DISTRIBUTED BEARING ESTIMATION VIA MATRIX COMPLETION

Andrew Waters[1] and Volkan Cevher[1,2]

[1]Rice University
[2] Ecole Polytechnique Federale de Lausanne

ABSTRACT

We consider bearing estimation of multiple narrow-band plane waves impinging on an array of sensors. For this problem, bearing estimation algorithms such as minimum variance distortionless response (MVDR), multiple signal classification, and maximum likelihood generally require the array covariance matrix as sufficient statistics. Interestingly, the rank of the array covariance matrix is approximately equal to the number of the sources, which is typically much smaller than the number of sensors in many practical scenarios. In these scenarios, the covariance matrix is low-rank and can be estimated via matrix completion from only a small subset of its entries. We propose a distributed matrix completion framework to drastically reduce the inter-sensor communication in a network while still achieving near-optimal bearing estimation accuracy. Using recent results in noisy matrix completion, we provide sampling bounds and show how the additive noise at the sensor observations affects the reconstruction performance. We demonstrate via simulations that our approach sports desirable tradeoffs between communication costs and bearing estimation accuracy.

Index Terms— Array signal processing, Covariance analysis, Direction of arrival estimation, Matrix completion

1. INTRODUCTION

Many algorithms for narrow-band direction finding with sensor arrays, parameter estimation of multiple sinusoidal signals superimposed with noise, and discriminating multiple overlapped echoes require the covariance matrix of the observations as sufficient statistics for estimation. Representative members include minimum variance distortionless response (MVDR), multiple signal classification (MUSIC), and maximum likelihood (ML) [1]. Although these algorithms have rigorous estimation guarantees, they have an important disadvantage when the sensors are untethered: centralized process-

ing. Hence, wireless communications among the sensors can create many bottlenecks and critical points of failure.

In a practical system, estimating the array covariance matrix in a distributed fashion can provide the ability to scale, reduce adversarial vulnerability, decrease communication costs, and share processing responsibilities among individual sensors as the number of sensors increase. Distributed processing over the network implies that the sensors only communicate with their immediate neighbors on the communication graph with a message passing scheme that uses constant bandwidth. In this case, the greatest challenge becomes finding a message passing scheme and an associated algorithm with provable estimation guarantees.

In this paper, we propose propagating local correlation estimates over the edges of a communication graph to partially fill in the entries of the sensor array's covariance matrix. We pay particular attention to tree structured graphs generated by Prim's algorithm and Dijkstra's algorithm [2]. Both of these algorithms take in a set of nodes and construct a tree structured graph that minimizes a cost function, such as one based on communication costs. We further propose extensions to this simple communication model by greedily propagating additional signal information through the tree in order to increase the number of observed matrix entries when needed.

For the bearing estimation problem, we discuss why the array covariance matrix is low rank for most practical scenarios. Then, we exploit recent results in matrix completion theory to recover the full covariance matrix from the subset of observed entries [3, 4]. We provide analytic bounds on the covariance estimation error as a function of the number of time snapshots, the additive noise variance, and the number of filled matrix entries. While doing so, we restate the matrix sampling bounds in [4] within the context of bearing estimation. We subsequently investigate the accuracy of the bearing estimates based on the recovered covariance matrix using MVDR and compare the results with the estimates based on the fully sampled covariance matrix. We show via simulations that our approach sports desirable tradeoffs between communication costs and bearing estimation accuracy.

The paper is organized as follows. In Section 2 we provide the necessary background on signal processing for bearing estimation and matrix completion. In Section 3 we provide details of our distributed estimation approach to determine the

		Form Approved OMB No. 0704-0188

Report Documentation Page

1. REPORT DATE **2009**	2. REPORT TYPE	3. DATES COVERED **00-00-2009 to 00-00-2009**

4. TITLE AND SUBTITLE **Distributed Bearing Estimation Via Matrix Completion**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Rice University,Washington,DC**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**4**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

array covariance matrix and provide performance bounds. In Section 4 we provide simulation results to demonstrate the validity of our approach with both synthetic and real-world data. We conclude in Section 5.

