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

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

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 sensor network with a fusion center. 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 which is determined by the network topology. Each network node delivers estimates for the subset of channel frequency responses, which then are communicated to a fusion center and 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 estimation errors comparable to fully centralized low-cost frequency-domain algorithms and a high-performing Quasi-Newton algorithm.

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 impulse 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-Netwon algorithm [17] on the other hand is a time-domain algorithm that utilizes the BFGS method to estimate the Hessian of the problem.

TODO: This is also a time-domain algorithm, should a comparison even be made? MCQN in plots

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.

TODO: cite

While there exist time-domain algorithms as introduced in [18, 19, 20], 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) [21] 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 subproblem 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 [22, 23].

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. Sensor Network

We consider a WSN consisting of M sensor nodes, each acquiring a signal and equipped with a processing unit, and define the index set $i \in \mathcal{M} \triangleq \{1, \ldots, M\}$ representing these. Each node shares information with a set of neighboring nodes $\mathcal{S}_i$ with $i \in \mathcal{M}$, and receives information by a set of nodes $\mathcal{R}_i$ with $i \in \mathcal{M}$. Fig. 1 shows a network where each node shares its signal information with one or more neighbors.

TODO: Which type of topology should be introduced. With fusion center? ring topology, "star" topology, depends on how consensus is introduced...
Make figure nicer.

*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. Source code available at `https://github.com/SOUNDS-RESEARCH/iwaenc2022-admm-bsi`.

†This work was supported in part by the Research Foundation - Flanders (FWO) grant 12ZD622N.

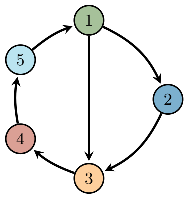


Fig. 1: WSN

2.2. Signal Model

We consider an acoustic SIMO system with the input signal $\mathbf{s}(n) = \begin{bmatrix} s(n) & s(n-1) & \dots & s(n-2L+2) \end{bmatrix}^{\mathrm{T}}$ and the M output signals $\mathbf{x}_i(n) = \begin{bmatrix} x_i(n) & x_i(n-1) & \dots & x_i(n-L+1) \end{bmatrix}^{\mathrm{T}}$ aquired by the sensor nodes of the WSN. Each output $\mathbf{x}_i(n)$ is the convolution of $\mathbf{s}(n)$ 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 $\mathbf{s}(n)$. The signal model is described by

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

where $\mathbf{H}_i$ is the $L \times (2L-1)$ linear convolution matrix of the ith channel using the elements of $\mathbf{h}_i$ of length L which represents the system to be identified. For this paper, the length of the impulse responses L is assumed to be known.

2.3. 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. 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

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

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 $\mathbf{R}_{ij}(n) = \mathrm{E}\left\{\mathbf{x}_i(n)\mathbf{x}_j^{\mathrm{T}}(n)\right\}$ and combining all cross-relations (2) (see e.g. [12] for a more thourough derivation) yields the system of equations

$$\mathbf{R}\mathbf{h} = \mathbf{0} \tag{3}$$

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

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

and $\mathbf{h} = \begin{bmatrix} \mathbf{h}_1^{\mathrm{T}} & \cdots & \mathbf{h}_M^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$. Formulating the problem in the frequency domain, the derivation is analogous. We denote all frequency-domain variables in bold cursive (e.g. $\boldsymbol{R}$) 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

$$\boldsymbol{R}\boldsymbol{h} = \mathbf{0} \tag{5}$$

where the $ML \times ML$ matrix is recursivly estimated by

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

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

$$\tilde{\boldsymbol{R}}_{ij}(m) = \boldsymbol{X}_i(m)\boldsymbol{X}_j^{\mathrm{H}}(m) \tag{7}$$

with

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

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

$$\boldsymbol{W}^{01}_{L\times 2L} = \mathbf{F}_{L\times L}\mathbf{W}^{01}_{L\times 2L}\mathbf{F}^{-1}_{2L\times 2L} \tag{9}$$

$$\boldsymbol{W}^{10}_{2L\times L} = \mathbf{F}_{2L\times 2L}\mathbf{W}^{10}_{2L\times L}\mathbf{F}^{-1}_{L\times L} \tag{10}$$

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 $\mathbf{W}^{01}_{L\times 2L} = \begin{bmatrix}\mathbf{0}_{L\times L} & \mathbf{I}_{L\times L}\end{bmatrix}$ and $\mathbf{W}^{10}_{2L\times L} = \begin{bmatrix}\mathbf{I}_{L\times L} & \mathbf{0}_{L\times L}\end{bmatrix}^{\mathrm{T}}$ denote the time-domain overlap-save matrices.

