Compressive Direction-of-Arrival Estimation Off the Grid

Shermin Hamzehei
Department of Electrical and Computer Engineering
University of Massachusetts
Amherst, MA 01003
shamzehei@umass.edu

Marco F. Duarte
Department of Electrical and Computer Engineering
University of Massachusetts
Amherst, MA 01003
mduarte@ecs.umass.edu

***Abstract*—While most literature in compressive sensing mostly concentrates on recovering a sparse signal from a reduced number of measurements, parameter estimation problems have recently been studied under this acquisition framework. In this paper, we focus on the problem of direction-of-arrival (DOA) estimation from compressive measurements taken at each antenna in a receiver array. In contrast with the common assumption that the DOAs are contained within a grid to obtain sparsity, we consider a gridless setting for the parameter space and introduce two algorithmic approaches for this setup. The first approach leverages a parametric estimation algorithm to design a suitable denoiser to be used in approximate message passing. The second approach uses a multiple measurement vector model for a sequence of snapshots followed by the same parametric estimation algorithm applied on the estimated signals. Our experimental results show that the proposed algorithms can significantly outperform existing approaches in terms of the average DOA estimation error and the sparsity-undersampling tradeoff.**

***Index Terms*—Compressive Sensing, Direction of Arrival, Spectral Estimation, Approximate Message Passing, Multiple Signal Classification (MUSIC)**

I. Introduction

Compressive sensing (CS) has recently attracted significant attention in the field of signal processing. CS enables a reduction in the number of measurements needed to recover a signal by exploiting its sparsity [1, 2], i.e., the fact that the signal possesses only a few nonzero or significant coefficients in a suitable transform domain. Even though most existing work in this area focuses on signal recovery from few measurements, some literature in the field of parameter estimation using sparse signal models is available as well [3–8]. While such models assume that only a few nonzero values exist in a signal's representation, more recent models for compressive parameter estimation inspired by sparsity assume that a small number of parameters suffices to completely describe the signal of interest. Some examples of such sparse parameter estimation models include frequency estimation, localization, and bearing estimation [5–16]. In such applications, one does not aim to recover the signal itself, but rather to leverage the parametric model in order to identify the signal from a reduced number of measurements.

We focus on the specific application of bearing estimation, also known as direction of arrival (DOA) estimation. DOA estimation refers to the process of retrieving the angular location of several far-field electromagnetic sources from the outputs of a number of receiving antennas that form a sensor array. DOA estimation is an important problem in array signal processing and has a variety of applications including radar, sonar, and wireless communications [17].

Generally, sparse methods for DOA estimation are classified into three categories: *on-grid, off-grid*, and *gridless* [18]. In *on-grid* sparse methods, the DOAs are assumed to lie on a prescribed grid; therefore, the continuous DOA domain is replaced by a given set of grid points. Hence, grid selection is an important problem in the recovery methods from this category, which affects the practical DOA estimation accuracy, computational speed, and the theoretical analysis. For example, there is a high likelihood of mismatch between the adopted discrete grid point values and the true continuous DOAs. To resolve this problem, a new class of *off-grid* approaches to parameter estimation has been recently introduced, e.g., [8–13]. In these approaches, a grid is still required to perform sparse estimation; however, the DOA estimates are not restricted to be on the grid. Therefore, the samples included in the grid need to have sufficient density and coverage to allow for accurate off-grid estimation. Off-grid algorithms commonly rely on nonconvex optimization or interpolation, and hence can only provide local convergence guarantees. As an alternative to on- and off-grid approaches for sparse DOA estimation, *gridless* approaches do not require gridding of the DOA parameter space. This type of algorithms directly operate in the continuous parameter domain and, hence, completely resolve the grid mismatch problem. Since the resulting problems are convex, the algorithms provide strong theoretical guarantees. Nonetheless, they are only applicable to settings featuring measurements from uniform or sparse linear arrays.

