NOVA

University of Newcastle Research Online

nova.newcastle.edu.au

Hyder, Md Mashud, Mahata, Kaushik; "Coherent spectral analysis of asynchronously sampled signals". Published in IEEE Signal Processing Letters Vol. 18, Issue 2, p. 126-129 (2011)

Available from: <http://dx.doi.org/10.1109/LSP.2010.2094607>

© 2011 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Accessed from: <http://hdl.handle.net/1959.13/1065349>

Coherent spectral analysis of asynchronously sampled signals

Md Mashud Hyder and Kaushik Mahata

Abstract—We propose a joint sparse signal recovery approach to coherent spectral analysis of irregularly sampled signals. These signals share the same frequencies, and are sampled asynchronously. Two types of solution procedures are considered. First one is a convex optimization approach, which optimizes a mixed $\ell_{2,1}$ -norm. The other method minimizes an approximation of $\ell_{2,0}$ -norm and the resulting algorithm can be implemented using a few FFTs and IFFTs. We demonstrate the effectiveness of the sparse recovery approach using simulation experiments. In particular, the $\ell_{2,0}$ approximation approach is very fast. In addition, it offers increased resolution, improved robustness to noise, and works well with limited number of data samples.

Index Terms—Coherence spectrum, joint sparse, sparse representation, spectral analysis.

I. INTRODUCTION

Frequency estimation is an extensively researched problem, see e.g. [1]. In a related problem one is interested in estimating the frequencies present in a set of signals, which share the same frequencies [2], [3]. In this paper, we assume the observed signals are irregularly sampled in an asynchronous manner. Instance of irregular sampling is common [4], [5], and asynchronous sampling is unavoidable when different signals are sampled using different data acquisition systems. Spectral analysis methods from unevenly sampled signals often use some least squares Fourier transform estimator [2], [3], which suffer from poor resolution performance. The segmented iterative adaptive approach (SIAA-MSC) [6] can handle two signals, and offers better resolution if the number of samples is sufficiently large. A generalized version of coherent spectral estimation from multiple signals is proposed in [7], but it requires synchronous uniform sampling.

Sparse recovery approaches have been shown to achieve high resolution in spectral analysis [4], [8]. In this paper we extend this idea to solve the coherent spectral analysis problem. First, we show that the coherent frequency estimation problem can be posed as the problem of estimating a jointly row sparse matrix X . Subsequently, we explore two avenues to solve the joint sparse recovery problem. We apply the popular approach of minimizing the joint $\ell_{2,1}$ norm of X [9]. We also extend a joint $\ell_{2,0}$ norm approximation (JLZA) algorithm [10] for minimizing an approximation of the $\ell_{2,0}$ norm of X . This method belongs to the class of smoothed ℓ_0 (SLO) approximation algorithms (see [11] and references therein), and requires a few FFTs and IFFTs. This makes the extended

JLZA (EJLZA) algorithm is very fast. Furthermore, it offers high resolution and robustness to measurement noise.

II. JOINT SPARSE MODEL

Consider r coherent signals $\{y_k(t)\}_{k=1}^r$ of the form

$$y_k(t) = \sum_{\ell=1}^n a_{k,\ell} e^{i\Omega_\ell t}. \quad (1)$$

Define the vector of sampling instants for $y_k(t)$ as

$$T_k = [t_{k,1} \ t_{k,2} \ \cdots \ t_{k,M}], \quad t_{k,i} < t_{k,i+1}. \quad (2)$$

We assume that T_k 's have the same length for all k to keep the description simple. However, the proposed algorithms can be extended trivially to a more general case. Let $t_0 = \min_k t_{k,1} = \min_{k,\ell} t_{k,\ell}$, and δ be the largest positive real number such that there are nonnegative integers $d_{k,i}$ satisfying

$$t_{k,i} = t_0 + d_{k,i}\delta, \quad k = 1, \dots, r, \quad i = 1, \dots, M, \quad (3)$$

Note that if $k_* = \arg \min_k t_{k,1}$ then $d_{k_*,1} = 0$. Then by defining $\omega_\ell = \delta\Omega_\ell$, and $b_{k,\ell} = a_{k,\ell} e^{i\Omega_\ell t_0}$, and using (1) we can write for any nonnegative integer m that

$$y_k(t_0 + m\delta) = \sum_{\ell=1}^n b_{k,\ell} e^{i\omega_\ell m}, \quad k = 1, \dots, r; \quad m = 0, 1, 2, \dots$$