2. BACKGROUND

2.1. Narrow-band bearing estimation problem

The problem of bearing estimation of multiple narrow-band plane waves with sensor arrays can be reduced to the following model [1, 5]:

$$\boldsymbol{y}(t_n) = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{x}(t_n) + \boldsymbol{w}(t_n), \quad n = 1, \ldots, N; \quad (1)$$

where $\boldsymbol{y} \in \mathbb{C}^{P \times 1}$ (complex) is the noisy data vector of the sensor network, $\boldsymbol{x} \in \mathbb{C}^{K \times 1}$ is the vector of unknown signal amplitudes, and $\boldsymbol{w} \in \mathbb{C}^{P \times 1}$ is an additive noise.

In (1), $\boldsymbol{A}(\boldsymbol{\theta}) \in \mathbb{C}^{P \times K}$ has the following structure

$$\boldsymbol{A}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{a}(\theta_1) & \ldots & \boldsymbol{a}(\theta_K) \end{bmatrix}, \quad (2)$$

where $\boldsymbol{a}(\theta_k) \in \mathbb{C}^{P \times 1}$ is called the steering vector, $\theta_k \in \mathbb{R}$ is the bearing of the k-th source, and $\boldsymbol{\theta} = \begin{bmatrix} \theta_1 & \ldots & \theta_K \end{bmatrix}$. We denote $\boldsymbol{X} \in \mathbb{C}^{P \times N}$ as the source matrix, where $\boldsymbol{X} = [\boldsymbol{x}(t_1), \ldots, \boldsymbol{x}(t_N)]$. The array data matrix $\boldsymbol{Y} \in \mathbb{C}^{P \times N}$ and the noise matrix $\boldsymbol{W} \in \mathbb{C}^{P \times N}$ are similarly defined.

The objective of the bearing estimation problem is to determine $\boldsymbol{\theta}$ given the noisy observations in (1). Without loss of generality, we focus on 2-D bearing estimation; hence, the p-th entry in the steering vector $\boldsymbol{a}(\theta_k)$ for the k-th source has the following well-known expression:

$$[\boldsymbol{a}(\theta_k)]_p = \exp\left\{-j\omega\rho_p \cos\left(\psi_p - \theta_k\right)\right\}, \quad (3)$$

where (ρ_p, ψ_p) is the p-th sensor position in polar coordinates (angle is measured with respect to the horizontal axis) and ω is the known narrow-band frequency of the sources.

In the sequel, we make the following assumptions:
A1 (*Sensor Network*): $P \gg K$. The number of sensors P is large compared to the number of sources K. Note that $P > K$ is a necessary condition for the uniqueness of the bearing estimates.
A2 (*Source Signals*): $\mathbb{E}\left\{[\boldsymbol{x}(t_i)]_k\right\} = 0$, $\mathbb{E}\left\{[\boldsymbol{x}(t_i)]_k'\,[\boldsymbol{x}(t_i)]_l\right\} = 0$ $(k \neq l)$, and $\mathbb{E}\left\{[\boldsymbol{x}(t_i)]_k'\,[\boldsymbol{x}(t_j)]_k\right\} = 0$ $(i \neq j)$. Source signals are independent from each other and are uncorrelated in time. The symbols $'$ and $\mathbb{E}$ denote the Hermitian and expectation operators, respectively.
A3 (*Noise*): $[\boldsymbol{w}(t_i)]_p \sim \mathcal{N}\left(0, \sigma^2\right)$. The noise has an independent and identically distributed (iid) Gaussian distribution with zero mean and known variance σ^2.