In our previous work [16], we introduced a compressive parameter estimation approach based on approximate message passing (AMP) [19], a modified version of the traditional, widely-used iterative soft thresholding algorithm for CS recovery [20]. AMP obtains an estimate of the signal, polluted by additive white Gaussian noise (AWGN), in each of its iterations by leveraging an "Onsager" correction term in its formulation. The algorithm then runs this estimate through a soft thresholding step, which can be shown to be an optimal denoiser for sparse signals embedded in additive white Gaussian noise [19]. To solve the compressive parameter estimation problem, we replace this denoising step throughout the execution of AMP with what we call an *analog denoiser*: a concatenation of a statistical parameter estimation algorithm and a signal synthesis step [16, 21, 22].

In this paper, we propose two algorithms for sparse DOA estimation. Our first algorithm belongs in the gridless category and relies on the design of an analog denoiser for the DOA estimation problem to be integrated within AMP. Our second algorithm belongs in the off-grid category and uses a multiple measurement vector model for a sequence of snapshots to perform signal recovery, followed by the straightforward application of a DOA estimation algorithm on the recovery output. Our experimental results show that the proposed algorithms can significantly outperform existing approaches in terms of the average DOA estimation error.

This paper is organized as follows. Section II provides additional background. In Section III, we present our approaches to leverage parameter estimation algorithms for the compressive DOA estimation problem. In Section IV, we focus our study of these algorithms on aspects introduced by the distributed nature of sensing in sensor arrays. Section V presents experimental results indicating the performance of the proposed approaches. Finally, we provide conclusions and some suggestions for future work in Section VI.

II. Background and Related Work

A. Compressive Sensing

Consider a discrete-time K-sparse signal $\mathbf{x} \in \mathbb{C}^N$, i.e. $\mathbf{x}$ has at most K nonzero elements, and a column-normalized measurement matrix $\Phi \in \mathbb{C}^{M \times N}$ with independent and identically distributed elements chosen from a complex Gaussian distribution. Considering the measurement vector of the signal $\mathbf{y} = \Phi\mathbf{x} \in \mathbb{C}^M$, when $M \ll N$, we attempt to recover $\mathbf{x}$ from $\mathbf{y}$ given Φ. This can be done using algorithms based on optimization [23] (such as basis pursuit) or greedy iterative algorithms such as iterative soft/hard thresholding [20], which can be succinctly stated as follows:

$$\mathbf{x}^{t+1} = \eta_K(\Phi^H(\mathbf{y} - \Phi\mathbf{x}^t) + \mathbf{x}^t), \quad (1)$$

starting from $\mathbf{x}^0 = \mathbf{0}$. Here, $\mathbf{x}^t \in \mathbb{C}^N$ denotes the signal estimate at iteration t, and $\eta_K(.)$ is the corresponding soft/hard thresholding function that provides the optimal K-sparse approximation of the input signal, in terms of the Euclidean distance.

B. Approximate Message Passing

Recently, Donoho et al. suggested a modification in the traditional iterative soft thresholding algorithm, adding an "Onsager" correction term to the iterative soft thresholding algorithm (1) [19]. The resulting first-order approximate message passing algorithm (AMP) proceeds as follows:

$$\mathbf{x}^{t+1} = \eta_K(\Phi^H\mathbf{z}^t + \mathbf{x}^t), \quad (2)$$

$$\mathbf{z}^t = \mathbf{y} - \Phi\mathbf{x}^t + \frac{1}{\delta}\mathbf{z}^{t-1}\langle\eta'_{t-1}(\Phi^H\mathbf{z}^{t-1} + \mathbf{x}^{t-1})\rangle, \quad (3)$$

where $\mathbf{z}^t$ denotes a residual, $\eta'_K(s) = \frac{\partial}{\partial s}\eta_K(s)$ is the entry-wise derivative of the soft thresholding function $\eta_K(.)$, $\delta = M/N$ is the measurement rate, and for a vector $\mathbf{u} = \left[u(1) \dots u(N)\right]$ we denote $\langle\mathbf{u}\rangle = \frac{1}{N}\sum_{i=1}^N u(i)$. It can be shown that the Onsager term added in (3) significantly reduces the number of measurements required for signal recovery with respect to iterative soft thresholding [19].

