

FINITE DIMENSIONAL FRI

Jon Oñativia[†], *Yue M. Lu*^{*} and *Pier Luigi Dragotti*[†]

[†] Communications and Signal Processing Group (CSP), Imperial College London, UK
{ jon.onativia, p.dragotti } @imperial.ac.uk

^{*} Signals, Information, and Networks Group (SING), Harvard University, USA
yuelu@seas.harvard.edu

ABSTRACT

Traditional Finite Rate of Innovation (FRI) theory has considered the problem of sampling continuous-time signals. This framework can be naturally extended to the case where the input is a discrete-time signal. Here we present a novel approach which uses both the traditional FRI sampling scheme, based on the annihilating filter method, and the fact that in this new setup the null space of the problem to be solved is finite dimensional.

In the noiseless scenario, we show that this new approach is able to perfectly recover the original signal at the critical sampling rate. We also present simulation results in the noisy scenario where this new approach improves performances in terms of the mean squared error (MSE) of the reconstructed signal when compared to the canonical FRI algorithms and compressed sensing (CS).

Index Terms— Finite rate of innovation, sampling theory, sparsity, annihilating filter.

1. INTRODUCTION

FRI sampling theory [1–4] has shown that it is possible to sample and reconstruct classes of non-bandlimited signals. Streams of Diracs are the canonical example of FRI signals. They are not bandlimited but are sparse in time and have a finite number of degrees of freedom per unit of time, which is known as the rate of innovation. The final goal of FRI methods is to retrieve the exact location and amplitude of the Diracs from a set of samples. In the continuous-time setup, the signal is filtered and then sampled in order to obtain a discrete sequence. In [1] the signal is filtered using the sinc kernel. Authors in [5] present the use of polynomial or exponential reproducing kernels with the advantage of achieving perfect reconstruction with compact support kernels. This framework has recently been extended to arbitrary sampling kernels with the penalty of not achieving perfect reconstruction [6]. To some extent, these methods are all based on the

fact that the Fourier transform of a sum of Diracs is given by a sum of exponentials. The reconstruction is then based on estimating exponentials from a sequence of samples, which is a classical problem in spectral estimation [7, 8].

This framework can be naturally extended to discrete-time signals. In this case, the sampling process can be modelled with a matrix multiplication. The input signal is given by a high dimensional vector with few non-zero elements. The acquired signal is a vector of lower dimension which is given by the product of a fat matrix with the input signal. Note that since the acquisition matrix is fat, the dimension of its null space is strictly positive. The goal is to reconstruct the sparse input vector from the acquired samples. Some preliminary work has already been published where the FRI framework is applied in the discrete-time setup [9] and is compared to ℓ_1 -minimization techniques, which is the traditional reconstruction approach in the compressed sensing (CS) framework [10, 11]. In this paper we present a novel method that is based on the annihilating filter that is used in the traditional FRI framework, but we take advantage of the fact that in this new context the null space has finite dimension. The annihilating method requires a root finding step that becomes unstable in high noise scenarios. This new approach avoids this root finding step. We show simulation results where the new finite dimensional FRI method outperforms traditional FRI and CS.

In Section 2 we describe the annihilating filter method which is inherently valid for the continuous-time and the discrete-time cases. If the input signal is discrete, the solution is mapped to a discrete time grid. In Section 3 we present our new approach, we first establish the uniqueness of the solution in the noiseless case, we then present an extension of the algorithm when noise is present. In Section 4 we present simulation results and then conclude in Section 5.

