A distributed algorithm for robust LCMV beamforming

Sherson, Thomas; Kleijn, W. Bastiaan; Heusdens, Richard

Citation (APA)
Sherson, T., Kleijn, W. B., & Heusdens, R. (2016). A distributed algorithm for robust LCMV beamforming. In M. Dong, & T. F. Zheng (Eds.), *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP): Proceedings* (pp. 101-105). IEEE . https://doi.org/10.1109/icassp.2016.7471645

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

A DISTRIBUTED ALGORITHM FOR ROBUST LCMV BEAMFORMING

Thomas Sherson[1]   *W. Bastiaan Kleijn*[1,2]   *Richard Heusdens*[1]

[1] Faculty of EEMCS, Delft University of Technology, Netherlands
[2] School of Engineering and Computer Science, Victoria University of Wellington, New Zealand

ABSTRACT

In this paper we propose a distributed reformulation of the linearly constrained minimum variance (LCMV) beamformer for use in acoustic wireless sensor networks. The proposed distributed minimum variance (DMV) algorithm, for which we demonstrate implementations for both cyclic and acyclic networks, allows the optimal beamformer output to be computed at each node without the need for sharing raw data within the network. By exploiting the low rank structure of estimated covariance matrices in time-varying noise fields, the algorithm can also provide a reduction in the total amount of data transmitted during computation when compared to centralised solutions. This is particularly true when multiple microphones are used per node. We also compare the performance of DMV with state of the art distributed beamformers and demonstrate that it achieves greater improvements in SNR in dynamic noise fields with similar transmission costs.

Index Terms— Distributed, LCMV, acoustic, beamforming.

1. INTRODUCTION

In speech processing, a common problem of interest is the recovery of a speech signal from within a set of recordings containing noise and interference. This task is often addressed via the use of beamformers (BFs), algorithms that exploit the spatial correlation of recordings to favour particular source locations over others [19]. While such systems are traditionally comprised of physically connected arrays of microphones, recent improvements in sensor and battery technologies have made it practical to use wireless sensor networks (WSNs) for the same application [5]. However, the decentralised nature of data acquisition in WSNs makes designing statistically optimal BFs a challenging procedure. Thus, distributed signal processing is an important tool in the implementation of WSN-based BFs, particularly in the presence of time-varying noise fields.

Existing distributed BFs are based on a range of different approaches. In [22] a distributed delay and sum BF utilising randomised gossip is proposed. While operating in any undirected network, this method only minimises the presence of uncorrelated noise. In many practical applications this results in a sub-optimal BF response.

In [6], a message passing approach is adopted to compute a minimum variance distortionless response (MVDR) BF for a known covariance matrix. However, this method requires that the sparsity pattern of the covariance matrix matches that of the network adjacency matrix. Either this sparsity pattern dictates the WSN connectivity and thus the network topology or an approximation of the covariance matrix needs to be made to enforce this match. Such an approximation again leads to a sub-optimal response.

In [14], a diffusion adaptation approach was used to design an approximate MVDR BF by constructing local covariance matrices based on the exchange of data between nodes. Unfortunately this approach requires the transmission of a solution vector between nodes which scales with network size, increasing each nodes power consumption and memory requirements. Additionally as this technique does not capture the true covariance of the centralised problem it also results in a suboptimal BF.

In contrast, the BFs in [1, 11] choose to restrict the WSN networks to be either fully connected or acyclic (tree shaped). By exploiting the ease of data aggregation in such networks, these algorithms allow for the combination of local BFs based on linearly constrained minimum variance (LCMV) [1] and generalised sidelobe canceller (GSC) [11] topologies, to approximate a global BF response. Both algorithms achieve fully scalable distributed implementations which, in static noise fields, converge to the optimal BF over multiple updates. However, in the case of varying noise fields, the rate of adaptation of these solutions can again result in suboptimal performance due to their inherently iterative nature. Furthermore the requirement of tree-shaped or fully connected networks is an unrealistic constraint for the likes of ad-hoc or self configuring networks. This in turn limits the practicality and use of these algorithms in real world applications.