C. Denoising-Based AMP

The power of the Onsager correction term is that at each iteration of the AMP algorithm, the input to the thresholding step in (2) resembles in distribution the original signal $\mathbf{x}$ embedded in AWGN [19, 22]. In subsequent work, Donoho et al. have shown that one can replace the traditionally used iterative soft thresholding function at each iteration of the AMP algorithm with an optimal AWGN denoiser for the class of signals of interest, noting that soft thresholding provides such an optimal denoiser for sparse signals [22]. This fact enables us to infuse additional knowledge of the signal model and application in the recovery algorithm. However, the drawback is that high-performance denoisers are usually data-dependent, and therefore it might be impossible or highly complex to explicitly express the Onsager correction term for such denoisers. Fortunately, Metzler et al. have shown that one can leverage a Monte Carlo method in order to obtain a numerical estimate of the Onsager correction term for any denoiser suitable for AMP [21].

D. DOA Estimation

In DOA estimation, an array of P sensors (usually microphones or antennas) can record one or multiple targets transmitting a signal to the array at specific bearing angles [24]. Assume that the p^{th} antenna is located at the coordinates (u_p, v_p) and that the antennas are configured as a uniform linear array (e.g., $u_p = u_0 + pd_x$, where d_x is the array inter-element spacing, and $v_p = 0$ for all p). The $P \times 1$ array snapshot vector $\mathbf{x}(q) = [x_1(q)\ x_2(q)\ \dots\ x_P(q)]^T$, containing observations from all antennas at each time $q = 1, \dots, Q$, can be modeled as $\mathbf{x}(q) = \mathbf{S}(\theta)\mathbf{a}(q) + \mathbf{n}(q)$, where $\theta = \left[\theta_1\ \dots\ \theta_K\right]^T$ is the $K \times 1$ vector of the signal DOAs, $\mathbf{S}(\theta) = \left[\mathbf{s}(\theta_1)\ \dots\ \mathbf{s}(\theta_K)\right]$ is the $P \times K$ signal steering matrix, $\mathbf{a}(q) = \left[a_1(q)\ \dots\ a_K(q)\right]^T$ is the $K \times 1$ vector collecting the scalar amplitudes of the received transmissions, and $\mathbf{n}(q)$ is the $P \times 1$ vector of antenna noise. Each $P \times 1$ steering vector can be expressed as

$$\mathbf{s}(\theta) = \begin{bmatrix} \exp(-j\frac{P-1}{2}\frac{2\pi}{\lambda}d_x\sin\theta) \\ \exp(-j\frac{P-3}{2}\frac{2\pi}{\lambda}d_x\sin\theta) \\ \vdots \\ \exp(j\frac{P-1}{2}\frac{2\pi}{\lambda}d_x\sin\theta) \end{bmatrix} = \begin{bmatrix} z^{-(P-1)/2} \\ z^{-(P-3)/2} \\ \vdots \\ z^{(P-1)/2} \end{bmatrix}, \quad (4)$$

where $z = \exp(j(2\pi/\lambda)d_x\sin\theta)$ and λ is the signal wavelength. We collect the multiple observations into the matrix equation $\mathbf{X} = \mathbf{S}(\theta)\mathbf{A} + \mathbf{N}$, with $\mathbf{X} = [\mathbf{x}(1)\ \dots\ \mathbf{x}(Q)]$, $\mathbf{A} = [\mathbf{a}(1)\ \dots\ \mathbf{a}(Q)]$, and $\mathbf{N} = [\mathbf{n}(1)\ \dots\ \mathbf{n}(Q)]$.

It is clear from (4) that the steering vectors $\mathbf{s}(\theta_k)$ will correspond to uniformly sampled complex exponentials with frequencies $f_k = \frac{d_x}{\lambda} \sin\theta_k$, and so the angles $\{\theta_k\}_{k=1}^K$ can be obtained by identifying the frequencies for the complex exponential components of the received (noisy) signals $\mathbf{x}(q)$. This frequency identification problem is well known in the signal processing literature as the line spectral estimation problem, for which many popular estimation algorithms exist [25].

