

SPECTRAL COMPRESSIVE SENSING WITH MODEL SELECTION

Zhenqi Lu, Rendong Ying, Sumxin Jiang, Zenghui Zhang, Peilin Liu, Wenxian Yu

Dept. of Electrical Engineering, Shanghai Jiao Tong University, Shanghai, P. R. China

ABSTRACT

The performance of existing approaches to the recovery of frequency-sparse signals from compressed measurements is limited by the coherence of required sparsity dictionaries and the discretization of frequency parameter space. In this paper, we adopt a parametric joint recovery-estimation method based on model selection in spectral compressive sensing. Numerical experiments show that our approach outperforms most state-of-the-art spectral CS recovery approaches in fidelity, tolerance to noise and computation efficiency.

Index Terms— Compressive sensing, frequency-sparse signal, model selection, parametric estimation, maximum likelihood estimator

1. INTRODUCTION

One of the recent research interests of compressive sensing (CS) has focused on the recovery of signals that are spectrally sparse from a reduced number of measurements [1–6]. A great many applications, including spectrum sensing [7] and wideband communication [3, 8], feature smooth or modulated signals that can be modelled as a superposition of a small number of sinusoids. Recovery of such *frequency-sparse signals* brings about a novel issue in the formulation of CS recovery problem: signal representations in frequency domain have a continuous parameter space, while recent CS researches [9–11] are rooted on signal decomposition in a discretized dictionary.

An intuitive solution to this problem is a denser sampling of the parameter space, which improves the compressibility of signal representations. But increasing the resolution of parameter sampling worsens the coherence between dictionary elements, which results in loss of sparsity and uniqueness of signal representations. Such ambiguity prevents certain algorithms [9, 10] from achieving the sparse representation successfully. Initial contributions to spectral CS recovery are concentrated on the recovery algorithm, the optimization problem formulation and the sparsity prior to combat the intricacy in signal representations [5, 12–16].

Estimation of the frequencies, amplitudes and phase shifts of sinusoids embedded in noise is a fundamental problem in

time series analysis and other more general statistical signal processing problems. Classical treatments of this problem are restricted to the Fourier frequencies. This implicit discretization barrier is overcome due to the introduction of the *minimum description length* (MDL) principle for *model selection* [17–19]. In this paper, we improve over existing approaches by applying model selection to spectral CS. We take into consideration *parametric* joint recovery-estimation methods, which determine parameters by minimizing a log-likelihood function of compressed measurements.

The novelty of our approach is that it performs estimation from compressed measurements rather than from signal samples. Since the log-likelihood minimization is in practice equivalent to an ℓ_2 -norm minimization, the estimation performance is guaranteed as long as the sensing matrix satisfies *restricted isometry property* (RIP), and thus distance-preserving [10]. It can be shown that random matrices from Gaussian, Rademacher, or more generally a sub-Gaussian distribution have the RIP with high probability under certain conditions [20]. We solve the optimization problem through an iterative greedy approach to outperform state-of-the-art spectral CS approaches. Experimental results show improved reconstruction fidelity against existing approaches, from both noiseless and noisy measurements. In addition, our approach is essentially greedy, and is thus more computationally efficient than optimization based approaches. Furthermore, compared to traditional model selection estimators [17–19], our approach reduces the number of samples needed and consequently the computational load of estimation.

2. PROBLEM FORMULATION AND RELATED PRIOR WORK

The problem of recovering frequency-sparse signals from compressed measurements is formulated as follows: Let $s(t)$ denote an unidimensional real-valued frequency-sparse signal composed of K sinusoids with unknown frequencies ω_j , amplitudes a_j , and phase shifts ϕ_j , and $x(t)$ be $s(t)$ corrupted by additive noise $\xi(t)$ with unknown noise level.

$$x(t) = s(t) + \xi(t) = \sum_{j=1}^K a_j \sin(\omega_j t + \phi_j) + \xi(t). \quad (1)$$

This work was partially supported by the National Natural Science Foundation of China under grant number 61171171 and 61102169.

