

DISTRIBUTED CROSS-RELATION-BASED FREQUENCY-DOMAIN BLIND SYSTEM IDENTIFICATION USING ONLINE-ADMM

M. Blochberger, F. Elvander, R. Ali, M. Moonen, J. Østergaard, J. Jensen, T. van Waterschoot

KU Leuven

Department of Electrical Engineering (ESAT)

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics
3001 Leuven, Belgium

ABSTRACT

In this paper, we propose a distributed cross-relation-based adaptive algorithm for blind identification of single-input multiple-output (SIMO) systems in the frequency domain using the alternating direction method of multipliers (ADMM) in a wireless sensor network (WSN). The network consists of a fixed number of nodes each equipped with a processing unit as well as a sensor that represents an output channel of the SIMO system. The proposed algorithm exploits the separability of the cross-channel relations by splitting the multichannel identification problem into sub-problems involving a subset of channels in a way that is determined by the network topology. Each node delivers estimates for the subset of channel frequency responses, which then are combined into a consensus estimate per channel using general-form consensus ADMM in an adaptive updating scheme. With numerical simulations, we show that it is possible to achieve convergence speeds and steady-state misalignment values comparable to fully centralized low-cost frequency-domain algorithms.

Index Terms— blind system identification, multichannel signal processing, distributed signal processing, ADMM, Online-ADMM

1. INTRODUCTION

The problem of blind system identification (BSI), which aims to estimate channel responses of an unknown system without knowing the input signal, has been the subject of extensive research over recent decades. It was introduced in [1] and various algorithms have been proposed since. Early algorithms used higher-order statistics [2, 3, 4] for channel estimation, however a high computational complexity has motivated research into algorithms using only second-order statistics. Such algorithms include the cross-relation (CR) algorithm [5, 6], subspace algorithms [7, 8, 9, 10] and maximum-likelihood algorithms [11]. Out of various multichannel BSI algorithms which have been proposed, adaptive cross-relation-based

least-mean-squares (LMS) algorithms in the time and frequency domain are the ones most widely used. The normalized multi-channel frequency-domain LMS (NMCFLMS) [12, 13] algorithm for instance is an efficient algorithm utilizing the fast Fourier transform (FFT) which has been extended to include constraints to improve robustness to noise and performance on acoustic impulse responses (RNMCFLMS [14], l_p -RNMCFLMS [15], phase-constrained- l_p -RNMCFLMS [16]). The quasi-Newton algorithm [17] on the other hand is a time-domain algorithm that utilizes the BFGS method to estimate the Hessian of the problem.

The expanding field of distributed signal processing in wireless sensor networks (WSN) brought forward algorithms for distributed noise control, echo cancellation, or beamforming, however, research into the task of BSI is limited. [18]

TODO: References

While there exist time-domain algorithms as introduced in [19, 20, 21], they do not pose apt comparisons to the frequency-domain algorithm being proposed. In this paper, we use the general-form consensus alternating direction method of multipliers (ADMM) [22] to distribute and solve the optimization problem posed by the task. We separate the inter-channel cross-relations of the BSI problem according to the network's topology, i.e. each node solves a sub-problem using data from its network neighbors, and the entirety of connected nodes subsequently reaches a consensus for the channel frequency responses. The ADMM update steps are applied in a block processing scheme forming an adaptive algorithm also referred to as Online-ADMM [23, 24].

As shown in numerical simulations, the resulting algorithm provides good estimation results as compared to state-of-the-art centralized frequency-domain algorithms.

2. PROBLEM STATEMENT

2.1. Signal Model

We consider an acoustic SIMO system with the input signal $s(n) = [s(n) \ s(n-1) \ \dots \ s(n-2L+2)]^T$ and the M output signals $\mathbf{x}_i(n) = [x_i(n) \ x_i(n-1) \ \dots \ x_i(n-L+1)]^T$ where $i \in \mathcal{M} \triangleq \{1, \dots, M\}$. Each output $\mathbf{x}_i(n)$ is the convolution of $s(n)$ with the respective channel impulse response \mathbf{h}_i and an additive noise term $\mathbf{v}_i(n)$, assumed to be zero-mean and uncorrelated to $s(n)$. The signal model is described by

$$\mathbf{x}_i(n) = \mathbf{H}_i \mathbf{s}(n) + \mathbf{v}_i(n) \quad (1)$$

where \mathbf{H}_i is the $L \times (2L-1)$ linear convolution matrix of the i th channel using the elements of \mathbf{h}_i of length L which represents the

