Field Inversion by Consensus and Compressed Sensing

Aurora Schmidt and José M. F. Moura
Department of Electrical and Computer Engineering
Carnegie Mellon University, Pittsburgh, PA 15213 USA
Email: aurorasc@andrew.cmu.edu, moura@ece.cmu.edu

***Abstract*— We study the inversion of a random field from pointwise measurements collected by a sensor network. We assume that the field has a sparse representation in a known basis. To illustrate the approach, consider the inversion of an acoustic field created by the superposition of a discrete number of propagating noisy acoustic sources. Our method combines compressed sensing (sparse reconstruction by ℓ_1-constrained optimization) with distributed average consensus (mixing the pointwise sensor measurements by local communication among the sensors). The paper describes the approach and demonstrates its good performance with synthetic data for several scenarios of practical interest.**

***Index Terms*— Consensus algorithm, compressed sensing, field inversion, field reconstruction, ℓ_1 optimization.**

I. Introduction

We consider the inversion of a random field that is monitored by a sensor network. To be concrete, we assume an acoustic field generated by an unknown number, S, of discrete noisy acoustic sources. The resulting field is the superposition of the wavefield propagated by each of these sources. To simplify the problem, we assume that the N sensors are placed on the nodes of a uniform $\sqrt{N} \times \sqrt{N}$ grid spanning the field. By field inversion, we mean estimating the number of sources, their intensity levels, and their locations in the physical space of interest.

A fusion center interrogates m of these sensors to invert the field. Because of networks constraints this number is much less than the total number of sensors. We can pose the question of which subset is most informative, so that the field inversion is as reliable as possible. However, since the sensors take pointwise measurements, i.e., they take measurements of the field only at their own location, this is a hopelessly difficult combinatorial problem. We reformulate this question by proposing a new algorithm where, first, the sensors mix their states through local communication, after which the fusion center samples m states with which to invert the field. As will be seen, with appropriate mixing, the issue of which sensors to interrogate becomes of secondary importance.

When N is large, our algorithm is more efficient than the fully centralized version, which is unfeasible for several reasons: 1) *Single point of failure.* If the fusion center fails, the network becomes inoperative; 2) *Scarce resources.* Transmitting N measurements taxes the communication and power constrained resources of the sensors and of the network. 3) *Computation burden.* Processing N measurements overburdens the fusion center.

Our solution avoids these difficulties by distributing the inversion load to all the sensors, allowing the fusion center to sample $m \ll N$ values. The two stages of the algorithm are: 1) *Mixing.* Akin to a consensus algorithm, [1], the sensors mix their current states with their neighbors in a distributed fashion that requires only local communication; and 2) *Sparse inversion.* Analogous to compressed sensing, [2], [3], a sparse field reconstruction inverts the field by an ℓ_1 constrained optimization from m sensor states, which are forwarded to a fusion center.

We comment briefly on the organization of the overall paper. Section II sets up the problem and reviews preliminary concepts. Section III gives background on compressed sensing and consensus. Section IV presents the proposed inversion algorithm. Section V presents a numerical study. Finally, Section VI concludes the paper and comments on generalizations of our approach.

II. Problem Formulation

A. Field Model

The field is defined on a finite $\sqrt{N} \times \sqrt{N}$ lattice, $\mathbb{L}$. We stack the field values on a N-dimensional vector, $\mathbf{x}$. The field is generated by a finite number S of unknown discrete sources with intensities S_i, for $i = 1, \ldots, S$, located at S nodes in the grid. The sources are collected in the N-dimensional vector, $\mathbf{v}$, where $v_i = S_i$ if source S_i is at node i and 0 if there is no source there. The field $\mathbf{x}$ is the superposition of the S propagating wavefields, i.e., $\mathbf{x} = D\mathbf{v}$, where D_{ij} is the propagation operator between nodes i and j. For example, in 3D free-space propagation, the values of D relate to a power of the distance between the N nodes in the grid.