Since the DOAs are not known in advance, it is common to pose a steering matrix or dictionary $\mathbf{S}$ corresponding to a sampling of the DOA parameter space instead of the generating matrix $\mathbf{S}(\theta)$. Since all the antennas are receiving signals from the same transmitters, and under the assumption that the observed DOAs are contained in the samples gathered in $\mathbf{S}$, the coefficient vector $\mathbf{a}(q)$ becomes a sparse vector $\mathbf{a}$ and the collected matrix $\mathbf{A}$ becomes a row-sparse matrix. Thus, when CS is applied to the measurements of each antenna, we have

$$\mathbf{Y} = \Phi\mathbf{X}^T = \Phi\mathbf{A}^T\mathbf{S}^T + \mathbf{W}, \tag{5}$$

where $\mathbf{Y} = \begin{bmatrix}\mathbf{y}(1) \ \dots \ \mathbf{y}(P)\end{bmatrix}$ and Φ and $\mathbf{W}$ are the sensing matrix and the measurement noise, respectively. It is worth noting that gridless methods for DOA from CS measurements will not require the design or use of a dictionary S.

III. Compressive DOA Estimation Algorithms

In this section, we consider the problem of DOA estimation from compressive measurements and leverage our prior work and related work described in Section II to formulate two alternative approaches for this problem that can be classified as gridless and off-grid, respectively.

A. Analog Denoiser for DOA Estimation

Previously, we studied the frequency estimation problem as an example of sparse parameter estimation leveraging analog denoisers [16]. An analog denoiser $\hat{\mathbf{x}} = \eta_{AD}(\mathbf{x})$ is a concatenation of a parameter estimation algorithm suitable for noisy observations of the given signal and a synthesis step for the corresponding parametric model. We note that parameter estimates are obtained as a byproduct of this analog denoising process. The proposed denoiser structure can also be applied to other parameter estimation problems as well. We create an analog denoiser for use within the AMP algorithm by leveraging an existing algorithm for DOA estimation (together with the transmitter magnitudes) from noisy observations as follows:

$$\{\hat{\theta}_k, \hat{\mathbf{a}}^k\}_{k=1}^K = \text{MUSIC}(\mathbf{X}, K), \tag{6}$$

$$\hat{\mathbf{X}} = \mathbf{S}(\hat{\theta})\hat{\mathbf{A}}. \tag{7}$$

Here, $\text{MUSIC}(\mathbf{X}, K)$ refers to the Root MUSIC algorithm [25] applied on the snapshots contained in $\mathbf{X}$, which estimates the DOAs $\{\hat{\theta}_k\}_{k=1}^K$ and the corresponding amplitude (column) vectors $\{\hat{\mathbf{a}}^k\}_{k=1}^K \in \mathbb{R}^Q$, and $\hat{\mathbf{A}} = \begin{bmatrix}\hat{\mathbf{a}}^1 \ \hat{\mathbf{a}}^2 \ \dots \ \hat{\mathbf{a}}^K\end{bmatrix}^T$. At each iteration of AMP, we leverage the above concatenation of the parametric DOA estimation step and the signal synthesizer as an analog denoiser $\hat{\mathbf{X}} = \eta_{AD}(\mathbf{X})$, noting that the estimates of the DOAs are obtained as a byproduct of the analog denoising process in each iteration. We also note that the Onsager correction term for the analog denoiser can be estimated using the numerical scheme described in Section II-C. We will refer to the resulting algorithm as AMP+MUSIC in the sequel.

B. Multiple Measurement Vector Recovery Model for DOA Estimation

Our second proposed method initially targets the recovery of the signal in the time domain, leveraging the multiple measurement vector (MMV) model [26]. In this method, the signal model and the measurements are given in (5). Recall that the matrix $\mathbf{A}$ is assumed to be row sparse, i.e., all the columns have the same sparse support due to the static locations of the transmitters throughout the data acquisition, which in turn fixes the frequencies present in each observed snapshot; nonetheless, the amplitudes of the transmitted signal may be different across snapshots to account for fluctuations in the magnitude of the transmitted signals.