This research work was carried out at the ESAT Laboratory of KU Leuven, in the frame of the SOUNDS European Training Network. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 956369. This research received funding in part by the Research Foundation - Flanders (FWO) grant 12ZD622N as well as from the European Union's Horizon 2020 research and innovation program / ERC Consolidator Grant: SONORA (No. 773268). This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information. Source code available at <https://github.com/SOUNDS-RESEARCH/iwaenc2022-admm-bsi>.

system to be identified. For this paper, the length of the impulse responses L is assumed to be known.

2.2. Cross-relation approach

The cross-relation approach for BSI aims to use only the output signals of the system to identify it. This can be achieved by exploiting the relative channel information when more than one channel is available, and the identifiability conditions [6] are satisfied. These conditions are: (i) the channel transfer functions have no common zeros (i.e. the polynomials are not co-prime), and (ii) the covariance matrix of the input signal $\mathbf{s}(n)$ is of full rank (i.e. the signal has a number of modes $\geq 2L + 1$).

The fundamental equality of this approach in the noiseless case $\mathbf{v}_i(n) = 0$, is

$$\mathbf{x}_i^T(n)\mathbf{h}_j = \mathbf{x}_j^T(n)\mathbf{h}_i, \quad i, j \in \mathcal{M}, i \neq j \quad (2)$$

which states that the channel output signal convolved with the impulse response of another is equal to the vice-versa as follows from the commutativity property of the convolution. Left-multiplication and applying the expectation operator to form the covariance matrix $\mathbf{R}_{ij}(n) = \mathbb{E}\{\mathbf{x}_i(n)\mathbf{x}_j^T(n)\}$ and combining all cross-relations (2) (see e.g. [12] for a more thorough derivation) yields the system of equations

$$\mathbf{R}\mathbf{h} = \mathbf{0} \quad (3)$$

where \mathbf{R} is an $ML \times ML$ matrix given by:

$$\mathbf{R} = \begin{bmatrix} \sum_{i \neq 1} \mathbf{R}_{ii} & -\mathbf{R}_{21} & \cdots & -\mathbf{R}_{M1} \\ -\mathbf{R}_{12} & \sum_{i \neq 2} \mathbf{R}_{ii} & \cdots & -\mathbf{R}_{M2} \\ \vdots & \vdots & \ddots & \vdots \\ -\mathbf{R}_{1M} & -\mathbf{R}_{2M} & \cdots & \sum_{i \neq M} \mathbf{R}_{ii} \end{bmatrix} \quad (4)$$

and $\mathbf{h} = [\mathbf{h}_1^T \dots \mathbf{h}_M^T]^T$. Formulating the problem in the frequency domain, the derivation is analogous. We denote all frequency-domain variables in bold italic (e.g. \mathbf{R}) compared to the time-domain bold upright (e.g. \mathbf{R}). The derivation involves the frame-based overlap-save technique, working with signal frames $\mathbf{x}_{i,2L}(m)$ of length $2L$, frame index m , leading to the system of equations

$$\mathbf{R}\mathbf{h} = \mathbf{0} \quad (5)$$

where $\mathbf{h} = [\mathbf{h}_1^T \dots \mathbf{h}_M^T]^T$ is a stacked vector of channel frequency responses and the $ML \times ML$ matrix \mathbf{R} is recursively estimated by

$$\hat{\mathbf{R}}(m) = \eta \hat{\mathbf{R}}(m-1) + (1-\eta) \tilde{\mathbf{R}}(m). \quad (6)$$

where $\eta \in [0, 1]$ is an exponential smoothing factor and $\tilde{\mathbf{R}}(m)$ is constructed analogously to (4), however with covariance matrices replaced by instantaneous cross-spectrum matrices

$$\tilde{\mathbf{R}}_{ij}(m) = \mathbf{X}_i(m)\mathbf{X}_j^H(m) \quad (7)$$

with

$$\mathbf{X}_i(m) = \mathbf{W}_{L \times 2L}^{01} \mathbf{D}_i(m) \mathbf{W}_{2L \times L}^{10}. \quad (8)$$