B. Measurement Model

The field is measured by N pointwise sensors, at the nodes of the lattice $\mathbb{L}$. Before sending m of the sensor values to the a fusion center, the sensors mix their measurements through local message passing in the network. The observations collected at the fusion center are no longer pointwise samples, but current sensor states as detailed in section III-A. The mixing operation is linear, and will be denoted as $\mathcal{M}$. We assume no noise during the communication process. However, there

This work was partially supported by the Office of Naval Research under MURI N000140710747.

may be additive noise in the initial field measurements, so the measurement equation is as given in (1).

$$\mathbf{y} \quad = \quad Q\mathcal{M}\left(D\mathbf{v}+\mathbf{z}\right) \tag{1}$$

The matrix, $Q \in \Re^{m\times N}$, is a row selection matrix, made of m unique rows from the N-by-N identity matrix. The additive noise vector is represented by $\mathbf{z}$. The goal is to recover $\mathbf{v}$ from $\mathbf{y}$. Since the number of observations is smaller than the dimensions of the field, the sensing matrixis underdetermined. We propose to carry out an ℓ_1-inversion. In the noiseless case, this amounts to solving the convex optimization, (P_1).

$$(P_1) : \min_{\mathbf{v}} \|\mathbf{v}\|_1 \text{ subject to } \mathbf{y} = Q\mathcal{M}D\mathbf{v} \tag{2}$$

When there is additive noise, with $\|\mathbf{z}\|_2 \le \epsilon$, we will change the constraints of the optimization to be as shown in (P_2).

$$(P_2) : \min_{\mathbf{v}} \|\mathbf{v}\|_1 \text{ subject to } \|y - Q\mathcal{M}D\mathbf{v}\|_2 \le \epsilon \tag{3}$$

Compressed sensing theory finds that there are circumstances in which the solution to the ℓ_1 optimization coincides with the solution to the optimal problem; the ℓ_0 minimization. However, certain incoherence properties of the sensing matrices must be satisfied and will be discussed in III-B.

III. Background: Compressed Sensing and Consensus

The optimization in (2) is easily recognizable as a compressed sensing recovery algorithm. Using common notation in compressed sensing, e.g., [3], the representation matrix, Ψ, for our problem is the propagation matrix D, and the sensing matrix, Φ, corresponds to $Q\mathcal{M}$. For a reliable inversion of the field, we need the sensing matrix, Φ, to be incoherent with Ψ.

An efficient way to generate a linear operator, $\mathcal{M}$, is by local message passing between adjacent sensors. We propose a distributed iterative mixing algorithm, which is derived from average consensus algorithms, to accomplish this. The next two sections summarize background on consensus and compressed sensing.

A. Consensus algorithm

There is an extensive literature on average consensus that finds application in many sensor networks and multiagent coordination problems. Average consensus is a distributed, iterative algorithm that computes the average of a large number of quantities, each of which is available at a node of the network, we refer the reader to the recent overview, [1], and to the references therein. Our own work on consensus includes [4], [5], [6], [7] that study distributed inference under a variety of network conditions and the impact of the topology and other network parameters on the convergence rate and performance of the algorithm. In its simplest formulation, sensors update iteratively their state by a linear weighted combination of their current state with the states of their neighbors. This iterative algorithm converges at each sensor, under a broad set of conditions, to the average of the initial sensor states. Formally, stacking the states $x_n(k)$ at the N sensors of the network and at time k in the vector $\mathbf{x}_k$, average consensus iterates as

$$\mathbf{x}_{k+1} = W\mathbf{x}_k = W^{k+1}\mathbf{x}_0, \;\; k \ge 0, \tag{4}$$

Where W is the (sparse) matrix of weights and $\mathbf{x}_0$ is the vector of initial states. In average consensus, under broad conditions,

$$\lim_{k\to\infty} W^k = \frac{1}{N}\mathbf{1}\mathbf{1}^T, \tag{5}$$

where $\mathbf{1}$ is a vector of 1's. The matrix of weights W reflects the topology, or neighborhood structure, of the network, i.e., with which sensors does each sensor communicate. When weights are chosen to have a uniform value α over all links, W may be expressed in terms of the identity matrix I and the discrete Laplacian of the network L

$$W = I - \alpha L$$

The optimal value for the constant α is given in [8].

B. Compressed sensing

Compressed sensing shows that when an orthonormal basis Φ is chosen such that it displays a low coherence, as defined in (6), with an orthonormal Ψ, a relatively small number of selected measurements will, with overwhelming probability, yield perfect reconstructions of $\mathbf{v}$ (and therefore $\mathbf{x}$), [3], [2], [9]. In these cases the solution to (P_1) coincides with the optimal ℓ_0 minimization.

$$\begin{aligned} \mu\left(\Phi,\Psi\right) \quad &= \quad \sqrt{N}\cdot \max_{1\le i,j\le N} | < \phi_i, \psi_j > | \qquad (6) \\ & \quad 1 \le \mu \le \sqrt{N} \end{aligned}$$

Where the ϕ_i^T are the rows of Φ and ψ_j the columns of Ψ. Lower coherence implies a smaller number of measurements needed for reliable inversion, [3].

