

# Guaranteed Performance in the FRI Setting

Xiaoyao Wei and Pier Luigi Dragotti

**Abstract**—Finite Rate of Innovation (FRI) sampling theory has shown that it is possible to sample and perfectly reconstruct classes of non-bandlimited signals such as streams of Diracs. In the case of noisy measurements, FRI methods achieve the optimal performance given by the Cramér-Rao bound up to a certain PSNR and breaks down for smaller PSNRs.

To the best of our knowledge, the precise anticipation of the breakdown event in FRI settings is still an open problem. In this paper, we address this issue by investigating the subspace swap event which has been broadly recognised as the reason for performance breakdown in SVD-based parameter estimation algorithms. We work out at which noise level the absence of subspace swap is guaranteed and this gives us an accurate prediction of the breakdown PSNR which we also relate to the sampling rate and the distance between adjacent Diracs. Simulation results validate the reliability of our analysis.

## I. INTRODUCTION

Recently it has been shown that specific classes of non-bandlimited signals which are completely described by a finite number of parameters, called signals with finite rate of innovation (FRI), can be sampled and perfectly reconstructed using specific classes of sampling kernels. These include the sinc or Gaussian function [1], [2] or exponential reproducing kernels [3], [4], [5], [6]. Currently the most robust and widely used FRI parameter estimation algorithms are SVD-based methods, which are centered around splitting the measurement space into an estimated signal-subspace and an orthogonal-subspace. The knowledge of the orthogonal subspace allows the unique reconstruction of the FRI signal. Examples of such algorithms include Cadzow iterative algorithm [7] and matrix pencil [8].

Previous works in FRI, e.g. in [4], [2], [9], have shown that FRI algorithms achieve optimal results given by Cramér-Rao bound up to certain signal-to-noise ratio (SNR). The estimation departs from Cramér-Rao bound when SNR falls below a certain threshold (see Fig. 1). It is well known that this threshold effect appears in non-linear estimation problems, e.g. direction of arrival, and has been studied in the past [10], [11], [12]. Knowing the breakdown is crucial for ensuring the estimation algorithm operates in the asymptotic region predicted by the Cramér-Rao bound. To the best of our knowledge, however, precise anticipation of the breakdown in FRI framework is still an open question. In this paper, we explain the breakdown event by *subspace swap*, which has been broadly recognised as the reason of performance breakdown in SVD-based parameter estimation algorithms. For subspace swap we mean the situation when due to

noise the orthogonal subspace mixes with the signal subspace making the retrieval of the signal unreliable. Based on that we predict the breakdown in the case where the sampled signal is a stream of Diracs and the kernel is an exponential reproducing kernel.

Deriving a tighter lower bound (e.g. Barankin bound [13], [14], [15], [16]) for breakdown estimation may be an alternative approach, but a specific estimator in general does not achieve the bound. The breakdown point provided is therefore usually over-optimistic. In comparison, our estimated asymptotic region is directly linked to specific settings, for example, sampling rate and distance between adjacent Diracs, moreover, it is simpler to compute.

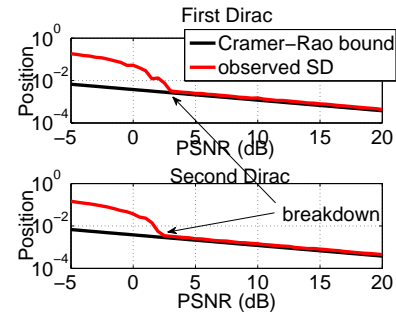


Fig. 1. Standard deviation (averages over 1000 realizations) of the retrieved locations of a FRI signal (a stream of two Diracs) compared to Cramér-Rao lower bounds. The estimation algorithm breaks down when SNR falls below a threshold.

Our contribution is two-fold. First, we explain the breakdown phenomenon in the FRI sampling method by subspace swap. Second, we provide a reliable method to anticipate the no-breakdown region in terms of PSNR, sampling rate and minimum distance between adjacent Diracs.

The paper is organized as follows. We give an overview of sampling streams of Diracs in Section 2. In Section 3 we explain the subspace swap event and provide a necessary condition for the event to happen. This leads to our key result which is the prediction of the breakdown event. We then show the reliability of our breakdown prediction by numerical results in Section 4. Finally we conclude in Section 5.