Analogous to the time-domain formulation, the $ML \times 1$ vector $\boldsymbol{h}$ 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 $\boldsymbol{R}$ is of full rank and may not have any. Therefore it is best solved by posing it as a least-mean-squares minimization problem [6, 12] with a non-triviality constraint to avoid the zero solution:

$$\hat{\boldsymbol{h}} = \arg\min_{\boldsymbol{h}} \quad \boldsymbol{h}^{\mathrm{H}}\hat{\boldsymbol{R}}\boldsymbol{h} \tag{11}$$

$$\text{s.t.} \quad \boldsymbol{h}^{\mathrm{H}}\boldsymbol{h} = 1. \tag{12}$$

In the following section, we introduce a distributed adaptive algorithm to find a solution to this problem.

3. PROPOSED ALGORITHM

3.1. General-Form Consensus ADMM

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

$$\tilde{J}(\boldsymbol{h}) = \sum_{i\in\mathcal{M}}\tilde{J}_i(\boldsymbol{w}_i) = \sum_{i\in\mathcal{M}}\boldsymbol{w}_i^{\mathrm{H}}\hat{\boldsymbol{P}}_i\boldsymbol{w}_i \tag{13}$$

where $\boldsymbol{w}_i$ is defined as the stacked vector of estimated frequency responses analogous to $\boldsymbol{h}$ and $\hat{\boldsymbol{P}}$ is the frequency-domain cross-relation matrix analogous to $\boldsymbol{R}$ as introduced in Sec. 2.3, however, both using the respective set of channels $\mathcal{R}_i$ available at node i. Minimization problems with separable cost functions as introduced here can be solved by the well-established method of consensus ADMM [21]

$$\text{minimize} \quad \sum_{i\in\mathcal{R}}\tilde{J}_i(\boldsymbol{w}_i) \tag{14}$$

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

$$\boldsymbol{h}^{\mathrm{H}}\boldsymbol{h} = 1 \tag{16}$$

where $\boldsymbol{w}_i$ will be referred to as the local variable, $\boldsymbol{h}$ as the global variable or consensus. $\mathcal{G}(i,j)$ denotes the mapping of L local variable components $(\boldsymbol{w}_i)_j$, i.e. one of the frequency responses in the stacked vector, to the corresponding global variable components $(\boldsymbol{h})_g$. For brevity, the mapped global variables $(\tilde{\boldsymbol{h}}_i)_j = \boldsymbol{h}_{\mathcal{G}(i,j)}$ are introduced. The equality constraint between local variables $\boldsymbol{w}_i$ and global variable $\boldsymbol{h}$ enforces the 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

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

The ADMM then consists of the steps:

$$\boldsymbol{w}_i^{k+1} = \underset{\boldsymbol{w}_i}{\operatorname{argmin}}\, \mathcal{L}_\rho(\boldsymbol{w},\boldsymbol{h}^k,\boldsymbol{u}^k), \tag{18}$$

$$\boldsymbol{h}^{k+1} = \underset{\boldsymbol{h},\|\boldsymbol{h}\|=1}{\operatorname{argmin}}\, \mathcal{L}_\rho(\boldsymbol{w}^{k+1},\boldsymbol{h},\boldsymbol{u}^k) \tag{19}$$

$$\boldsymbol{u}_i^{k+1} = \boldsymbol{u}_i^k + \rho\left(\boldsymbol{w}_i^{k+1} - \tilde{\boldsymbol{h}}_i^{k+1}\right). \tag{20}$$