Randomly generated Φ tend to be incoherent with all bases. A random orthonormal basis created by drawing entries from an i.i.d. Gaussian distribution and orthonormalizing has high probability of having a coherence around $\sqrt{2\log N}$, [3]. Motivated by this result, our algorithm draws consensus weights from independent Gaussian distributions, to create sensing matrices with low coherence.

IV. Field Inversion by Consensus and Compressed Sensing (FICCS)

This section presents the FICCS (read as "fix") algorithm for unknown sparse field inversion. Sensors in the network collect noisy measurements of the field, mix their measurements using weighted adjacency matrices W_i, and then transmit a small subset of M sensor values to a centralized location.

Assuming noisy sensor measurements, the observation vector $\mathbf{y}$, after mixing, is given by

$$\begin{aligned} \mathbf{y} \quad &= \quad Q\left(\prod_{i=1}^{k} W_i\right)\left(D\mathbf{v}+\mathbf{z}\right) \qquad (7) \\ &= \quad Q\mathcal{M}\left(D\mathbf{v}+\mathbf{z}\right) \qquad (8) \end{aligned}$$

To perform the ℓ_1 optimization, the fusion center needs to know the values of the propagation, weights, and selection matrices, D, $\{W_i\}_{1\leq i\leq k}$ (k is the number of mixing steps), and Q, respectively. This can be accomplished by coordination among the sensors prior to deployment, where they choose a set of random weights to use during operation and ensure that they are normalized such that the rows of W_i have unit-norm. Our numerical studies will compare alternate strategies regarding the mixing weights: different weights at each mixing, $\{W_i\}_{1\leq i\leq k}$, versus constant weights, $W_i \equiv W$, $1 \leq i \leq k$.

The additive noise, $\mathbf{z}$, is modeled as i.i.d. and Gaussian with variance σ_z^2. As a recovery algorithm, the fusion center performs the convex optimization (P_2), defined in (3), choosing a threshold based on σ_z^2, using the inverse chi-squared function in Matlab, as in (9).

$$\epsilon = \sqrt{\text{chi2inv}(0.99, m)\sigma_z^2} \tag{9}$$

Performing this optimization amounts to solving a second-order cone program,[3]. This convex problem can be handled by efficient numerical solvers. We used the cvx package for Matlab, [10], in our particular tests. The bounded constraint in (3) makes (P_2) fit the semidefinite program framework (SDP), [11]. For the threshold ϵ given in (9), 99% of the noise vectors will satisfy $\|z\|_2 \leq \epsilon$. Although not every noise vector satisfies the bound, we will find the performance of (P_2) to be satisfactory.

Results from Robust Compressed Sensing, [12], [13], show that, when an additive noise vector, $\mathbf{e}$, as in $\mathbf{y} = A\mathbf{x} + \mathbf{e}$, has a bounded ℓ_2-norm, and the matrix A has low mutual coherence, the recovery program, (P_2), leads to low ℓ_2-error solutions. Furthermore, these results show that the ℓ_2-error of the recovered $\mathbf{x}$ is proportional to the noise energy bound ϵ, where the proportionality constant is a function of the sparsity (number of sources) S,[12],

$$\|\mathbf{x} - \hat{\mathbf{x}}\|_2 \leq C_S \cdot \epsilon \tag{10}$$

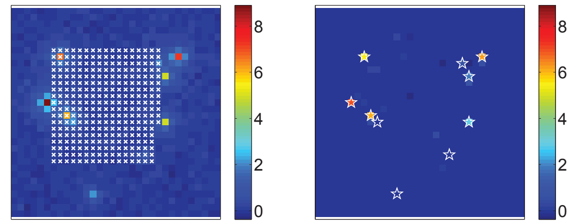
The mutual coherence is defined in [13] as

$$M_A \quad := \max_{1\leq k,j\leq n} \left|[A^T A]_{k,j}\right| \tag{11}$$

Our experimental results confirm, as indicated by the results in [13], that the performance of the inversion algorithm in the presence of noise is inversely related to the mutual coherence M_A of $A = (Q\mathcal{M}D)$.