Let $\mathbf{s} = \{s_t\}_{t=1}^N$ and $\mathbf{x} = \{x_t\}_{t=1}^N$ be the observed samples of $s(t)$ and $x(t)$ at discrete times $\{t_j = j\}_{j=1}^N$, and $\mathbf{m} = \Phi \mathbf{x} \in \mathbb{R}^M$ be the compressed measurements of \mathbf{x} , where $\Phi \in \mathbb{R}^{M \times N}$ is the sensing matrix.

Given the compressed measurements \mathbf{m} , the problem is to recover the K sinusoids in (1). Many different approaches have been suggested in the literature for spectral CS recovery. Under certain conditions for matrix Φ , one can recover signal \mathbf{s} through an ℓ_1 -norm optimization problem [9, 10], denoted as ℓ_1 -synthesis,

$$\hat{\mathbf{s}} = \mathbf{F} \hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\eta}} = \arg \min_{\boldsymbol{\eta} \in \mathbb{R}^N} \|\boldsymbol{\eta}\|_1 \text{ s.t. } \|\mathbf{m} - \Phi \mathbf{F} \boldsymbol{\eta}\|_2 \leq \epsilon, \quad (2)$$

where ϵ is an appropriately chosen bound on the noise level, \mathbf{F} is the orthonormal DFT basis, and $\boldsymbol{\eta}$ is the DFT coefficients of signal samples \mathbf{s} . The optimal recovery of signal \mathbf{s} by optimizing (2) is feasible provided that the decomposition of \mathbf{s} in the DFT basis \mathbf{F} is K -sparse, i.e. $\|\boldsymbol{\eta}\|_0 = K$ [9, 10]. Unfortunately, not only the DFT coefficients of frequency-sparse signals are *not sparse*, but even worse, they are *just barely compressible*. One way to remedy this problem would be to employ a redundant DFT frame

$$\begin{aligned} \Psi(c) &:= [e(0) \ e(\Delta) \ \cdots \ e(2\pi - \Delta)], \Delta := 2\pi/cN, \\ e(\omega) &:= \frac{1}{\sqrt{N}} \left[1 \ e^{j\omega} \ e^{j2\omega} \ \cdots \ e^{j\omega(N-1)} \right]^T, \end{aligned} \quad (3)$$

as a substitute for \mathbf{F} . But the redundant DFT frame violates the incoherence requirement for the dictionary [5].

It has recently been shown that the incoherence condition of dictionary \mathbf{D} is not necessary concerning the recovery of signal \mathbf{x} , provided that the frame coefficients $\mathbf{D}^H \mathbf{x}$ are sufficiently sparse [12], where $(\cdot)^H$ designates the Hermitian operation. Under this circumstance, ℓ_1 -analysis yields good recovery result for signal \mathbf{x} . However, the redundant DFT frame coefficients of frequency-sparse signals $\Psi(c)^H \mathbf{s}$ do not have the sparsity property.

An alternative approach is to benefit from structured sparsity by using a coherence inhibition signal model [5]. The resulting Structured Iterative Hard Thresholding (SIHT) algorithm is able to recover frequency-sparse signals by selecting elements with low coherence in an redundant DFT frame. Other algorithms with similar flavor to SIHT include Band-excluded Orthogonal Matching Pursuit (BOMP), which takes advantage of band-exclusion [15]. However, the reconstruction fidelity of SIHT and BOMP is substantially limited due to the simplicity in the formulation of algorithms.

One way to remedy the discretization of parameter space is the *polar interpolation* approach, and the corresponding algorithm is named Continuous Basis Pursuit (CBP) [16]. Like other optimization based algorithms, CBP suffers from its high computational complexity. A novel algorithm, Band-excluded Interpolating Subspace Pursuit (BISP), combining the merits of band-exclusion and polar interpolation, has been proposed more recently [14, 21]. By incorporating polar interpolation with greedy algorithm, BISP improves the

convergence rate of CBP while only inducing an amenable reduction in performance.

