JOINT RECOVERY OF SPARSE SIGNALS AND PARAMETER PERTURBATIONS WITH PARAMETERIZED MEASUREMENT MODELS

Erik C. Johnson, Douglas L. Jones

Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign, Urbana, IL, 61801, USA

ABSTRACT

Many applications involve sparse signals with unknown, continuous parameters; a common example is a signal consisting of several sinusoids of unknown frequency. Applying compressed sensing techniques to these signals requires a highly oversampled dictionary for good approximation, but these dictionaries violate the RIP conditions and produce inconsistent results. We consider recovering both a sparse vector and parameter perturbations from an initial set of parameters. Joint recovery allows for accurate reconstructions without highly oversampled dictionaries. Our algorithm for sparse recovery solves a series of linearized subproblems. Recovery error for noiseless simulated measurements is near zero for coarse dictionaries, but increases with the oversampling. This technique is also used to reconstruct Radio Frequency data. The algorithm estimates sharp peaks and transmitter frequencies, demonstrating the potential practical use of the algorithm on real data.

Index Terms— Sparse Reconstruction, Sparse Signal, Parameterized Model, Parameter Perturbations, Frequency Estimation

1. INTRODUCTION

A wide range of signal processing applications require recovery of signals with unknown, continuous parameters. Quite often these signals are also sparse. One example is a time-domain signal consisting of multiple sinusoids of unknown frequency, which is frequency-sparse.

In the field of Compressed Sensing (CS), there has been a great deal of research into the reconstruction of sparse signals from few measurements. A k-sparse signal $\mathbf{x} \in \mathbb{R}^N$ is one with only k nonzero elements. By enforcing signal sparsity, one can recover, under certain conditions, the sparse solution to an undetermined system ([1] and [2] among many others). Given a linear measurement model $\mathbf{b} = \mathbf{A} * \mathbf{x} + \mathbf{e}$ and appropriate conditions on $\mathbf{A}$, one can often recover a k-sparse $\mathbf{x}$ by solving an ℓ_1 minimization problem known as Basis Pursuit (BP) [3] or using greedy algorithms such as Orthogonal Matching Pursuit [4].

One of the major assumptions of CS recovery is that the signal dictionary $\mathbf{A}$ is exactly known. For sparse signals with unknown parameters, however, the dictionary is not given. In this case, one could sample and discretize the parameter space to form a signal dictionary. Intuitively a dense sampling is needed to obtain an accurate reconstruction. A dense discretization, however, can result in poor recovery as the highly correlated dictionary violates the RIP condition [1].

The case of completely unknown signal dictionaries has been explored in Blind CS [5]. Sparsity-cognizant Total Least-Squares [6] considers perturbing a known $\mathbf{A}$ by an unstructured perturbation term. The authors of [7] and [8] consider a structured perturbation that is parameterized, but consider only a linearized form. The authors of [9] formulate an iterative thresholding algorithm specifically for parametric spectral estimation.

In this paper, we consider a general parameterized model for the measurement matrix, $\mathbf{A} = \mathbf{A}(\omega + \mathbf{d}\omega)$. Here $\omega \in \mathbb{R}^N$ is a set of parameters used to generate the sensing matrix. Given an initial vector of parameters ω, the goal is to jointly recover a sparse vector $\mathbf{x}$ and parameter perturbations $\mathbf{d}\omega$. This problem is solved by a new nonlinear programming algorithm to find critical points of the problem. This recovery algorithm, Successive Linearized Programming for Sparse Recovery (SLPSR), solves successive linearized problems within a trust-region [10].

Our work makes several contributions. The formulation of our parameter perturbation recovery is more general than the linear structure in [7], and contains the linear structure as a subset. The SLPSR algorithm is, to the best of our knowledge, novel and is applicable to a far wider range of models than the alternating algorithm in [7]. Unlike the spectral-estimation algorithm in [9], SLPSR is applicable to many different parameterized measurement models. Finally, we apply our recovery method to real-world data, helping to demonstrate the practical potential of jointly recovering sparse signals and parameter perturbations.

In this paper, we use bold, lower-case letters for vectors and bold, upper-case letters for matrices. The $*$ superscript

We would like to acknowledge the funding from NSF EFRI-BSBA Award No. 0938007 and NSF IGERT Grant No. 0903622T. We would also like to thank Professor Steve Franke of the University of Illinois for his help collecting the HF Spectrum dataset.

denotes the solution to an optimization problem. Subscripts indicate an element or subset of a vector, e.g. x_3 and $\mathbf{x}_{\mathbf{x}>0}$.