V. Numerical Studies

To test the FICCS approach, we simulated sparse inversion experiments for a 32×32 ($N = 1024$) grid of sensors. We characterize the inversion performance (vertical axes in the plots) by: 1) the probability of recovery, when the measurements are noiseless; or 2) the ratio of ℓ_2 error to the error bound ϵ. We plot the number of selected measurements, m, as the independent variable, and study the impact on performance of: 1) weights strategy–Gaussian changing mixing weights $\{W_i\}$ versus constant mixing weights W, set to the optimal uniform consensus weights, [8]; 2) communication graph topology–a Erdös-Rényi random graph topology versus a nearest 8 neighbors topology, both with the same number of communication channels (graph edges); 3) the number k of mixing steps–if k is small, the mixing level may be insufficient because each sensor state carries little information about the other sensor measurements, while if k is very large, the mixing leads to average consensus, which is undesirable.

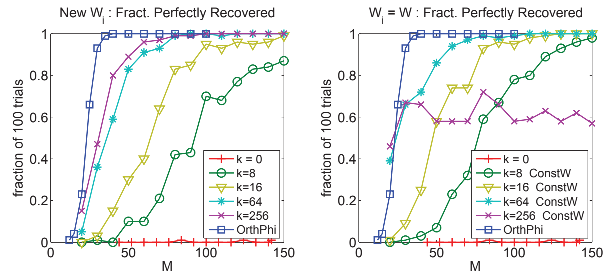


(a) Obs. and Sensor Selections (b) Reconstructed $\mathbf{v}$

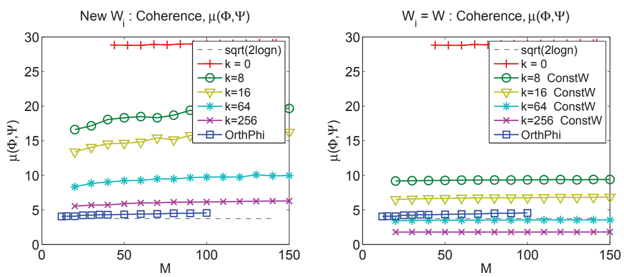
Fig. 1. *Field inversion example.*

Fig. 1 shows on the left the image sensed by the full array of $N = 1024$ sensors from 10 random sources; superimposed at the center are white crosses marking the locations of the $m = 300$ selected sensors used to perform the inversion. The additive measurement noise has a variance of $\sigma_z^2 = 10^{-2}$. The full grid of sensors are connected with a nearest 8-neighbor graph with $7,812$ edges, much smaller than the roughly 10^6 edges of the complete graph. The right panel shows the result of the field inversion $\hat{\mathbf{v}}$, after $k = 16$ mixings with Gaussian mixing weights. The sources' locations and intensities are recovered with an $\ell_2 = 5.18$ error. Comparing to the recovery error without mixing, which is 11.83, we see an improvement using local sensor coordination; and the oracle error (knowing the source locations and estimating intensities) is 2.40 due to noise. Although the selected sensor locations are away from most sources, we find that if sufficient mixing is performed, the field recovery performance is quite robust to the locations of the sensors communicating with the fusion center.

Figure 2 shows a study of mixing weight strategies over 100 Monte Carlo runs with randomly generated source vectors and selection matrices. The communication graph is again the nearest 8 neighbors and there are 3 unknown sources. The upper plots compare the fraction of perfect recoveries to the number of seleted measurements, m, while lower plots show the measured coherences. The keys show the number of mixing iterations: $k = 0$ means no mixing and corresponds to the worst performance observed. The leftmost line is an upper bound on the performance, using a random orthonormal basis, Φ, which is unachievable by local mixing. The performance due to mixing improves faster using the optimal consensus weights, since they are designed to mix very quickly. This is encouraging since less prior coordination is required when sensors can use uniform constant weights. However, the performance of the consensus strategy suffers when there is too much mixing ($k = 256$), due to a loss of linear independence of the measurements. Although the results of this study are shown for randomly selecting the reporting sensors, when we used the strategy of selecting a contiguous middle region, as



(a) Gaussian W_i (b) $W_i = W = I - \alpha^* L$



(c) avg. coherence Gaussian W_i (d) avg. coherence $W_i = W$

Fig. 2. *Impact of weights strategies, random weights (left) vs. constant optimal consensus weights (right), at various mixing levels, k. The corresponding coherences of sensing and propagation matrices are shown below.*

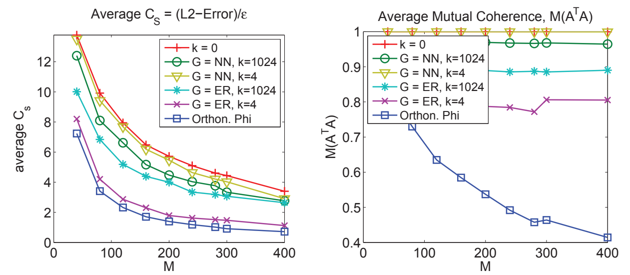
in fig. 1, we found the consensus weights were less robust to selection strategy than the non-constant Gaussian weights.