The work presented in this paper approaches the development of a distributed BF by casting a robust version of the standard LCMV BF as a distributed convex optimisation problem [3, 24]. Unlike existing distributed implementations, the proposed algorithm does not require multiple updates to compute the optimal BF response. Additionally, the BF computation can be performed via a range of existing distributed tools [2, 13, 16, 21, 17] for both cyclic and acyclic networks. This flexibility makes it possible to use the proposed algorithm for optimal beamforming without the impractical restriction of enforcing tree-shaped networks, as required in [1, 11].

In the remainder of this paper we introduce our distributed reformulation of the robust LCMV BF, termed the distributed minimum variance (DMV) BF, for which we provide both cyclic and acyclic implementations. The performance of the DMV algorithm is then compared to existing state of the art approaches in the case of a per-block updating scheme. This shows how our algorithm provides an improved rate of adaptation in dynamic noise fields compared to existing solutions with a similar overall communication cost.

2. THE DISTRIBUTED MINIMUM VARIANCE (DMV) BEAMFORMER

In this section we develop a distributed reformulation of the diagonally loaded (robust) LCMV BF. Subsection 2.1 describes the separable conversion of the robust LCMV algorithm; subsection 2.2 constructs a distributed dual problem to eliminate the global constraint functions; and finally, in subsections 2.3 and 2.4 we demonstrate

ICASSP 2016

DMV implementations for acyclic and cyclic networks respectively.

2.1. Constructing a separable robust LCMV beamformer

Consider the use of a WSN of N nodes equipped with a total of M microphones in forming a wide-band acoustic BF. Based on a linear propagation model, each microphone records a noise corrupted version of the target signal (x) given by $y_m(t) = d_m(t) * x(t) + n_m(t)$ where n_m and d_m denote local additive noise and the acoustic transfer function (ATF) from the target source to microphone m respectively. We assume that these ATFs are known and that the nodes are synchronised by a global clock.

Let k be a time index of overlapping windowed sections of audio. If the maximum delay introduced by the ATFs is less than the windowing length, then the STFTs of the microphone signals can be approximated by $Y_{m(k)} = D_{m(k)}U_{(k)} + N_{m(k)}$.

For such a system, consider the optimal estimation of $U_{(k)}$ via a standard LCMV BF with an additional regularisation term to control the energy of the BF weight vector w. As noted in [4, 8, 9], this regularisation term improves the robustness of the BF to errors in the source location such as those which occur when these are estimated rather than known. Let α scale the regularisation term and thus control its importance relative to the LCMV objective. For each frequency bin, the optimisation criterion of this BF is given by

$$\min_{w} \quad \frac{1}{2}w^H R w + \frac{\alpha}{2}||w||^2$$
$$\text{s.t.} \quad D^H w = s. \tag{1}$$

Here, $R \in \mathbb{C}^{M \times M}$ is the signal covariance matrix, $D \in \mathbb{C}^{M \times P}$ denotes the set of frequency domain ATFs from P target sources and $s \in \mathbb{C}^P$ denotes the desired response in the direction of these sources.

In practice, R is estimated via an unbiased sample covariance matrix

$$\hat{R} = \frac{1}{L}\sum_{l=1}^{L} Y_{(l)}Y_{(l)}^H, \tag{2}$$

where L is the number of samples used in this estimation process and each $Y_{(l)}$ denotes the microphone signals in the frequency domain in section $k - l + 1$. We also assume that the WSN used can be represented via an undirected graph G with vertex set V and edge set E, denoting the sets of nodes and communication channels between nodes respectively.

By substituting (2) into (1) and splitting the objective into node-based variables (denoted by a subscript), we can rewrite (1) as

$$\min_{w_i \forall i \in V} \quad \sum_{l=1}^{L}\left(\frac{1}{2L}|\sum_{i \in V}Y_{i(l)}^H w_i|^2\right) + \sum_{i \in V}\left(\frac{\alpha}{2}||w_i||^2\right)$$
$$\text{s.t.} \quad \sum_{i \in V} D_i^H w_i = s, \tag{3}$$

where $w_i \in \mathbb{C}^{m_i}$, $Y_{i(l)} \in \mathbb{C}^{m_i}$, $D_i \in \mathbb{C}^{p \times m_i}$ and m_i denote the weightings, windowed signals, ATFs and number of microphones at node i respectively. To create a separable objective function we then introduce a set of additional variables $\hat{x}_{i(l)} \forall i \in V, l = 1, ..., L$ subject to the constraint that