2.2. Covariance based bearing estimation algorithms

Many algorithms to estimate $\boldsymbol{\theta}$ exist in the literature [1, 5]. The defining characteristics of these algorithms is their reliance on the the sample covariance matrix $\boldsymbol{R}$, given below,

as sufficient statistics:

$$\boldsymbol{R} = \frac{1}{N}\boldsymbol{Y}\boldsymbol{Y}' = \mathbf{Z} + \sigma^2\boldsymbol{I}, \quad (4)$$

where $\mathbf{Z} = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{X}\boldsymbol{X}'\boldsymbol{A}'(\boldsymbol{\theta})/N$ is a rank K positive semidefinite matrix. Example algorithms include minimum variance distortionless response (MVDR), multiple signal characterization (MUSIC), and maximum likelihood (ML).

To determine the source bearings, the MVDR algorithm calculates a power vs. bearing pattern via

$$P_{\text{MVDR}}(\theta; \boldsymbol{R}) = \left[\boldsymbol{a}'(\theta)\boldsymbol{R}^{-1}\boldsymbol{a}(\theta)\right]^{-1}, \quad (5)$$

whose local maxima indicates the locations. In contrast, the MUSIC algorithm relies on a partial eigenvector expansion of $\boldsymbol{R}$ (in a similar fashion to (5)) corresponding to the noise subspace by assuming K. The ML solution minimizes the mean squared data error, which can be directly written as a function of $\boldsymbol{R}$ [1].

2.3. Matrix completion from incomplete data

Here, we review the recent results on noisy matrix completion that will be necessary to quantify the performance of our approach. We start by assuming that we only observe a partial subset of the entries of a matrix $\mathbf{Z}$ ($P \times P$). Clearly, in the general case, we would be at a loss to provide any meaningful estimate of the missing entries. However, in the special case where $\mathbf{Z}$ is of low rank and the observations are noiseless, $M = \mathcal{O}\left(P\log^2 P\right)$ randomly chosen entries are sufficient to recover all the entries of $\mathbf{Z}$ with probability at least $1 - P^{-3}$ via the following convex optimization problem [3]

$$\widehat{\mathbf{Z}} = \arg\min \|\mathbf{Z}'\|_\star \text{ s.t. } \mathcal{P}_\Omega(\mathbf{Z}) = \mathcal{P}_\Omega(\mathbf{Z}'), \quad (6)$$

where $\|\cdot\|_\star$ is the nuclear norm, Ω is the set of observed elements in the matrix $\mathbf{Z}$, and $\mathcal{P}_\Omega(\mathbf{Z})$ retrieves the elements of Ω from $\mathbf{Z}$.

In this paper, we are directly interested in the case when the observed entries have errors,

$$[\mathbf{U}]_{ij} = [\mathbf{Z}]_{ij} + [\mathbf{E}]_{ij}, \quad (7)$$

where $[\mathbf{E}]_{ij}$ is iid zero mean Gaussian with variance τ^2. For this case, Candes and Yaniv [4] empirically show that stable recovery is possible when the size of the set Ω is a constant times the degrees of freedom. For a $P \times P$ correlation matrix $\mathbf{Z}$ of rank K, the number of required entries is $\mathcal{O}\left(K\left(2P - K\right)\right)$. The matrix completion is then accomplished by solving:

$$\widehat{\mathbf{Z}} = \arg\min \frac{1}{2}\|\mathcal{P}_\Omega(\mathbf{Z}' - \mathbf{U})\|_{\text{F}}^2 + \mu\|\mathbf{Z}'\|_\star, \quad (8)$$

where $\|\cdot\|_{\text{F}}$ is the Frobenius norm. Moreover, the relaxation parameter is chosen as $\mu = \sqrt{2m/P}\tau$ so that $\|\mathcal{P}_\Omega(\widehat{\mathbf{Z}} - \mathbf{U})\|_{\text{F}} \leq \delta$ where $\delta = \sqrt{m + \sqrt{8m}}\tau$, where $m = |\Omega|$. The estimation error can then be approximately bounded by $\|\mathbf{Z} - \widehat{\mathbf{Z}}\| \leq 4\sqrt{\frac{P^3}{m}}\delta$; see [4].