Recent advances in convex geometry has proved that frequency-sparse signals can be recovered from random subsamples via *atomic norm minimization* [13], which can be implemented as a semidefinite program (SDP) [22]. Though atomic norm has several appealing properties, SDP is in practice computationally expensive. Moreover, the formulation of SDP is limited to random subsampling matrix, and no discussion for arbitrary measurement settings is provided.

3. MODEL SELECTION FOR SPECTRAL COMPRESSIVE SENSING

In this paper, we adopt a parametric joint recovery-estimation method, which estimates the unknown frequencies, amplitudes, and phase shifts. Under the assumption of white Gaussian noise, with the number of sinusoids K *a priori* known, a common method to estimate the $3K$ parameters is by maximizing the likelihood function \mathcal{L} of observed data \mathbf{x} [17]

$$\mathcal{L}(\boldsymbol{\theta}_K, \mathbf{x}) = \prod_{t=1}^N e^{-|x_t - \sum_{j=1}^K a_j \sin(\omega_j t + \phi_j)|^2}, \quad (4)$$

where $\boldsymbol{\theta}_K = \{a_j, \omega_j, \phi_j\}_{j=1}^K$ contains the $3K$ parameters of the K sinusoids. In spectral CS recovery problem, the signal samples \mathbf{x} is not available, and as a substitute the compressed measurements \mathbf{m} is observed. Thus the estimation method is reformulated as the minimization of the log-likelihood function of \mathbf{m}

$$\hat{\boldsymbol{\theta}}_K = \arg \min_{\boldsymbol{\theta}_K} -\ln \mathcal{L}(\boldsymbol{\theta}_K, \mathbf{m}), \quad (5)$$

which is equivalent to an ℓ_2 -norm minimization problem

$$\hat{\boldsymbol{\theta}}_K = \arg \min_{\boldsymbol{\theta}_K} \left\| \mathbf{m} - \sum_{j=1}^K \Phi \mathbf{s}_j \right\|_2^2, \quad (6)$$

where $\mathbf{s}_j = \{a_j \sin(\omega_j t + \phi_j)\}_{t=1}^N$ is the sample vector of K sinusoids with estimated $3K$ parameters. Obviously, (6) can be reformulated as the following ℓ_2 -norm minimization problem

$$\hat{\boldsymbol{\vartheta}}_K = \arg \min_{\boldsymbol{\vartheta}_K} \left\| \mathbf{m} - \sum_{j=1}^K \Phi (a_{1,j} \mathbf{sin}_{\omega_j} + a_{2,j} \mathbf{cos}_{\omega_j}) \right\|_2^2, \quad (7)$$

where $\boldsymbol{\vartheta}_K = \{\omega_j, a_{1,j}, a_{2,j}\}_{j=1}^K$ contains the reformulated $3K$ parameters with $a_{1,j} = a_j \cos \phi_j$, $a_{2,j} = a_j \sin \phi_j$ designating the amplitudes of sine and cosine sinusoids, and $\mathbf{sin}_{\omega_j} = \{\sin n\omega_j\}_{n=1}^N$, $\mathbf{cos}_{\omega_j} = \{\cos n\omega_j\}_{n=1}^N$ are the samples of sinusoids with frequency ω_j .

In recent parametric estimation works, the best K matching sinusoids are iteratively recovered [17]. In this paper,

we reformulate previous methods as a spectral CS recovery method as shown in Algorithm 1, which deviates from previous approaches in that the input is compressed measurements rather than signal samples. In each iteration, the estimated compressed measurements of other $K - 1$ sinusoids are trimmed from input compressed measurements \mathbf{m} , and then the parameters of the best matching sinusoid to the residual measurements \mathbf{r} are estimated through function $\mathcal{R}(\Phi, \mathbf{r})$.

