This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LCOMM.2016.2587674, IEEE Communications Letters

An Iterative Dictionary Learning-based algorithm For DOA Estimation

Hojatollah Zamani, Hadi Zayyani, and Farrokh Marvasti, *Senior Member, IEEE*

***Abstract*—This letter proposes a dictionary learning algorithm for solving the grid mismatch problem in direction of arrival (DOA) estimation from both the array sensor data and from the sign of the array sensor data. Discretization of the grid in the sparsity-based DOA estimation algorithms is a problem which leads to a bias error. To compensate this bias error, a dictionary learning technique is suggested which is based on minimizing a suitable cost function. We also propose an algorithm for estimation of DOA from the sign of the measurements. It extends the Iterative Method with Adaptive Thresholding (IMAT) algorithm to the one bit compressed sensing framework. Simulation results show the effectiveness of the dictionary learning based algorithms in comparison to the counterpart algorithms in DOA estimation both from the sensors' data and from the sign of the sensors' data.**

***Index Terms*—Direction of Arrival, Compressed sensing, Sign of the measurements, Dictionary learning, Steepest-descent.**

I. Introduction

DIRECTION of arrival (DOA) estimation is a well-known problem which has many applications in wireless communications [1], radar [2] and sonar [3]. Some of the classical algorithms for DOA estimation are conventional beamformer [4], Minimum Variance Distortionless Response (MVDR) [5] and MUSIC [6]. Sparsity-based algorithms are also suggested for DOA estimation which exploit the spatial sparsity of the sources in a discrete grid [7]-[10]. These algorithms suffer from the problem of grid mismatch, i.e. the true DOAs are not on the discretized sampling grid. To deal with this problem, an off-grid DOA estimation is suggested which uses sparse Bayesian inference [11]. [12] uses the errors in variables (EIV) model which treats the grid mismatch as an additive error matrix and proposes a block sparse estimator for grid matching and sparse recovery. A super-resolution compressed sensing algorithm is also suggested for joint parameter learning and sparse signal recovery [13].

The first contribution of this letter is to suggest to use a dictionary learning algorithm to solve the grid mismatch problem for DOA estimation from the measurements of the array sensor output. It is performed by an alternate dictionary update plus a sparse recovery algorithm. This idea is not new and the similar idea is also used in [13], [14] and [15]. The current work and [13] have different update strategies. The current work uses a two-stage alternating optimization scheme, whereas [13] uses an iterative reweighted scheme for joint dictionary and sparse signal refinement. The latter scheme is less likely to be stuck in undesirable local minima, as pointed out in [13]. A Total Least Squares (TLS) algorithm is used for sparse recovery in [14], where perturbation appear in both data vector as well as in the regression matrix. [15] solves the problem of basis mismatch by an iterative, biconvex search algorithm. In contrast to these counterpart algorithms, our algorithm uses a simple steepest-descent algorithm for dictionary update. In addition, we use an Iterative Method with Adaptive Thresholding (IMAT) [16], [17] for the sparse recovery step.

H. Zamani is with the Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran (e-mail: hzamani@ee.sharif.edu).

H. Zayyani is with the Department of Electrical and Computer Engineering, Qom University of Technology, Qom, Iran (e-mail: zayyani@qut.ac.ir).

F. Marvasti is with the Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran (e-mail: marvasti@sharif.edu).

On the other hand, the one bit compressed sensing has been extensively investigated recently [18]-[21]. According to compressed sensing (CS) theory, a sparse signal can be reconstructed from a number of linear measurements which could be much smaller than the signal dimension [22]-[23]. In the one bit compressed sensing framework, it is proved that an accurate and stable recovery can be achieved by using only the sign of linear measurements [19].

The main novelty and second contribution of this paper is to estimate the DOAs using the sign of the measurements of an array output. This is done previously in [24] before the advent of the one bit compressed sensing framework. [24] presented two one bit DOA estimator based on covariance matrix reconstruction and empirical probability estimation. Both of these two methods need multiple snapshots of array output data. On the other hand, in this letter, we use the one bit compressed sensing framework for DOA estimation which needs only one snapshot of the array output data. At first, a sparse recovery algorithm called IMAT [16], [17] is generalized to use for one bit DOA estimation. Similar to sparsity-based algorithms for DOA estimation, this algorithm suffers from the problem of grid mismatch which means that the true DOAs are not on the discretized sampling grid. Therefore, secondly, an iterative dictionary learning algorithm is used for solving the problem of grid mismatch.