2. RECOVERY PROBLEM FORMULATION WITH PARAMETERIZED MODEL

We consider recovering a sparse vector $\mathbf{x} \in \mathbb{C}^N$ and parameter perturbations $\mathbf{d}\omega \in \mathbb{R}^N$, given a set of initial parameters $\omega \in \mathbb{R}^N$ and a model $\mathbf{a}(\omega_i)$ that maps a parameter value to a vector in $\mathbb{C}^N$. In essence, ω discretizes the space of possible parameters. The goal is to recover the unknown signal $\mathbf{x}$ from a set of noisy measurements given by $\mathbf{b} \in \mathbb{C}^M$. In order to jointly recover a sparse vector of weights and a sparse vector of parameter perturbations, we solve the following problem, referred to as Perturbed Sparse Recovery (PSR).

$$\begin{aligned} \min_{\mathbf{x},\mathbf{d}\omega} \quad & \|\mathbf{x}\|_1 + \lambda * \|\mathbf{d}\omega\|_1 \\ \text{subject to} \quad & \|\mathbf{b} - \mathbf{A}(\omega + \mathbf{d}\omega) * \mathbf{x}\|_2 \leq \epsilon \end{aligned}$$

where the measurement model $\mathbf{A}(\omega + \mathbf{d}\omega)$ maps $\mathbb{R}^N$ to $\mathbb{C}^{M \times N}$ is given by

$$\mathbf{A}(\omega + \mathbf{d}\omega) = [\mathbf{a}(\omega_1 + d\omega_1) | \mathbf{a}(\omega_2 + d\omega_2) | \cdots | \mathbf{a}(\omega_N + d\omega_N)]$$

In this formulation, the parameter $\lambda \in \mathbb{R}$ weights the trade-off between sparsity of the recovered weights and the size of the perturbations and the parameter $\epsilon \in \mathbb{R}$ allows for deviation from the noisy measurements. The ℓ_1 norm is used to promote sparsity in the final solution. The elements of the model $\mathbf{a}(\omega_i)$ must be differentiable with respect to ω_i.

3. SUCCESSIVE LINEARIZED PROGRAMMING FOR SPARSE RECOVERY

In this section, we propose an algorithm for finding local minima of the PSR problem by solving linearized subproblems. The algorithm, Successive Linearized Programming for Sparse Recovery (SLPSR), takes an initial, feasible point $\mathbf{x}_0, \mathbf{d}\omega_0$ and generates a series of feasible iterates $\mathbf{x}_k, \mathbf{d}\omega_k$. The iterates are generated as the solutions to linearized versions of the PSR problem generated by a first-order Taylor series approximation, using the Jacobian ($\mathbf{J}(\mathbf{x}_k, \mathbf{d}\omega_k)$) of the function $\mathbf{b} - \mathbf{A}(\omega + \mathbf{d}\omega_k) * \mathbf{x}_k$ at $\mathbf{x}_k, \mathbf{d}\omega_k$. A step $\mathbf{s}_\mathbf{x} = \mathbf{x} - \mathbf{x}_k, \mathbf{s}_{\mathbf{d}\omega} = \mathbf{d}\omega - \mathbf{d}\omega_k$ from the current iterate is found by solving the linearized problem. As the linearization only holds around the point $\mathbf{x}_k, \mathbf{d}\omega_k$, the step is restricted to a Trust-Region [10] around the point $\mathbf{x}_k, \mathbf{d}\omega_k$. The linearized subproblem, $MinSubproblem$, is defined as:

$$\begin{aligned} \arg\min_{\mathbf{s}_x, \mathbf{s}_{d\omega}} \quad & \|\mathbf{x}_k + \mathbf{s}_x\|_1 + \lambda * \|\mathbf{d}\omega_k + \mathbf{s}_{\mathbf{d}\omega}\|_1 \\ \text{subject to} \quad & \|\mathbf{b} - \mathbf{A}(\omega + \mathbf{d}\omega_k) * \mathbf{x}_k + \mathbf{J} * (\begin{bmatrix} \mathbf{s}_\mathbf{x} \\ \mathbf{s}_{\mathbf{d}\omega} \end{bmatrix})\|_2 \leq \epsilon \\ & \|\begin{bmatrix} \mathbf{s}_\mathbf{x} \\ \mathbf{s}_{\mathbf{d}\omega} \end{bmatrix}\|_2 \leq \Delta_k \end{aligned}$$

