

DISTRIBUTED SPARSE MVDR BEAMFORMING USING THE BI-ALTERNATING DIRECTION METHOD OF MULTIPLIERS

Matt O'Connor^{1,2}, W. Bastiaan Kleijn¹ and Thushara Abhayapala²

¹ School of Engineering and Computer Science, Victoria University of Wellington, New Zealand

² Research School of Information Sciences and Engineering, Australian National University, Australia

ABSTRACT

Until now, distributed acoustic beamforming has focused on optimizing for a beamformer over an entire network, with each node contributing to the beamformer output. We present a novel approach that introduces sparsity to this beamformer computation, where we attempt to optimize for a subset of nodes within the network that produce SNR gains roughly equivalent to that of the optimal MVDR case. Due to the physical nature of sound, this approach trades a small loss in SNR for a large reduction in communication power and iterations required to produce a beamformer output by reducing the active node set of our network. Our approach operates in a fully distributed and asynchronous manner and does not require a high update iteration rate to produce an output at each sample.

Index Terms— Distributed, sparse, beamforming, sensor networks

1. INTRODUCTION

Beamforming, or spatial filtering, is a signal processing technique used to enhance the transmission or reception of electromagnetic (EM) or acoustic signals by exploiting the interference properties of waves. Arrays of transmitters or receivers are required to produce a range of observations signals followed by a weighted averaging that optimally combines the various observations [1]. This results in, for example, higher fidelity speech recordings when using an array of microphones in the presence of noise and interference [2]. However, as these arrays become large the cost of centrally optimizing and computing our beamformer output becomes prohibitive.

With recent advancements in sensor networks [3, 4, 5, 6, 7] a natural progression for beamforming is its distribution over a set of self-contained nodes, each equipped with microprocessors, wireless EM communications and acoustic sensors. These nodes are assumed to be within communication range of only a local subset of the total nodes in the network. A distributed beamforming system should reduce total transmission energy required within the network, provide robustness to the addition or removal of nodes, eliminate the problem of a central master node failure, lower the coordination or calibration cost associated with setting up a large distributed network, and facilitate scalability of the beamformer to arbitrary network sizes [8].

One of the earliest distributed beamforming implementations for wireless sensor networks was by Bertrand and Moonen [9, 10] who distributed the processing of a linearly constrained minimum variance beamformer, which is a generalisation of the minimum variance distortionless response (MVDR) beamformer. Their method does not assume prior knowledge of the noise covariance matrix and can handle a full covariance matrix. However, their system does assume the network is fully connected (or connected in a tree topology, depending on the implementation method) and that each node

k is equipped with multiple sensors. Additionally an ordering of the computations in the nodes is required.

The distributed delay-and-sum beamformer developed by Zeng and Hendriks [11] iteratively broadcasts information between neighbouring nodes using a randomized gossip protocol [12]. The approach requires no restriction on the network topology and may perform updates independently across all nodes. However, the noise at all nodes must be uncorrelated and assumed known, resulting in suboptimal performance in the presence of general correlated interference. The message-passing based MVDR beamformer of [13] operates on scalar values in an asynchronous manner to perform weight vector optimization, does not require global convergence or collection phase for weight updates, and may be used in arbitrarily connected networks. However, the performance of the resulting MVDR beamformer is limited by the network topology as the interferences of non-neighbouring nodes are assumed to be uncorrelated. This is compounded by the requirement of the covariance matrix to be diagonally dominant, which is accomplished in [13] by an artificial damping of all off-diagonal elements. The diffusion-based MVDR beamformer of [14] is adaptive to varying interference statistics, only requires a single update per iteration, and approximates a full MVDR centralized beamformer with two-hop covariances. However, each node ultimately ends up with a large vector containing every node in the network's weight value as well as require the projection of this vector onto the linear constraint subspace. This limits the algorithm's true distributed nature, particularly in very large networks. Additionally, convergence can be slow due to the gradient optimization employed. All three of the above methods also require a global averaging at each time sample to produce a beamformer output.