Simulation results both in DOA estimation from the array sensor output or from the sign of the measurements of the array sensor output, show that the dictionary learning can improve the accuracy of DOA estimations.

II. System model

To obtain the general model of DOA estimation, consider K sources in direction angles of $\theta_k, k = 1, \cdots, K$ in far-field impinging independent narrowband signals $s_k(t), k = 1, \cdots, K$ into an array in an isometric environment. The array include M omni-directional sensor placed in a line with uniform distribution known as Uniform Linear Array (ULA). The output vector of the array $\mathbf{y}(t) = [y_1(t), \cdots, y_M(t)]^T$ at each time snapshot t can be modeled as:

$$\mathbf{y}(t) = \tilde{\mathbf{A}}(\theta)\tilde{\mathbf{s}}(t) + \mathbf{n}(t) \tag{1}$$

where $\tilde{\mathbf{s}}(t) = [s_1(t), \cdots, s_K(t)]^T$ is the source vector and $\mathbf{n}(t) = [n_1(t), \cdots, n_M(t)]^T$ is the sensor array noise vector. The array manifold matrix is $\tilde{\mathbf{A}}(\theta) = [\mathbf{a}(\theta_1), \cdots, \mathbf{a}(\theta_K)]_{M\times K}$ and $\mathbf{a}(\theta_k) = [1, e^{-j\frac{2\pi}{\lambda}d\sin(\theta_k)}, \cdots, e^{-j\frac{2\pi}{\lambda}(M-1)d\sin(\theta_k)}]^T$ is the steering vector which provides the delay information of the kth source to the all sensors based on the geometry of the array. The parameter d is the distance between adjacent elements and $\lambda = \frac{c}{f}$ represents the wavelength corresponding to frequency f, and c is the velocity of wave propagation. The array manifold $\tilde{\mathbf{A}}(\theta)$ include K columns of steering vectors related to K sources. By discretizing the spatial space into finite angle points and settle the related steering vectors of nonexistent of sources angles into the array manifold, the extended array manifold is obtained and also by extending the vector $\mathbf{s}(t)$ by adding zeros corresponding to the nonexistent source angles, the sparse form of the problem is formulated as

$$\mathbf{y}(t) = \mathbf{A}(\theta)\mathbf{s}(t) + \mathbf{n}(t), \tag{2}$$

where $\mathbf{A}(\theta)$ is $M \times L$ extended array manifold and $\mathbf{s}(t)$ is $L \times 1$ extended source vector. L is the number of finite angle points in the grids such that $L \gg K$ and $\mathbf{s}(t)$ is K-sparse which means only K elements of it is nonzero.
The One-bit form of the (2) can be obtained by replacing the $\mathbf{y}(t)$ with $\mathbf{csgn}\{\mathbf{y}(\mathrm{t})\}$, i.e.

$$\mathbf{y}(t) = \mathbf{csgn}\left\{\mathbf{A}(\theta)\mathbf{s}(\mathrm{t}) + \mathbf{n}(\mathrm{t})\right\} \tag{3}$$

where $\mathbf{csgn}\{\mathbf{x}\} = \mathrm{sign}(\mathrm{Re}(\mathbf{x}))+\mathrm{jsign}(\mathrm{Im}(\mathbf{x}))$ is the complex sign of $\mathbf{x}$.

III. The Proposed Algorithms

In this section, we introduce a dictionary learning technique for grid mismatch problem of estimated DOAs for both formulation of (2) and (3). Also, we introduce an efficient method to estimate the DOAs based on the model (3).