The trust-region is enforced by limiting the norm of the step $\mathbf{s}_\mathbf{x}, \mathbf{s}_{\mathbf{d}\omega}$ with the parameter Δ_k. This ensures iterates reduce the objective function. After calculating the step for the linearized subproblem, the new point $\mathbf{x}_k + \mathbf{s}_x, \mathbf{d}\omega_k + \mathbf{s}_{\mathbf{d}\omega}$ is only a tentative solution as it may not be feasible. A correction step is calculated by fixing $\mathbf{d}\omega + \mathbf{s}_{\mathbf{d}\omega}$ and calculating the minimum change in $\delta_\mathbf{x}$ to produce a feasible solution. The function $FeasibleProjection$ is defined as the minimum norm $\delta_\mathbf{x}$ that satisfies $\|\mathbf{b} - \mathbf{A}(\omega + \mathbf{d}\omega_k + \mathbf{s}_{\mathbf{d}\omega}) * (\mathbf{x}_k + \mathbf{s}_x + \delta_\mathbf{x})\|_2 \leq \epsilon$. Given this new point $\mathbf{x}_k + \mathbf{s}_x + \delta_\mathbf{x}, \omega + \mathbf{d}\omega_k + \mathbf{s}_{\mathbf{d}\omega}$, the new objective function value, $\|\mathbf{x}_k + \mathbf{s}_x + \delta_\mathbf{x}\|_1 + \lambda * \|\omega + \mathbf{d}\omega_k + \mathbf{s}_{\mathbf{d}\omega}\|_1$, is calculated. If this function value is an improvement over the previous k, the new point is accepted and the trust region radius, Δ_{k+1}, increased. If it is not an improvement, the algorithm remains at $\mathbf{x}_k, \mathbf{d}\omega_k$ and decreases the trust region radius Δ_k. This continues until the first-order necessary conditions are satisfied or a maximum number of iterations are reached. To recover complex-valued signals, the subproblems are cast as Second-Order Conic Programs. Solving for the feasible projection is simply the projection onto a convex set.

Algorithm 1 Successive Linearized Programming for Sparse Recovery (SLPSR)

Require: $0 < \nu < 1$
while $\mathbf{x}, \mathbf{d}\omega$ is not First-Order Critical **do**
  $\mathbf{J}_k \leftarrow \mathbf{J}(\mathbf{x}, \mathbf{d}\omega)$
  $\mathbf{A}_k \leftarrow \mathbf{A}(\omega + \mathbf{d}\omega)$
  $\mathbf{s}_\mathbf{x}, \mathbf{s}_{\mathbf{d}\omega} \leftarrow MinSubproblem$
  $\delta_x \leftarrow FeasibleProjection$
  if $\|\mathbf{x} + \mathbf{s}_\mathbf{x} + \delta_\mathbf{x}\|_1 + \lambda * \|\mathbf{d}\omega + \mathbf{s}_{\mathbf{d}\omega}\|_1 < \|\mathbf{x}\|_1 + \lambda * \|\mathbf{d}\omega\|_1$ **then**
    $\mathbf{x} \leftarrow \mathbf{x} + \mathbf{s}_\mathbf{x} + \delta_\mathbf{x}$
    $\mathbf{d}\omega \leftarrow \mathbf{d}\omega + \mathbf{s}_{\mathbf{d}\omega}$
    $\Delta \leftarrow \frac{1}{\nu} * \Delta$
  else
    $\mathbf{x} \leftarrow \mathbf{x}$
    $\mathbf{d}\omega \leftarrow \mathbf{d}\omega$
    $\Delta \leftarrow \nu * \Delta$
  end if
end while

3.1. First-Order Necessary Conditions for Inequality Constraints

To provide a stopping criterion for SLPSR, we must be able to characterize first-order critical points of the PSR problem. The ℓ_1 norm, however, is not differentiable at any point where an element equals zero. It is possible, however, to characterize first-order critical points using the subdifferential [11] of the objective function, $\partial f(\mathbf{x}, \mathbf{d}\omega)$. The first-order critical condition at a feasible point $\mathbf{x}, \mathbf{d}\omega$ is $0 \in \partial f(\mathbf{x}, \mathbf{d}\omega) + J(\mathbf{x}, \mathbf{d}\omega)^T * (\mathbf{b} - \mathbf{A}(\omega + \mathbf{d}\omega) * \mathbf{x}) * \gamma$ for some lagrange multiplier $\gamma \geq 0$. This can be extended to complex-valued variables using the well-known $\mathbb{CR}$ calculus [12] and writing the real-valued objective function in terms of the real and complex parts of the complex variables.