More recently, the distributed privacy-protecting beamformer of [15] requires each node to hold a vector (with length on the order of the network size) of covariances with all other nodes, which is impractical and potentially impossible in very large networks. The LCMV beamformer of [16] assumes a fully connected network, but does contain details on how to extend the approach for more general networks. However, this extension requires node coordination to prune the partially connected network to a tree topology which introduces costly system overheads. Suboptimal approaches such as the time-frequency masking procedure in [17] and the related pseudo-coherence beamformer of [18] require less transmission power when compared with the above methods by utilizing sub-arrays of the total network. However, these approaches lack the flexibility to change their sub-array size depending on whether higher fidelity or lower power consumption is desired.

The contributions of this paper are a novel approach to distributed beamforming using sparsity, and a fully asynchronous and independent method for accomplishing this. Instead of computing a full network-wide MVDR beamformer we aim to find a subset of nodes and the associated beamformer that are optimal for a given

sparsity level. We therefore describe a distributed sparse beamformer that approximates a full MVDR beamformer with a sparsity tradeoff parameter, using the bi-alternating direction method of multipliers [19, 20, 21] for distributed optimization. The traditional MVDR beamforming cost function is regularized with an l_1 penalty of the weight vector to encourage sparsity in the final optimized weight vector. This reduces the number of nodes that contribute to the final beamformer output, simplifying the aggregation step required at each time sample and greatly improving the practicality of the distributed beamforming system.

2. SYSTEM MODEL AND BACKGROUND

We consider a network of nodes denoted by the set \mathcal{N} with cardinality $N = |\mathcal{N}|$. The network is connected by a set of edges \mathcal{E} with cardinality $E = |\mathcal{E}|$. If there exists an edge between two nodes k and l we say $(k, l) \in \mathcal{E}$. All nodes are considered self-connected. The node and edge sets together form our network graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$. The set of all nodes sharing an edge with a node k is denoted as its neighbourhood \mathcal{N}_k with cardinality $N_k = |\mathcal{N}_k|$, i.e. $\mathcal{N}_k = \{l | (k, l) \in \mathcal{E}\}$.

We denote our wideband signal of interest as $s_0(t, \omega)$, with t and ω indexing the time sample and frequency subband, respectively. Let $u_k(t, \omega)$ denote node k 's observed noisy signal at time t and subband ω that is a combination of the desired signal $s_0(t, \omega)$, interferences $s_p(t, \omega)$ and additive noise $v_k(t, \omega)$. $d_{k0}(\omega)$ and $d_{kp}(\omega)$ represent the complex acoustic transfer functions that scale and phase shift each subband of the source and interference signals, respectively, which are constant over time. We may then express the linear receive model at node k of our source signal, P interferers, and noise as

$$\begin{aligned} u_k(t, \omega) &= d_{k0}(\omega)s_0(t, \omega) + \sum_{p=1}^P d_{kp}(\omega)s_p(t, \omega) + v_k(t, \omega), \\ &= d_k(\omega)s(t, \omega) + n_k(t, \omega). \end{aligned} \quad (1)$$

The noise and interference can be combined into the variable $n_k(t, \omega)$ for convenience, allowing us to drop the 0 subscripts on $d_{k0}(\omega)s_0(t, \omega)$. When multiple sensors, such as microphones, are used we may represent the acoustic transfer functions to each sensor as the vector $\mathbf{d}(\omega) \in \mathbf{C}^N$, allowing us to express a vector of noisy observations across these sensors as

$$\mathbf{u}(t, \omega) = \mathbf{d}(\omega)s(t, \omega) + \mathbf{n}(t, \omega), \quad (2)$$

where $\{\mathbf{u}(t, \omega), \mathbf{n}(t, \omega)\} \in \mathbf{C}^N$ are the vectors of observations and interferences from all N sensors, respectively.