Figure 3 shows a study between two graph topologies with small amounts of measurement noise, $\sigma_z^2 = 0.01$. There are 3 unknown sources and each curve is the result of 100 Monte Carlo trials. The performance measure (left) is the ratio of the ℓ_2 estimation error to the noise energy bound, ϵ, as a function of the number of selected measurements, m. In the key, NN, represents a nearest 8 neighbor topology, while ER is for the randomly generated Erdös-Rényi graph (drawn uniformly from connected graphs with the same number of edges as the NN graph), and k gives the number of mixing iterations.

The Erdös-Rényi graph yields lower C_S's and therefore smaller ℓ_2-errors than the nearest neighbor topology. Contrasting with the noiseless case, where larger mixing levels lead to better performance, here $k = 4$ is better than the $k = 0$ (no mixing) and $k = 1024$ (heavy mixing) strategies. No mixing again gives the worst performance, showing that moderate mixing is required, while large amounts of mixing can be detrimental. The right panel of 3 shows the measured mutual coherence of $A = \Phi\Psi$. Overall, when this quantity is low, the observed ratio C_S is also low, agreeing with the relationship predicted by robust compressed sensing. The closeness of the error ratio to that of the random orthonormal basis shows that the mixing procedures and Erdös-Rényi topology effectively create conditions for reliable recovery.

VI. Conclusion and Generalizations

This paper describes an algorithm to invert a random field defined over a finite lattice, $\mathbb{L}$. The algorithm combines the ideas of local communications from consensus and sparse



(a) avg. C_S ratio (b) avg. mutual coherence

Fig. 3. *Impact of topology. Measured ratios, C_S, for 3 sources and $\sigma_z^2 = 0.01$ (left). Average mutual coherence of the product $A = \Phi\Psi$ matrices (right).*

recovery from compressed sensing to reduce the number of measurements to send to a fusion center. The fusion center performs an ℓ_1-inversion to recover the sparse input signal. Simulation tests demonstrate the good performance of local mixing for creating adequate sensing bases.

Preliminary tests show the strategy of drawing new random weights for each mixing iteration to yield sensing matrices with good properties for sparse signal recovery. Future work will include investigating ways to design weights for mixing matrices to maximize compressed sensing properties as well as adjacency graphs with good mixing properties for sparse recovery.

References

[1] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *IEEE Proceedings*, vol. 95, no. 1, pp. 215–233, January 2007.

[2] D. Donoho and Y. Tsaig, "Extensions of compressed sensing," *Signal Processing*, vol. 86, no. 3, pp. 533–548, March 2006.

[3] E. Candès and M. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, March 2008.

[4] S. Kar, S. A. Aldosari, and J. M. F. Moura, "Topology for distributed inference on graphs," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2609–2613, June 2008.

[5] S. Kar and J. M. F. Moura, "Sensor networks with random links: Topology design for distributed consensus," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3315–3326, July 2008.

[6] ——, "Distributed consensus algorithms in sensor networks: Link failures and channel noise," 2008, accepted for publication. [Online]. Available: http://arxiv.org/abs/0711.3915

[7] U. A. Khan, S. Kar, and J. M. F. Moura, "Distributed sensor localization in random environments using minimal number of anchor nodes," 2008, accepted for publication. [Online]. Available: http://arxiv.org/abs/0802.3563v2

[8] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Proceedings of IEEE Conference on Decision and Control*, pp. 4997–5002, December 2003.

[9] E. J. Candès, "Compressive sampling," *Int. Congress of Mathematics*, vol. 3, no. 7, pp. 1433–1452, 2006.

[10] S. Boyd, M. Grant, and Y. Ye, "cvx code package," *http://www.stanford.edu/~boyd/cvx/cvx_usrguide.pdf*, 2008.

[11] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[12] E. J. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1233, August 2006.

[13] D. L. Donoho, M. Elad, and V. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. on Information Theory*, vol. 52, no. 1, pp. 6–18, January 2006.