Choose $N > \max_k d_{k,M}$, set up a frequency grid

$$\mathbb{G} = \{\bar{\omega}_\ell = 2\pi\ell/N : \ell = 0, 1, \dots, N-1\}, \quad (4)$$

and define the N -point DFT matrix $\Psi \in \mathbb{C}^{N \times N}$ such that

$$[\Psi]_{k,\ell} = \exp[i(k-1)\bar{\omega}_{\ell-1}], \quad k, \ell \in \{1, \dots, N\}. \quad (5)$$

There may not exist any $\ell \in \{0, 1, \dots, N-1\}$ such that a given $\omega_k = \bar{\omega}_\ell$, i.e., it is likely that ω_k may not lie on the grid $\mathbb{G}$. However, making N large enough, the grid $\mathbb{G}$ becomes sufficiently dense. Then for every $k \in \{1, \dots, n\}$, we have $\omega_k \approx \bar{\omega}_\ell$ for some $\ell \in \{0, 1, \dots, N-1\}$, and we can write

$$z_k := [y_k(t_0) \ y_k(t_0 + \delta) \ \cdots \ y_k(t_0 + (N-1)\delta)]' = \Psi x_k + \bar{e}_k,$$

where A' denotes the transpose of A , and $\bar{e}_k$ accounts for measurement noise and unmodelled residuals. The components of x are mostly zero. The ℓ -th component $x_{k,\ell}$ is non-zero only if $\bar{\omega}_\ell \approx \omega_k$ for some $k \in \{1, \dots, n\}$. This makes x_k a sparse vector, and the locations of the dominant peaks in x_k indicate the frequencies present in the data, while the dominant values give the complex-valued amplitudes. Furthermore, every $y_k(t)$ have the same frequencies. Hence if the ℓ -th component of x_1

is non-zero, then the ℓ -th component of x_k is nonzero for all k . Hence the matrix

$$X = [x_1 \ x_2 \ \cdots \ x_r]$$

is joint row-sparse, i.e. it has only a few non-zero rows.

By (3) the sampled measurements of the signals are given by

$$\mathbf{y}_k := [y_k(t_{k,1}) \ \cdots \ y_k(t_{k,M})]' = \Theta_k \Psi x_k + e_k, \quad (6)$$

where Θ_k is a matrix such that its i -th row is the same as the $d_{k,i} + 1$ -th row of the $N \times N$ identity matrix, and $e_k = \Theta_k \bar{e}_k$. Therefore the problem of coherent spectral analysis is the problem of estimating a joint sparse X from $\{\mathbf{y}_k\}_{k=1}^r$, given the “measurement matrices” $\{\Theta_k\}_{k=1}^r$.

Suppose that it is possible to recover the joint sparse X from $\{\mathbf{y}_k\}_{k=1}^r$, then we have estimates of $\{\omega_\ell\}_{\ell=1}^n$. These can be used to uniquely identify the continuous-time frequencies $\{\Omega_\ell\}_{\ell=1}^n$ only if we know a *feasible* set $\mathbb{O}$ a priori, such that $\Omega_\ell \in \mathbb{O}, \ell = 1, \dots, n$. In addition, the intersection between $\mathbb{O}$ and any of its aliased version should be empty, i.e. $\mathbb{O} \cap (\mathbb{O} + 2\pi s/\delta) = \emptyset$ for every nonzero $s \in \mathbb{Z}$. We can allow $\mathbb{O}$ to be made up of several smaller mutually disjoint intervals. The simplest (and most popular) examples of $\mathbb{O}$ are $\mathbb{O} = [-\pi/\delta, \pi/\delta]$ and $\mathbb{O} = [0, 2\pi/\delta]$. However, more complicated cases like $\mathbb{O} = [0, \pi/\delta] \cup (3\pi/\delta, 4\pi/\delta)$ can also be allowed. On the other hand, $\mathbb{O} = [-2\pi/\delta, 2\pi/\delta]$ is an example where the identifiability is lost due to aliasing.