Beamforming aims to estimate the source signal $s(t, \omega)$ by combining the noisy observations in a weighted sum to produce a single scalar output

$$z(t, \omega) = \mathbf{w}^H(\omega)\mathbf{u}(t, \omega), \quad (3)$$

where $z(t, \omega)$ is the beamformed output, $(\cdot)^H$ represents Hermitian transposition, and $\mathbf{w}(\omega) \in \mathbf{C}^N$ is the weighting vector used to combine the observations. Since the signal is assumed to be independent over each frequency bin ω and for all time t we will simplify notation by omitting these indices henceforth.

The traditional MVDR beamformer [1, 2] is obtained by minimizing the energy of the beamformed output signal subject to unity gain in the direction of the source, which may be expressed by the constrained optimization problem

$$\begin{aligned} &\text{minimize} && \mathbb{E}[|\mathbf{w}^H \mathbf{u}|^2], \\ &\text{subject to} && \mathbf{d}^H \mathbf{w} = 1, \end{aligned} \quad (4)$$

which can be solved using Lagrange multipliers as

$$\mathbf{w} = \frac{R^{-1} \mathbf{d}}{\mathbf{d}^H R^{-1} \mathbf{d}}, \quad (5)$$

where $R = \mathbb{E}[\mathbf{n}\mathbf{n}^H]$ is the noise covariance matrix, and the source signal is assumed independent of the noise. The MVDR beamformer requires only knowledge of the source location (or equivalently the acoustic transfer functions \mathbf{d} at each node) and knowledge of the noise statistics in the form of the covariance matrix R .

Throughout the remainder of the paper we will assume a common mapping from the complex domain \mathbf{C}^N to the real domain of twice the dimensionality \mathbf{R}^{2N} in order to focus on the distributed convex optimization techniques employed. For the sake of brevity we omit the conversion of our problem, and for notational simplicity we will refer to our problem as one in \mathbf{R}^N . For details see, for example, [22].

3. DISTRIBUTED BIADMM BEAMFORMER

In this section we present the motivation for a sparse distributed beamformer, followed by the derivation of our system.

3.1. Physical Motivation for Distributed Sparse Beamforming

Previously, emphasis has been placed on creating fully distributed solutions to the traditionally centralized problem of MVDR beamforming. These methods ideally aim to produce a beamformer output that is equal to the centralized case, but approximations are often introduced to deal with the restriction of the network topology or the nature of the distributed algorithm being used. The most common approximations are limitations on the structure of the noise covariance matrix, such as the requirement of diagonal dominance or a sparsity structure representative of the underlying network graph. However, as of the time of this paper no work has yet dealt with the approach of optimizing for a sparse *weight vector*.

In very large networks, nodes distant from a source may practically capture none of the signal of interest, yet nearly all of the distributed beamforming methods reviewed require a mixing (or averaging) process across the entire network to produce a beamformer output at *each signal sample*. For networks on the order of 50 nodes (roughly the number used for simulation in the literature) the number of iterations required to reach convergence per sample is not a problem, but for network sizes where these distributed beamforming algorithms would actually be beneficial (at least on the order of thousands of nodes) the iteration time and associated communication cost required become infeasible.

One way of reducing the costly global mixing process is to encourage sparsity in our beamforming weight vector. Since those nodes closest to our source location are likely to be in the presence of a relatively high signal-to-noise ratio (due to the physical fall-off of sound power over distance), a sparsely optimized weight vector will naturally produce zero entries in nodes far from our source. Once we arrive at our sparse weight vector we may produce our beamformer output in one of three ways: we may perform the costly mixing process across the whole network; we may form a subnetwork with the nodes containing nonzero weight vector entries and perform a more efficient mixing over this far smaller network; or we may simply designate a single node (such as the node closest to the source) as our collection node for the current sample window and instruct all nodes with nonzero weight values to transmit to this collection node. Since all nonzero nodes will, in most cases, be physically near to the source the transmission distances required will generally not be prohibitive.