The function $\{\hat{\omega}, \hat{a}, \hat{\phi}\} = \mathcal{R}(\Phi, \mathbf{r})$, as shown in Algorithm 2, estimates the parameters $\hat{\theta} = \{\hat{\omega}, \hat{a}, \hat{\phi}\}$ of the best matching sinusoid from residual CS measurement vector \mathbf{r} . The algorithm solves the parametric estimation by minimizing the following log-likelihood function

$$\hat{\theta} = \arg \min_{\theta} -\ln \mathcal{L}(\theta, \mathbf{r}), \quad (8)$$

which is equivalent to an ℓ_2 -norm minimization problem

$$\hat{\vartheta} = \arg \min_{\vartheta} \|\mathbf{r} - \Phi(a_1 \sin_{\omega} + a_2 \cos_{\omega})\|_2^2, \quad (9)$$

where $\vartheta = \{\omega, a_1, a_2\}$ contains the reformulated sinusoid parameters. (9) can be rewritten in matrix form

$$\hat{\vartheta} = \arg \min_{\vartheta} \|\mathbf{r} - \mathbf{A}_{\omega} \mathbf{a}\|_2^2, \quad (10)$$

where $\mathbf{A}_{\omega} = [\Phi \sin_{\omega}, \Phi \cos_{\omega}]$ is composed of the compressed measurements of sinusoid samples at frequency ω , and $\mathbf{a} = [a_1, a_2]^T$ contains the amplitudes of sine and cosine sinusoids, respectively. Given frequency ω a priori known, (10) is converted to the estimation of amplitudes \mathbf{a} by solving an ℓ_2 -norm minimization problem

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a} \in \mathbb{R}^2} \|\mathbf{r} - \mathbf{A}_{\omega} \mathbf{a}\|_2^2, \quad (11)$$

to which the solution is given in a simple form $\hat{\mathbf{a}} = \mathbf{A}_{\omega}^{\dagger} \mathbf{r}$, where $\mathbf{A}_{\omega}^{\dagger} = (\mathbf{A}_{\omega}^H \mathbf{A}_{\omega})^{-1} \mathbf{A}_{\omega}^H$ denotes the pseudoinverse of matrix \mathbf{A}_{ω} . Applying (11) on (10) yields the following two-step optimization problem, which estimates frequency ω and amplitudes \mathbf{a} consecutively.

$$\hat{\omega} = \arg \min_{\omega \in [0, \pi]} \|\mathbf{r} - \mathbf{A}_{\omega} \mathbf{A}_{\omega}^{\dagger} \mathbf{r}\|_2^2, \hat{\mathbf{a}} = \mathbf{A}_{\hat{\omega}}^{\dagger} \mathbf{r}. \quad (12)$$

In Algorithm 2, the estimation of frequency ω and amplitudes \mathbf{a} are carried out in an iterative form to gradually converge to the numerical solution to (9). In each iteration, the minimum ℓ_2 -norm error is calculated for each frequency point sampled from predetermined frequency range with equal interval, and then, the frequency range is shrunk to the neighborhood of the frequency point with least ℓ_2 -norm error, through which the estimation precision is improved.

4. NUMERICAL EXPERIMENTS

We compared the reconstruction performance of our algorithm, denoted as MDS, to state-of-the-art methods including

Algorithm 1: Model Selection for Spectral Compressive Sensing

Input: CS matrix $\Phi \in \mathbb{R}^{M \times N}$, CS measurement vector $\mathbf{m} \in \mathbb{R}^M$, sparsity K

Output: Reconstructed frequency-sparse signal $\hat{\mathbf{s}}$

Initialize: $\hat{\mathbf{s}}^{(j)} = 0, j = 1, \dots, K$

while halting criterion false **do**

for $i = 1$ to K **do**

 {form residual measurement}

$\mathbf{r} \leftarrow \mathbf{m} - \sum_{j=1, j \neq i}^K \Phi \hat{\mathbf{s}}^{(j)}$