3. DISTRIBUTED COVARIANCE ESTIMATION ON GRAPHS

Although the bearing estimation algorithms discussed in Section 2 have rigorous estimation guarantees [1], they have an important disadvantage: they require centralized processing where a fusion center collects $\boldsymbol{Y}$ in order to calculate $\boldsymbol{R}$. The centralized nature of these algorithms result in increased communication costs with many bottlenecks and critical points of failure. In contrast, we seek distributed algorithms since they provide the ability to scale, reduce vulnerability, decrease communication, and share processing responsibilities among individual sensors as the number of sensors increase.

Suppose now we are given a communication graph $\mathcal{G}$, where the edges correspond to communication links and the vertices correspond to the sensors with some pre-determined hierarchical ordering of the sensors, e.g., a routing scheme. Distributed processing over $\mathcal{G}$ implies that the sensors only communicate with their immediate neighbors on the graph that share an edge with a message passing scheme that use constant bandwidth. The graph structure $\mathcal{G}$ could be constructed with multiple objectives in mind, such as minimizing communication costs, satisfying a minimum connectivity probability, etc. In general, the number of edges is much smaller than $\binom{P}{2}$ for practical purposes.

3.1. Covariance estimation on trees

To demonstrate the ideas more concretely, we focus on tree structured graphs $\mathcal{T}$, obtained by, say, Prim's Algorithm or Dijkstra's algorithm [2]. We use a bottom-up communication hierarchy where the children nodes communicate to their parents their observations as well as the inherited matrix entries from their own children. In this case, a parent sensor can immediately compute the cross-correlation terms for the matrix $\boldsymbol{R}$ corresponding to itself and its children nodes, along with all pairwise relations between children. This scheme continues recursively until the root node is reached.

It is easy to see that the number of observed entries in a covariance matrix is lower bounded by $3P-2$, corresponding to a chain graph. Under random deployment, Prim or Dijkstra algorithms provide much higher connectivity than this lower bound (empirically enough entries for two targets). Higher target numbers, however, may require a larger number of entries due to the increased rank of the covariance matrix. When this occurs, signal information can be passed up through multiple levels of the tree instead of only to the parent. Controlling the amount of information sent through the tree allows us to create a tradeoff between the communication cost and the number of observed entries in the covariance matrix. Passing all signal information in the tree reveals all entries of the covariance matrix but incurs the maximum communication cost.

3.2. Recovering the full covariance matrix

As a result of the tree message passing scheme, the root node observes a partial set of entries Ω of the covariance matrix $\boldsymbol{R}$ where Ω is determined by the edges of the communication tree $\mathcal{T}$. The number of observed entries is defined by $m = |\Omega|$. The following lemmas characterize the observed entries of $\boldsymbol{R}$ and the recovered covariance matrix $\widehat{\mathbf{R}}$. Proofs of these lemmas can be found in [6].

Lemma 1 (Diagonal entries of $\boldsymbol{R}$). *Define* $\mathbf{U} = \boldsymbol{R} - \sigma^2 \boldsymbol{I}$. $[\mathbf{U}]_{pp}$ $(p = 1, \ldots, P)$ *is approximately zero mean Gaussian distributed with variance* $\tau_{pp}^2 \leq 4N^{-1} [\boldsymbol{R}]_{pp} \sigma^2$.

Lemma 2 (Off-diagonal entries of $\boldsymbol{R}$). $[\mathbf{R}]_{pl}$ $(p, l = 1, \ldots, P;$ $p \neq l)$ *is approximately zero mean Gaussian distributed with variance* $\tau_{pl}^2 \leq N^{-1} \left([\boldsymbol{R}]_{pp} + [\boldsymbol{R}]_{ll} \right) \sigma^2 + \sigma^4$.

Lemma 3 (Frobenius norm of $\mathcal{P}_\Omega \left(\boldsymbol{E} \right)$). *Define* $\delta = \| \mathcal{P}_\Omega \left(\boldsymbol{E} \right) \|_F$. *Then, w.h.p.,* $\delta \leq 2\sigma N^{-1/2} \max_p \sqrt{[\boldsymbol{R}]_{pp} \left(m + \sqrt{8m} \right)}$.