As this is still denoted as an iterative algorithm, 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 time-dependent data matrix. It therefore follows that the data term in (17) is also time-dependent, which from here on will be denoted with the additional superscript time index m as $\boldsymbol{w}_i^{\mathrm{H}} \hat{\boldsymbol{P}}_i^m \boldsymbol{w}_i$. A thorough overview of time-varying data terms in ADMM can be found in [22, 23] 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 index m

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

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

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

$$\tilde{\boldsymbol{V}}_i^m = \operatorname{diag}\left\{ \left(\operatorname{diag}\left\{ \hat{\boldsymbol{P}}_i^m \right\} + \rho\mathbf{1} \right)^{-1} \right\}, \tag{22}$$

similar to NMCFLMS [13], which is straightforward to compute.

For the update step of the consensus variable $\boldsymbol{h}$, it may be readily verified (cf. [21]) that the solution to (19) is given by

$$\boldsymbol{h}^{m+1} = \frac{\bar{\boldsymbol{w}}^{m+1} + \frac{1}{\rho}\bar{\boldsymbol{u}}^m}{\left\|\bar{\boldsymbol{w}}^{m+1} + \frac{1}{\rho}\bar{\boldsymbol{u}}^m\right\|}. \tag{23}$$

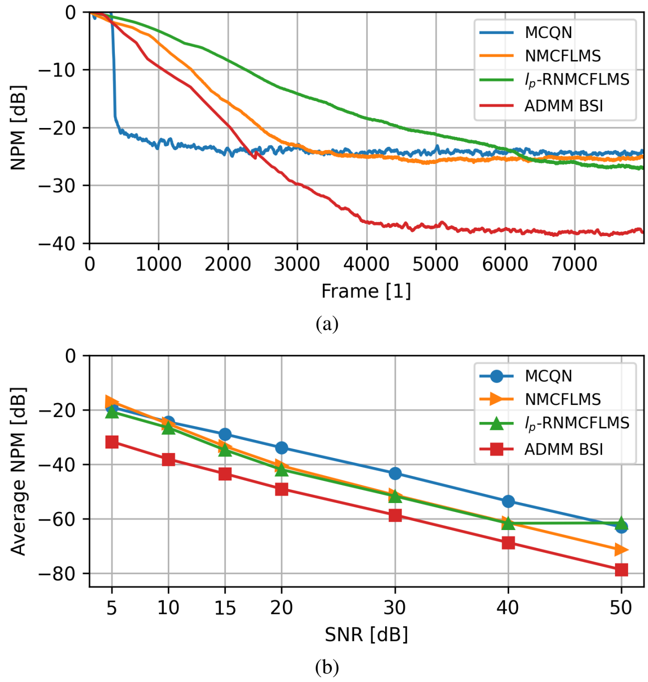


Fig. 2: Comparison of ADMM-BSI convergence behaviour with $L = 64$ at different SNR values. Shown are (a) NPM at SNR $=$ 10 dB over frame index m for convergence speed comparison and (b) the steady-state NPMs over different SNRs.

where the $ML \times 1$ vectors $\bar{\boldsymbol{w}}^{m+1}, \bar{\boldsymbol{u}}^{m+1}$ are computed as the mapped averages

$$(\bar{\boldsymbol{w}}^{m+1})_g = \frac{1}{N_g} \sum_{\mathcal{G}(i,j)=g} (\boldsymbol{w}_i^{m+1})_j, \quad g,i,j \in \mathcal{M}, \tag{24}$$

and

$$(\bar{\boldsymbol{u}}^{m})_g = \frac{1}{N_g} \sum_{\mathcal{G}(i,j)=g} (\boldsymbol{u}_i^{m})_j, \quad g,i,j \in \mathcal{M}. \tag{25}$$

This results in a computationally inexpensive update step forcing the norm of the consensus to have unit value.

TODO: Okay, so (23) is the major hurdle for selling this algorithm as distributed. This step still requires some sort of fusion center. But there must be a way to avoid that, maybe can be figured out this week?