 {estimate sinusoid parameters}

$\{\hat{\omega}_i, \hat{a}_i, \hat{\phi}_i\} \leftarrow \mathcal{R}(\Phi, \mathbf{r})$

 {form sinusoid estimate}

$\hat{\mathbf{s}}^{(i)} \leftarrow \left\{ \hat{a}_i \sin(\hat{\omega}_i t + \hat{\phi}_i) \right\}_{t=1}^N$

end for

end while

return $\hat{\mathbf{s}} \leftarrow \sum_{j=1}^K \hat{\mathbf{s}}^{(j)}$

Algorithm 2: Sinusoid Parametric Estimation $\mathcal{R}(\Phi, \mathbf{r})$

Input: CS matrix $\Phi \in \mathbb{R}^{M \times N}$, residual CS measurement vector $\mathbf{r} \in \mathbb{R}^M$

Output: Sinusoid parameter estimates $\hat{\omega}, \hat{a}, \hat{\phi}$

Initialize: $\alpha = 0, \beta = \pi, S = \infty$

while halting criterion false **do**

 {frequency estimate}

$\{\omega_i\}_{i=0}^N \leftarrow \{\alpha + i(\beta - \alpha)/N\}_{i=0}^N$

for $i = 0$ to N **do**

 {calculate compressed measurement}

$\mathbf{A}_{\omega_i} \leftarrow [\Phi \sin_{\omega_i}, \Phi \cos_{\omega_i}]$

 {calculate square error}

$S_{\omega_i} \leftarrow \|\mathbf{r} - \mathbf{A}_{\omega_i} \mathbf{A}_{\omega_i}^{\dagger} \mathbf{r}\|_2^2$

if $S_{\omega_i} < S$ **then**

$S \leftarrow S_{\omega_i}, j \leftarrow i$

 {update parameter estimate}

$\hat{\mathbf{a}} = [\hat{a}_1, \hat{a}_2]^T \leftarrow \mathbf{A}_{\omega_i}^{\dagger} \mathbf{r}, \hat{\omega} \leftarrow \omega_i$

end if

end for

 {frequency range refinement}

$\alpha \leftarrow \max\{\omega_{j-1}, \alpha\}, \beta \leftarrow \min\{\omega_{j+1}, \beta\}$

end while

return $\hat{\omega}, \hat{a} \leftarrow (\hat{a}_1^2 + \hat{a}_2^2)^{\frac{1}{2}}, \hat{\phi} \leftarrow \arctan(\hat{a}_2/\hat{a}_1)$

ℓ_1 -analysis, ℓ_1 -synthesis, SIHT, BISP, BOMP, SDP and CBP, from both noisy and noise-free measurements¹. We chose the normalized ℓ_2 -norm error as major performance measure,

¹The authors would like to thank Marco F. Duarte, Karsten Fyhn, Boaz Nadler and Gongguo Tang (listed in alphabetical order) for providing the implementation of their algorithms.

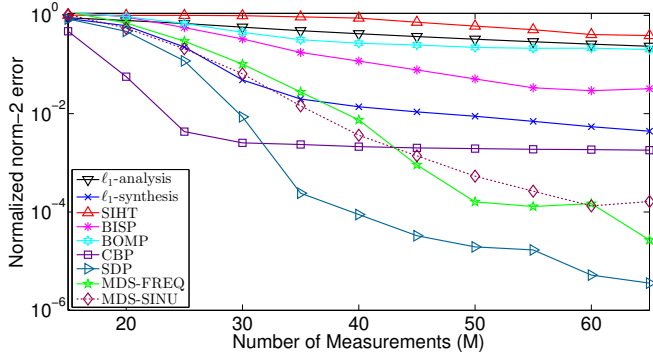


Fig. 1: Signal reconstruction performance in noiseless case.