A. Dictionary learning based algorithms

Most of DOA estimators need to search on the discretized grid and the problem is that the true angle may not fall into grid, therefore, the grid mismatch occurs. To solve the off-grid problem, we propose an iterative technique similar to dictionary learning algorithms [25]. The suggested algorithm includes two steps. At the first step, the dictionary $\mathbf{A}(\theta)$ is fixed and sparse vector $\mathbf{s}$ will be estimated by sparse recovery algorithms such as the well known Orthogonal Matching Pursuit algorithm (OMP) or IMAT [16], [17]. At the second step, we fix $\mathbf{s}$ and then update the dictionary $\mathbf{A}(\theta)$ or equivalently the angle vector θ. To update the dictionary, we propose to minimize the following cost function

$$\min_{\theta} \|\mathbf{y} - \mathbf{A}(\theta)\mathbf{s}\|_2^2 . \tag{4}$$

The cost function defined as $F(\theta) \triangleq (\mathbf{y} - \mathbf{A}(\theta)\mathbf{s})^H(\mathbf{y} - \mathbf{A}(\theta)\mathbf{s})$. Therefore, using the steepest decent, we can iteratively estimate the true DOA by learning the array manifold matrix $\mathbf{A}$. The steepest-descent iteration is

$$\theta^{(i+1)} = \theta^{(i)} - \mu\nabla_\theta F(\theta) \tag{5}$$

In the appendix A, it is proved that the final recursion for updating the estimated angle vector θ at iteration $i+1$ is

$$\theta^{(i+1)} = \theta^{(i)} - \mu\mathcal{R}\boldsymbol{e}\{c_0\mathbf{e}^H[\mathbf{B}(\theta^{(i)}) \odot \mathbf{A}(\theta^{(i)})]\mathbf{s}\} \tag{6}$$

where μ is the step size parameter, $c_0 = -j\frac{2\pi d}{\lambda}$, $\mathbf{e} = \mathbf{A}(\theta)\mathbf{s} - \mathbf{y}$ is the error, $\mathbf{B}$ is the derivative of $\mathbf{A}$ with respect to θ and is given in the appendix A, and $\odot$ is element-wise product.
In the one-bit framework, We use a similar two step algorithm. At the first step, the dictionary $\mathbf{A}(\theta)$ is fixed and sparse vector $\mathbf{s}$ will be estimated by one bit compressed sensing algorithms such as BIHT [19] or binary IMAT which will be suggested in section III-B. At the second step, we fix $\mathbf{s}$ and then update the dictionary $\mathbf{A}(\theta)$ or equivalently the angle vector θ. To update the dictionary, we propose to minimize the following cost function

$$\min_{\theta} \|\mathbf{y} - \mathbf{csgn}(\mathbf{A}(\theta)\mathbf{s})\|_2^2 . \tag{7}$$

the cost function defined as $H(\theta) \triangleq (\mathbf{y} - \mathbf{csgn}\{\mathbf{A}(\theta)\mathbf{s}\})^H(\mathbf{y} - \mathbf{csgn}\{\mathbf{A}(\theta)\mathbf{s}\})$. Also $\mathbf{csgn}\{\mathrm{x}\}$ is approximated by a continuous function as $\mathbf{csgn}\{\mathrm{x}\} \approx \mathrm{S}(\mathrm{Re}(\mathrm{x})) + \mathrm{jS}(\mathrm{Im}(\mathrm{x}))$ where the non-continuous sign function is approximated by a continuous S-shaped function $\mathrm{S}(x) = \frac{1-\exp(-x)}{1+\exp(-x)}$. This approximation is also used in [26] for similar purpose. In the appendix B, it is proved that the final recursive formula for updating the estimated angle is as follows

$$\theta^{(i+1)} = \theta^{(i)} - \mu\mathcal{R}\boldsymbol{e}\left\{c_0\mathbf{e}^H[\mathbf{B}(\theta^{(i)}) \odot \mathbf{A}(\theta^{(i)})]\mathbf{s} \odot \mathbf{d}(\theta^{(i)})\right\} \tag{8}$$

where μ is the step size parameter, $c_0 = -j\frac{2\pi d}{\lambda}$, $\mathbf{e} = \mathbf{csgn}\{\mathbf{A}(\theta)\mathbf{s}\} - \mathbf{y}$ is the error, $\mathbf{B}$ is the derivative of $\mathbf{A}$ with respect to θ and is given in the appendix A and $\mathbf{d}(\theta^{(i)})$ is the derivative vector defined in appendix B.