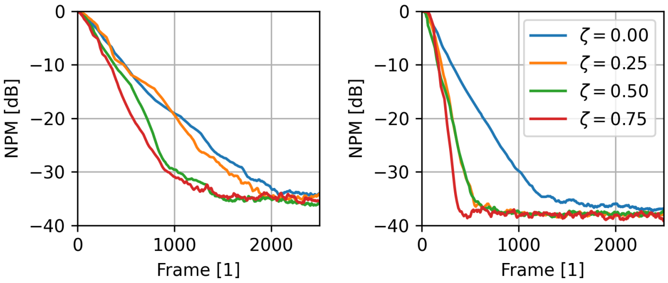
4. NUMERICAL EVALUATION

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

$$\mathrm{NPM}(m) = 20\,\log_{10}\left(\frac{\left\| \mathbf{h}(m) - \frac{\mathbf{h}^{\mathrm{T}}(m)\mathbf{h}_{\mathrm{t}}}{\mathbf{h}_{\mathrm{t}}^{\mathrm{T}}\mathbf{h}_{\mathrm{t}}}\mathbf{h}(m) \right\|_2}{\|\mathbf{h}(m)\|_2} \right) \tag{26}$$

where $\mathbf{h}$ is the stacked vector of impulse response estimates (cf. Sec. 2.3), which are the inversely Fourier-transformed estimated frequency responses stacked in $\boldsymbol{h}$ and $\mathbf{h}_{\mathrm{t}}$ is the ground truth. The signal-to-noise ratio (SNR) for the simulations is defined as

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



(a) $M = 4$, $P = 4$ (b) $M = 8$, $N = 8$

Fig. 3: Comparison of ADMM-BSI convergence behaviour with $L = 16$ for different values of the sub-problem overlap ζ for SNR $=$ 20 dB.

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-channel system ($M = 5$, $N = 5$). The impulse responses are drawn from a Normal 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: $\mu_{\text{MCQN}} = 0.5$, $\mu_{\text{NMCFLMS}} = 0.4$, $\mu_{l_p-\text{NMCFLMS}} = 0.3$, $\mu_{\text{ADMM}} = 0.6$, $\rho = 1$, $\eta = 0.98$. Fig. 2 shows the median of 30 Monte-Carlo runs where the averaged NPM of the last 100 frames is plotted. It is observable that the proposed algorithm yields a lower steady-state NPM than the compared NMCFLMS, RNMCFLMS, and l_p-RNMCFLMS algorithms.

The second simulation is a small-scale assessment of the influence of the overlap of sub-problems. Two base scenarios, $M = 4$, $N = 4$ and $M = 8$, $N = 8$, are evaluated using short ($L = 16$) random impulse responses and different values for the sub-problem overlap parameter ζ which is defined as the ratio of ones to zeros of only considering elements of $\mathbf{G}$ that are not part of the main diagonal, superdiagonal and the first element of the last row (marked gray on the inset of Fig. 3a). Random patterns satisfying this definition for $\zeta \in \{0.0, 0.25, 0.5, 0.75\}$ are generated. Fig. 3 shows the median of 30 Monte-Carlo runs for the two setups, which shows the proportional relation of convergence speed, channel number M and sub-problem overlap ζ while the steady-state error is not dependent on ζ.

TODO: Rewrite this. Confusing.

5. CONCLUSIONS

In this paper, an adaptive ADMM algorithm for blind system identification was developed. The algorithm separates the BSI problem into lower-dimensional sub-problems to reduce computational complexity and allows for parallel or distributed processing while maintaining steady-state error performance and convergence speed. Preliminary results using white Gaussian noise and randomly generated impulse responses have demonstrated improved steady-state error measures compared to state-of-the-art algorithms and that steady-state performance is not affected by separating the BSI problem into sub-problems while convergence speed is.

6. REFERENCES

[1] Y. Sato, "A Method of Self-Recovering Equalization for Multilevel 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 multichannel least mean square and Newton algorithms for blind channel identiÿcation," *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 $\ell _p$-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] 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.

[19] 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.

[20] 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.

[21] Stephen Boyd, Neal Parikh, Eric Chu, and Jonathan Eckstein, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Foundations and Trends{\textregistered} in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[22] Huahua Wang and Arindam Banerjee, "Online Alternating Direction Method (longer version)," *arXiv:1306.3721 [cs, math]*, July 2013.

[23] 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.