The matrix $\mathbf{D}_i(m) = \text{diag}\{\text{FFT}_{2L}\{\mathbf{x}_{i,2L}(m)\}\}$ contains the signal spectrum on its diagonal and

$$\mathbf{W}_{L \times 2L}^{01} = \mathbf{F}_{L \times L} \mathbf{W}_{L \times 2L}^{01} \mathbf{F}_{2L \times 2L}^{-1} \quad (9)$$

$$\mathbf{W}_{2L \times L}^{10} = \mathbf{F}_{2L \times 2L} \mathbf{W}_{2L \times L}^{10} \mathbf{F}_{L \times L}^{-1} \quad (10)$$

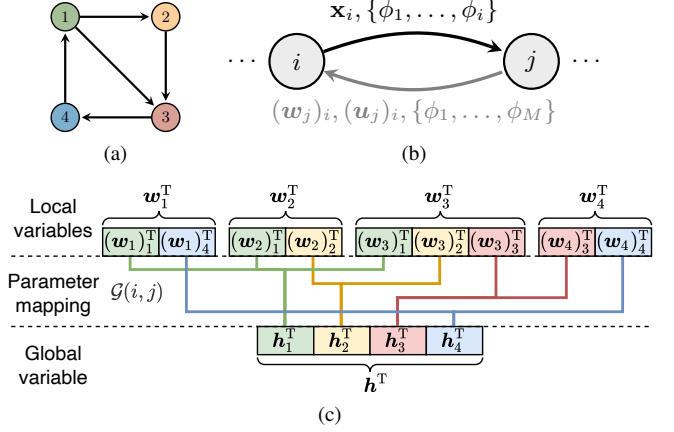


Fig. 1: (a) shows an example of a network topology with 4 nodes (\circ). Arrows indicate exchange of signal information. (b) visualizes the data transmitted between nodes i and j for each frame m . (c) shows the equivalent parameter mapping of the local variable components $(\mathbf{w}_i)_j$ to the global variable components $(\mathbf{h}_j)_i$ via $\mathcal{G}(i, j)$ for the network in (a).

are the frequency-domain overlap-save matrices where $\mathbf{F}_{L \times L}$ and $\mathbf{F}_{2L \times 2L}$ are the discrete Fourier transform (DFT) matrices for sizes L and $2L$ respectively and $\mathbf{W}_{L \times 2L}^{01} = [\mathbf{0}_{L \times L} \quad \mathbf{I}_{L \times L}]$ and $\mathbf{W}_{2L \times L}^{10} = [\mathbf{I}_{L \times L} \quad \mathbf{0}_{L \times L}]^T$ denote the time-domain overlap-save matrices.

Analogously to the time-domain formulation, the $ML \times 1$ vector \mathbf{h} is a stacked vector of complex-valued frequency responses. The null-space problem in (5) can not be solved by computing the eigenvector corresponding to the zero-valued eigenvalue because, in the presence of noise, the system matrix \mathbf{R} is of full rank and may not have any. Therefore it is best solved by posing it as a quadratic minimization problem [6, 12] with a non-triviality constraint to avoid the zero solution:

$$\hat{\mathbf{h}} = \arg \min_{\mathbf{h}} \quad \mathbf{h}^H \hat{\mathbf{R}} \mathbf{h} \quad (11)$$

$$\text{s.t.} \quad \mathbf{h}^H \mathbf{h} = 1. \quad (12)$$

2.3. Wireless Sensor Network

We consider a WSN consisting of M sensors each acquiring an output signal $\mathbf{x}_i(n)$. We define a set $\mathcal{T}_i \subseteq \mathcal{M}$ for each node i with $i \in \mathcal{M}$, which contains the indices of all nodes which have access to the signals $\mathbf{x}_i(n)$, i.e. the signal information is transmitted to. Further, we define a set $\mathcal{R}_i \subseteq \mathcal{M}$ for each node with $i \in \mathcal{M}$, which contains all indices of nodes from which the node receives signal information. Note that $i \in \mathcal{T}_i$ and $i \in \mathcal{R}_i$ for simplifying notation in later sections, but evidently, no actual transmission is necessary in that case, as the signal $\mathbf{x}_i(n)$ is available locally at node i . Fig. 1a shows a network where each node shares its signal information with one or more neighbors. For this work, we assume synchronicity of all nodes (i.e. no sampling rate offset) and perfect transmission of data with no time delay.