## II. SAMPLING STREAMS OF DIRACS

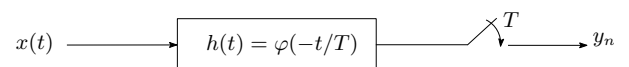


Fig. 2. Sampling set-up. Here  $x(t)$  is the input signal,  $h(t)$  is the impulse response of the acquisition device and  $T$  is the sampling period. The samples are given by  $y_n = \langle x(t), \varphi(t/T - n) \rangle$ .

In this section we give a brief overview of the theory of sampling signals with FRI and we refer to [1], [2], [3] for

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

The authors are with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK (email: ivy.wei@imperial.ac.uk; p.dragotti@imperial.ac.uk).

more details. Here emphasis is given only to key aspects of the method which will be used in Section 3 to predict when noisy FRI recovery fails.

We consider the acquisition set-up of Fig. 2, where the signal  $x(t)$  is acquired using the sampling kernel  $\varphi(t)$  and is then uniformly sampled with sampling period  $T$ . Under this model the observed samples are given by

$$y_n = \langle x(t), \varphi(t/T - n) \rangle, \quad n = 0, 1, \dots, N-1. \quad (1)$$

We assume  $x(t)$  is a stream of  $K$  Diracs:

$$x(t) = \sum_{k=0}^{K-1} a_k \delta(t - t_k), \quad (2)$$

with  $t_k \in [0, 1]$ ,  $k = 0, 1, \dots, K-1$  and the goal is to estimate the locations  $\{t_k\}_{k=0}^{K-1}$  and the amplitudes  $\{a_k\}_{k=0}^{K-1}$  from the  $N$  samples  $y_n$  (sampling rate  $T = 1/N$ ). Exact retrieval of these parameters is possible when the sampling kernel  $\varphi(t)$  is properly chosen and an important family of kernels that allows perfect reconstruction and has compact support is given by the family of exponential reproducing functions. A function  $\varphi(t)$  is able to reproduce exponentials  $e^{\alpha_m t}$ ,  $m = 0, 1, \dots, P$  when a linear combination of  $\varphi(t)$  with its integer shifts satisfies:

$$\sum_{n \in \mathbb{Z}} c_{m,n} \varphi(t - n) = e^{\alpha_m t} \quad m = 0, 1, \dots, P \quad (3)$$

for a proper choice of coefficients  $c_{m,n}$ ,  $m = 0, 1, \dots, P$ .

Assume we now have a  $P$ -th order exponential reproducing kernel  $\varphi(t)$  that can reproduce  $P+1 = 2M+1$  exponentials  $e^{\alpha_m t} = e^{m' \beta t}$  with  $m' = m - M$  and  $m = 0, \dots, 2M$ . We use this kernel to take the  $N$  samples  $y_n$  of  $x(t)$ . By linearly combining the samples  $y_n$  with the coefficients  $c_{m,n}$  in (3) we obtain the following  $2M+1$  new measurements, called exponential moments:

$$\begin{aligned} \tau_m &= \sum_n c_{m,n} y_n = \langle x(t), \sum_n c_{m,n} \varphi(t/T - n) \rangle \\ &= \int_{-\infty}^{\infty} x(t) e^{\alpha_m t/T} dt = \sum_{k=0}^{K-1} a_k e^{(\beta t_k/T) m'} \\ &= \sum_{k=0}^{K-1} \underbrace{a_k e^{-(\beta t_k/T) M}}_{\hat{a}_k} \underbrace{e^{(\beta t_k/T) m}}_{u_k^m} \\ &= \sum_{k=0}^{K-1} \hat{a}_k u_k^m, \quad m = 0, 1, \dots, 2M. \end{aligned} \quad (4)$$

Retrieving  $\hat{a}_k$  and  $u_k$  from  $\tau_m$  is now a classical problem in spectral estimation and can be solved using the annihilating filter method (a.k.a. Prony's method). We consider a filter  $\{h_m\}_{m=0 \dots K}$  which has the property that the roots of its  $z$ -transform  $\hat{h}(z)$  correspond to the locations  $u_k = e^{\beta t_k/T}$ . It can be shown that this specific filter can annihilate  $\tau_m$ , i.e.  $h_m * \tau_m = 0$ , and this can be written in matrix form as  $\mathbf{A} \mathbf{h} = \mathbf{0}$ , where  $\mathbf{A}$  is a  $(K+1) \times (K+1)$  Toeplitz matrix built from  $2K+1$  consecutive  $\tau_m$ . In noiseless situation,  $\mathbf{A}$  is rank deficient (rank  $K$ ). Therefore, the null space has dimension one and we can solve for  $\mathbf{h}$ . Then from the roots of  $\hat{h}(z)$  we

retrieve the locations exactly. Given  $\{u_k\}_{k=0}^{K-1}$ , the amplitudes are found by solving (4).