Therefore, the overall dictionary learning based algorithm for DOA estimation from the array output and from the sign of the array output is summarized in Algorithm 1 and Algorithm 2.

Algorithm 1: Dictionary learning based DOA estimation from the array sensor output

input :
  Array output $\mathbf{y} \in \mathbb{R}^{M\times 1}$ based on model (2)
output:
  DOA estimation $\widehat{\mathbf{s}} \in \mathbb{R}^L$
  Learned angle vector $\hat{\theta}$
$\mathbf{s}^{(0)} \leftarrow 0$;
$\theta^{(1)} \leftarrow \theta$;
for $i = 1$ **to** DL_Iter$_{max}$ **do**
  Fixed $\mathbf{A}(\theta^{(\mathbf{i})})$: $\mathbf{s}^{(i)} \leftarrow \mathrm{SparseRecovery}(\mathbf{y}, A(\theta^{(\mathbf{i})}))$;
  Fixed $\mathbf{s}^{(i)}$: Dictionary learning based on (6);
end
$\widehat{\mathbf{s}} \leftarrow \mathbf{s}^{(\mathrm{DL-Iter}_{max})}$;
$\widehat{\theta} \leftarrow \theta^{(\mathrm{DL-Iter}_{\mathbf{max}}+\mathbf{1})}$;

Algorithm 2: Dictionary learning based DOA estimation from the sign of the array sensor output

input :
  Array output $\mathbf{y} \in \mathbb{R}^{M\times 1}$ based on model (3)
output:
  DOA estimation $\widehat{\mathbf{s}} \in \mathbb{R}^L$
  Learned angle vector $\hat{\theta}$
$\mathbf{s}^{(0)} \leftarrow 0$;
$\theta^{(1)} \leftarrow \theta$;
for $i = 1$ **to** DL_Iter$_{max}$ **do**
  Fixed $\mathbf{A}(\theta^{(\mathbf{i})})$:
  $\mathbf{s}^{(i)} \leftarrow \mathrm{Onebit\_SparseRecovery}(\mathbf{y}, A(\theta^{(\mathbf{i})}))$;
  Fixed $\mathbf{s}^{(i)}$: Dictionary learning based on (8);
end
$\widehat{\mathbf{s}} \leftarrow \mathbf{s}^{(\mathrm{DL-Iter}_{max})}$;
$\widehat{\theta} \leftarrow \theta^{(\mathrm{DL-Iter}_{\mathbf{max}}+\mathbf{1})}$;

B. One-Bit IMAT algorithm

In order to solve the DOA estimation problem based on one-Bit model (3), we propose an iterative method namely one-Bit Iterative Method and Adaptive Thresholding (One-Bit IMAT or Binary IMAT). It extends the IMAT algorithm [16], [17] for sparse recovery to the case of one bit sparse recovery which finds the sparse vector $\mathbf{s}$ based on the sign of the data as in the model (3). Therefore, the nonzero positions of vector $\hat{\mathbf{s}}(t)$, are the estimated DOAs. The details of the algorithm are illustrated in Algorithm 3. The thresholding operator $T^k(.)$ is defined as [16], [17]:

$$T^k(z) = \begin{cases} 0 & |z| < \theta(k), \\ z & Othewise, \end{cases} \tag{9}$$

where $\theta(k) = \theta(0)\exp(-k\alpha)$ is the threshold sequence in which $\theta(0)$ is a large initial threshold and α is the decay factor [16], [17].

