 and three examples, at SNR $= 30$, 15 and 5 dB, are illustrated in Figure 4. In the figure, graphs 4(a), 4(b) and 4(c) compare the noiseless FRI samples and the original Diracs to those recovered using the Dichotomic algorithm. Whereas graphs 4(d), 4(e) and 4(f) compare the noisy FRI samples to the reconstructions achieved using the Dichotomic algorithm and the BIC with Cadzow denoising.

The results illustrate three main points: first, in benign noise conditions, the Dichotomic algorithm is capable of determining the true RI, and in turn the true FRI signal. This should not be surprising as it is built around the model-fitting algorithm that was shown to reach the CR bounds in Section 4.1. Second, the Dichotomic algorithm always obtains a lower estimate of the number of Diracs and hence a sparser estimate of the noisy FRI samples. Finally, the graphs in Figure 4 demonstrate that the Dichotomic algorithm continues to estimate the Diracs it find accurately and hence allows for a good quality approximation of the FRI samples.

7. CONCLUSIONS

In this paper, we proposed a novel framework to find the minimum rate of innovation of a noisy FRI signal. The framework is based on using a MSE criterion to assess the recovery of an FRI signal in a model-fitting algorithm. The idea is to find the sparsest model that fits the noisy FRI samples whilst satisfying the MSE criterion. To achieve this, we adapted an existing model-fitting algorithm so that it reliably met the MSE criterion when the RI was correct. We also demonstrated that the algorithm reached the Cramér-Rao bounds. The key element is that the model-fitting method acts as a binary test for arbitrary RI values - the criterion is either met or it is not. Accordingly, we presented a Dichotomic algorithm that used the model-fitting method to perform an efficient binary search to determine the minimum RI of a noisy FRI signal. Finally, we showed that the algorithm is capable of obtaining the correct RI value when no noise is present and obtaining the sparsest estimate of the RI when noise corruption occurs.

8. 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, 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, 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, 2007.

[4] 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, 2010.

[5] 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.

[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, 2012.

[7] S. Mulleti and C. S. Seelamantula, “Periodic non-uniform sampling for FRI signals,” in *Proc. Int. Acoust., Speech and Signal Process. (ICASSP)*, Brisbane, Australia, 2015, pp. 5942–5946.

[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] Ivan Dokmanić and Yue M. Lu, “Sampling spherical finite rate of innovation signals,” in *Proc. Int. Acoust., Speech and Signal Process. (ICASSP)*, Brisbane, Australia, 2015, pp. 5962–5966.

[10] 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.

[11] H. Pan, T. Blu, and P. L. Dragotti, “Sampling curves with finite rate of innovation,” *IEEE Trans. Signal Process.*, vol. 62, no. 2, pp. 458–471, 2014.

[12] A. Nair and P. Marziliano, “Noisy channel detection using the common annihilator with an application to electrocardiograms,” in *Proc. Int. Acoust., Speech and Signal Process. (ICASSP)*, Brisbane, Australia, 2015, pp. 5972–5976.

[13] G. Ongie and M. Jacob, “Super-resolution MRI using finite rate of innovation curves,” in *Proc. IEEE Int. Symp. Biomed. Imag. (ISBI)*, Brooklyn, NY, USA, 2015, pp. 1248–1251.

[14] 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.

[15] 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.

[16] Z. Doğan, C. Gilliam, T. Blu, and D. Van De Ville, “Reconstruction of finite rate of innovation signals with model-fitting approach,” *IEEE Trans. Signal Process.*, vol. 63, no. 22, pp. 6024–6036, 2015.

[17] X. Wei and P. L. Dragotti, “Universal sampling of signals with finite rate of innovation,” in *Proc. Int. Acoust., Speech and Signal Process. (ICASSP)*, Florence, Italy, 2014, pp. 1803–1807.

[18] C. Gilliam and T. Blu, “Fitting instead of annihilation: Improved recovery of noisy FRI signals,” in *Proc. Int. Acoust., Speech and Signal Process. (ICASSP)*, Florence, Italy, 2014, pp. 51–55.

[19] P. Stoica and Y. Selen, “Model-order selection: a review of information criterion rules,” *IEEE Signal Process. Mag.*, vol. 21, no. 4, pp. 36–47, 2004.

[20] 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.

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

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

[23] L. Condat and A. Hirabayashi, “Cadzow denoising upgraded: A new projection method for the recovery of dirac pulses from noisy linear measurements,” *Sampling Theory in Signal and Image Process.*, vol. 14, no. 1, pp. 17–47, 2015.

[24] 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.

[25] 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, 2005.

[26] 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.

[27] 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.

[28] A. Erdozain and P. Crespo, “A new stochastic algorithm inspired on genetic algorithms to estimate signals with finite rate of innovation from noisy samples,” *Signal Process.*, vol. 90, no. 1, pp. 134 – 144, 2010.

[29] X. Wei and P.L. Dragotti, “Guaranteed performance in the FRI setting,” *IEEE Signal Process. Letters*, vol. 22, no. 10, pp. 1661–1665, 2015.

[30] A. Hirabayashi, T. Iwami, S. Maeda, and Y. Hironaga, “Reconstruction of the sequence of Diracs from noisy samples via maximum likelihood estimation,” in *Proc. Int. Acoust., Speech and Signal Process. (ICASSP)*, Kyoto, Japan, 2012, pp. 3805–3808.

[31] A. Hirabayashi, Y. Hironaga, and L. Condat, “Sampling and recovery of continuous sparse signals by maximum likelihood estimation,” in *Proc. Int. Acoust., Speech and Signal Process. (ICASSP)*, Vancouver, Canada, 2013, pp. 6058–6062.

[32] 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.

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