The aforementioned model for the CS observations allows us to pose a simple off-grid compressive DOA estimation algorithm. In this method, a group ℓ_1-norm minimization algorithm (Gℓ_1) can be applied to recover the coefficient matrix $\hat{\mathbf{A}}$ from the measurements $\mathbf{Y}$ [27, 28]. In group ℓ_1-norm minimization, we assume that the matrix $\mathbf{A}$ containing the sparse representation coefficients for multiple signals will be row sparse. We then estimate the coefficient matrix $\tilde{\mathbf{A}}$ via the optimization

$$\hat{\mathbf{A}} = \underset{\tilde{\mathbf{A}}}{\text{argmin}} \|\tilde{\mathbf{A}}\|_{2,1}, \ \text{s.t.} \ \mathbf{Y} = \mathbf{\Phi}\tilde{\mathbf{A}}^T\mathbf{S}^T,$$

where $\|\mathbf{A}\|_{2,1}$ denotes the mixed $(2,1)$ matrix norm for $\mathbf{A}$ and is equal to the sum of the ℓ_2 norms of the rows of $\mathbf{A}$. Recall that the DOAs observed in $\mathbf{X}$ may not correspond to the samples gathered by the dictionary $\mathbf{S}$. Thus, once the estimate $\hat{\mathbf{A}}$ is obtained, a DOA parameter estimation algorithm (such as Root MUSIC, cf. (6)) is applied on the recovered signal $\hat{\mathbf{X}} = \hat{\mathbf{S}}\hat{\mathbf{A}}$ to estimate the DOAs and determine the location of each transmitter; this assumes that $\hat{\mathbf{X}} \approx \mathbf{X}$. We will refer to this approach as Gℓ_1→MUSIC in the sequel.

IV. Analog Denoisers in Distributed Sensing

The introduction of distributed acquisition settings brings additional difficulties to the integration of analog denoisers within AMP. As an example, in the DOA estimation setup of Section II, it is natural to assume that each antenna will perform CS only of the samples it acquires, e.g., those contained in one row of the $P \times Q$ matrix $\mathbf{X}$, cf. (5). This assumption is applied in existing work integrating DOA estimation and CS [4, 6, 29]. One can vectorize the matrix equation (5) by stacking the columns of the measurement matrix $\mathbf{Y}$ into a single column vector $\bar{\mathbf{y}} \in \mathbb{R}^{PM}$ and the transposed rows of the signal matrix $\mathbf{X}$ (e.g., the observations from each of the antennas) sequentially into a single column vector $\bar{x} \in \mathbb{R}^{PQ}$. The distributed acquisition process can then be written in terms of the equation $\bar{\mathbf{y}} = (\mathbf{I} \otimes \Phi)\bar{\mathbf{x}}$, where the Kronecker product $\mathbf{I} \otimes \Phi$ represents a block-diagonal matrix containing P copies of the CS matrix Φ in the diagonal. The structure of this matrix encodes the dependence of each measurement on samples obtained only by a single antenna, and has been studied extensively in the context of distributed CS [26, 30, 31]. The resulting block-diagonal matrix stands in contrast with that assumed in the formulation and initial analysis of the AMP algorithm, which is the standard random matrix with independent and identically distributed (i.i.d.) Gaussian entries [19]. Nonetheless, we see experimentally that despite the mismatch in the matrix model used, the use of analog denoisers still provides significant performance advantages in CS when compared to methods based on discrete signal models or on standard subsampling.

V. Experimental Results

We test the performance of several DOA estimation algorithms for signals acquired via CS. We consider a setup with $P = 128$ antennas (e.g., beamforming in massive MIMO) recording observations of length $Q = 128$ for each antenna via CS, where the same measurement matrix Φ having M rows is used in each of the antennas (i.e., each antenna records the same number of CS measurements) with i.i.d. entries following a zero-mean Gaussian distribution with variance $\sigma^2 = 1/M$. We measure the DOA estimation error by computing the cost of the Hungarian matching between the vectors containing the bearing angle values and their estimates. In our experiments, we compare the performance of AMP+MUSIC and Gℓ_1→MUSIC to that of three alternative baselines: (i) simultaneous recovery of all snapshots using ℓ_1-norm minimization followed by standard DOA estimation (ℓ_1→MUSIC); (ii) band-excluding interpolating subspace pursuit (BISP) [8], a coherence-controlling sparsity-based algorithm; and (iii) subsampling, i.e., acquisition from M antennas with $Q = 128$ snapshots, followed by standard DOA estimation. Note that no CS takes place in this last case. For the algorithms requiring a sparsity dictionary $\mathbf{S}$, we build a parametric dictionary containing antenna observations for transmitters located at various angles $\theta = \Delta \cdot i$, with $\Delta = 0.5°$ and $i = -\frac{90°}{\Delta}, \ldots, \lceil \frac{90°}{\Delta} \rceil - 1$.