Algorithm 3: One Bit IMAT Algorithm

input :
 Array output $\mathbf{y} \in \mathbb{R}^{M\times 1}$
output:
 DOA estimation $\widehat{\mathbf{s}} \in \mathbb{R}^{L}$
$\mathbf{s}^0 \leftarrow 0$;
for $k = 1$ **to** Iter_{max} **do**
 $\mathbf{r} \leftarrow \mathbf{y} - \mathbf{csgn}\left\{\mathbf{A}\mathbf{s}^{\mathrm{k}-1}\right\}$;
 $\mathbf{s}^k \leftarrow T^k(\mathbf{s}^{k-1} + \lambda\mathbf{A}^H\mathbf{r})$;
end
$\widehat{\mathbf{s}} \leftarrow \mathbf{s}^{\mathrm{Iter}_{max}}$;

IV. Simulation Results

This section presents the simulation results. In the simulations, two experiments were performed to show the efficacy of the proposed dictionary learning based DOA estimation algorithms. At first experiment, DOAs are estimated from the array sensor output, while at the second experiment, they are estimated from the sign of the array sensor output. We considered three sources ($K = 3$) at angles $\theta_1 = -12.50°$, $\theta_2 = 43.85°$ and $\theta_3 = 76.80°$. The number of array elements are assumed to be $M = 25$. For the discrete grid, the angle interval $[-90°, 90°]$ is divided into 90 equal bins with the step of 2°. The sensor array noise vector $\mathbf{n}(t)$ is considered to be complex independent white Gaussian noise with zero mean. The Signal to Noise Ratio (SNR) is defined as $\mathrm{SNR(dB)} = 10\log(\frac{E\{|\mathbf{As}|^2\}}{E\{|\mathbf{n}|^2\}})$. For the performance metric, Mean Square Error (MSE) of estimated angles is used which is defined as $\mathrm{MSE} = \sqrt{\frac{1}{K}\sum_{i=1}^{K}(\theta_i - \hat{\theta}_i)^2}$ which are averaged over 1000 independent monte carlo runs. The number of iterations of the proposed dictionary learning based algorithm is selected as 100.

At the first experiment, we used IMAT algorithm [10] for the sparse recovery in our proposed dictionary learning based algorithm. Therefore, the performance of the Dictionary Learning based IMAT (DL-IMAT) is compared to MUSIC [6], OMP, ℓ^1-SVD [8], WSS-TLS[14] and IMAT [16], at different SNRs. We used five snapshots of the array output for all the algorithms because two of the competing algorithms need multiple snapshots. These two algorithms are MUSIC and ℓ^1-SVD. For the other three single snapshot based algorithms (proposed algorithm, OMP and IMAT), average of the estimated angles are used as the final estimation. For the proposed algorithm, the step size parameter is selected as $\mu = 0.00001$. For IMAT, we used the initial threshold equal to $\theta(0) = T_{\max} = 20$, the decay rate equal to $\alpha = 0.05$ and maximum number of 85 iterations. For ℓ^1-SVD, the parameter is selected high enough $\beta = 10$ to yield sharp peaks [8]. The maximum number of iterations of OMP is equal to the number of sources which is $K = 3$. The results are shown in Fig 1. It shows that the proposed dictionary learning based algorithm enhances the performance of DOA estimation.

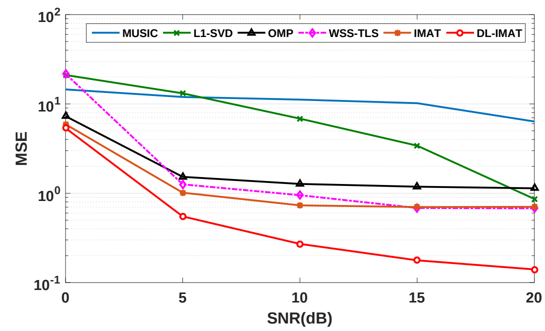

Fig. 1. MSE versus SNR for DOA estimation from the array sensor output.

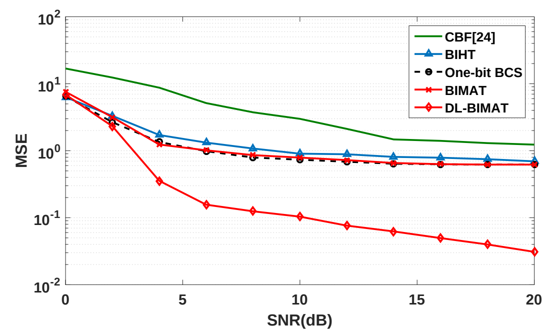

Fig. 2. MSE versus SNR for DOA estimation from the sign of the array sensor output.