3.2. Derivation of the sparse distributed BiADMM beamformer

The proposed distributed sparse beamformer requires two separate operations. Firstly, the network will optimize (in a fully distributed manner) the weight values at all nodes while encouraging sparsity in these weights. This process will be independent and asynchronous at each node and requires communication only between neighbouring nodes. Additionally, this optimization process can be performed at any update rate without requiring network-wide convergence in a short time (such as between signal samples). Given these sparse weight values, we next require an aggregation operation to collect the weighted observations into a single scalar output. As mentioned previously, this may be done in multiple ways but we will restrict ourselves to single-node collection for the purpose of demonstrating the effectiveness of the sparse BiADMM beamformer.

3.2.1. Optimization of the Weighting Vector

We begin by optimizing for a scaled version of the weight vector $\mathbf{x}^* = R^{-1}\mathbf{d}$, where $(\cdot)^*$ refers to an optimal point, since the beamformer output is then simply the ratio of two averages [23]

$$z^* = \frac{\mathbf{x}^{*T}\mathbf{u}}{\mathbf{x}^{*T}\mathbf{d}} = \frac{\frac{1}{N}\sum_{k \in \mathcal{N}}[\mathbf{x}^*]_k^T[\mathbf{u}]_k}{\frac{1}{N}\sum_{k \in \mathcal{N}}[\mathbf{x}^*]_k^T[\mathbf{d}]_k}, \quad (6)$$

where $(\cdot)^T$ represents transposition and $[\cdot]_k$ denotes the k th element of a vector. Additionally, we would like to encourage sparsity in our optimized vector through the addition of an l_1 penalty on the resulting vector \mathbf{x} . Therefore, in order to obtain the optimal sparse vector \mathbf{x}^* we construct the unconstrained, l_1 -regularized quadratic program

$$\text{minimize } f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T R \mathbf{x} - \mathbf{d}^T \mathbf{x} + \alpha \|\mathbf{x}\|_1, \quad (7)$$

where α is our regularization parameter and $\|\cdot\|_1$ is the l_1 norm, widely known to encourage sparsity in our optimization solution [24, 25]. Next, we make an approximating assumption that nodes more than two communication hops apart are uncorrelated allowing us to decompose the quadratic term (as in [14]), resulting in

$$\begin{aligned} \text{minimize } & \sum_{k \in \mathcal{N}} \left(\frac{1}{2} \mathbf{x}_k^T R_k \mathbf{x}_k - \mathbf{d}_k^T \mathbf{x}_k \right) \\ \text{subject to } & A_{k \rightarrow l} \mathbf{x}_k + A_{l \rightarrow k} \mathbf{x}_l = 0 \quad \forall (k, l) \in \mathcal{E}, \end{aligned} \quad (8)$$

where the local vector $\mathbf{x}_k \in \mathbf{R}^{N_k}$ is node k 's estimate of the elements of \mathbf{x} belonging to its neighbourhood \mathcal{N}_k , $R_k = (C^{\dagger 2} \circ R) \in \mathbf{R}^{N_k \times N_k}$ is the covariance matrix for the neighbourhood \mathcal{N}_k that we assume to be estimated within each neighbourhood, $\mathbf{d}_k \in \mathbf{R}^{N_k}$ is a vector containing all zeros apart from the k th entry that is equal to node k 's scalar ATF d_k , the matrices $A_{k \rightarrow l} \in \mathbf{R}^{2 \times N_k}$ and $A_{l \rightarrow k} \in \mathbf{R}^{2 \times N_l}$ contain entries of 1, -1 or 0 to enforce consistency (with one row each for the consensus of node k and l 's primary elements and their copies) and $C^{\dagger 2}$ is the protected elementwise inverse of the square of the adjacency matrix [26, 14].

In order to include the l_1 -regularization term and to facilitate the computation of the Fenchel conjugate of our cost function, we consider each node k to have a virtual pair node $k+N$, connected only to node k , that holds the l_1 penalty for node k . Since the l_1 norm is separable across each element of \mathbf{x} we may simply assign each virtual node the absolute value of the scalar $[\mathbf{x}]_k$ as their penalty function. These virtual nodes will have consistency between the primary element of node k and their own scalar variable x_{k+N} enforced via consensus constraints, similar to those discussed above for the quadratic

decomposition. For notational simplicity we denote \mathcal{N} as the set of all physical nodes responsible for the quadratic penalty terms, and \mathcal{V} (with cardinality $V = |\mathcal{V}|$) as the set of all virtual nodes responsible for the l_1 -regularization penalty terms.