Note that the subspace spanned by the  $K$  singular vectors related to the non-zero singular values of  $\mathbf{A}$  is called the signal subspace. Also note that an extended Toeplitz matrix  $\mathbf{T}$ , which is built from  $\tau_m$ ,  $m = 0, 1, \dots, 2M$  and is of size  $(2M+1-L) \times (L+1)$  with  $L \geq K$ , is also of rank  $K$  and there are  $L-K+1$  independent vectors  $\{h_i\}_{i=0}^L$  which can annihilate  $\tau_m$ . This is shown in matrix form as follows:

$$\mathbf{T} \mathbf{h}_{L+1} = \mathbf{0}. \quad (5)$$

Now we assume the measurements  $y_n$  are corrupted with additive noise, and we have access to

$$\tilde{y}_n = y_n + \epsilon_n, \quad n = 0, 1, \dots, N-1, \quad (6)$$

where  $\epsilon_n$  are i.i.d Gaussian random variables with zero mean and standard deviation  $\sigma_y$ . The moments  $\tau_m$  become noisy as shown below:

$$\tilde{\tau}_m = \sum_n c_{m,n} \tilde{y}_n = \tau_m + \underbrace{\sum_n c_{m,n} \epsilon_n}_{b_m}, \quad m = 0, 1, \dots, 2M. \quad (7)$$

The noisy  $\mathbf{T}$ , denoted by  $\tilde{\mathbf{T}}$ , is now full rank and (5) is not satisfied any more. We can look for its total least square (TLS) solution that can be found by performing singular value decomposition (SVD) on  $\tilde{\mathbf{T}}$  with  $L = K$ . The eigenvector corresponding to the smallest eigenvalue is the desired estimation of  $\mathbf{h}$ . This approach can be further improved by an initial denoising (model matching) step, called Cadzow iterative algorithm [7], [2]. The main idea of Cadzow algorithm is to recover the structured low-rank matrix from the full-rank matrix  $\tilde{\mathbf{T}}$ . First, we consider the Toeplitz matrix  $\tilde{\mathbf{T}}$  with  $L = M$  for effective denoising and perform a SVD on  $\tilde{\mathbf{T}}$ :  $\tilde{\mathbf{T}} = \mathbf{U} \mathbf{S} \mathbf{V}^*$ . Then we truncate it to rank- $K$  approximation  $\mathbf{T}'$  by forcing to zero the  $M+1-K$  smallest singular values. Now  $\mathbf{T}'$  is no longer Toeplitz but its best Toeplitz approximation can be obtained by averaging the diagonals of  $\mathbf{T}'$ . A few these iterations lead to a denoised set of moments.

Cadzow denoising works under the assumption that  $\mathbf{T}$  is well approximated by the  $K$  signal-subspace singular vectors of  $\tilde{\mathbf{T}}$  and that discarding information in the orthogonal-subspace of  $\tilde{\mathbf{T}}$  removes certain amount of noise. However, when the noise on  $\tau_m$  is very high we may experience a subspace swap. Specifically, in the high level of noise  $\tilde{\mathbf{T}}$  is better approximated by some orthogonal-subspace singular vectors rather than by the weakest signal-subspace singular vector. In this case, the truncation of the  $M+1-K$  smallest singular values of  $\tilde{\mathbf{T}}$  in the first iteration of Cadzow will lead to an unrecoverable removal of information which belongs to the signal-subspace. Consequently we are not able to retrieve the locations by TLS and it is widely recognised that this is when performance breakdown is observed. Moreover, we conjecture that as long as the subspace swap does not happen in the first iteration, the following iterations would separate the subspaces correctly. Hence the analysis on the SVD of the original noisy  $\tilde{\mathbf{T}}$  is sufficient for predicting when subspace swap event happens and is given in Section 3.

### III. NECESSARY CONDITION FOR SUBSPACE SWAP EVENT

In this section we are going to work out at which noise level the orthogonal-subspace singular vectors substitute the position of the weakest signal-space singular vector in SVD of the noisy data matrix  $\tilde{\mathbf{T}}$  with  $L = M$ .