Our first experiment generates a phase transition plot for DOA estimation, inspired by the recovery-based counterparts from [22, 32] and mimicking that introduced in [16] for line spectral estimation. The phase transition plot of a given recovery algorithm finds the maximum value of the normalized sparsity[1] for which the algorithm successfully recovers a sparse signal at least 50% of the time for a set of signals drawn at random from a uniform distribution over K-sparse signals from the continuous model. The plot is usually interpreted as showing the division between the (δ, ρ) region for which the success probability goes to one as $Q \to \infty$ (below the curve) from the (δ, ρ) region for which the success probability goes to zero as $Q \to \infty$ (above). Thus, curves with higher values of ρ for a given value of δ are better.

[1] Note that since the number of resolvable transmitters (i.e., the sparsity K) is upper bounded by the number of antennas P, we do not normalize the sparsity by the total number of measurements MP as usually done in phase transition plots.

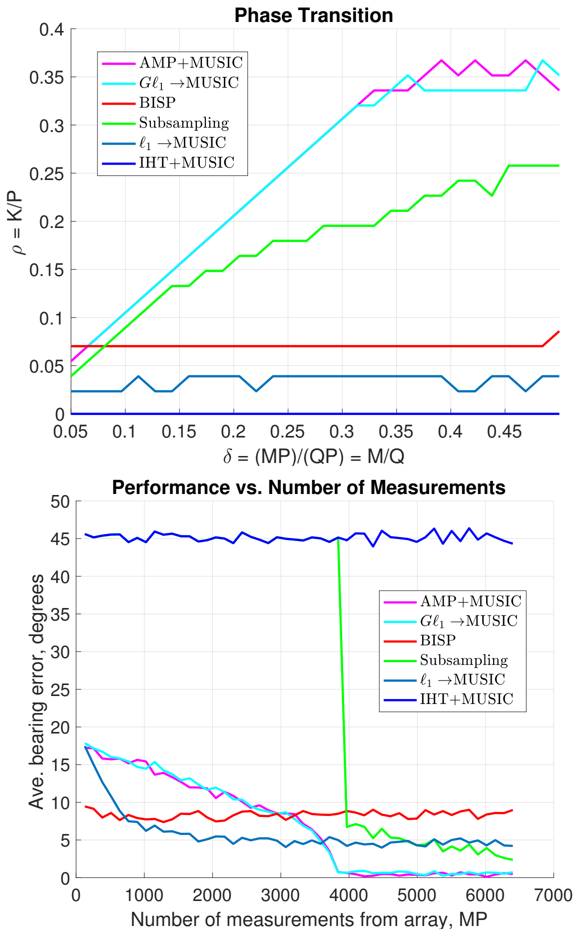

Fig. 1: Top: *Phase transition plot for compressive line spectral estimation. The line shows the maximum value of the sparsity ratio $\rho = K/P$ for which at least 50% of the trial DOA estimations under the measurement rate $\delta = M/Q$ are successful (i.e., within $1°$) for each compressive DOA estimation algorithm. The performance of AMP+MUSIC and Gℓ_1→ MUSIC is significantly better than that of all baseline counterparts.* Bottom: *Average frequency estimation error for several compressive DOA estimation algorithms for $K = 30$ targets. Once again, the performance of the proposed algorithms is significantly better. than that of its baseline counterparts. Note that since there are $K = 30$ transmitters in this experiment, we do not expect good performance (low average bearing estimation error) for values of $M < 30$, i.e., $MP < 3840$. As seen in the results, this intuition is in consistence with the numerical experiments.*

For the compressive DOA estimation algorithms' phase transition plots, we define success as having an average DOA estimation error (over the K bearing angles) of up to $1°$. For each value of the (δ, ρ) duplet, we execute 100 trials with randomly drawn bearing angles (uniformly at random in $[-90°, 90°)$, with arbitrary resolution), amplitudes (uniformly at random in $[0, 1]$), and measurement matrices. Fig. 1 (top) shows the DOA estimation phase transition for our proposed algorithms and the aforementioned baselines, where AMP+MUSIC and Gℓ_1→MUSIC achieve noticeably better performance, i.e., much higher ρ for each value of δ.