We now arrive at two distinct local cost function forms that capture the quadratic and l_1 -regularization terms of our original problem, i.e.

$$f_k(\mathbf{x}_k) = \begin{cases} \frac{1}{2} \mathbf{x}_k R_k \mathbf{x}_k - \mathbf{d}_k^T \mathbf{x}_k & \text{for } k \in \mathcal{N} \\ \alpha |x_k| & \text{for } k \in \mathcal{V}. \end{cases} \quad (9)$$

These result in an equivalent problem to (7) when summed over all real and virtual nodes:

$$\begin{aligned} \text{minimize } & \sum_{k=1}^{N+V} f_k(\mathbf{x}_k) \\ \text{subject to } & A_{k \rightarrow l} \mathbf{x}_k + A_{l \rightarrow k} \mathbf{x}_l = 0 \quad \forall (k, l) \in \mathcal{E}. \end{aligned} \quad (10)$$

Finally, in order to use the BiADMM algorithm [21] we require the Fenchel conjugates for our local cost functions of (9). These may be derived independently (due to the formulation of our problem using virtual nodes) as

$$f_k^*(A_k^T \boldsymbol{\lambda}_k) = \begin{cases} \frac{1}{2} \boldsymbol{\lambda}_k A_k R_k^{-1} A_k^T \boldsymbol{\lambda}_k + \mathbf{d}_k^T R_k^{-1} A_k^T \boldsymbol{\lambda}_k + \frac{1}{2} \mathbf{d}_k^T R_k^{-1} \mathbf{d}_k & \text{for } k \in \mathcal{N} \\ \|A_k^T \boldsymbol{\lambda}_k\|_\infty & \text{for } k \in \mathcal{V}, \end{cases} \quad (11)$$

where $A_k^T \boldsymbol{\lambda}_k = \sum_{k \in \mathcal{N}_k} A_{k \rightarrow l}^T \boldsymbol{\lambda}_{k|l}$. Using the local primal cost functions (9) and the local conjugate cost functions (11), we may now derive the BiADMM primal and dual variable update equations, as described in [21]. The primal update equations are

$$\mathbf{x}_k^{i+1} = \begin{cases} \left(\sum_{l \in \mathcal{N}_k} A_{k \rightarrow l}^T A_{k \rightarrow l} + R_k \right)^{-1} \left(\sum_{l \in \mathcal{N}_k} A_{k \rightarrow l}^T (\boldsymbol{\lambda}_{l|k}^i - A_{l \rightarrow k} \mathbf{x}_l^i) + \mathbf{d}_k \right) & \text{for } k \in \mathcal{N} \\ \frac{(-b_k + \text{sgn}(b_k) \min\{|b_k|, \alpha\})}{(A_{k \rightarrow l}^T A_{k \rightarrow l})} & \text{for } k \in \mathcal{V}, \end{cases} \quad (12)$$

where $b_k = (2\boldsymbol{\lambda}_{l|k}^i A_{l \rightarrow k}^T A_{k \rightarrow l} - A_{k \rightarrow l}^T \boldsymbol{\lambda}_{l|k}^i)$, $\text{sgn}(\cdot)$ is the sign function, and $\min\{\cdot\}$ outputs the minimum of the set $\{\cdot\}$. The dual update equations are

$$\boldsymbol{\lambda}_{k|l}^{i+1} = \begin{cases} \boldsymbol{\lambda}_{l|k}^i - A_{l \rightarrow k} \mathbf{x}_k^i - A_{k \rightarrow l} R_k^{-1} (\mathbf{d}_k + \mathbf{z}_k) & \text{for } k \in \mathcal{N} \\ P_\infty^\alpha (2\boldsymbol{\lambda}_{l|k}^i \boldsymbol{\lambda}_{k|l} - \mathbf{x}_k^{iT} A_{l \rightarrow k}^T) & \text{for } k \in \mathcal{V}. \end{cases} \quad (13)$$