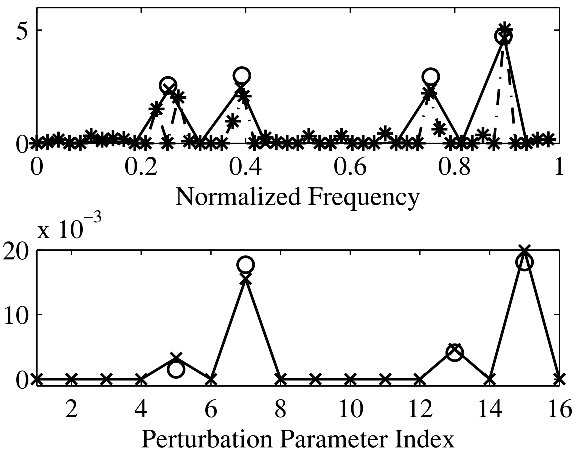

Fig. 1. The top figure shows sparse signal recovery of four complex sinusoids recovered from 16 measurements at 10dB SNR of simulated frequency-sparse signals. The BP recovery with a 3-times oversampled dictionary is plotted with $*$, and the SLPSR recovery with x's. The lower figure shows the recovery of the parameter perturbation. The true weights and parameter perturbations are marked with circles.

3.2. Algorithm Convergence and Parameter Selection

Because of the nonlinear form of the constraint, it is not immediately obvious that the SLPSR algorithm converges to a first-order critical point of the PSR problem. Given minor constraints on $\mathbf{A}(\cdot)$, it can be shown [13] that there exists a $\Delta > 0$ such that the step of the SLPSR algorithm reduces the objective function unless the current iterate is already a first-order critical point. This implies that the SLPSR algorithm will reduce the objective function until a first-order critical point is reached. The parameters $\lambda, \epsilon, \Delta_k$ and ω must all be selected for the SLPSR algorithm. Choosing ω will depend on the model. In the following section, ω represents digital frequency and is chosen using evenly-spaced samples on the interval $\{0, 2\pi\}$. The parameter Δ_k is adjusted throughout the procedure using the parameter ν. A good choice of ν is 0.5. The parameter λ corrects for the scale between changes in parameters and changes in the signal weights. Therefore, it is critical to choose $0 < \lambda \leq \frac{\Delta_{\mathbf{x}}}{\Delta_{\mathbf{d}\omega}}$ where $\Delta_{\mathbf{x}}$ is an approximation of the change in the ℓ_1 norm of $\mathbf{x}$ possible by perturbing the initial parameters. Similarly, $\Delta_{\mathbf{d}\omega}$ is an approximation of the increase in the ℓ_1 norm of $\mathbf{d}\omega$. The parameter ϵ should be chosen to be slightly larger than the expected norm of the noise term.

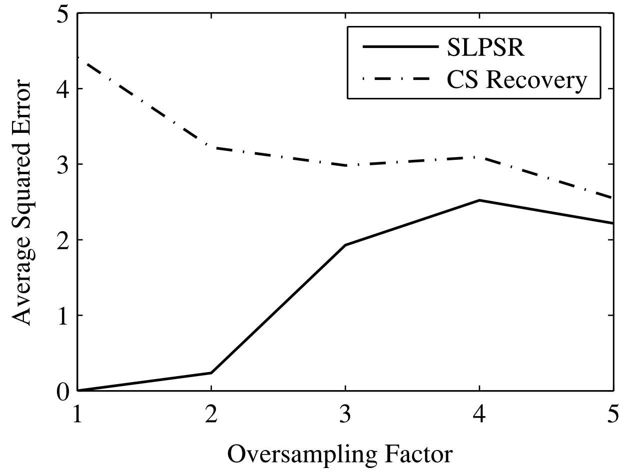

Fig. 2. Algorithm recovery error as a function of frequency oversampling. Reconstructions are done on length-16 time signals with no added noise. The optimal weights are rounded to the nearest frequency in the recovery and the ℓ_2 norm of the error calculated. The SLPSR solution has minimal error at coarse discretizations. As the oversampling factor increases, the performance of the two algorithms becomes similar.