which is defined as $\|x - \hat{x}\|_2 / \|x\|_2$, between the original signal x and the recovered signal \hat{x} . To evaluate and compare to other algorithms the frequency estimation performance of MDS, we generated frequency-sparse signals of length $N = 128$ composed of $K = 3$ real-valued sinusoids with frequencies selected uniformly at random, unit amplitudes, and zero phase shifts. The frequencies were well-separated so that no two tones were closer than π/N to keep compatible with algorithms based on coherence inhibition signal model, including SIHT, BOMP, and BISP. Such separation is reasonable due to recent theoretical advances in relation between signal recoverability and minimum separation between spectral spikes [13, 23]. Additionally, we generated frequency-sparse signals with frequencies as in the former setting but amplitudes and phase shifts selected uniformly at random to evaluate the parametric estimation performance of MDS. The two parameter settings are denoted in our experiment as MDS-FREQ and MDS-SINU, respectively. We performed Monte Carlo experiments and averaged over 600 trials. The sensing matrix Φ was chosen as the Gaussian random matrix², and the redundant DFT frame was with $c = 5$.

In the first experiment, we evaluated the reconstruction performance from noiseless measurements with M varying from 15 to 65. We set $\epsilon = 10^{-10}$ for relevant algorithms. The results of numerical experiment are shown in Figure 1. In the noiseless case, SDP obtains the best result. When the number of measurements is sufficiently large, the frequency estimation performance of MDS outperforms CBP, whereas for M smaller than 40 it is worse than CBP and ℓ_1 -synthesis, while still better than other algorithms. The parametric estimation performance of MDS is similar to the frequency estimation. Among other algorithms, CBP outperforms ℓ_1 -synthesis and remains static precision level for a wide range of M . Though the redundant DFT coefficients recovered by ℓ_1 -synthesis is actually not sparse and exhibits severe frequency mismatch phenomenon, the signal is in practice reconstructed accurately. The performance of ℓ_1 -analysis, BOMP, and SIHT is

²For the SDP algorithm we used a random subsampling matrix, as the algorithm is only defined for such a sensing matrix.

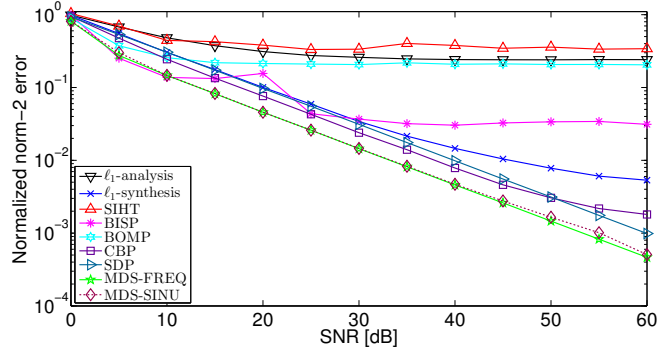


Fig. 2: Signal reconstruction performance in noisy case.

Methods	Noiseless	Noisy
ℓ_1 -analysis	232.5267	329.2279
ℓ_1 -synthesis	9.6257	12.8537
SIHT	1.0950	1.1129
SDP	39.2334	59.1094
BOMP	0.0398	0.0396
CBP	133.7628	115.8614
BISP	17.7901	15.7219
MDS-FREQ	0.9900	0.9499
MDS-SINU	1.1507	1.0233

Table 1: Average computation times in seconds.

the worst among the algorithms tested.

In the second experiment, we included additive Gaussian noise in the signal model. We fixed $M = 64$ and explored a wide range of signal-to-noise ratio (SNR) value from 0 to 60 dB. The resulting performance curves are shown in Fig. 2. In the noisy case, the frequency estimation performance of our algorithm outperforms other algorithms, and when random amplitudes and phase shifts are involved, only negligible deteriorations are induced. This is because model selection relies less on signal sparsity, and more on the matching to superposition of sinusoids. Among other algorithms, CBP and SDP obtain the best result, and ℓ_1 -synthesis exhibits high fidelity when SNR level is sufficiently high. Despite its satisfactory performance in low SNR level, the performance of BISP is mediocre when noise level is low.