where I_{N_k} is the $N_k \times N_k$ identity matrix, $P_\infty^\alpha(c)$ is Euclidean projection of c onto the rectangular set $\{c | -\alpha \leq c \leq \alpha\}$ [27], and

$$\mathbf{z}_k = \left(\sum_{l \in \mathcal{N}_k} A_{k \rightarrow l}^T A_{k \rightarrow l} R_k^{-1} + I_{N_k} \right)^{-1} \sum_{l \in \mathcal{N}_k} A_{k \rightarrow l}^T (\boldsymbol{\lambda}_{l|k}^i - A_{l \rightarrow k} \mathbf{x}_k^i - A_{k \rightarrow l} R_k^{-1} \mathbf{d}_k). \quad (14)$$

In practice, when a node k randomly triggers for updating it would perform both the quadratic primal/dual updates, followed by its virtual node $(k+N)$'s l_1 primal/dual updates since these are both physically the same node. The primal and dual variable update equations are asynchronous, fully distributed, and may be performed continuously over time at any rate deemed appropriate.

3.2.2. Computation of the Beamformer Output

For each signal sample we require a collection phase where the noisy observations from throughout the network are weighted and summed to produce our beamed output signal. At each time sample t we therefore must perform sharing and averaging iterations many times for information to mix throughout the network. However, as we will show in our simulation results, information at distant nodes within the network often provide negligible performance gains. For the sake of our sparse BiADMM beamformer we simply transmit the weighted observations of our active nodes to a single collection node (designated as the node closest to the source).

4. EXPERIMENTAL RESULTS

In this section we describe our simulation setup and present some experimental results.

4.1. Experimental Setup

We simulated a network with $N = 50$ microphone nodes and a source signal randomly placed in a $100\text{ m} \times 100\text{ m} \times 100\text{ m}$ free-space environment. The results are averaged over 20 realizations. The distances from node k to all other nodes were assumed to be known and neighbours were truncated to fall within a transmission distance of 50 m. The acoustic transfer function for each node was generated using the free-space model. The signal of interest was a 20 s speech sample randomly chosen from a 60 s recording. The interference was a randomly placed zero-mean Gaussian point source with power equal to -5 dB , 0 dB , and 5 dB when compared to the source signal. Estimation of the partial covariance matrix was assumed beforehand. The sample rate at each node was $f_s = 16\text{ kHz}$ and processing was carried out on 25 ms Hann windowed blocks with a 50% overlap. The weight vector optimization process was performed asynchronously.

4.2. Results

Figure 1 shows the convergence of the BiADMM weight vector optimization for various signal-to-noise ratios, where these iterations were performed once per signal sample. We observe a linear decrease in weight vector mean squared error until the noise floor begins to slow convergence.

Figure 2 shows the effect of sparsity on our output SNR and on the communication power required for beamformer outputs per sample, when compared with the fully active MVDR node set (which uses distributed averaging for sample output). Due to our local covariance matrix approximation we see a drop of around 3 dB from the optimal case when all nodes are active, with a loss of at most 7 dB as our active node set falls to below 10 nodes. Communication power required for output (using [28] for only power amplifier transmission with free space parameters) in our sparse system using single node collection is initially higher than distributed averaging due to the expense of long range wireless communication with distant active nodes. However, our sparse system rapidly becomes more efficient as these distant (low fidelity nodes) are excluded, falling to around 15 % of the power required for distributed network-wide averaging.