Lemma 4 (Estimation guarantee on $\widehat{\boldsymbol{R}}$). *With δ defined as above and with $\widehat{\boldsymbol{E}} = \boldsymbol{R} - \widehat{\boldsymbol{R}}$, we have the following reconstruction guarantee,* $\| \widehat{\boldsymbol{E}} \|_F \leq 4\sqrt{\frac{2P^3 + mP}{m}} \delta + 2\delta$.

In our approach, the root node solves the optimization problem in (8) and reports $\widehat{\boldsymbol{R}} + \sigma^2 \boldsymbol{I}$. We use $\tau^2 = \max(\tau_{pp}^2, \tau_{pl}^2)$ as characterized in the lemmas 1 and 2 when choosing the relaxation parameter μ.

We can now bound the performance of the beamformer operating on the reconstructed correlation matrix. We further assume that $\| \boldsymbol{R}^{-1} \widehat{\boldsymbol{E}} \|_F < 1$. This condition implies that $\boldsymbol{R}$ is sufficiently well conditioned a posteriori via diagonal loading, which is common in the application of the MVDR beamformer. We denote $\boldsymbol{R}' = \boldsymbol{R} + \gamma \mathbf{I}$ as the regularized covariance matrix. As shown in [6], γ satisfying $\gamma > \| \boldsymbol{E} \|_F \leq 4\sqrt{\frac{2P^3 + mP}{m}} \delta + 2\delta$ is sufficient to guarantee $\| \boldsymbol{R}^{-1} \widehat{\boldsymbol{E}} \|_F < 1$. Often lower values of γ can be safely used. Defining $\kappa(\boldsymbol{R}')$ as the condition number of $\boldsymbol{R}'$ we can show that $\kappa(\boldsymbol{R}') \approx \frac{P/K + \gamma + \sigma^2}{\gamma + \sigma^2}$.

With the above background, we can state the following lemma using some results from matrix perturbation theory [7]:

Lemma 5. *Defining $\kappa(\boldsymbol{R}')$ as the condition number of $\boldsymbol{R}'$ and assuming that $\| \boldsymbol{R}'^{-1} \widehat{\boldsymbol{E}} \|_F < 1$. Then we have the following bound on beamformer deviation:*

$$\left| P_{MVDR}(\theta; \boldsymbol{R}') - P_{MVDR}(\theta; \widehat{\boldsymbol{R}'}) \right| \leq \frac{\kappa \| \boldsymbol{R}'^{-1} \widehat{\boldsymbol{E}} \|_F}{1 - \| \boldsymbol{R}'^{-1} \widehat{\boldsymbol{E}} \|_F} P_{MVDR}(\theta; \boldsymbol{R}').$$

Our experiments with synthetic and real data show that this worst-case bound is somewhat pessimistic and that the actual deviation from the ideal beampattern is often quite small. We conjecture that the errors introduced in the matrix completion process has an underlying structure that should be exploited to obtain tighter bounds.

4. SIMULATION RESULTS

4.1. Communication Tradeoffs

As explained in Section 3, using a tree based routing approaches in conjunction with matrix completion we can form a tradeoff between communication costs and the number of entries observed in our autocorrelation matrix, $\mathbf{R}$. This, in turn, affords a tradeoff between the overall communication cost and the resulting accuracy of our estimator. Here, we present simulation results to demonstrate this tradeoff for various values of the array signal to noise ratio. Each simulation trial uses a random network of $P = 30$ sensors which we link together via Dijkstra's algorithm. The desired number of observed matrix entries determines how much information is passed up the tree. Observing the full matrix corresponds to passing all signals up to the fusion center, with a communication cost equal to the full cost of the Dijkstra tree. Each simulation point is averaged over 1000 trials.
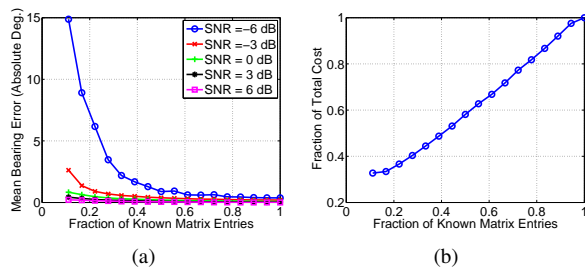