At the second experiment, the performance of the Dictionary Learning based Binary IMAT (DL-BIMAT) is compared to BIHT [19], BIMAT without dictionary learning (proposed in this paper), One bit BCS [21] and Conventional BeamFormer (CBF) [24] for different SNRs. Single snapshot is used for all the algorithms. For the proposed algorithm, the step size parameter is selected as $\mu = 0.0001$. The maximum number of iterations of BIHT is equal to 100. For BIMAT, we used the initial threshold $\theta(0) = T_{\max} = 30$ and decay factor $\alpha = 0.1$ and maximum number of iterations is equal to 100. For one bit BCS, the parameters are selected as $\mathrm{Tol} = 1e^{-6}$ and maximum number of iterations equal to 500 as suggested in [21]. The results are illustrated in Fig. 2. It shows that the proposed dictionary learning based algorithm is significantly outperforms the other algorithms. It also demonstrates that one bit BCS is slightly better than BIMAT algorithm and BIMAT is slightly better than BIHT.

V. Conclusion

We have proposed new iterative dictionary learning based algorithms for DOA estimation from both the array sensor

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LCOMM.2016.2587674, IEEE Communications Letters

output and the sign of the array sensor output. These algorithms solve the grid mismatch problem which occurs for sparsity-based DOA estimation algorithms. Also, for DOA estimation from the sign of the array sensor output, we use the framework of one bit compressed sensing. For the sparse recovery from the one bit measurements, a binary IMAT algorithm is suggested which extends the IMAT algorithm to the one bit compressed sensing framework. Simulation results show that dictionary learning improves the accuracy of DOA estimation.

APPENDIX A
FINAL RECURSION: FROM THE SENSOR DATA

By some calculation, the cost function can be formed as:

$$\begin{aligned}F(\theta) &= (\mathbf{y}-A(\theta)\mathbf{s})^H(\mathbf{y}-A(\theta)\mathbf{s})\\ &= \mathbf{y}^H\mathbf{y}-\mathbf{y}^H(\mathbf{A}(\theta)\mathbf{s})-(\mathbf{A}(\theta)\mathbf{s})^H\mathbf{y}-(\mathbf{A}(\theta)\mathbf{s})^H(\mathbf{A}(\theta)\mathbf{s})\\ &= \mathbf{y}^H\mathbf{y}-2\mathrm{Re}\left\{\mathbf{y}^{\mathrm{H}}\mathbf{A}(\theta)\mathbf{s}\right\}-\mathbf{s}^{\mathrm{H}}\mathbf{A}(\theta)^{\mathrm{H}}\mathbf{A}(\theta)\mathbf{s}.\end{aligned}$$

The partial derivative $\frac{\partial}{\partial\theta}F(\theta)$ is as follows:

$$-\frac{\partial}{\partial\theta}\left(\mathbf{y}^H\mathbf{A}(\theta)\mathbf{s}\right)-\frac{\partial}{\partial\theta}\left(\mathbf{A}(\theta)\mathbf{s}^H\mathbf{y}\right)+\frac{\partial}{\partial\theta}\left(\mathbf{A}(\theta)\mathbf{s}^H\mathbf{A}(\theta)\mathbf{s}\right)$$