2. TRADITIONAL FRI IN DISCRETE-TIME

Let $\mathbf{x} \in \mathbb{C}^N$ be a discrete time signal formed by a stream of K Diracs. The expression in time is given by

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

$$x[n] = \sum_{k=1}^K a_k \delta[n - n_k], \quad n = 0, 1, \dots, N-1, \quad (1)$$

where $n_k \in [0, N-1]$ and $a_k \in \mathbb{C} \setminus \{0\}$, for $1 \leq k \leq K$, are unknown integer delays and complex valued amplitudes, respectively. We assume that all delays are distinct and therefore the signal \mathbf{x} has $2K$ degrees of freedom. Let us also assume that we have access to $M < N$ coefficients of the unitary discrete Fourier transform (DFT) of \mathbf{x} . We can express the sampling process in matricial form as follows

$$\mathbf{y} = \mathbf{D} \mathbf{x}, \quad (2)$$

where $\mathbf{y} \in \mathbb{C}^M$ are the available samples and $\mathbf{D} \in \mathbb{C}^{M \times N}$ is a partial Fourier matrix: $(\mathbf{D})_{m,n} = \exp(-j2\pi mn/N) / \sqrt{N}$, with $m = 0, \dots, M-1$ and $n = 0, \dots, N-1$.

The unitary DFT of \mathbf{x} , denoted by $\hat{\mathbf{x}} \in \mathbb{C}^N$, consists of the sum of K exponentials

$$\hat{x}[m] = \frac{1}{\sqrt{N}} \sum_{k=1}^K a_k e^{-j\omega_k m}, \quad m = 0, 1, \dots, N-1, \quad (3)$$

where $\omega_k = \frac{2\pi}{N} n_k$. The annihilating filter $h[m]$ is a $K+1$ taps filter such that $(\hat{\mathbf{x}} * h)[m] = 0$. The z -transform of $h[m]$ is given by $H(z) = \sum_{m=0}^K h[m] z^{-m} = \prod_{k=1}^K (1 - u_k z^{-1})$, where $u_k = e^{-j\omega_k}$. $H(z)$ is a polynomial of order K in the complex variable z with roots at $z = u_k$, that is, $H(z)|_{z=u_k} = 0$. Note that the K roots are distinct and non-zero, we thus always have that $h[0] \neq 0$ and $h[K] \neq 0$. It can easily be shown that the sequence $\hat{x}[m]$ is annihilated by this filter:

$$\begin{aligned} (\hat{\mathbf{x}} * h)[m] &= \sum_{l=0}^K \hat{x}[m-l] h[l] \\ &= \frac{1}{\sqrt{N}} \sum_{l=0}^K \sum_{k=1}^K a_k u_k^{m-l} h[l] \\ &= \frac{1}{\sqrt{N}} \sum_{k=1}^K a_k u_k^m \underbrace{\left(\sum_{l=0}^K h[l] u_k^{-l} \right)}_{H(z)|_{z=u_k}} = 0. \end{aligned} \quad (4)$$

The annihilating filter method consists of finding the filter coefficients $(h[m])_{m=0}^K$, and estimating the frequencies ω_k from the roots of the polynomial $H(z)$. The filter can be obtained by solving the following linear system

$$\begin{bmatrix} \hat{x}[K] & \hat{x}[K-1] & \dots & \hat{x}[0] \\ \hat{x}[K+1] & \hat{x}[K] & \dots & \hat{x}[1] \\ \vdots & \vdots & \ddots & \vdots \\ \hat{x}[2K] & \hat{x}[2K-1] & \dots & \hat{x}[K] \end{bmatrix} \begin{bmatrix} h[0] \\ h[1] \\ \vdots \\ h[K] \end{bmatrix} = 0 \quad (5)$$

It can be shown that if the coefficients $\hat{x}[k]$ satisfy (3), the Toeplitz matrix in (5) has rank K [12]. Thus, the system has a unique solution up to an amplitude factor. If we impose $h[0] = 1$ we can drop a row of the system and find the unique solution from only $2K$ consecutive values of $\hat{x}[k]$. From the knowledge of ω_k , and since coefficients $\hat{x}[k]$ are linear in a_k , the amplitudes a_k are computed from K samples of (3).

This approach is able to perfectly reconstruct the K -sparse signal \mathbf{x} from only $2K$ coefficients of the DFT of \mathbf{x} , but requires a nonlinear step to find the roots of the annihilating filter. These can be obtained from the eigenvalues of the companion matrix. There exist an alternative approach [12] that do not involve root finding and is based on the matrix pencil method [13]. It is in essence based on the same principle that is used in the ESPRIT algorithm [14] for the estimation of directions of arrival of signals in arrays of antennas. This approach does not explicitly find the roots of a polynomial, but involves solving a generalized eigenvector problem and presents similar performances.