4. APPLICATION TO RECOVERY OF FREQUENCY-SPARSE SIGNALS

In order to demonstrate the joint recovery of sparse vectors and parameter perturbations, we consider the reconstruction of frequency-sparse signals from M time-domain measurements. The parameters to be estimated are the digital frequencies present in the time signal and the weights are the complex-valued amplitudes corresponding to those frequencies. The elements of the parameterized model are $a_j(\omega_i) = \frac{1}{\sqrt{M}} * e^{\imath * \omega_i * (j-1)},\ j = 1, 2, \ldots, M$. Then $\mathbf{A}(\omega)$ maps the vector of frequencies $\omega \in \mathbb{R}^N$ to a matrix in $\mathbb{C}^{M \times N}$ where each column is a complex sinusoid of frequency ω_i and length M.

Given an initial set of parameters ω_0 and measurements $\mathbf{b} \in \mathbb{C}^N$, the goal is to recover a set of sparse weights $\mathbf{x} \in \mathbb{C}^N$ and parameter perturbations $\mathbf{d}\omega \in \mathbb{R}^N$. In this section, the SLPSR algorithm is initialized using $k * M$ frequency parameters, spaced equally between 0 and 2π, where k is the oversampling factor. We compare the recovery of our algorithm with the BP solution with the sensing matrix fixed at $\mathbf{A}(\omega_0)$.

To test the recovery, 4 complex sinusoids are generated with frequencies ω^* uniformly randomly distributed between 0 and 2π. If generated frequencies are closer than $\frac{2\pi}{M}$, we regenerate ω^*. The real and imaginary parts of the weights are drawn independently from uniform random distributions on $[1, -1]$.

For all results in this section, SLPSR subproblems are solved with Second-order Conic Programming. The param-

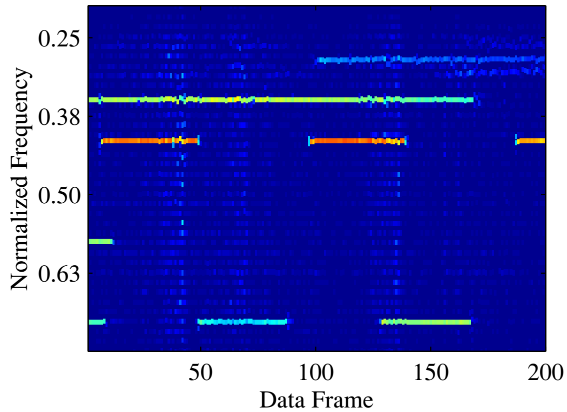

Fig. 3. SLPSR recovery of Radio Frequency data on non-overlapping, length-64 frames of data. The initial parameters are simply the DFT frequencies. The weights are plotted at the corresponding recovered parameter. The algorithm sharply recovers the transmitter amplitude and estimates the transmitter frequency without an oversampled dictionary.

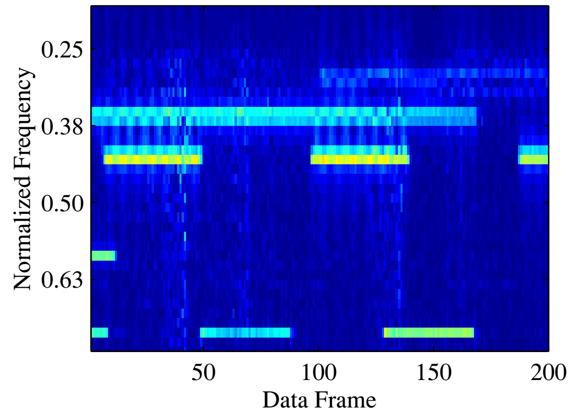

Fig. 4. Discrete Fourier Transform of Radio frequency data with length-64 frames of data. There is no overlap between frames and no window was applied.

eters λ, ϵ were found by a search around initial estimates. All standard BP reconstructions were performed using Second-order Conic Programming formulations with inequality constraints for noisy measurements and equality constraints for noiseless measurements.

To demonstrate recovery in under noisy conditions, complex gaussian noise is added to the signal. Figure 1 demonstrates recovery of the complex sinusoids at 10dB SNR. Recovery using SLPSR is not exact, but still quite reasonable. Interestingly, the reconstruction with standard BP techniques using a 3-times oversampled dictionary demonstrate odd recovery artifacts with the left-most sinusoid. The parameter perturbations are able to recover the perturbation necessary to recover the unknown frequency patterns to within a small error despite the coarse sampling grid.