We first look at the noiseless  $\mathbf{T} \in \mathbb{C}^{(M+1) \times (M+1)}$  and we rewrite it as follows:

$$\mathbf{T} = (M+1)\mathbf{G} \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_K \end{bmatrix} \mathbf{G}^*, \quad (8)$$

where

$$\mathbf{G} = \frac{1}{\sqrt{M+1}} [\mathbf{g}(t_0) \quad \dots \quad \mathbf{g}(t_{K-1})] \quad (9)$$

and

$$\mathbf{g}(t) = [e^{\beta t/T} \quad e^{2\beta t/T} \quad \dots \quad e^{(M+1)\beta t/T}]^T. \quad (10)$$

$\mathbf{G}$  is Vandermonde and has full rank  $K$  since the locations  $t_k$  are distinct, hence in the noiseless case  $\mathbf{T}$  has rank  $K$ . SVD of  $\mathbf{T}$  can be written as  $\mathbf{T} = \mathbf{U}\Sigma_s\mathbf{V}^*$ , where  $\Sigma_s = \text{diag}(s_1, s_2, \dots, s_K, \underbrace{0, \dots, 0}_{M+1-K})$ ,  $s_1 \geq s_2 \geq \dots \geq s_K > 0$ . The unitary matrix  $\mathbf{U}$  has size  $(M+1) \times (M+1)$  and can be partitioned into  $\mathbf{U}_s = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_K)$  and  $\mathbf{U}_0 = (\mathbf{u}_{K+1}, \mathbf{u}_{K+2}, \dots, \mathbf{u}_{M+1})$ . The matrix  $\mathbf{U}_s$  generates the signal-subspace and the orthogonal-subspace covered by  $\mathbf{U}_0$  completes the space. The subspace  $\mathbf{U}_0$  is used to identify the locations of Diracs as discussed in the previous section.

When there is noise, the new moments  $\tilde{\tau}_m$  are noisy as described in (7), which can also be written in matrix form as follows:

$$\begin{aligned} \tilde{\boldsymbol{\tau}} &= \mathbf{C}\mathbf{y} + \mathbf{C}\boldsymbol{\epsilon} \\ &= \boldsymbol{\tau} + \mathbf{b}, \end{aligned} \quad (11)$$

where  $\mathbf{C}$  is composed of entries  $c_{m,n} = c_{m,0} e^{(m-M)\beta n}$  at position  $(m, n)$ . Note that the statistics of the noise  $\mathbf{b}$  depends directly on the distribution of the sample noise  $\boldsymbol{\epsilon}$  and on the coefficients  $\mathbf{C}$ . A stable matrix  $\mathbf{C}$  is therefore desired. It is known that the most stable  $\mathbf{C}$  has condition number one and this can be achieved by choosing  $\beta = j\frac{2\pi}{N}$  and  $|c_{m,0}| = 1/\sqrt{N}$  for  $m = 0, 1, \dots, 2M$  [4]. In our context we use the exponential reproducing kernel which leads to this specific  $\mathbf{C}$ . More details on how to design this kernel are in [4]. Since by construction  $\mathbf{C}^*\mathbf{C} = \mathbf{I}$  we have that the covariance matrix of the moment noise equals that of the sample noise:  $\mathbf{R}_b = \mathbf{R}_\epsilon = \sigma_y^2 \mathbf{I}$ . Note that for  $\beta$  other than  $j\frac{2\pi}{N}$ ,  $\mathbf{R}_b$  is not exactly diagonal, but we have experimental evidence that the derivation in the following part still approximately applies.

The matrix  $\mathbf{T}$  becomes  $\tilde{\mathbf{T}} = \mathbf{T} + \mathbf{N}$ , where  $\mathbf{N} \in \mathbb{C}^{(M+1) \times (M+1)}$  is a Toeplitz matrix built with  $\{b_m\}_{m=0}^{2M}$  in (7).  $\tilde{\mathbf{T}}$  has full rank and will not lie entirely within the signal-subspace. When noise is strong,  $\tilde{\mathbf{T}}$  will lie far from the signal-subspace and at certain point the subspace swap happens.

Now we describe a worst subspace swap scenario and show that this gives us the necessary condition for a swap event to happen. When this condition is not satisfied, absence of subspace-swap is guaranteed and the standard deviation of the retrieved locations can be predicted by Cramér-Rao bound.