In what follows we present a new linear approach for FRI signal reconstruction. The proposed method is also based on the annihilating filter method, but it takes advantage of the fact that the underlying vector \mathbf{x} is finite dimensional.

3. FINITE DIMENSIONAL FRI: NEW APPROACH

In the previous section we have shown that we can compute the annihilating filter coefficients if the number of available samples M is greater than or equal to $2K$. In this section we assume that the coefficients of the annihilating filter have already been computed, but we want to avoid the root finding step. We now present an algebraic approach to reconstructing the K -sparse vector \mathbf{x} .

3.1. Perfect reconstruction in the noiseless scenario

Equation (2), where \mathbf{x} is unknown, is an underdetermined system and therefore the solution is not unique. Among all the possible solutions we want to find the one that is K -sparse.

Proposition 1. *Let $\mathbf{x} \in \mathbb{C}^N$ and $\mathbf{y} \in \mathbb{C}^M$ be given as in (1) and (2) respectively. If $M \geq 2K$ the solution to (2) is unique and can be found by solving two linear inversions.*

Proof. Since $M \geq 2K$, we estimate \mathbf{h} from (5) but now we do not compute its roots. We have that \mathbf{D} is of full rank M , since it is obtained from the first M rows of the Fourier matrix \mathbf{F}_N . Using the pseudoinverse \mathbf{D}^\dagger , we can retrieve \mathbf{x} up to the null space of \mathbf{D} . Note that since \mathbf{F}_N is an orthogonal matrix, that is $\mathbf{F}_N \mathbf{F}_N^H = \mathbf{F}_N^H \mathbf{F}_N = \mathbf{I}_M$, the pseudoinverse of \mathbf{D} is directly \mathbf{D}^H , where the superscript H denotes the Hermitian transpose of a matrix. We can therefore write

$$\mathbf{x} = \mathbf{D}^H \mathbf{y} + \sum_{l=1}^L \beta_l \mathbf{n}_l \quad (6)$$

where β_l are unknown coefficients, $L = N - M$ is the size of the null space and \mathbf{n}_l are L orthonormal vectors that span the null space of \mathbf{D} , that is $\mathbf{n}_l \in N(\mathbf{D})$ where $N(\mathbf{D}) = \{\mathbf{n} \in \mathbb{C}^M \mid \mathbf{D}\mathbf{n} = \mathbf{0}\}$. The natural choice for \mathbf{n}_l , $l = 1, \dots, L$, is to pick the last $N - M$ columns of \mathbf{F}_N^H . If we premultiply equation (6) by \mathbf{F}_N we obtain the Fourier transform of \mathbf{x}

$$\hat{\mathbf{x}} = \mathbf{F}_N \mathbf{x} = \mathbf{F}_N \mathbf{D}^H \mathbf{y} + \sum_{l=1}^L \beta_l \mathbf{F}_N \mathbf{n}_l. \quad (7)$$

It can easily be verified that $\mathbf{F}_N \mathbf{D}^H = \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{L,M} \end{bmatrix}$ and $\mathbf{F}_N \mathbf{n}_l = \mathbf{e}_{M+l}$, where \mathbf{e}_i is the i -th vector of the canonical basis. Note that coefficients β_l are exactly the missing coefficients of the Fourier transform of \mathbf{x} . Let define $\mathbf{z} = \mathbf{F}_N \mathbf{D}^H \mathbf{y}$.