In the following section, we introduce a distributed adaptive algorithm to find a solution to (11) utilizing the sensor nodes.

3. PROPOSED ALGORITHM

3.1. General-Form Consensus ADMM

Defined by \mathcal{R}_i , each node $i \in \mathcal{M}$ has its own signal information and additionally signal information of its neighbors available. This is used to solve a BSI problem of the same form as in (11), which however only involves the subset of channels. We define the local variables $\mathbf{w}_i = \left[\{(\mathbf{w}_i)_j^T\}_{j \in \mathcal{R}_i} \right]^T$, which is a stacked vector of locally estimated frequency responses (cf. Fig. 1c). Further, the global variable $\mathbf{h} = [\mathbf{h}_1^T \dots \mathbf{h}_M^T]^T$ is the stacked vector of all M channel response consensus variables. Lastly, we define the matrix $\hat{\mathbf{P}}_i$, which is the frequency-domain cross-relation matrix analogous to $\hat{\mathbf{R}}$ as introduced in Sec. 2.2, however only using the signals \mathbf{x}_j with $j \in \mathcal{R}_i$.

We replace the cost function minimized in (11) with the separable cost function

$$\tilde{J}(\mathbf{w}) = \sum_{i \in \mathcal{M}} \tilde{J}_i(\mathbf{w}_i) = \sum_{i \in \mathcal{M}} \mathbf{w}_i^H \hat{\mathbf{P}}_i \mathbf{w}_i. \quad (13)$$

with $\mathbf{w} \triangleq (\mathbf{w}_1, \dots, \mathbf{w}_M)$ as the local stacked estimates. Minimization problems with separable cost functions, as introduced here, can be solved by the well-established method of consensus ADMM [22]

$$\underset{\mathbf{w}, \mathbf{h}}{\text{minimize}} \quad \sum_{i \in \mathcal{M}} \tilde{J}_i(\mathbf{w}_i) \quad (14)$$

$$\text{subject to} \quad (\mathbf{w}_i)_j = \mathbf{h}_{\mathcal{G}(i,j)}, \quad i \in \mathcal{M}, j \in \mathcal{R}_i \quad (15)$$

$$\mathbf{h}^H \mathbf{h} = 1 \quad (16)$$

where $\mathcal{G}(i, j)$ denotes the mapping of L local variable components $(\mathbf{w}_i)_j$, i.e. one of the frequency responses in the stacked vector, to the corresponding global variable components \mathbf{h}_j (cf. Fig. 1c). For brevity, the mapped global variables $(\tilde{\mathbf{h}}_i)_j = \mathbf{h}_{\mathcal{G}(i,j)}$ are introduced. This defines $\tilde{\mathbf{h}}_i = \left[\{(\mathbf{h}_j^T)_{j \in \mathcal{R}_i}\}^T \right]^T$. The equality constraint between local variables \mathbf{w}_i and global variable \mathbf{h} enforces consensus, i.e. a common solution taking into account estimates of all nodes that share data of the same channel.

The augmented Lagrangian for this particular general-form consensus problem is

$$\mathcal{L}_\rho(\mathbf{w}, \mathbf{h}, \mathbf{u}) = \sum_{i \in \mathcal{M}} \left(\mathbf{w}_i^H \hat{\mathbf{P}}_i \mathbf{w}_i + 2\Re \left(\mathbf{u}_i^H (\mathbf{w}_i - \tilde{\mathbf{h}}_i) \right) + \rho \left\| \mathbf{w}_i - \tilde{\mathbf{h}}_i \right\|^2 \right) \quad (17)$$