III. ESTIMATION OF X

In the following we denote $\Phi_k = \Theta_k \Psi$ for short. Estimation of a joint sparse X via convex optimization has been considered in literature (see [9] and references therein), where

$$\Phi_1 = \Phi_2 = \cdots = \Phi_r. \quad (7)$$

The most successful approach is the minimization of mixed $\ell_{2,1}$ norm of X . This method has a straightforward extension to the case where (7) does not hold. We solve

$$\min_X \sum_{k=1}^N \|X(k, :)\|_2 \text{ subject to } \sum_{k=1}^r \|\mathbf{y}_k - \Phi_k \mathbf{x}_k\|_2^2 \leq \nu \quad (8)$$

where $X[k, :]$ indicates the k -th row of X . The choice of ν depends on the noise level [12]. In our experiments, we set $\nu = 7$ for a signal to noise ratio 0 dB to 20 dB of the observed data. However, this relaxation approach is quite slow compared to some other alternatives. In addition, this convex relaxation strategy may perform poorly in presence of noise in the data. These motivate an alternative $\ell_{2,0}$ norm minimization approach described below.

The joint $\ell_{2,0}$ approximation (JLZA) algorithm [10] is one of the most successful methods for joint sparse recovery, which has been extended in [13] to account for the case where (7) does not hold. Here we approximate the zero norm by sum of Gaussian functions [10], [11]. The idea is to take a small real number σ , and solve

$$X_*(\sigma) := \arg \min_X L_\sigma(X), \\ L_\sigma(X) := -\sum_{j=1}^N e^{-\frac{\|X(j, :)\|_2^2}{2\sigma^2}} + \frac{\lambda}{2} \sum_{k=1}^r \|\mathbf{y}_k - \Phi_k \mathbf{x}_k\|_2^2 \quad (9)$$

TABLE I: Extended JLZA algorithm for minimizing $L_{\sigma_0}(X)$

Initialization
1. Form $X^{(0)} = \Psi^* [\Theta_1^* \mathbf{y}_1 \ \cdots \ \Theta_r^* \mathbf{y}_r] / N$.
2. Set $\sigma = 4 \max_j (\ X_0[j, :]\ _1)$, $\lambda \in [1, 100]$ and $\rho = 0.5, \eta = 0.5, \gamma = 0.5$.
repeat
3. Set $\beta = 1$.
4. while $L_\sigma\{\beta \zeta(X^{(i)}) + (1 - \beta)X^{(i)}\} > L_\sigma(X^{(i)})$ $\beta = \gamma\beta$.
5. $X^{(i+1)} = \beta \zeta(X^{(i)}) + (1 - \beta)X^{(i)}$
6. If $\tau^{(i)} = \ X^{(i+1)} - X^{(i)}\ _2 < \eta\sigma$ then $\sigma = \rho\sigma$.
7. $i = i + 1$.
while $\sigma > \sigma_0$.
Output: $X_* = X^{(i)}$

In principle, one should take $\sigma \rightarrow 0$. However in that case we have to deal with an NP hard problem. Hence one solves (9) for $\sigma = \sigma_0$, where σ_0 is a sufficiently small positive number. The value of σ_0 can be chosen using a criterion derived in [11]. A sequential optimization method is used to minimize L_{σ_0} . It can be shown that as $\sigma \rightarrow \infty$,

$$X_*(\infty) := [\Phi_1^*(\Phi_1 \Phi_1^*)^{-1} \mathbf{y}_1 \ \cdots \ \Phi_r^*(\Phi_r \Phi_r^*)^{-1} \mathbf{y}_r] \\ = \Psi^* [\Theta_1^* \mathbf{y}_1 \ \cdots \ \Theta_r^* \mathbf{y}_r] / N \quad (10)$$

is the minimizer of L_σ . Hence, we take a large σ initially and solve (9) by initializing an iterative convex-concave procedure (see below) at $X_*(\infty)$ [11]. The resulting solution is then used to initialize the algorithm for solving (9) with σ reduced by a small amount. The procedure is repeated so long $\sigma > \sigma_0$.

A Gauss-Newton type convex-concave procedure is used to minimize L_σ for a fixed σ which builds on the following Lemma proved in [13], where further details can be found.