The computation time is of equal importance, and the average computation times are listed in Table 1³. The table shows that the excellent performance of frequency estimation and parameter estimation of MDS is enhanced by its high computational efficiency. Moreover, it is observed that the distinguished performance of CBP and SDP are restrained by their high computational expense. In addition, MDS only requires the sensing matrix to have the RIP. This flexibility on measurement scheme increases its performance advantage over SDP.

³We set $M = 64$ for the noiseless case and SNR = 30 for the noisy case.

5. REFERENCES

- [1] E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse problems*, vol. 23, no. 3, p. 969, 2007.
- [2] J. Tropp, J. Laska, M. Duarte, J. Romberg, and R. Baraniuk, "Beyond Nyquist: Efficient sampling of sparse bandlimited signals," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 520–544, Jan. 2010.
- [3] M. Mishali and Y. C. Eldar, "From theory to practice: Sub-nyquist sampling of sparse wideband analog signals," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 375–391, Apr. 2010.
- [4] M. Wakin, S. Becker, E. Nakamura, M. Grant, E. Sovero, D. Ching, J. Yoo, J. Romberg, A. Emami-Neyestanak, and E. Candès, "A nonuniform sampler for wideband spectrally-sparse environments," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 2, no. 3, pp. 516–529, Sept. 2012.
- [5] M. F. Duarte and R. G. Baraniuk, "Spectral compressive sensing," *Appl. and Computational Harmonic Anal.*, vol. 35, no. 1, pp. 111 – 129, 2013.
- [6] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4053–4085, Sept. 2011.
- [7] J. Meng, W. Yin, H. Li, E. Hossain, and Z. Han, "Collaborative spectrum sensing from sparse observations in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 327–337, Feb. 2011.
- [8] M. Mishali, Y. C. Eldar, and A. J. Elron, "Xampling: Signal acquisition and processing in union of subspaces," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4719–4734, Oct. 2011.
- [9] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Sept. 2006.
- [10] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
- [11] R. G. Baraniuk, "Compressive sensing [lecture notes]," *IEEE Signal Processing Mag.*, vol. 24, no. 4, pp. 118–121, Jul. 2007.
- [12] E. J. Candès, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Appl. and Computational Harmonic Anal.*, vol. 31, no. 1, pp. 59–73, 2011.
- [13] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *submitted for publication*.
- [14] K. Fyhn, H. Dadkhahi, and M. Duarte, "Spectral compressive sensing with polar interpolation," in *2013 IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, Vancouver, Canada, May 2013, pp. 6225–6229.
- [15] A. Fannjiang and W. Liao, "Coherence pattern-guided compressive sensing with unresolved grids," *SIAM J. on Imaging Sci.*, vol. 5, no. 1, pp. 179–202, 2012.
- [16] C. Ekanadham, D. Tranchina, and E. P. Simoncelli, "Recovery of sparse translation-invariant signals with continuous basis pursuit," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4735–4744, Oct. 2011.
- [17] B. Nadler and A. Kontorovich, "Model selection for sinusoids in noise: Statistical analysis and a new penalty term," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1333–1345, Apr. 2011.
- [18] E. J. Hannan, "Determining the number of jumps in a spectrum," *Develop. Time Series Anal.*, pp. 127–138, 1993.
- [19] L. Kavalieris and E. Hannan, "Determining the number of terms in a trigonometric regression," *J. Time Series Anal.*, vol. 15, no. 6, pp. 613–625, 1994.
- [20] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.
- [21] K. Fyhn, M. F. Duarte, and S. H. Jensen, "Compressive parameter estimation for sparse translation-invariant signals using polar interpolation," *submitted for publication*.
- [22] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The convex geometry of linear inverse problems," *Foundations of Computational Math.*, vol. 12, no. 6, pp. 805–849, 2012.
- [23] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Commun. Pure and Appl. Math.*, 2013.