To explore the difference between the SLPSR formulation and BP further, 50 examples of noiseless time signals were generated as described above. The SLPSR recovery is compared with BP recovery for different oversampling factors. Figure 2 shows the mean 2-norm error between the optimal weights and the recovered sparse signal. The optimal weights are rounded to the nearest estimated frequency in the sparse signal. Our technique recovers the optimal weights with very low error at coarse discretizations. As the oversampling factor increases, both SLPSR and CS recovery demonstrate similar error. This result demonstrates that SLPSR is capable of estimating frequency sparse signals from coarse discretizations.

Finally, we apply our nonlinear sparse recovery algorithm to radio-frequency data, specifically the HF spectrum. The data was downconverted from 7.1 Mhz to complex baseband and sampled at 250 kHz. The complex baseband data is band-pass filtered and downsampled to extract a band of interest. In this band, there are multiple transmitters apparently broadcasting morse code. The data is reconstructed using non-overlapping blocks of 64 samples. Figure 3 shows the SLPSR recovery using the DFT frequencies. To display the results, the sparse weights are interpolated to the nearest display frequency. Figure 4 shows the DFT reconstruction. The sparse recovery significantly sharpens the visible peaks of the transmitters. Similarly, the parameter perturbations give a more accurate estimation of the transmitter frequencies. This result suggests that SLPSR can successfully reconstruct real-world signals.

5. CONCLUSIONS AND FUTURE WORK

We can jointly estimate a sparse vector and the corresponding parameter vector using our SLPSR algorithm. This problem formulation supports a wide variety of parameterized measurement models and is more general than previous recovery methods for CS with perturbed models. This scheme can estimate both the amplitudes and the frequencies present in frequency-sparse data. Critically, the SLPSR approach can recover frequency-sparse signals from coarse initial discretizations. The experiments with Radio Frequency data suggest that this recovery is more robust than highly oversampled BP recoveries while still providing accurate parameter estimates.

Moving forward, it is necessary to explore the resolution limits of this technique. This recovery algorithm also has interesting implications for sequential recovery problems or online dictionary learning with parameterized models. By sequentially updating the parameters, we can preform online tracking of the dictionary.

6. REFERENCES

[1] E.J. Candès, “Compressive sampling,” in *Proceedings oh the International Congress of Mathematicians: Madrid, August 22-30, 2006: invited lectures*, 2006, pp. 1433–1452.

[2] D.L. Donoho, “Compressed sensing,” *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289–1306, 2006.

[3] 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.

[4] J.A. Tropp and A.C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” *Information Theory, IEEE Transactions on*, vol. 53, no. 12, pp. 4655–4666, 2007.

[5] S. Gleichman and Y.C. Eldar, “Blind compressed sensing,” *Information Theory, IEEE Transactions on*, vol. 57, no. 10, pp. 6958–6975, 2011.

[6] H. Zhu, G. Leus, and G.B. Giannakis, “Sparsity-cognizant total least-squares for perturbed compressive sampling,” *Signal Processing, IEEE Transactions on*, vol. 59, no. 5, pp. 2002–2016, 2011.

[7] Zai Yang, Cishen Zhang, and Lihua Xie, “Robustly stable signal recovery in compressed sensing with structured matrix perturbation,” *Signal Processing, IEEE Transactions on*, vol. 60, no. 9, pp. 4658 –4671, sept. 2012.

[8] Zai Yang, Cishen Zhang, and Lihua Xie, “Stable signal recovery in compressed sensing with a structured matrix perturbation,” in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, march 2012, pp. 2737 –2740.

[9] Marco F. Duarte and Richard G. Baraniuk, “Spectral compressive sensing,” *Applied and Computational Harmonic Analysis*, Submitted.

[10] A.R. Conn, N.I.M. Gould, and P.L. Toint, *Trust region methods*, vol. 1, Society for Industrial Mathematics, 1987.

[11] R. T. Rockafellar, “Directionally lipschitzian functions and subdifferential calculus,” *Proceedings of the London Mathematical Society*, vol. s3-39, no. 2, pp. 331–355, 1979.

[12] J. Eriksson, E. Ollila, and V. Koivunen, “Essential statistics and tools for complex random variables,” *Signal Processing, IEEE Transactions on*, vol. 58, no. 10, pp. 5400 –5408, oct. 2010.

[13] E. Johnson, “Recovery of sparse signals and parameter perturbations from parameterized signal models,” M. sc. thesis, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL, May 2013.