$$\begin{aligned}&= -2\mathcal{R}\boldsymbol{e}\left\{\mathbf{y}^H\frac{\partial(\mathbf{A}(\theta)\mathbf{s})}{\partial\theta}\right\}+\left(\frac{\partial(\mathbf{A}(\theta)\mathbf{s})^H}{\partial\theta}\right)(\mathbf{A}(\theta)\mathbf{s})\\ &+(\mathbf{A}(\theta)\mathbf{s})^H\left(\frac{\partial(\mathbf{A}(\theta)\mathbf{s})}{\partial\theta}\right)\\ &= -2\mathcal{R}\boldsymbol{e}\left\{\mathbf{y}^H\frac{\partial(\mathbf{A}(\theta)\mathbf{s})}{\partial\theta}\right\}+\left((\mathbf{A}(\theta)\mathbf{s})^H\left(\frac{\partial(\mathbf{A}(\theta)\mathbf{s})}{\partial\theta}\right)\right)^H\\ &+\left((\mathbf{A}(\theta)\mathbf{s})^H\left(\frac{\partial(\mathbf{A}(\theta)\mathbf{s})}{\partial\theta}\right)\right)\\ &= -2\mathcal{R}\boldsymbol{e}\left\{\mathbf{y}^H\frac{\partial(\mathbf{A}(\theta)\mathbf{s})}{\partial\theta}\right\}+2\mathcal{R}\boldsymbol{e}\left\{(\mathbf{A}(\theta)\mathbf{s})^H\left(\frac{\partial(\mathbf{A}(\theta)\mathbf{s})}{\partial\theta}\right)\right\}\\ &= 2\mathcal{R}\boldsymbol{e}\left\{\left((\mathbf{A}(\theta)\mathbf{s})^H-\mathbf{y}^H\right)\frac{\partial(\mathbf{A}(\theta)\mathbf{s})}{\partial\theta}\right\}\\ &= 2\mathcal{R}\boldsymbol{e}\left\{\mathbf{e}^H\frac{\partial(\mathbf{A}(\theta)\mathbf{s})}{\partial\theta}\right\}.\end{aligned}$$

Also the derivation of array manifold with respect to θ is:

$$\frac{\partial}{\partial\theta}(\mathbf{A}(\theta)\mathbf{s}) = c_0[\mathbf{B}(\theta)\odot\mathbf{A}(\theta)]\mathbf{s}.$$

where $c_0 = -j\frac{2\pi d}{\lambda}$, $\mathbf{B} = [\mathbf{b}(\theta_1),\mathbf{b}(\theta_2),\cdots,\mathbf{b}(\theta_L)]$, and $\mathbf{b}(\theta_i) = [0,\ \cos(\theta_i),\ 2\cos(\theta_i),\ \cdots,(M-1)\cos(\theta_i)]^T$. Finally, put all things together, the final recursion in (6) is obtained.

APPENDIX B
FINAL RECURSION: FROM THE SIGN OF THE SENSOR DATA

For one-bit DOA estimation, With the same procedure of appendix A, the derivation of cost function with respect to θ is as follow:

$$\frac{\partial}{\partial\theta}H(\theta) = 2\mathcal{R}\boldsymbol{e}\left\{\mathbf{e}^H\frac{\partial(\mathbf{csgn}\left\{\mathbf{A}(\theta)\mathbf{s}\right\})}{\partial\theta}\right\}.$$

and the derivative of array manifold with respect to θ can be obtained as:

$$\frac{\partial}{\partial\theta}(\mathbf{csgn}\left\{\mathbf{A}(\theta)\mathbf{s}\right\}) = c_0[\mathbf{B}(\theta)\odot\mathbf{A}(\theta)]\mathbf{s}\odot\mathbf{d}(\theta).$$

where $c_0 = -j\frac{2\pi d}{\lambda}$, $\mathbf{B}(\theta) = [\mathbf{b}(\theta_1),\mathbf{b}(\theta_2),\cdots,\mathbf{b}(\theta_L)]$, $\mathbf{b}(\theta_i) = [0,\ \cos(\theta_i),\ 2\cos(\theta_i),\ \cdots,(M-1)\cos(\theta_i)]^T$ and $\mathbf{d}(\theta) = \mathrm{S}^{'}(\mathrm{Re}(\mathbf{A}(\theta)\mathbf{s}))+j\mathrm{S}^{'}(\mathrm{Im}(\mathbf{A}(\theta)\mathbf{s}))$ is the derivative vector where $\mathrm{S}^{'}(x) = \frac{2\exp(-x)}{1+\exp(-x)}$.

REFERENCES

[1] L. Godara, "Application of antenna arrays to mobile communications. ii. beam-forming and direction-of-arrival considerations," *Proceedings of the IEEE*, vol. 85, no. 8, pp. 1195–1245, 1997.