$$\sum_{i \in V} \hat{x}_{i(l)} = N\sum_{j \in V} Y_{j(l)}^H w_j \tag{4}$$

We can therefore construct an alternative optimisation problem to (3) which is given by

$$\min_{w_i, \hat{x}_{i(l)}} \quad \sum_{i \in V}\left(\sum_{l=1}^{L}\left(\frac{|\hat{x}_{i(l)}|^2}{2LN}\right) + \frac{\alpha}{2}||w_i||^2\right)$$
$$\text{s.t.} \quad \sum_{i \in V} D_i^H w_i = s$$
$$NY_{(l)}^H w = \sum_{i \in V}\hat{x}_{i(l)} \quad \forall l = 1, ..., L. \tag{5}$$

Proposition 1. *Problems* (3) *and* (5) *are equivalent.*

Proof. Consider the modified Lagrangian ($\mathcal{L}$) of (5) given by

$$\mathcal{L}(w, \hat{x}, \nu, \mu) = \sum_{i \in V}\left(\sum_{l=1}^{L}\left(\frac{|\hat{x}_{i(l)}|^2}{2LN} - \frac{\nu_{(l)}^*}{2}\left(NY_{i(l)}^H w_i - \hat{x}_{i(l)}\right)\right)\right.$$
$$-\frac{\nu_{(l)}}{2}\left(NY_{i(l)}^T w_i^* - \hat{x}_{i(l)}^*\right)\right) - \frac{\mu^H}{2}\left(D_i^H w_i - \frac{s}{N}\right)$$
$$\left.-\frac{\mu^T}{2}\left(D_i^T w_i^* - \frac{s^*}{N}\right) + \frac{\alpha}{2}||w_i||^2\right), \tag{6}$$

with dual variables $\nu_{(l)} \forall l = 1, ..., L$ and μ. This modified Lagrangian ensures that (6) is a real valued function [7]. We are interested in finding the values of $\hat{x}_{i(l)}$ corresponding to the stationary points of (6). However, as (6) is not an analytic function of $\hat{x}_{i(l)}$, its derivative is undefined. As demonstrated in [12, 18], the stationary points of (6) can alternatively be found by treating $\hat{x}_{i(l)}$ and $\hat{x}_{i(l)}^*$ as independent variables. The stationary points then occur at the intersecting zeros of the partial derivatives $\frac{\partial \mathcal{L}}{\partial \hat{x}_{i(l)}}$ and $\frac{\partial \mathcal{L}}{\partial \hat{x}_{i(l)}^*}$ such that $(\hat{x}_{i(l)})^* = \hat{x}_{i(l)}^*$. Therefore, at optimality, we find that

$$\frac{\partial \mathcal{L}}{\partial \hat{x}_{i(l)}} = \frac{(\hat{x}_{i(l)})^*}{2LN} + \frac{\nu_{(l)}^*}{2} = 0$$
$$\frac{\partial \mathcal{L}}{\partial \hat{x}_{i(l)}^*} = \frac{\hat{x}_{i(l)}}{2LN} + \frac{\nu_{(l)}}{2} = 0. \tag{7}$$

Given (4) and (7) it can be shown that $\hat{x}_{i(l)} = Y_{(l)}^H w \,\forall i \in V$ and thus, by inspection, (3) and (5) are equivalent. ∎

2.2. An equivalent distributed dual problem

As (6) has a separable structure it can be converted to the dual domain to obtain a fully distributed optimisation problem. For the sake of simplicity, we first rewrite (6) as

$$\mathcal{L}(y, \lambda) = \sum_{i \in V}\left(\frac{1}{2}||z_i||_{A_i}^2 - \frac{\lambda^H}{2}\left(B_i^H z_i + C\right)\right.$$
$$\left.-\frac{\lambda^T}{2}\left(B_i^T z_i^* + C^*\right)\right). \tag{8}$$