Let $\mathbf{S} \in \mathbb{C}^{(N-K) \times (K+1)}$ be the Toeplitz matrix where the first row is given by coefficients $\hat{x}[K], \hat{x}[K-1], \dots, \hat{x}[0]$ and the first column by $\hat{x}[K], \hat{x}[K+1], \dots, \hat{x}[N-1]$. Note that \mathbf{S} is fully specified by vector $\hat{\mathbf{x}}$. In what follows, we denote by $\text{Toe}_K \{\cdot\}$ the operation of building the $(N-K) \times (K+1)$ Toeplitz matrix from a $N \times 1$ vector in the way we define matrix \mathbf{S} . Since the N coefficients $\hat{x}[k]$ satisfy (3), the annihilating filter's system from (5) is also satisfied:

$$\mathbf{S} \mathbf{h} = \mathbf{0}. \quad (8)$$

If we replace the Fourier transform of \mathbf{x} with the expression obtained in (7), matrix \mathbf{S} can be expressed as

$$\mathbf{S} = \mathbf{Z} + \sum_{l=1}^L \beta_l \mathbf{E}_{M+l}, \quad (9)$$

where $\mathbf{Z} = \text{Toe}_K \{\mathbf{z}\}$ and $\mathbf{E}_i = \text{Toe}_K \{\mathbf{e}_i\}$. We then replace (9) in (8), and given that \mathbf{h} is known, we can turn the system of equations in a new problem with L unknowns:

$$\begin{bmatrix} \mathbf{E}_{M+1} \mathbf{h} & \dots & \mathbf{E}_N \mathbf{h} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_L \end{bmatrix} = -\mathbf{Z} \mathbf{h}. \quad (10)$$

The first $M - K$ rows of this system are trivially equal to zero because matrices \mathbf{E}_{M+l} are all zero in these rows and rows of $\mathbf{Z} \mathbf{h}$ are also equal to zero because they contain the convolution of exactly $K + 1$ consecutive coefficients of $\hat{\mathbf{x}}$ with the annihilating filter. Let define \mathbf{A} as the remaining $N - M$ rows of matrix $\begin{bmatrix} \mathbf{E}_{M+1} \mathbf{h} & \dots & \mathbf{E}_N \mathbf{h} \end{bmatrix}$. Matrix \mathbf{A} is of size $L \times L$ and is given by

$$\mathbf{A} = \begin{bmatrix} h[0] & 0 & \dots & \dots & \dots & 0 & 0 \\ h[1] & h[0] & \dots & \dots & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & & & \vdots & \vdots \\ h[K] & \dots & h[1] & h[0] & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & \dots & h[K] & \dots & h[1] & h[0] \end{bmatrix}. \quad (11)$$

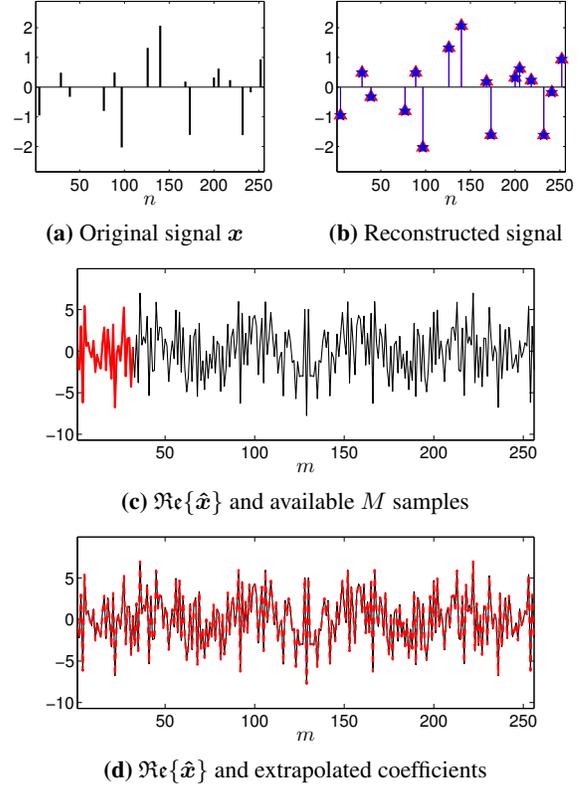


Fig. 1: $N = 256$, $M = 32$ and $K = 16$. (a) Signal with $K = 16$ Diracs. (b) Perfect reconstruction of the signal, in red the original signal and in blue the reconstruction. (c) Real part of the Fourier coefficients in black and available samples in red. (d) Real part of the Fourier coefficients in black and extrapolated coefficients using the method in Proposition 1, in red.