[2] M. Greco, F. Gini, A. Farina, and L. Timmoneri, "Direction-of-arrival estimation in radar systems: moving winsdow against approximate maximum likelihood estimator," *IET Radar, Sonar Navigation*, vol. 3, no. 5, pp. 552–557, 2009.

[3] J. Thompson, P. Grant, B. Mulgrew, and R. Rajagopal, "Generalized algorithm for DOA estimation in a passive sonar," *IEE Proceedings F, Radar and Signal Processing*, vol. 140, no. 5, pp. 339–340, 1993.

[4] D. H. Johnson, and D. E. Dudgeon, *Array signal processing: Concepts and Techniques*, 1992.

[5] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, no. 8, pp. 1408–1418, 1969.

[6] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[7] J. Zheng and M. Kaveh, 'Sparse spatial spectral estimation: A covariance fitting algorithm, performance and regularization," *IEEE Transactions on Signal Processing*, vol. 61, no. 11, pp. 2767-2777, 2013.

[8] D. Malioutov, M. Cetin, and A. Willsky, "A sparse signal recounstruction prespective for source localization with sensor arrays," *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 3010–3022, 2005.

[9] A. Gurbuz, V. Cevher, and J. McClellan, "Bearing estimation via spatial sparsity using compressive sensing," *IEEE Trans. Aerospace and Electronic Systems*, vol. 48, no. 2, pp. 1358–1369, 2012.

[10] M. Rossi, A. Haimovich, and Y. Eldar, "Spatial compressive sensing for MIMO radar," *IEEE Trans. on Signal Proc.*, vol. 62, no. 2, pp. 419–430, Jan 2014.

[11] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse bayesian inference," *IEEE Transactions on Signal Processing*, vol. 61, no. 1, pp. 38–43, 2013.

[12] R. Jagannath and K. Hari, "Block sparse estimator for grid matching in single snapshot doa estimation," *IEEE Signal Processing Letters*,vol. 20, no. 11, pp. 1038–1041, 2013.

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

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

[15] J. M. Nichols, A. K. Oh, and R. M. Willet, "Reducing basis mismatch in harmonic signal recovery via alternating convex search," *IEEE Signal Processing Letters*,vol. 21, no. 8, pp. 1007–1011, 2014.

[16] F. Marvasti, et al, "A unified approach to sparse signal processing," *EURASIP Journal on Advances in Signal Processing*, vol. 44, 2012.

[17] M. Azghani,and F. Marvasti, "Sparse signal processing," *Chapter 8, New perspectives on approximation and sampling theory*, pp. 189–213, 2014.

[18] P. Boufounos and R. Baraniuk, "1-bit compressive sensing," *in proceeding 42nd Annu. Conf. Inf. Sci. Sys.*, Princeton, pp. 16–21, Mar 2008.

[19] L. Jacques, J. Laska, P. Boufounos, and R. Baraniuk, "Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2082–2102, April 2013.

[20] C. H. Chen and J. Y. Wu, "Amplitude-aided 1-bit compressive sensing over noisy wireless sensor networks," *IEEE Wireless Communoication Letters*, vol. 4, no. 5, pp. 473–476, 2015.

[21] F. Li, J. Fang, H. Li, and L. Huang, "Robust one-bit Bayesian compressed sensing with sign-flip errors," *IEEE Signal Processing Letters*, vol. 22, no. 7, pp. 857–861, 2015.

[22] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

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

[24] O. Bar-shalom, and A. J. Weiss, "DOA estimation using one-bit quatized measurements," *IEEE Trans. Aerospace and Electronic Systems*, vol. 38, no. 3, pp. 868–884, 2002.

[25] I. Tosic, and P. Frossard, "Dictionary Learning," *IEEE Signal Proc. Magazine*, vol. 28, no. 2, pp. 27–38, March 2011.

[26] H. Zayyani, M. Korki, and F. Marvasti, "Dictionary learning for blind one bit compressed sensing," *IEEE Signal Processing Letters*, vol. 23, no. 2, pp. 187–191, Feb 2016.