Lemma 1. Let $X_*(\sigma) = [\mathbf{x}_{1*} \ \mathbf{x}_{2*} \ \cdots \ \mathbf{x}_{r*}]$. Then $\mathbf{x}_{k*} = \zeta_k\{X_*(\sigma)\}$, $k \in \{1, \dots, r\}$, where

$$\zeta_k(X) = \lambda [W(X)/\sigma^2 + \lambda \Phi_k^* \Phi_k]^{-1} \Phi_k^* \mathbf{y}_k, \quad (11)$$

and $W(X)$ is an $N \times N$ diagonal matrix defined as

$$W(X) = \text{diag} \left\{ e^{-\frac{\|X(1, :)\|_2^2}{2\sigma^2}} \ \cdots \ e^{-\frac{\|X(N, :)\|_2^2}{2\sigma^2}} \right\}.$$

Furthermore, if $\zeta(X) := [\zeta_1(X) \ \zeta_2(X) \ \cdots \ \zeta_r(X)]$, then for any X there exists a real-valued scalar $\kappa \geq 0$ such that

$$L_\sigma[\kappa \zeta(X) + (1 - \kappa)X] \leq L_\sigma(X). \quad (12)$$

Since $L_\sigma(X)$ decreases along the direction $\zeta(X) - X$, we minimize L_σ using $\zeta(X) - X$ as the descent direction, and a local convergence is guaranteed. JLZA algorithm [13] in Table I is a sequential optimization procedure, where we repeatedly minimize L_σ while decreasing the value of σ by a factor ρ after each minimization. We initialize at $X_*(\infty)$ in (10). Simulations experiments suggest that by taking ρ large (close to one) we may expect to avoid the local minima, and track the path $X_*(\sigma)$ with σ varying from a large value to σ_0 . The parameter λ depends on the noise level. See [14] for a discussion on methods for choosing λ . For our experiments we set $\sigma_0 = 0.01$ and $\lambda = 10$.

As can be seen in Table I, the major fraction of the computation is involved in computing $\zeta(X)$. However using the matrix inversion lemma in (11) one can verify that

$$\zeta_k(X) = W^{-1} \Psi^* \Theta_k^* [I/(\lambda \sigma^2) + \Theta_k \Psi W^{-1} \Psi^* \Theta_k^*]^{-1} \mathbf{y}_k. \quad (13)$$

Since W is a diagonal matrix, it follows using (5) that

$$[\Psi W^{-1} \Psi^*]_{k,\ell} = \sum_{j=1}^N \mathbf{w}_j e^{i(k-\ell)\bar{\omega}_j - 1} = [\Psi \mathbf{w}]_{k-\ell},$$

where $\mathbf{w} = [1/W_{1,1} \ 1/W_{2,2} \ \dots \ 1/W_{N,N}]'$. Consequently, $\Psi W^{-1} \Psi^*$ is a Toeplitz matrix. Also, $\Psi \mathbf{w}$ can be computed efficiently using IFFT. As Θ_k is a row-selector, $R_k := \Theta_k \Psi W^{-1} \Psi^* \Theta_k^*$ is a just small block of $\Psi W^{-1} \Psi^*$. Next, we solve the $M \times M$ positive definite system of equations

$$[R_k + I/(\lambda \sigma^2)] z_k = \mathbf{y}_k. \quad (14)$$

Subsequently, to compute $\zeta_k(X)$ in (13), we calculate $\hat{z}_k = \Psi^*(\Theta_k^* z_k)$ by computing the FFT of $\Theta_k^* z_k$. Note that forming $\Theta_k^* z_k$ does not require any multiplication. Finally, as W is diagonal, we need N multiplications to compute $W^{-1} \hat{z}_k$. Note that for uniform sampling $R_k + I/\kappa$ is Toeplitz, and we can solve (14) using a fast Toeplitz system solver.