where

$$z_i = [\hat{x}_{i(1)}, \hat{x}_{i(2)}, ..., \hat{x}_{i(L)}, w_{i(1)}, w_{i(2)}, ..., w_{i(m_i)}]^T$$
$$\lambda = [\nu_{(1)}, \nu_{(2)}, ..., \nu_{(L)}, \mu]^T$$
$$A_i = \text{diag}\left([\frac{1}{LN}, \frac{1}{LN}, ..., \frac{1}{LN}, \alpha, \alpha, ..., \alpha]^T\right)$$
$$B_i = \begin{bmatrix} -I_{L \times L} & & & 0_{L \times P} \\ NY_{i(1)} & \cdots & NY_{i(L)} & D_i \end{bmatrix}$$
$$C = [0, 0, ..., 0, \frac{s^T}{N}]^T. \tag{9}$$

The stationary points and thus dual of (8) can then be found in the same manner as (7) via the use of complex partial derivatives. These stationary points can be shown to occur when

$$z_i = A_i^{-1} B_i \lambda. \tag{10}$$

The dual problem can be found by substituting (10) into (8) giving

$$\min_{\lambda} \quad \frac{1}{2} \lambda^H \left(\sum_{i \in V} B_i^H A_i^{-1} B_i \right) \lambda - \sum_{i \in V} \frac{(\lambda^H C + \lambda^T C^*)}{2}. \tag{11}$$

Finally, by introducing a set of local variables $\lambda_i \, \forall i \in V$ and imposing that, at consensus, $\lambda_i = \lambda_j \, \forall (i,j) \in E$, we obtain a distributed form of (3) given by

$$\min_{\lambda_i \, \forall i \in V} \quad \sum_{i \in V} \left(\frac{1}{2} \lambda_i^H B_i^H A_i^{-1} B_i \lambda_i - \frac{(\lambda_i^H C + \lambda_i^T C^*)}{2} \right)$$
$$\text{s.t.} \quad \lambda_i - \lambda_j = 0 \quad \forall (i,j) \in E. \tag{12}$$

Through the combination of (10) and (12), we can calculate the optimal beamformer output $\hat{x}_{i(1)}$ and local weight vector w_i at each node in a fully distributed manner.

2.3. Implementing DMV in acyclic WSNs

Equations (11) and (12) can be used to define a number of different implementations of DMV for various network topologies. Firstly, we consider the more restrictive and less practical case of DMV operating in an acyclic network to allow for comparison with other state of the art algorithms.

In this context we note that to solve (11) we only need to calculate the matrix $\sum_{i \in V} B_i^H A_i^{-1} B_i$ as the vector C is already known at each node. This can be achieved by aggregating data via the max-sum algorithm [20] where each node i generates messages $m_{ij} \in \mathbb{C}^{(L+P) \times (L+P)}$, for their set of neighbours $\mathcal{N}(i) = \{j | (i,j) \in E\}$, given by

$$m_{ij} = B_i^H A_i^{-1} B_i + \sum_{k \in \mathcal{N}(i), k \neq j} m_{ki}. \tag{13}$$

Here, each message is a positive definite matrix with $\frac{(L+P)^2}{2} + \frac{L+P}{2}$ unique variables. Thus as the max-sum algorithm converges in $2N - 1 - K$ transmissions for an acyclic network [15], where K denotes the number of leaf nodes, this provides a tight bound on the amount of data which needs to be transmitted to reach consensus.

In the presence of a dynamic noise field, where the DMV algorithm needs to be solved for each windowed section of audio, we can reduce the total data transmitted per message. This is due to the structure of each $B_i^H A_i^{-1} B_i$ matrix and in particular, the reuse of $L - 1$ sections of audio between successive DMV problems. As such, only $L + P$ new variables need to be transmitted per message for stationary sources. The inclusion of dynamic source locations requires a further $P(L - 1) + \frac{P^2}{2} + \frac{P}{2}$ variables to be transmitted per message.

By reusing the repeated matrix entries, the transmission costs of the block-updated acyclic DMV implementation (for each frequency bin) are given by

$$(L + P)(2N - 1 - K) \quad \textit{Static}$$
$$\left(L(1 + P) + \frac{P^2}{2} + \frac{P}{2} \right)(2N - 1 - K) \quad \textit{Non-static} \tag{14}$$

2.4. Implementing DMV in WSNs containing cyclic paths

Generally speaking, practical WSNs will contain cyclic paths which do not allow us to exploit the efficient data aggregation techniques used in subsection 2.3. These cycles are particularly common in ad-hoc or self configuring systems which are often used in real world applications.