We denote the singular values of  $\mathbf{N}$  with  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{M+1}$ . The worst situation is when the strongest noise component with strength  $\lambda_1$  is aligned with one of the elements in  $\mathbf{U}_0$ , and the second strongest noise component with strength  $\lambda_2$  is aligned with the weakest signal component with strength  $s_K$  and sums destructively with it making it become even weaker ( $s_K - \lambda_2$ ). Hence subspace swap happens when

$$\lambda_1 > s_K - \lambda_2. \quad (12)$$

This is a *necessary condition* since any other swap event requires stronger conditions on the amplitude of noise singular values.

It is known that the maximum singular value of a  $n \times n$  random symmetric Toeplitz matrix whose entries come from a sequence of i.i.d random Gaussian variables with variance  $\sigma^2$  is smaller than  $\sqrt{2\sigma^2 n \ln n}$  [17]. Given this, we can relate the maximum singular value of  $\mathbf{N}$  to the order  $2M$  of the sampling kernel and the noise level as follows:

$$\lambda_1 < \sqrt{2\sigma_y^2(M+1) \ln(M+1)}. \quad (13)$$

Now the least favourable scenario is when both  $\lambda_1$  and  $\lambda_2$  equals the maximum:  $\lambda_2 = \lambda_1 = \sqrt{2\sigma_y^2(M+1) \ln(M+1)}$  which yields the following necessary condition for a subspace swap:

$$\begin{aligned} \lambda_1 > s_K - \lambda_2 &\iff \lambda_1 > s_K/2 \\ &\iff \sqrt{2\sigma_y^2(M+1) \ln(M+1)} > s_K/2. \end{aligned} \quad (14)$$

This can be further derived to:

$$\sigma_y^2 > \frac{s_K^2}{8(M+1) \ln(M+1)}. \quad (15)$$

**K = 1** : When we have 1 Dirac with amplitude  $a$ , then  $s_K = s_1 = |a|(M+1)$ . We define PSNR =  $10 \log_{10} \frac{a^2}{\sigma_y^2}$ . Hence we can relate breakdown PSNR to the order  $P = 2M$  of the kernel as follows:

$$\text{PSNR} < 10 \log_{10} \frac{8 \ln(M+1)}{(M+1)}. \quad (16)$$

The condition in (16) is shown in Fig. 3 by solid curve. Note that (16) is a necessary condition for breakdown, hence breakdown may happen for settings in the area below the blue curve and it is *guaranteed* no-breakdown will happen for the area above the curve.

**K = 2** : When there are 2 Diracs and we assume both Diracs are with same amplitude  $a$ , (8) now can be written as  $\mathbf{T} = a(M+1)\mathbf{G}\mathbf{G}^*$ . Its singular values  $s_1$  and  $s_2$ , i.e. the non-zero eigenvalues of  $\mathbf{T}$ , can be derived easily by computing the eigenvalues of the two-by-two matrix  $a(M+1)\mathbf{G}^*\mathbf{G}$ . The expressions are as follows:

$$\begin{aligned} s_1 &= |a| (M+1 + |\langle \mathbf{g}(t_0), \mathbf{g}(t_1) \rangle|) \\ \text{and } s_2 &= |a| (M+1 - |\langle \mathbf{g}(t_0), \mathbf{g}(t_1) \rangle|). \end{aligned} \quad (17)$$

Further manipulations gives

$$\begin{aligned} s_2 &= |a| (M+1 - |\langle \mathbf{g}(t_0), \mathbf{g}(t_1) \rangle|) \\ &= |a| \left( M+1 - \left| \frac{\sin(\frac{\beta}{2}(M+1)\Delta t/T)}{\sin(\frac{\beta}{2}\Delta t/T)} \right| \right), \end{aligned} \quad (18)$$

where  $\Delta t = t_1 - t_0$ . By substituting (18) to (15) and defining PSNR with  $10 \log_{10} \frac{\sigma_s^2}{\sigma_y^2}$ , we can relate the breakdown PSNR to the sampling period  $T$ , the distance  $\Delta t$  between the two Diracs, the frequency interval  $\beta$  and the order of the kernel  $P = 2M$ :

$$\text{PSNR} < 10 \log_{10} \frac{8(M+1) \ln(M+1)}{\left(M+1 - \left| \frac{\sin(\frac{\beta}{2}(M+1)\Delta t/T)}{\sin(\frac{\beta}{2}\Delta t/T)} \right| \right)^2}. \quad (19)$$

The threshold PSNR in (19) for different number of samples  $N$  is illustrated by blue curves in Fig. 4, where we use the fact that we set  $T = 1/N$ .