Our second experiment compares the performance of the different algorithms among randomly drawn signals under the same probability model as the first experiment. We repeat the setup from our phase transition experiment while fixing the number of emitters to $K = 10$, and evaluate the average DOA estimation error as a function of the number of measurements from the array MP over the same 100 trials for each of the compressive parameter estimation algorithms. Fig. 1 (bottom) shows that the performance of the proposed algorithms is significantly improved over those of its baseline counterparts.

VI. Conclusions and Future Work

In this paper, we studied the problem of DOA estimation from sparse measurements, while considering uniform linear array setup for the receiving antennas. In order to recover the unknown DOAs, we proposed two approaches: the first approach leverages the use of line spectral estimation to implement an analog denoiser within the AMP algorithm, obtaining parameter estimates as a byproduct of denoising. The second approach uses a group ℓ_1-norm minimization algorithm to exploit the fact the the matrix of snapshots is row-sparse since each antenna should receive information from transmitters from the same locations; we then perform standard parametric estimation on the recovered signals.

Our experimental results indicate that the proposed algorithms outperform those available in the literature, both from the aspects of phase transition and average recovery error. This is particularly surprising for our second approach, since the recovery step used there relies on a gridding of the DOA parameter space. We expect further work in the direction of compressive DOA estimation to focus on whether the performance guarantees available for AMP can translate to the proposed AMP-based compressive parameter estimation algorithms. Additionally, it would be interesting to pursue an analytical study of the effects of distributed sensing on the performance of the proposed algorithms.

Acknowledgement

We thank Dror Baron, Yanting Ma, and Junan Zhu for many helpful comments on our work.

References

[1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Info. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[2] E. J. Candès, "Compressive sampling," in *Int. Congress of Mathematicians*, vol. 3, Madrid, Spain, Aug. 2006, pp. 1433–1452.

[3] M. F. Duarte and R. G. Baraniuk, "Spectral compressive sensing," *Appl. Comput. Harmon. Anal.*, vol. 35, no. 1, pp. 111–129, 2013.

[4] V. Cevher, A. C. Gurbuz, J. H. McClellan, and R. Chellappa, "Compressive wireless arrays for bearing estimation," in *IEEE Int. Conf. Acoustics, Speech, and Signal Proc. (ICASSP)*, Las Vegas, NV, Apr. 2008, pp. 2497–2500.

[5] V. Cehver, M. F. Duarte, and R. G. Baraniuk, "Distributed target localization via spatial sparsity," *European Signal Processing Conference*, pp. 1–5, Aug. 2008.

[6] M. F. Duarte, "Localization and bearing estimation via structured sparsity models," *IEEE Statistical Signal Processing Workshop (SSP)*, pp. 333–336, 2012.

[7] D. Mo and M. F. Duarte, "Compressive parameter estimation with earth mover's distance via k-median clustering," in *Wavelets and Sparsity XV*, ser. Proc. SPIE, vol. 8858, San Diego, CA, Aug. 2013.

[8] K. Fyhn, H. Dadkhahi, and M. F. Duarte, "Spectral compressive sensing with polar interpolation," in *IEEE Int. Conf. Acoustics, Speech, and Signal Proc. (ICASSP)*, Vancouver, BC, May 2013, pp. 6225–6229.

[9] Z. Tan, P. Yang, and A. Nehorai, "Joint sparse recovery method for compressed sensing with structured dictionary mismatches," *IEEE Trans. Signal Proc.*, vol. 62, no. 19, pp. 4997–5008, 2014.

[10] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse bayesian inference," *IEEE Trans. Signal Proc.*, vol. 16, no. 1, pp. 38–43, 2013.

[11] J. Fang, J. Li, Y. Shen, and H. Li, "Super-resolution compressed sensing: An iterative reweighted algorithm for joint parameter learning and sparse signal recovery," *IEEE Signal Proc. Letters*, vol. 21, no. 6, pp. 761–765, 2014.