For cyclic WSNs, (12) has been formulated to be directly solvable via a range of iterative distributed signal processing techniques including the alternating direction method of multipliers (ADMM [2]) and the primal dual method of multipliers (PDMM, formally BiADMM [23]). Therefore, we can construct a cyclic implementation of DMV using PDMM by introducing a set of dual-dual variables $\gamma_{ij} \, \forall (i,j) \in E$ and adopting a simplified updating scheme, as described in [23]. These edge related variables reflect the communication channels between nodes where data is only able to be transmitted between neighbouring pairs, $(i,j) \in E$, in the form of "primal" ($\lambda_i^{(t)}$) and "dual" ($\gamma_{ij}^{(t)}$) estimates. For the sake of brevity, the primal and dual updates, which need to be solved per iteration, are included below without derivation.

$$\lambda_i^{(t+1)} = \left(B_i^H A_i^{-1} B_i + \sum_{j \in \mathcal{N}(i)} R_{pij} \right)^{-1}$$
$$\left(C + \sum_{j \in \mathcal{N}(i)} \left(\frac{i-j}{|i-j|} \gamma_{ji}^{(t)} + R_{pij} \lambda_j^{(t)} \right) \right)$$
$$\gamma_{ij}^{(t+1)} = \gamma_{ji}^{(t)} - \frac{i-j}{|i-j|} R_{pij} \left(\lambda_i^{(t+1)} - \lambda_j^{(t)} \right). \tag{15}$$

Here t is used to index the iterations of the algorithm whilst each $R_{pij} = R_{pji} \succ 0$ is a unique penalty term for the edge $(i,j) \in E$. When operated in an asynchronous fashion, a single node is activated for each t which updates as per (15). This node then transmits its $\lambda_i^{(t+1)} \in \mathbb{C}^{L+P}$ and $\gamma_{ij}^{(t+1)} \in \mathbb{C}^{L+P}$ estimates to its neighbouring nodes. If a variable should not be updated in either step it simply retains its previous value e.g. $\lambda_j^{(t+1)} = \lambda_j^{(t)}$. This process, which is run for each frequency bin and section of audio, continues until the network reaches consensus at which point the algorithm terminates.

3. DMV PERFORMANCE IN DIFFERENT NETWORK TOPOLOGIES

We now demonstrate the characteristics of (13) and (15) in computing the DMV response. Subsection 3.1 focuses on the iterative convergence of DMV using PDMM while subsection 3.2 compares the performance of acyclic DMV with other state of the art algorithms, in particular distributed LCMV (DLCMV) [1] and distributed GSC (DGSC) [11].

3.1. DMV in graphs containing cyclic paths

As nodes will often be equipped with finite battery supplies, power consumption, which is largely contributed to by the transmission of data, is a primary consideration of WSN based systems. Thus, although DMV computes an optimal BF response we are interested in how quickly it converges to the optimal solution. In the case of (15), this rate of convergence is highly dependent on the selection of the penalty terms R_{pij} for which, at this time, an optimal choice is unknown. Therefore, for the sake of demonstration, a random instance of this algorithm was simulated for a network of 5 nodes for various topologies with $R_{pij} = \frac{1}{N}(B_i + B_j)^H A_i^{-1}(B_i + B_j) \, \forall (i,j) \in E$.

We have found that this particular choice often improves convergence rate over other options such as scaled identity matrices. The results of this simulation are included below in Figure 1.
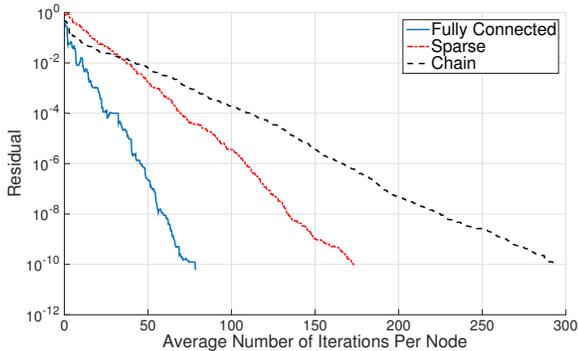