We can observe that when the relative distance  $\Delta t/T$  increases, the level of noise we can handle increases too. Once the PSNR has reached a minimum, it then oscillates near the minimum PSNR. Interestingly, when we use our favourable setting  $P+1 = N$ , then the first local minimum is exactly at  $\Delta t = 2/(N+1)$ . This implies that if we guarantee  $N+1$  to be larger than  $2/\Delta t$  we can achieve in general the most robust estimation result in the sense of low breakdown PSNR.

When there are more than  $K = 2$  Diracs, the breakdown point given by subspace swap condition for the case of  $K = 2$  still approximately applies when picking the two closest Diracs.

#### IV. SIMULATIONS

In this section, we verify our proposed no-breakdown condition by comparing it to the empirical breakdown points obtained by the method as in Fig. 1. We show that the estimation algorithm<sup>1</sup> breaks down at a PSNR just below our predicted no-breakdown PSNR, confirming the reliability of our derivation. For simplicity, we only show the verification result when we use the most robust sampling kernel as mentioned in Section III, i.e. a kernel of order  $P$  which reproduce  $P+1 = N$  exponentials  $\{e^{m'\beta t/T}\}_{m'=-M}^M$  with  $\beta = j\frac{2\pi}{N}$  and its exponential reproducing coefficient  $|c_{m,0}| = 1/\sqrt{N}$ ,  $m = 0, 1, \dots, 2M$ . We emphasize that for other settings our calculated necessary condition for subspace swap (16) and (19) still accurately anticipate the breakdown event.

##### A. $K = 1$

The guaranteed no-breakdown PSNR (16) for different kernel order  $P = 2M$  is shown by the blue curve in Fig. 3. The measured breakpoints shown by red markers in general agree with our proposed condition.

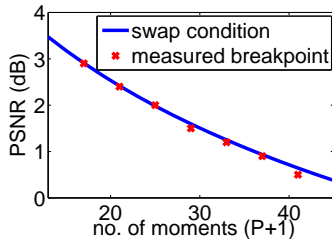


Fig. 3. Subspace swap condition compared to the empirical breakdown points for  $K = 1$  with the settings  $N = P + 1$ ,  $T = 1/N$  and  $\beta = j2\pi/N$ .

<sup>1</sup>Note that Cadzow iterative algorithm and matrix pencil lead to very similar performances.

##### B. $K = 2$

The guaranteed no-breakdown PSNR (19) for different number of samples  $N$  and Diracs interval  $\Delta t$  is shown by the blue curves in Fig. 4. We can see that the observed breakdown points, which are shown by red markers, are well predicted by our proposed condition.

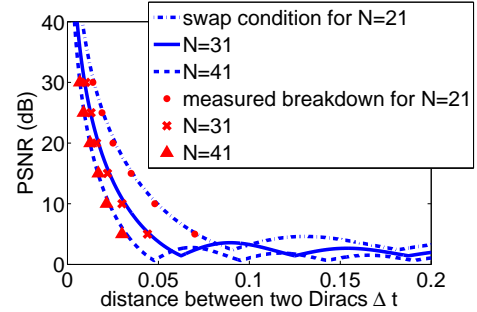


Fig. 4. Subspace swap condition compared to the empirical breakdown points for  $K = 2$  with the settings  $N = P + 1$ ,  $T = 1/N$  and  $\beta = j2\pi/N$ .

##### C. $K > 2$

In this simulation, we show that when there are more than  $K = 2$  Diracs, the breakpoint given by subspace swap condition for  $K = 2$  still approximately applies. For example, from Fig. 4 we see that the smallest possible breakdown PSNR for  $K = 2$  and  $N = 41$  is roughly 0dB and is achieved when  $\Delta t \geq 2/(N+1)$ . We now demonstrate that the guideline is reliable even for  $K > 2$ . We show in Fig. 5 that we accurately reconstruct 17 Diracs from  $N = 41$  samples in noise of PSNR = 0dB, where the distance between every two Diracs is greater than  $2/(N+1)$ .