This is a lower triangular matrix where all the elements in the main diagonal are non-zero. Moreover, we have imposed $h[0] = 1$ when computing the annihilating filter, thus, the elements of the main diagonal are all equal to one. Matrix \mathbf{A} has therefore rank $N - M$ and the system has a unique solution. \square

Remark 1. Proposition (1) shows that it is possible to perfectly reconstruct the K -sparse signal \mathbf{x} from the critical number of samples $M = 2K$ without having to compute explicitly the roots of the annihilating filter.

Remark 2. Samples in \mathbf{y} correspond to the lower frequencies of $\hat{\mathbf{x}}$, but this can be generalized to any $M = 2K$ consecutive elements of $\hat{\mathbf{x}}$. Moreover, due to the periodicity of the DFT the samples only need to be consecutive modulo (N) .

Figure 1 illustrates the perfect reconstruction of a sparse signal with $K = 16$ Diracs. The number of available samples is equal to the number of degrees of freedom of the signal:

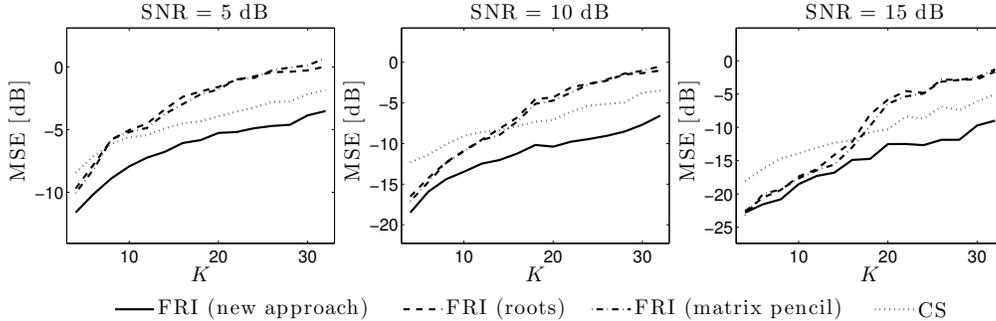


Fig. 2: Simulation results showing new finite dimensional FRI approach outperforming traditional FRI methods (root finding of the annihilating filter or matrix pencil) and compressed sensing reconstruction. The dimensionality of the input vector is $N = 256$ and the number of available samples is $M = 64$. Simulations performed at different levels of noise (SNR of 5, 10 and 15 dB) and different levels of sparsity K (horizontal axis). The vertical axis shows the normalized average value of the MSE of the reconstructed sparse vector compared to the true \mathbf{x} .

$M = 2K = 32$. The $\beta_l, l = 1, \dots, L$ coefficients in the proof of Proposition (1) correspond to the missing coefficients of the Fourier transform of \mathbf{x} and their real parts are illustrated in Figure 1-(d).

3.2. Noisy case

To make the algorithm more robust to noise we have to increase the number of available samples M and apply some denoising algorithm to the samples. In the presence of noise the Toeplitz matrix $\text{Toe}_K\{\mathbf{y}\}$ is full rank instead of being rank deficient with rank K . We first denoise vector \mathbf{y} applying an iterative algorithm that finds the closest Toeplitz matrix of rank K . This iterative algorithm is known as Cadzow denoising [15] and have been successfully applied in the FRI context [2].

The annihilating filter is then estimated computing the total least squares (TLS) solution that minimizes $\|\mathbf{Y}\mathbf{h}\|^2$ subject to $\|\mathbf{h}\| = 1$. In order to reconstruct the sparse vector we have then to estimate the β_l coefficients. In the noiseless case, the first $M - K$ rows of (10) are equal to zero, but this does not hold when noise is present. We thus have an overdetermined system that we also solve by computing the TLS solution. Note that because we have a larger M than the critical value of $2K$, the dimension of the null space $L = N - M$ is reduced, and thus there are less coefficients β_l to be estimated making the overall algorithm stable in noisy scenarios. From the estimated coefficients β_l and the available samples \mathbf{y} , we build vector $\hat{\mathbf{x}}$ with equation (7) and apply again the Cadzow denoising algorithm to $\text{Toe}_K\{\hat{\mathbf{x}}\}$. The sparse vector \mathbf{x} is then obtained by taking the inverse Fourier transform of $\hat{\mathbf{x}}$.