where $\mathbf{u} \triangleq (\mathbf{u}_1, \dots, \mathbf{u}_M)$ are the stacked local dual variables (Lagrange multipliers) $\mathbf{u}_i = \left[\{(\mathbf{u}_i)_j^T\}_{j \in \mathcal{R}_i} \right]^T$ with $(\mathbf{u}_i)_j$ denoting the dual variable for channel j at node i . These variables follow from the equality constraints in (15). The ADMM then consists of the steps [22]:

$$\mathbf{w}_i^{k+1} = \underset{\mathbf{w}_i}{\text{argmin}} \mathcal{L}_\rho(\mathbf{w}, \mathbf{h}^k, \mathbf{u}^k), \quad (18)$$

$$\mathbf{h}^{k+1} = \underset{\mathbf{h}, \|\mathbf{h}\|=1}{\text{argmin}} \mathcal{L}_\rho(\mathbf{w}^{k+1}, \mathbf{h}, \mathbf{u}^k) \quad (19)$$

$$\mathbf{u}_i^{k+1} = \mathbf{u}_i^k + \rho (\mathbf{w}_i^{k+1} - \tilde{\mathbf{h}}_i^{k+1}). \quad (20)$$

where k is the iteration index. Evidently, this is an iterative algorithm, therefore we will introduce the online/adaptive aspect of the proposed algorithm in the following.

3.2. Online ADMM-BSI

We introduced the original problem in (5) with a time-dependent estimate $\mathbf{R}(m)$ of the matrix $\mathbf{R}(m)$. It therefore follows that the data term in (17) can also be considered time-dependent, which from here on will be denoted with the additional superscript time index m as $\mathbf{w}_i^H \hat{\mathbf{P}}_i^m \mathbf{w}_i$. A thorough review of ADMM with time-varying data terms can be found in [23, 24] where it is referred to as ‘‘Online-ADMM’’. We transform the iterative batch processing method (18)-(20) into an adaptive one by computing only a (small) finite number of iterations with each time-frame- m specific data term. Here specifically, we apply one iteration per time frame, which manifests itself as simply replacing the iteration index k with the time frame index m .

The minimization problem for the local variable \mathbf{w}_i (18) can be solved by various algorithms, in this case however we perform the update step as

$$\mathbf{w}_i^{m+1} = \mathbf{w}_i^m - \mu \mathbf{V}_i^m \left(\hat{\mathbf{P}}_i^m \mathbf{w}_i^m + \mathbf{u}_i^m + \rho (\mathbf{w}_i^m - \tilde{\mathbf{h}}_i^m) \right) \quad (21)$$

where μ , ($0 < \mu \leq 1$), is a step size and $\mathbf{V}_i^m = \left(\hat{\mathbf{P}}_i^m + \rho \mathbf{I} \right)^{-1}$ is the regularized inverse Hessian of the problem. The implicit overlap-save matrices in \mathbf{V}_i^m make this a constrained update step which is costly to compute, so in order to reduce computational complexity, we introduce the approximation

$$\tilde{\mathbf{V}}_i^m = \text{diag} \left\{ \text{diag} \left\{ \hat{\mathbf{P}}_i^m \right\} + \rho \mathbf{1} \right\}^{-1} \quad (22)$$

which is straightforward to compute.

It may be readily verified (cf. [22]) that the solution to (19) is given by an update step which computes

$$\mathbf{h}_i^{m+1} = \frac{\tilde{\mathbf{h}}_i^{m+1}}{\sqrt{\sum_{j \in \mathcal{T}_i} \phi_i}} \quad (23)$$

for each node $i \in \mathcal{M}$ with the local unnormalized consensus

$$\tilde{\mathbf{h}}_i^{m+1} = \bar{\mathbf{w}}_i^{m+1} + \frac{1}{\rho} \bar{\mathbf{u}}_i^m. \quad (24)$$

and the squared norm $\phi_i = \|\bar{\mathbf{h}}_i^{m+1}\|^2$. Here, $\bar{\mathbf{w}}_i = \frac{1}{N_i} \sum_{j \in \mathcal{T}_i} (\mathbf{w}_j)_i$ and $\bar{\mathbf{u}}_i = \frac{1}{N_i} \sum_{j \in \mathcal{T}_i} (\mathbf{u}_j)_i$, with $N_i = |\mathcal{T}_i|$, are neighborhood averages of the channel response estimate and the dual variable of channel i computed at node i . The values of all ϕ_i , $i \in \mathcal{M}$ are transmitted and relayed through the network until can compute the denominator in (23).