(a)   (b)

Fig. 1. *Communication and accuracy tradeoffs.*

4.2. Bearing estimation experiments

We present bearing estimation results for both synthetic and field data. For synthetic data, we simulate two targets moving on a circular track and estimate bearing angles at several time steps. We deploy a random array of $P = 30$ and connect the network via Prim's algorithm. We then pass inter-sensor messages as described in Section 3 and observe 200 out of a possible 900 matrix entries. The root node performs matrix completion to obtain $\widehat{\mathbf{R}}$ which then serves as the input to its beamformer. We use $\mathbf{R}' = \mathbf{R} + 0.1\mathbf{I}$ when beamforming. Our results, displayed in Figure 2, show that bearing estimates with our approach are quite close to the full data case.
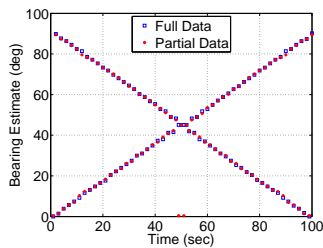

Fig. 2. Two-target bearing tracks.

Next, we demonstrate our approach on acoustic field data corresponding to a multi-vehicle convoy moving along an elliptical track observed by a network of 10 sensors. The sensors are distributed in a circle of 3 m diameter with the reference sensor in the center of the ring. Because both Dijkstra's and Prim's algorithm would force full connectivity for this deployment, we enforce sparsity as shown in 3(a) resulting in observing 40 out of the 100 entries of the correlation matrix. No diagonal loading was used in this example. Estimate results are displayed in Figure 3(b).
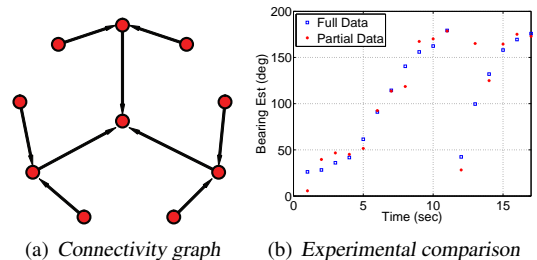

(a) *Connectivity graph*   (b) *Experimental comparison*

Fig. 3. *Field data results.*

5. CONCLUSION

We discussed distributed low-rank covariance estimation over tree-based graphs via matrix completion. Our approach exploited the ability to stably reconstruct low rank matrices from highly incomplete partial observations. We quantified how additive white Gaussian noise in the sensor observations directly affect the matrix completion performance. We observed empirically that the noise is well controlled and has very little effect on the bearing estimation of the MVDR beamformer. As future work, we will investigate other graph structures, such as the Bethe lattice [8] and loopy graphs.

6. REFERENCES

[1] P. Stoica and N. Arye, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, no. 5, pp. 720–741, 1989.

[2] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, MIT Press and McGraw-Hill, 2001.

[3] E. J. Candes and T. Tao, "The power of convex relaxation: Near-optimal matrix completion," *preprint*, 2009.

[4] E. J. Candès and Y. Plan, "Matrix completion with noise," *CoRR*, vol. abs/0903.3131, 2009.

[5] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*, Prentice Hall, 1993.

[6] A. Waters and V. Cevher, "Proof of lemmas for matrix completion beamformer," Tech. Rep., Rice University, 2009, Available from the first author's website.

[7] G. W. Stewart and J. Sun, *Matrix Perturbation Theory*, Academic Press, 1990.

[8] H. A. Bethe, "Statistical theory of superlattices," *Proceedings of the Royal Society, Series A*, pp. 552–575, 1935.