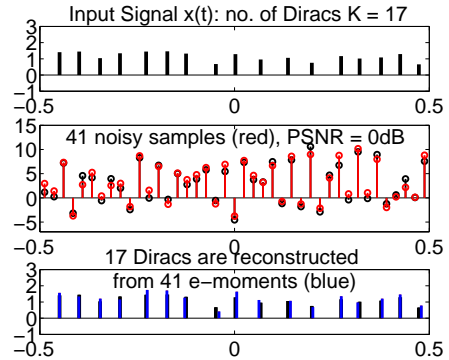


Fig. 5. Reconstruction of 17 Diracs with minimum Dirac separation  $\Delta t \geq 2/(N+1)$  from 41 samples with PSNR = 0dB.

#### V. CONCLUSION

In this paper, we have studied when noisy FRI estimation algorithms break down using the subspace swap criterion. We have then derived the breakdown region and related it to the sampling rate, the minimum distance between two Diracs and PSNR. The reliability of our predicted breakdown region has then been confirmed by simulation results.

## REFERENCES

- [1] M. Vetterli, P. Marziliano, and T. Blu, "Sampling Signals with Finite Rate of Innovation," *IEEE Transactions on Signal Processing*, 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," *Signal Processing Magazine, IEEE*, 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 Transactions on Signal Processing*, vol. 55, no. 5, Part 1, pp. 1741–1757, 2007.
- [4] J. A. Urigüen, T. Blu, and P. L. Dragotti, "FRI Sampling with Arbitrary Kernels," *IEEE Transactions on Signal Processing*, vol. 61, no. 21, pp. 5310–5323, 2013.
- [5] R. Tur, Y. C. Eldar, and Z. Friedman, "Innovation Rate Sampling of Pulse Streams with Application to Ultrasound Imaging," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1827–1842, 2011.
- [6] P. L. Dragotti, M. Vetterli, and T. Blu, "Exact Sampling Results for Signals with Finite Rate of Innovation Using Strang-Fix Conditions and Local Kernels," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05)*. IEEE, 2005, vol. 4, pp. 233–236.
- [7] J. A. Cadzow, "Signal Enhancement-A Composite Property Mapping Algorithm," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 36, no. 1, pp. 49–62, 1988.
- [8] Y. Hua and T. K. Sarkar, "Matrix Pencil Method for Estimating Parameters of Exponentially Damped/Undamped Sinusoids in Noise," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, no. 5, pp. 814–824, 1990.
- [9] P. L. Dragotti and F. Homann, "Sampling Signals with Finite Rate of Innovation in the Presence of Noise," in *IEEE International Conference on Acoustics, Speech and Signal Processing, 2009. ICASSP 2009*. IEEE, 2009, pp. 2941–2944.
- [10] J. K. Thomas, L. L. Scharf, and D. W. Tufts, "The Probability of a Subspace Swap in the SVD," *IEEE Transactions on Signal Processing*, vol. 43, no. 3, pp. 730–736, 1995.
- [11] D. W. Tufts, A. C. Kot, and R. J. Vaccaro, "The Threshold Analysis of SVD-based Algorithms," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, 1988.*, Apr 1988, pp. 2416–2419 vol.4.
- [12] M. Hawkes, A. Nehorai, and P. Stoica, "Performance Breakdown of Subspace-based Methods: Prediction and Cure," in *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2001. (ICASSP'01)*. IEEE, 2001, vol. 6, pp. 4005–4008.
- [13] R. McAulay and E. M. Hofstetter, "Barankin Bounds on Parameter Estimation," *IEEE Transactions on Information Theory*, vol. 17, no. 6, pp. 669–676, Nov 1971.
- [14] A. Pinkus and J. Tabrikian, "Barankin Bound for Range and Doppler Estimation Using Orthogonal Signal Transmission," in *2006 IEEE Conference on Radar*, April 2006.
- [15] L. Knockaert, "The Barankin Bound and Threshold Behavior in Frequency Estimation," *IEEE Transactions on Signal Processing*, vol. 45, no. 9, pp. 2398–2401, 1997.
- [16] Y. Barbotin, *Parametric Estimation of Sparse Channels: Theory and Applications*, Ph.D. thesis, École Polytechnique Fédérale de Lausanne, 2013.
- [17] M. W. Meckes, "On the Spectral Norm of a Random Toeplitz Matrix," *Electronic Communications in Probability*, vol. 12, pp. 315–325, 2007.