The dual variables are straightforward to compute as

$$\mathbf{u}_i^{m+1} = \mathbf{u}_i^m + \rho (\mathbf{w}_i^{m+1} - \tilde{\mathbf{h}}_i^{m+1}). \quad (25)$$

The computations of local, consensus and dual variables, in (21), (23) and (25) respectively, require data transmitted by other nodes (cf. Fig. 1b). For this paper, we assume that all transmissions are done instantaneously whenever variables are required by neighboring nodes.

4. NUMERICAL EVALUATION

The performance of the proposed algorithm is assessed via numerical simulations. As an error measure, we use the normalized projection misalignment (NPM) [13]

$$\text{NPM}(m) = 20 \log_{10} \left(\left\| \mathbf{h}(m) - \frac{\mathbf{h}^T(m) \mathbf{h}_t}{\mathbf{h}_t^T \mathbf{h}_t} \mathbf{h}(m) \right\| / \|\mathbf{h}(m)\| \right) \quad (26)$$

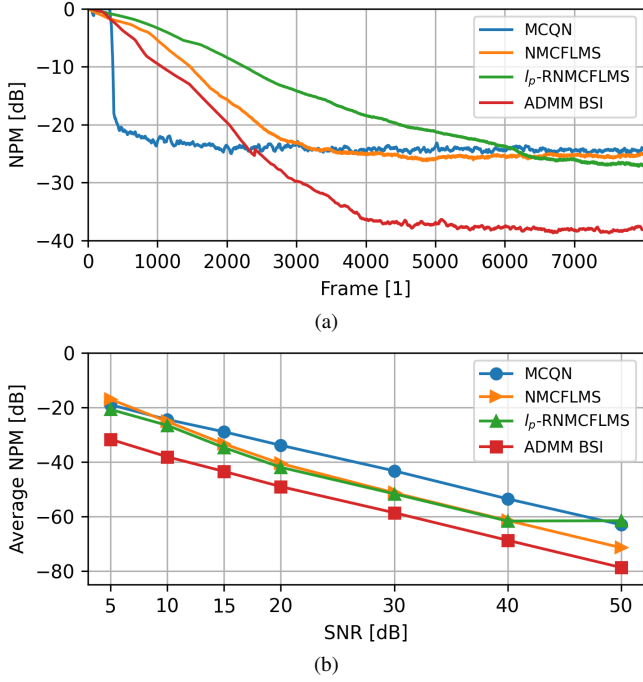


Fig. 2: Comparison of ADMM-BSI convergence behaviour with $L = 64$ at different SNR values. Shown are (a) NPM at SNR = 10 dB over frame index m for convergence speed comparison and (b) the steady-state NPMs over different SNRs.

where \mathbf{h} is the stacked vector of impulse response estimates (cf. Sec. 2.2), which are the inversely Fourier-transformed estimated frequency responses \mathbf{h}_i stacked in \mathbf{h} and \mathbf{h}_t is the ground truth. The signal-to-noise ratio (SNR) for the simulations is defined as

$$\text{SNR} = 10 \log_{10} \left(\frac{\sigma_s^2 \|\mathbf{h}_t\|^2}{M \sigma_v^2} \right) \quad (27)$$

where σ_s^2 and σ_v^2 are the variance of signal and noise respectively which both are modelled by (channel-independent) white Gaussian noise (WGN).