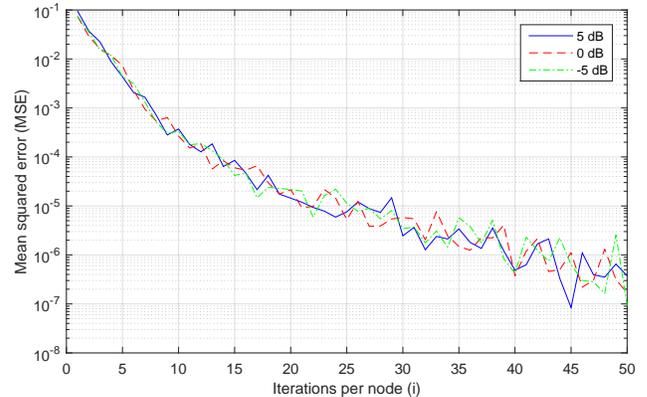


Fig. 1. Convergence to optimal MVDR weight vector for -5 dB , 0 dB , and 5 dB raw signal SNR with no sparsity.

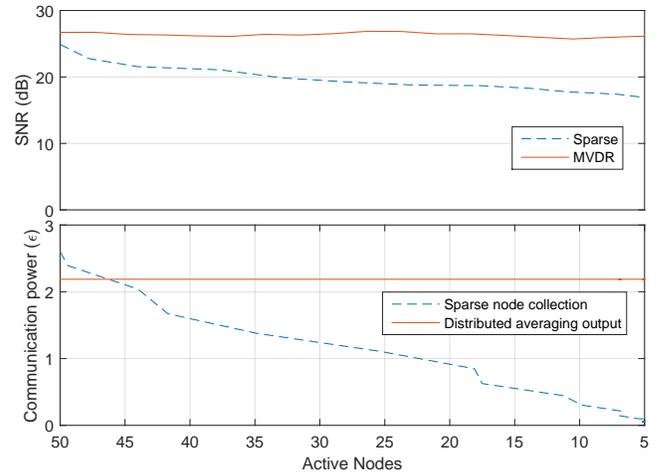


Fig. 2. Top: SNR as a function of active nodes; Bottom: Communication Power as a function of active nodes.

5. CONCLUSION

We have motivated, designed and tested a sparse distributed beamforming system that trades SNR performance for reduced inter-node power consumption. The l_1 norm is a natural first application of regularization to the problem due to the physical fall-off of sound power over distance and the resulting node fidelity pattern, particularly for large and sparse networks. Our system optimization performs updates that are asynchronous, independent, and do not rely on global collection phases. Additionally, since distant nodes generally have their weights optimized to zero the sampling synchronization errors encountered in large networks, required for beamformer output, are reduced. Further work would include a convergence analysis for regularized BiADMM over graphs and research into more appropriate regularization functions that, for example, explicitly take into account the fidelity of each node and their distance from the source.