Fig. 1: The convergence of PDMM based DMV for a fully connected network, a random sparse network and a chain network

The impact of the differing network topologies in Figure 1 on the conversion of DMV highlights one of the major challenges of these cyclic methods: it is difficult to bound the transmission costs of distributed algorithms for arbitrary networks. However, such limitations also effect other existing BFs [6, 22] which, unlike DMV, do not achieve a statistically optimal response. Additionally, as DMV is not dependent on a particular solver, should improved or finite convergence acyclic solvers be developed in the future then these can be utilised to develop new BF implementations without further modification of the base algorithm.

3.2. A comparison between acyclic DMV and the state of the art

We now compare the bounded transmission costs (Tx) of acyclic DMV with other state of the art algorithms, namely DGSC and DLCMV. Additionally, the performance of a centralised implementation based on a simple data aggregation method was considered. The costs of these algorithms, under the assumption of stationary source positions, are included in Table 1.

Table 1: Transmission costs of different distributed BFs operated with a per-block updating scheme in an acyclic network.

Algorithm	Tx per frequency bin
DMV	$(L+P)(2N-1-K)$
DLCMV [1]	$(L+P)(2N-1-K)$
DGSC [11]	$(1+P)(2N-1-K)$
Centralised	$M(N-1)$

As per Table 1, DLCMV and DMV require the same number of transmissions to update their BFs with both providing an improvement over the centralised implementation in large networks. This stems from the fact that for dynamic noise environments, $2(L+P) < M$ due to the de-correlation of audio sections with time which limits the practical size of L. In contrast, as DGSC uses an LMS updating scheme, rather than constructing a covariance matrix, it again

reduces data transmission per block. However, this reduction comes at the cost of a decrease in the rate of adaptation of DGSC to changes in the noise field. As such, in practical applications DGSC requires a more frequent updating scheme to improve its ability to track noise correlation, increasing the total power it consumes [10].

The iterative nature of DLCMV also effects its rate of adaptation, particularly in the case of its tree-shaped variant where the degrees of freedom of the local BFs are greatly reduced [1]. This results in slower convergence to the optimal BF even in the case of static noise fields. Additionally, as only one node updates per audio section in DLCMV, increasing the number of nodes in the network decreases the algorithms ability to adapt to changing noise fields [10]. To demonstrate this, we consider the case of a linear array with $m_i = 1 \, \forall i \in V$ targeting a single speaker positioned perpendicular to the array. A moving noise source approaches and passes to one side of the array in a perpendicular direction at a speed of 5 ms^{-1}, perhaps simulating someone biking past. For $N = 1, ..., 11$ the improvements in SNR (SNR$_{imp}$) over that of a single node recording were simulated for both DMV and DLCMV using a window length of 512 samples with a 50% overlap and a sampling frequency of 22kHz. The resulting data is included in Figure 2.
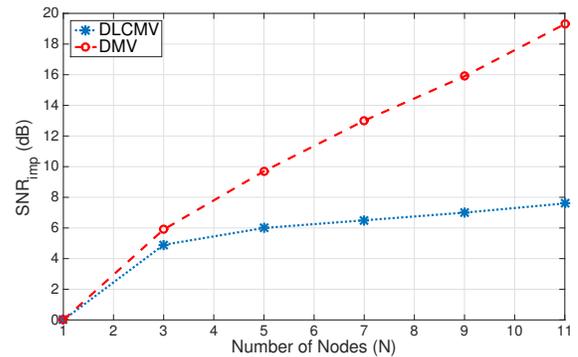

Fig. 2: A comparison of SNR$_{imp}$ of DLCMV and DMV for varying network size.

As expected, due to the iterative nature of DLCMV, the gain in SNR$_{imp}$ for an increasing number of nodes is reduced in contrast to that of DMV by as much as 12 dB for $N = 11$. For real WSNs in dynamic noise fields, this reduction highlights a main benefit of DMV. By exactly solving the robust LCMV problem at each step, whilst remaining competitive in the total transmission power required, DMV can achieve better rejection of correlated noise with the same rate of adaptation as the centralised solution.