[12] J. Fang, F. Wang, Y. Shen, H. Li, and R. Blum, "Super-resolution compressed sensing for line spectral estimation: An iterative reweighted approach," *IEEE Trans. Signal Proc.*, vol. 64, no. 18, pp. 4649–4662, 2016.

[13] H. Zhu, G. Leus, and G. B. Giannakis, "Sparsity-cognizant total least-squares for perturbed compressive sampling," *IEEE Trans. Signal Proc.*, vol. 59, no. 5, pp. 2002–2016, 2011.

[14] C. Feng, S. Valaee, and Z. Tan, "Multiple target localization using compressive sensing," in *Global Telecommunications Conf. (GLOBECOM)*, Honolulu, HI, Dec. 2009, pp. 1–6.

[15] S. Safavi and U. A. Khan, "Localization and tracking in mobile networks: Virtual convex hulls and beyond," 2015, preprint. Available at http://arxiv.org/1504.01207/.

[16] S. Hamzehei and M. F. Duarte, "Compressive parameter estimation via approximate message passing," in *IEEE Int. Conf. Acoustics, Speech, and Signal Proc. (ICASSP)*, Brisbane, Australia, Apr. 2015, pp. 3327–3331.

[17] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Simon and Schuster, 1992.

[18] Z. Yang, J. Li, P. Stoica, and L. Xie, "Sparse methods for direction-of-arrival estimation," Sep. 2016, preprint. Available at http://arxiv.org/1609.09596.

[19] D. L. Donoho, A. Maleki, and A. Montanari, "Message passing algorithms for compressed sensing," *Proceedings of the National Academy of Sciences*, vol. 106, no. 45, pp. 18 914–18 919, 2009.

[20] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 948–958, 2010.

[21] C. A. Metzler, A. Maleki, and R. G. Baraniuk, "From denoising to compressed sensing," *IEEE Trans. Info. Theory*, vol. 62, no. 9, pp. 5117–5144, Sep. 2016.

[22] D. L. Donoho, I. Johnstone, and A. Montanari, "Accurate prediction of phase transitions in compressed sensing via a connection to minimax denoising," *IEEE Trans. Info. Theory*, vol. 59, no. 6, pp. 3396–3433, June 2013.

[23] M. A. Davenport, M. F. Duarte, Y. C. Eldar, and G. Kutyniok, "Introduction to compressive sensing," in *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2011.

[24] D. Johnson and D. Dudgeon, "Array signal processing: Concepts and techniques," *Prentice-Hall Inc.*, 1993.

[25] P. Stoica and R. L. Moses, "Spectral analysis of signals," *Prentice Hall*, 2005.

[26] M. F. Duarte, M. B. Wakin, D. Baron, S. Sarvotham, and R. G. Baraniuk, "Measurement bounds for sparse signal ensembles via graphical models," *IEEE Trans. Info. Theory*, vol. 59, no. 7, pp. 4280–4289, 2013.

[27] L. Meier, S. V. D. Geer, and P. Bühlmann, "The group lasso for logistic regression," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 70, no. 1, pp. 53–71, 2008.

[28] J. Friedman, T. Hastie, and R. Tibshirani. (2010) A note on the group lasso and a sparse group lasso. [Online]. Available: arXiv:1001.0736

[29] A. C. Gurbuz, V. Cevher, and J. H. McClellan, "A compressive beamformer," in *IEEE Int. Conf. Acoustics, Speech, and Signal Proc. (ICASSP)*, Las Vegas, NV, Apr. 2008, pp. 2617–2620.

[30] J. Y. Park, H. L. Yap, C. J. Rozell, and M. B. Wakin, "Concentration of measure for block diagonal matrices with applications to compressive signal processing," *IEEE Trans. Signal Proc.*, vol. 59, no. 12, pp. 5859–5875, 2011.

[31] A. Eftekhari, H. L. Yap, C. J. Rozell, and M. B. Wakin, "The restricted isometry property for random block diagonal matrices," *Appl. Comput. Harmon. Anal.*, vol. 38, no. 1, pp. 1–31, 2015.

[32] D. L. Donoho and J. Tanner, "Precise undersampling theorems," *Proc. IEEE*, vol. 98, no. 6, pp. 913–924, 2010.