6. REFERENCES

- [1] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Acoustics, Speech, and Signal Processing Magazine*, vol. 5, no. 2, pp. 4–24, 1988.
- [2] S. Haykin, *Adaptive Filter Theory*. NJ: Prentice Hall, 1996.
- [3] K. Frampton, "Acoustic self-localization in a distributed sensor network," *IEEE Sensors Journal*, vol. 6, no. 1, pp. 166–172, Feb 2006.
- [4] N. Kantas, S. Singh, and A. Doucet, "Distributed maximum likelihood for simultaneous self-localization and tracking in sensor networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 10, pp. 5038–5047, Oct 2012.
- [5] A. Wiesel and A. Hero, "Distributed covariance estimation in gaussian graphical models," *Signal Processing, IEEE Transactions on*, vol. 60, no. 1, pp. 211–220, Jan 2012.
- [6] J. Schmalenstroeyer, P. Jebramcik, and R. Haeb-Umbach, "A gossiping approach to sampling clock synchronization in wireless acoustic sensor networks," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, May 2014, pp. 7575–7579.
- [7] M. Li, G. Zheng, and J. Li, "Clock self-synchronization protocol based on distributed diffusion for wireless sensor networks," *International Journal of Future Generation Communication and Networking*, vol. 7, 2014.
- [8] J. Her, T. Wysocki, and B. Wysocki, "Review of distributed beamforming," *Journal of Telecommunications and Information Technology*, pp. 78–88, 2011.
- [9] A. Bertrand and M. Moonen, "Distributed adaptive estimation of node-specific signals in wireless sensor networks with a tree topology," *Signal Processing, IEEE Transactions on*, vol. 59, no. 5, pp. 2196–2210, 2011.
- [10] —, "Distributed node-specific LCMV beamforming in wireless sensor networks," *Signal Processing, IEEE Transactions on*, vol. 60, no. 1, pp. 233–246, 2012.
- [11] Y. Zeng and R. C. Hendriks, "Distributed delay and sum beamformer for speech enhancement in wireless sensor networks via randomized gossip," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*. IEEE, 2012, pp. 4037–4040.
- [12] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized Gossip Algorithms," *IEEE Transactions on Information Theory*, vol. 52, no. 6, 2006.
- [13] R. Heusdens, G. Zhang, R. C. Hendriks, Y. Zeng, and W. B. Kleijn, "Distributed MVDR Beamforming for (Wireless) Microphone Networks Using Message Passing," *International Workshop on Acoustic Signal Enhancement*, 2012.
- [14] M. O'Connor and W. B. Kleijn, "Diffusion-Based Distributed MVDR Beamformer," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, May 2014, pp. 810–814.
- [15] Y. Zeng and R. C. Hendriks, "Distributed estimation of the inverse of the correlation matrix for privacy preserving beamforming," *Signal Processing*, vol. 107, pp. 109–122, 2015.
- [16] A. Bertrand and M. Moonen, "Distributed LCMV beamforming in a wireless sensor network with single-channel per-node signal transmission," *Signal Processing, IEEE Transactions on*, vol. 61, no. 13, pp. 3447–3459, 2013.
- [17] M. Taseska, S. Markovich-Golan, E. Habets, and S. Gannot, "Near-field source extraction using speech presence probabilities for ad hoc microphone arrays," in *Acoustic Signal Enhancement (IWAENC), 2014 14th International Workshop on*, Sept 2014, pp. 169–173.
- [18] V. Tavakoli, J. Jensen, M. Christensen, and J. Benesty, "Pseudo-coherence-based mvdr beamformer for speech enhancement with ad hoc microphone arrays," in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*, April 2015, pp. 2659–2663.
- [19] G. Zhang and R. Heusdens, "Bi-alternating direction method of multipliers," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, May 2013, pp. 3317–3321.
- [20] G. Zhang, R. Heusdens, and W. Kleijn, "On the convergence rate of the bi-alternating direction method of multipliers," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, May 2014, pp. 3869–3873.
- [21] G. Zhang and R. Heusdens, "Bi-alternating direction method of multipliers over graphs," *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*, April 2015.
- [22] J. K. Au, "An ab initio approach to the inverse problem-based design of photonic bandgap devices," Ph.D. dissertation, California Institute of Technology, 2007.
- [23] R. Heusdens, G. Zhang, R. C. Hendriks, Y. Zeng, and W. B. Kleijn, "Distributed MVDR beamforming for (wireless) microphone networks using message passing," in *Acoustic Signal Enhancement; Proceedings of IWAENC 2012; International Workshop on*, Sept 2012, pp. 1–4.
- [24] F. Bach, R. Jenatton, J. Mairal, G. Obozinski *et al.*, "Convex optimization with sparsity-inducing norms," *Optimization for Machine Learning*, pp. 19–53, 2011.
- [25] M. Schmidt, G. Fung, and R. Rosales, "Optimization methods for l1-regularization," *University of British Columbia, Technical Report TR-2009*, vol. 19, 2009.
- [26] F. Harary, *Graph Theory*. Reading, MA: Addison-Wesley, 1994.
- [27] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.
- [28] Q. Wang, M. Hempstead, and W. Yang, "A realistic power consumption model for wireless sensor network devices," in *Sensor and Ad Hoc Communications and Networks, 2006. SECON'06. 2006 3rd Annual IEEE Communications Society on*, vol. 1. IEEE, 2006, pp. 286–295.