The first simulation evaluates the performance using randomly generated impulse responses of length $L = 64$ under different signal-to-noise ratios (SNR) on a 5-node network (5-channel system, $M = 5$). The impulse responses are drawn from a zero-mean Gaussian distribution with unit variance, and the signal is 8×10^5 samples of WGN to ensure convergence of all algorithms. The step sizes are hand tuned for similar convergence speed: $\mu_{\text{MCQN}} = 0.5$, $\mu_{\text{NMCFLMS}} = 0.4$, $\mu_{l_p\text{-RNMCFMLS}} = 0.3$, $\mu_{\text{ADMM}} = 0.6$, $\rho = 1$, $\eta = 0.98$. Fig. 2a shows the median NPM of 30 Monte-Carlo runs over frame index m . Fig. 2b shows the median of 30 Monte-Carlo runs where the averaged NPM of 100 frames after convergence is plotted. It is observable that the proposed algorithm yields a lower steady-state NPM than the compared NMCFLMS, RNMCFMLS, and l_p -RNMCFMLS algorithms. The choice to present the median is based on the fact that randomly generated impulse responses may have close common zeros diminishing identifiability (for all algorithms) [25], here then considered outliers.

The second simulation is a small-scale assessment of the influence of the overlap of sub-problems. Two base scenarios, $M =$

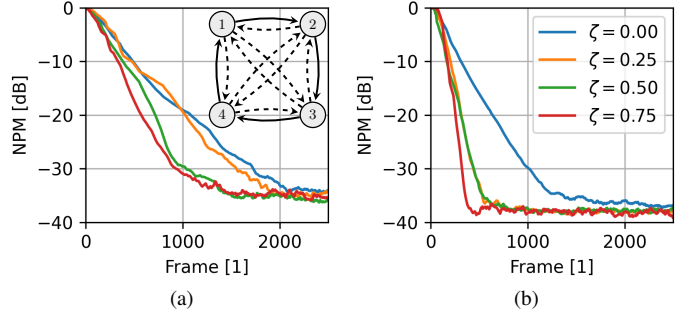


Fig. 3: Comparison of ADMM-BSI convergence behaviour with $L = 16$ for different values of ζ for SNR = 20 dB. NPM is plotted for systems of (a) 4 nodes/channels and (b) 8 nodes/channels. The inset of (a) shows the base ring topology (black arrows) and all other possible connections (dashed black arrows).

4 and $M = 8$ are evaluated using short ($L = 16$) random impulse responses and with semi-random topologies. Each topology is generated by forming a ring topology as marked with solid arrows on the inset of Fig. 3a and additionally randomly selecting a defined number of connections from all remaining possible connections, marked with dashed arrows. The parameter ζ is defined as the number of randomly selected connections to possible connections (excluding the ring topology). Topologies satisfying this definition for $\zeta \in \{0.0, 0.25, 0.5, 0.75\}$ were generated. Fig. 3 shows the median of 30 Monte-Carlo runs for the two setups, which suggests a proportional relation of convergence speed, channel number M and sub-problem overlap ζ while the steady-state error appears not dependent on ζ .

Attention: Some plots and values will change slightly.

5. CONCLUSIONS

In this paper, a distributed adaptive algorithm for blind system identification using ADMM in the context of sensor networks was developed. Each node in a WSN solves a BSI sub-problem estimating channel responses using signal data from neighboring nodes to form the cross-correlation problem with a subset of channels. Through

TODO: FINISH

6. REFERENCES

- [1] Y. Sato, "A Method of Self-Recovering Equalization for Multi-level Amplitude-Modulation Systems," *IEEE Trans. Commun.*, vol. 23, no. 6, pp. 679–682, June 1975.
- [2] D. Godard, "Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems," *IEEE Trans. Commun.*, vol. 28, no. 11, pp. 1867–1875, Nov. 1980.
- [3] L. Tong, G. Xu, and T. Kailath, "A new approach to blind identification and equalization of multipath channels," in *[1991] Conference Record of the Twenty-Fifth Asilomar Conference on Signals, Systems & Computers*, Pacific Grove, CA, USA, 1991, pp. 856–860, IEEE Comput. Soc. Press.
- [4] J.M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications," *Proc. IEEE*, vol. 79, no. 3, pp. 278–305, Mar. 1991.