IV. SIMULATION RESULTS

Our simulations are performed using MATLAB7 running under Mac OS X on a Intel Core 2 Duo, 2.8 GHz processor with 2GB RAM. We compare the proposed methods with a Welch-type LS-MSC estimator [3] and SIAA-MSC [6]. The settings recommended in [6] are used for MSC estimators. Two different values for the filter length L is used for SIAA-MSC, with 12 iterations for better results. The CVX interface [15] is used for $\ell_{2,1}$. The measurement noise is taken as a zero-mean white complex Gaussian process with a variance μ^2 . The value of μ is different in different simulations. In all simulations we set $|a_{k,\ell}| = 1$, for $\forall k, \ell$, and for this setting the frequency-wise signal to noise ratio (SNR) in dB is given by $\text{SNR} = -10 \log_{10}(\mu^2)$. The phases of $a_{k,\ell}$'s are iid random variables, uniform on $[0, 2\pi)$. The sampling instants $t_{k,i}$ are drawn uniformly at random from the set $\{0, 1, \dots, R\}$, where R can vary. Similarly the frequencies $\{\Omega_k\}_{k=1}^n$ are drawn uniformly at random from the set $\{2\pi k/1000, k = 1, \dots, 999\}$. In all experiments we set $N = 1024 = 2^{10}$ i.e., $\Phi \in \mathbb{C}^{M \times 1024}$. This allows using radix-2 FFT/IFFT for computing $\zeta_k(X)$. Note that none of the elements of $\{\Omega_k\}_{k=1}^n$ lie on the grid $\mathbb{G}$. The experiments where $T_1 = T_2 = \dots = T_r$, are denoted by 'E', e.g. SIAA-MSC:E. We use 'D' to indicate no two T_k 's are the same, e.g. SIAA-MSC:D.

Fig. 1, shows typical frequency spectrums for $n = 14$, $r = 2$, $M = 50$, $\text{SNR} = 5$ dB, and $R = 199$ obtained from a realization. Although SIAA-MSCs can resolve some of the frequencies, they also generate many spurious peaks. Both $\ell_{2,1}$ and EIJLZA can detect all frequencies. However, $\ell_{2,1}$ produces two spurious peaks at $0.9102 \times 2\pi$ and $0.9502 \times 2\pi$ rad/sec (in about 40% cases $\ell_{2,1}$ produces spurious peaks in this setting). In Fig. 2 (a), we examine the resolution performance of different algorithms. Here $n = 10$ and $\Omega_1 = 2\pi \times 0.05$ rad/sec.

Fig. 1: A typical frequency spectrum. $M = 50$, $\text{SNR} = 5$ dB and $n = 14$.

Fig. 2: Resolution performance of different algorithms for $M = 50$, $\text{SNR} = 5$ dB, $n = 10$. (a) Probability of detecting the frequencies as a function of $|\Omega_1 - \Omega_2|/(2\pi)$ with $R = 199$. (b) $|\Omega_1 - \Omega_2|/(2\pi)$ as a function of R .

We vary Ω_2 and plot the empirical probability of correctly detecting all ten frequencies as a function of $|\Omega_1 - \Omega_2|/(2\pi)$.

The plot is based on 100 Monte-Carlo simulations, where in each simulation the 8 frequency locations $[\Omega_3, \dots, \Omega_{10}]$, the noise realization and the phases of $a_{k,\ell}$'s are varied. For a signal with n frequency components, we say the frequencies are detected correctly when an algorithm produces n spikes,

TABLE II: Computation time comparison. SNR=5 dB, $r = 2$, $n = 10$, $R = 199$.

M	Time (sec)			
	EJLZA	$\ell_{2,1}$	SIAA: L=0.6M	SIAA: L=0.8M
30	0.033	5.8	16.41	10.1
40	0.038	7.98	24.64	15.33
50	0.045	11.15	34.59	21.78