4. SIMULATION RESULTS

Figure 2 shows simulations results where the new finite dimensional FRI approach outperforms traditional FRI methods and CS reconstruction. Two different approaches have

been tested for the traditional FRI setup. One approach is to evaluate the annihilating filter polynomial on a grid of N points and estimate the roots from the local minima of its absolute value. The second approach is to obtain the locations by applying the matrix pencil method. In both cases, the samples have been first denoised applying the Cadzow denoising algorithm. The CS reconstruction is obtained by applying basis pursuit denoising [16]. The solution is given by $\min_{\mathbf{x} \in \mathbb{C}^N} \|\mathbf{x}\|_1 + (1/2\rho) \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2$ and we have used the YALL1 implementation [17] with parameter $\rho = 10^{-3}$. We improve the CS solution by picking the K largest values. For each sparsity level K , 10 different distributions for the locations of the Diracs have been uniformly generated. For each instance of locations the amplitudes of the Diracs are drawn from a Gaussian distribution with parameters $\mathcal{N}(0, 1)$. Complex noise is added to vector \mathbf{y} . The real and imaginary parts are drawn from $\mathcal{N}(0, \sigma)$ where σ is adjusted to satisfy the different SNR levels in the measured samples \mathbf{y} . For each realization of the sparse vector, 100 realizations of noise have been generated. We thus have 1000 realizations for each level of sparsity.

The new finite dimensional FRI method clearly outperforms the other approaches especially in high noise scenarios (SNR = 5 dB) and with large values of K where the root finding approach becomes more unstable.

5. CONCLUSIONS

In this paper we have presented a novel method to reconstruct a finite dimensional sparse vectors from partial knowledge of its discrete Fourier transform. We have shown that in the noiseless scenario perfect reconstruction is achieved with the critical number of samples. In the noisy case, this method is more stable and outperforms traditional FRI approaches and CS because it takes advantage of the fact that the null space of the underdetermined system is finite dimensional.

6. 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, June 2002.
- [2] T. Blu, P. L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse sampling of signal innovations," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 31–40, March 2008.
- [3] J. Berent, P. L. Dragotti, and T. Blu, "Sampling piecewise sinusoidal signals with finite rate of innovation methods," *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 613–625, February 2010.
- [4] 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, April 2011.
- [5] 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, pp. 1741–1757, May 2007.
- [6] 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, November 2013.
- [7] B. D. Rao and K. S. Arun, "Model based processing of signals: A state space approach," *Proceedings of the IEEE*, vol. 80, no. 2, pp. 283–309, February 1992.
- [8] P. Stoica and R. Moses, *Spectral Analysis of Signals*, 1st ed. Prentice Hall, March 2005.
- [9] A. Hormati and M. Vetterli, "Annihilating filter-based decoding in the compressed sensing framework," in *Proceedings SPIE 6701, Wavelets XII*, vol. 6701, no. 670121-670121-10, August 2007.
- [10] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.
- [11] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, February 2006.
- [12] I. Maravić and M. Vetterli, "Sampling and reconstruction of signals with finite rate of innovation in the presence of noise," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2788–2805, August 2005.
- [13] 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, May 1990.
- [14] A. Paulraj, R. H. Roy, and T. Kailath, "Estimation of signal parameters via rotational invariance techniques – ESPRIT," in *19th Asilomar Conference on Circuits, Systems and Computers, 1985*, November 1985, pp. 83–89.
- [15] 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, January 1988.
- [16] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.
- [17] Y. Zhang, J. Yang, and W. Yin, "YALL1: Your Algorithms for L1," online at <http://yall1.blogs.rice.edu>, 2011.