4. CONCLUSIONS

In this paper we have presented a new method for computing a statistically optimal BF in distributed WSNs. The proposed DMV algorithm can be utilised in both cyclic and acyclic networks, with finite convergence and bounded performance guarantees in the latter. In quickly varying noise fields, where per-block calculation of the BF output is required, the acyclic DMV algorithm requires similar transmission costs to DLCMV while being truly optimal in each block. Furthermore, the ability of DMV to be applied in cyclic networks facilitates the use of optimal acoustic beamforming in WSNs regardless of the topology, a point not previously possible. This currently makes DMV the only optimal distributed BF able to be used in both cyclic and acyclic networks.

5. REFERENCES

[1] 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.

[2] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[3] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[4] L. Ehrenberg, S. Gannot, A. Leshem, and E. Zehavi, "Sensitivity analysis of MVDR and MPDR beamformers," in *Electrical and Electronics Engineers in Israel (IEEEI), 2010 IEEE 26th Convention of*. IEEE, 2010, pp. 000 416–000 420.

[5] D. Estrin, L. Girod, G. Pottie, and M. Srivastava, "Instrumenting the world with wireless sensor networks," in *Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP'01). 2001 IEEE International Conference on*, vol. 4. IEEE, 2001, pp. 2033–2036.

[6] 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*. VDE, 2012, pp. 1–4.

[7] D. Johnson, "Constrained optimization," *The Connexions Project (cnx. org), July 2003*, 2006.

[8] J. Li, P. Stoica, and Z. Wang, "On robust capon beamforming and diagonal loading," *Signal Processing, IEEE Transactions on*, vol. 51, no. 7, pp. 1702–1715, 2003.

[9] R. G. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *Signal Processing, IEEE Transactions on*, vol. 53, no. 5, pp. 1684–1696, 2005.

[10] S. Markovich-Golan, A. Bertrand, M. Moonen, and S. Gannot, "Optimal distributed minimum-variance beamforming approaches for speech enhancement in wireless acoustic sensor networks," *Signal Processing*, vol. 107, pp. 4–20, 2015.

[11] S. Markovich-Golan, S. Gannot, and I. Cohen, "Distributed multiple constraints generalized sidelobe canceler for fully connected wireless acoustic sensor networks," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 21, no. 2, pp. 343–356, 2013.

[12] D. Messerschmitt, "Stationary points of a real-valued function of a complex variable," *EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2006-93*, 2006.

[13] K. P. Murphy, Y. Weiss, and M. I. Jordan, "Loopy belief propagation for approximate inference: An empirical study," in *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 1999, pp. 467–475.

[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*. IEEE, 2014, pp. 810–814.

[15] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, 2014.

[16] S. S. Ram, A. Nedić, and V. V. Veeravalli, "Distributed stochastic subgradient projection algorithms for convex optimization," *Journal of optimization theory and applications*, vol. 147, no. 3, pp. 516–545, 2010.

[17] P. Richtárik and M. Takáč, "Distributed coordinate descent method for learning with big data," *arXiv preprint arXiv:1310.2059*, 2013.

[18] D. Sarason, *Complex function theory*. American Mathematical Soc., 2007.

[19] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE assp magazine*, vol. 5, no. 2, pp. 4–24, 1988.

[20] M. J. Wainwright, T. S. Jaakkola, and A. S. Willsky, "MAP estimation via agreement on trees: message-passing and linear programming," *Information Theory, IEEE Transactions on*, vol. 51, no. 11, pp. 3697–3717, 2005.

[21] E. Wei and A. Ozdaglar, "Distributed alternating direction method of multipliers," in *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*. IEEE, 2012, pp. 5445–5450.

[22] Y. Zeng and R. C. Hendriks, "Distributed delay and sum beamformer for speech enhancement via randomized gossip," *Audio, Speech, and Language Processing, IEEE/ACM Transactions on*, vol. 22, no. 1, pp. 260–273, 2014.

[23] G. Zhang and R. Heusdens, "Bi-alternating direction method of multipliers over graphs," in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 3571–3575.

[24] M. Zhu and S. Martínez, "On distributed convex optimization under inequality and equality constraints," *Automatic Control, IEEE Transactions on*, vol. 57, no. 1, pp. 151–164, 2012.