and the maximum absolute difference between any actual frequency and recovered frequency is $0.001/\pi$ rad/sec. We say a location in the spectrum is a spike when the location is a local maximum on the spectral plot and its absolute magnitude is above a threshold ϵ . The value of $\epsilon = 0.01$ for $\ell_{2,1}$ and EJLZA. However, SIAA generates many spurious peaks and hence we set $\epsilon = 0.1$ for SIAA. Fig. 2 (b), shows the frequency resolution achievable by different algorithms as a function of R . In this setting, $\Omega_1 = 0.05 \times 2\pi$ and Ω_2 is varied to control $|\Omega_1 - \Omega_2|$. The minimum $|\Omega_1 - \Omega_2|$ for which an algorithm detects all 10 frequencies with an empirical probability of 80% is recorded as a function of R . The resolution performance increases with increasing r . Surprisingly, the performance of SIAA degrades as R increases, particularly when $R > 250$. Fig. 3(a) shows the empirical probability of detecting the frequencies as a function of n , the number of frequencies in the signals. Here $R = 199$, $M = 50$, and SNR = 5 dB. The plots are based on 100 Monte-Carlo simulations with the frequency locations, noise-realization and phases of $a_{k,\ell}$ varying in each simulation. Also the criterion used for deciding whether the frequencies have been detected correctly remains as described above. As expected, the detection improves with increasing r . Moreover, when T_k 's in (2) are different, all algorithms give improved results. Note that $\ell_{2,1}$ and EJLZA outperform other algorithms, with EJLZA being somewhat better than $\ell_{2,1}$. Fig. 3(b) shows the maximum number of frequencies detected with an empirical probability of 80% by different algorithms as a function of SNR. In this setup, we increase n for a fixed SNR and apply different algorithms. As before, EJLZA outperforms all algorithms. Finally, as can be seen in Table II, EJLZA is at least 200 times faster than the other methods. Also $\ell_{2,1}$ is faster than SIAA.

V. CONCLUSION

Coherent spectral estimation using joint sparse representation is proposed in this work, which lead to a significant performance improvement at a much lower computational cost. In particular, the EJLZA algorithm outperforms other state-of-the-art approaches.

REFERENCES

- [1] P. Stoica and R. Moses, *Introduction to Spectral Analysis*, 2nd ed. Prentice-Hall, Upper Saddle River, NJ, 2004.
- [2] J. D. Scargle, "Studies in astronomical time series analysis. II. statistical aspects of spectral analysis of unevenly spaced data," *The Astronomical Journal*, vol. 263, pp. 835–853, 1982.
- [3] M. Schulz and K. Stattegger, "Spectrum: Spectral analysis of unevenly spaced paleoclimatic time series," *Comput. Geosci.*, vol. 23, pp. 929–945, 1997.
- [4] S. Bourguignon, H. Carfantan, and J. Idier, "A sparsity-based method for the estimation of spectral lines from irregularly sampled data," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 575–585, dec. 2007.

Fig. 3: Frequency recovery by different algorithms for $M = 50$, $R=199$. (a) Empirical probability of detecting frequencies, SNR = 5 dB, (b) Maximum number of frequency components is detected by different algorithms as a function of SNR.

- [5] G. Foster, "The cleanest fourier spectrum," *The Astronomical Journal*, vol. 109, pp. 1889–1902, April 1995.
- [6] N. Butt and A. Jakobsson, "Coherence spectrum estimation from nonuniformly sampled sequences," *IEEE Signal Processing Letters*, vol. 17, no. 4, pp. 339–342, april 2010.
- [7] D. Ramirez, J. Via, and I. Santamaria, "A generalization of the magnitude squared coherence spectrum for more than two signals: definition, properties and estimation," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, april 2008, pp. 3769–3772.
- [8] M. Hyder and K. Mahata, "An ℓ_0 norm based method for frequency estimation from irregularly sampled data," in *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2010, pp. 4022–4025.
- [9] Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," 2008. [Online]. Available: <http://www.citebase.org/abstract?id=oai:arXiv.org:0807.4581>
- [10] M. Hyder and K. Mahata, "A robust algorithm for joint-sparse recovery," *IEEE Signal Processing Letters*, vol. 16, no. 12, pp. 1091–1094, dec. 2009.
- [11] —, "An improved smoothed ℓ_0 approximation algorithm for sparse representation," *IEEE Transactions on Signal Processing*, vol. 58, no. 4, pp. 2194–2205, april 2010.
- [12] D. L. Donoho, M. Elad, and V. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Transactions on Information Theory*, vol. 52, pp. 6–18, Jan. 2006.
- [13] M. Hyder and K. Mahata, "Direction-of-arrival estimation using a mixed $\ell_{2,0}$ norm approximation," *IEEE Transactions on Signal Processing*, vol. 58, pp. 4646–4655, Sep 2010.
- [14] D. Malioutov, M. Cetin, and A. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 3010–3022, aug. 2005.
- [15] M. Grant and S. Boyd, "Cvx: Matlab software for diciplined convex programming," <http://stanford.edu/~boyd/cvx>, Feb 2009.