- [5] Lang Tong, Guanghan Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inform. Theory*, vol. 40, no. 2, pp. 340–349, Mar. 1994.
- [6] Guanghan Xu, Hui Liu, Lang Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. Signal Process.*, vol. 43, no. 12, pp. 2982–2993, Dec./1995.
- [7] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Process.*, vol. 43, no. 2, pp. 516–525, Feb./1995.
- [8] Sharon Gannot and Marc Moonen, "Subspace Methods for Multimicrophone Speech Dereverberation," *EURASIP J. Adv. Signal Process.*, vol. 2003, no. 11, pp. 769285, Dec. 2003.
- [9] Konstantinos I. Diamantaras and Theophilos Papadimitriou, "An Efficient Subspace Method for the Blind Identification of Multichannel FIR Systems," *IEEE Transactions on Signal Processing*, vol. 56, no. 12, pp. 5833–5839, Dec. 2008.
- [10] Qadri Mayyala, Karim Abed-Meraim, and Azzedine Zerguine, "Structure-Based Subspace Method for Multichannel Blind System Identification," *IEEE Signal Processing Letters*, vol. 24, no. 8, pp. 1183–1187, Aug. 2017.
- [11] Yingbo Hua, "Fast maximum likelihood for blind identification of multiple FIR channels," *IEEE Transactions on Signal Processing*, vol. 44, no. 3, pp. 661–672, Mar. 1996.
- [12] Yiteng Arden Huang and Jacob Benesty, "Adaptive multi-channel least mean square and Newton algorithms for blind channel identification," *Signal Processing*, p. 12, 2002.
- [13] Yiteng Huang and J. Benesty, "A class of frequency-domain adaptive approaches to blind multichannel identification," *IEEE Transactions on Signal Processing*, vol. 51, no. 1, pp. 11–24, Jan. 2003.
- [14] M. Hu, S. Doclo, D. Sharma, M. Brookes, and P. A. Naylor, "Noise robust blind system identification algorithms based on a Rayleigh quotient cost function," in *2015 23rd European Signal Processing Conference (EUSIPCO)*, Aug. 2015, pp. 2476–2480.
- [15] Hongsen He, Jingdong Chen, Jacob Benesty, and Tao Yang, "Noise Robust Frequency-Domain Adaptive Blind Multichannel Identification With ℓ_p -Norm Constraint," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 9, pp. 1608–1619, Sept. 2018.
- [16] Byeongho Jo and Paul Calamia, "Robust Blind Multichannel Identification based on a Phase Constraint and Different ℓ_p -norm Constraints," in *2020 28th European Signal Processing Conference (EUSIPCO)*, Jan. 2021, pp. 1966–1970.
- [17] Emanuel A.P. Habets and Patrick A. Naylor, "An online quasi-Newton algorithm for blind SIMO identification," in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, Dallas, TX, Mar. 2010, pp. 2662–2665, IEEE.
- [18] Alexander Bertrand, "Applications and trends in wireless acoustic sensor networks: A signal processing perspective," in *2011 18th IEEE Symposium on Communications and Vehicular Technology in the Benelux (SCVT)*, Nov. 2011, pp. 1–6.
- [19] Chengpu Yu, Lihua Xie, and Yeng Chai Soh, "Distributed blind system identification in sensor networks," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 5065–5069.
- [20] Ying Liu, Hao Liu, and Chunguang Li, "Distributed blind identification of sparse channels in sensor networks," in *2016 35th Chinese Control Conference (CCC)*, July 2016, pp. 5122–5127.
- [21] Rui Liu and Han-Fu Chen, "Distributed and recursive blind channel identification to sensor networks," *Control Theory Technol.*, vol. 15, no. 4, pp. 274–287, Nov. 2017.
- [22] Stephen Boyd, Neal Parikh, Eric Chu, and Jonathan 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.
- [23] Huahua Wang and Arindam Banerjee, "Online alternating direction method," in *29th International Conference on Machine Learning, ICML 2012*, 2012, pp. 1119–1126.
- [24] S. Hosseini, A. Chapman, and M. Mesbahi, "Online distributed ADMM via dual averaging," in *53rd IEEE Conference on Decision and Control*, Dec. 2014, pp. 904–909.
- [25] P. A. Naylor, X. Lin, and A. W. H. Khong, "Near-Common Zeros in Blind Identification of Simo Acoustic Systems," in *2008 Hands-Free Speech Communication and Microphone Arrays*, May 2008, pp. 21–24.