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

Matthias Blochberger, *Filip Elvander*, *Randall Ali, Toon van Waterschoot, Marc Moonen*

KU Leuven
Department of Electrical Engineering (ESAT)
STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics
3001 Leuven, Belgium

ABSTRACT

In this contribution, we propose a 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). The proposed algorithm exploits the separability of the cross-channel relations by splitting the multichannel identification problem into lower-dimensional sub-problems with reduced computational complexity, which can be solved in parallel. Each sub-problem yields estimates for a 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 comparable to low-cost frequency-domain algorithms and estimation errors better than a high-performing Quasi-Newton algorithm.

***Index Terms*—** blind system identification, multichannel 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. Suh 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 multichannel 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.

The alternating direction method of multipliers (ADMM) [18] solves convex optimization problems by splitting them into smaller sub-problems, each less complex to solve than the original one. In this paper, we use ADMM to separate the inter-channel cross-relations to form smaller sub-problems involving overlapping subsets of the entire channel set and use equality constraints to find a consensus on the overlapping estimated sub-problem parameters. The ADMM update steps are applied in a block processing scheme forming an adaptive algorithm also referred to as Online-ADMM [19, 20].

The resulting algorithm provides good estimation results while keeping the computational complexity low and providing scalability through distributed processing possibilities.

2. PROBLEM STATEMENT

2.1. Signal Model

We consider an acoustic SIMO system with the input signal $\mathbf{s}(n)$ and $i \in \mathcal{M}$ with $\mathcal{M} \triangleq \{1, \ldots, M\}$ outputs $\mathbf{x}_i(n)$ defined as

$$\mathbf{s}(n) = \begin{bmatrix} s(n) & s(n-1) & \ldots & s(n-2L+2) \end{bmatrix}^{\mathrm{T}} \quad (1)$$

$$\mathbf{x}_i(n) = \begin{bmatrix} x_i(n) & x_i(n-1) & \ldots & x_i(n-L+1) \end{bmatrix}^{\mathrm{T}}. \quad (2)$$

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

[*]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 work was supported in part by the Research Foundation - Flanders (FWO) grant 12ZD622N.

$\mathbf{v}_i(n)$, assumed to be zero-mean and uncorrelated. The signal model is described by

$$\mathbf{x}_i(n) = \mathbf{H}_i\mathbf{s}(n) + \mathbf{v}_i(n), \tag{3}$$

where $\mathbf{H}_i$ is the $L \times (2L-1)$ linear convolution matrix of the ith channel using the elements of $\mathbf{h}_i$ which represents the system to be identified.

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. are not coprime), and (ii) the covariance matrix of the input signal $\mathbf{s}(n)$ is of full rank (i.e. the signal fully excites the channels).

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

$$\mathbf{x}_i^{\mathrm{T}}(n)\mathbf{h}_j = \mathbf{x}_j^{\mathrm{T}}(n)\mathbf{h}_i, \quad i, j \in \mathcal{M},\, i \neq j \tag{4}$$

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 $\mathbf{R}_{ij} = \mathrm{E}\left\{\mathbf{x}_i\mathbf{x}_j^{\mathrm{T}}\right\}$ and combining all cross-relations (4) (see e.g. [12] for a more thourough derivation) yields the system of equations

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

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

$$\mathbf{R} = \begin{bmatrix} \sum_{i\neq 1}\mathbf{R}_{ii} & -\mathbf{R}_{21} & \cdots & -\mathbf{R}_{M1} \\ -\mathbf{R}_{12} & \sum_{i\neq 2}\mathbf{R}_{ii} & \cdots & -\mathbf{R}_{M2} \\ \vdots & \vdots & \ddots & \vdots \\ -\mathbf{R}_{1M} & -\mathbf{R}_{2M} & \cdots & \sum_{i\neq M}\mathbf{R}_{ii} \end{bmatrix} \tag{6}$$

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

$$\boldsymbol{R}(m)\boldsymbol{h} = \mathbf{0} \tag{7}$$

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

$$\boldsymbol{R}(m) = \eta\boldsymbol{R}(m-1) + (1-\eta)\hat{\boldsymbol{R}}(m) \tag{8}$$

instead of estimating the expectation by sample averages. Here, $\eta \in [0,1]$ is an exponential smoothing factor and $\hat{\boldsymbol{R}}(m)$ is constructed analogous to (6), however with covariance matrices replaced by cross-spectrum matrices

$$\boldsymbol{R}_{ij}(m) = \boldsymbol{X}_i^{\mathrm{H}}(m)\boldsymbol{X}_j(m) \tag{9}$$

with

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

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

$$\boldsymbol{W}_{L\times 2L}^{01} = \mathbf{F}_{L\times L}\mathbf{W}_{L\times 2L}^{01}\mathbf{F}_{2L\times 2L}^{-1} \tag{11}$$

$$\boldsymbol{W}_{2L\times L}^{10} = \mathbf{F}_{2L\times 2L}\mathbf{W}_{2L\times L}^{10}\mathbf{F}_{L\times L}^{-1} \tag{12}$$

are the frequency-domain overlap-save matrices where $\mathbf{F}_{L\times L}$ and $\mathbf{F}_{2L\times 2L}$ are the DFT matrices for sizes L and 2L respectively and

$$\mathbf{W}_{L\times 2L}^{01} = \begin{bmatrix}\mathbf{0}_{L\times L} & \mathbf{I}_{L\times L}\end{bmatrix} \tag{13}$$

$$\mathbf{W}_{2L\times L}^{10} = \begin{bmatrix}\mathbf{I}_{L\times L} & \mathbf{0}_{L\times L}\end{bmatrix}^{\mathrm{T}} \tag{14}$$

denote the time-domain overlap-save matrices.

Analogous to the time-domain formulation, the vector $\boldsymbol{h}$ is a stacked vector of complex-valued frequency responses. In the presence of noise, the system of equations (7) is best solved by posing it as a least-mean-squares minimization problem:

$$\hat{\boldsymbol{h}}(m+1) = \arg\min_{\boldsymbol{h}} \quad \boldsymbol{h}^{\mathrm{H}}(m)\boldsymbol{R}(m)\boldsymbol{h}(m) \tag{15}$$

$$\text{s.t.} \quad \boldsymbol{h}^{\mathrm{H}}(m)\boldsymbol{h}(m) = a \tag{16}$$

where a is a scaling to be chosen.

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

3. PROPOSED ALGORITHM

3.1. Problem Splitting

In state-of-the-art algorithms [15, 17], the minimization problem (15) is solved in its full form resulting in potentially high computational effort when the number of channels and length of impulse responses or number of DFT bins is large. Here however, we split the problem into $N \in \mathbb{N}$ sub-problems and denote the index set of these N sub-problems as $\mathcal{N} \triangleq \{1, \ldots, N\}$. Each sub-problem is defined by a subset of the full channel set $\mathcal{C}_i \subseteq \mathcal{M}$ with $i \in \mathcal{N}$. Following from that, we define the sets $\bar{\mathcal{C}}_j = \{i | j \in \mathcal{C}_i\}$ for $i, j \in \mathcal{M}$ where for each channel j a set represents the sub-problems that particular channel is part of. This channel-to-sub-problem relation can also be represented as a $M \times N$ matrix $\mathbf{G}$ where the sets $\mathcal{C}_i$ and $\bar{\mathcal{C}}_j$ are the indices of non-zero elements of the rows and columns respectively (see Fig. 1). Further, $M_i = |\mathcal{C}_i|$ and $N_j = \left|\bar{\mathcal{C}}_j\right|$ with $i, j \in \mathcal{M}$.

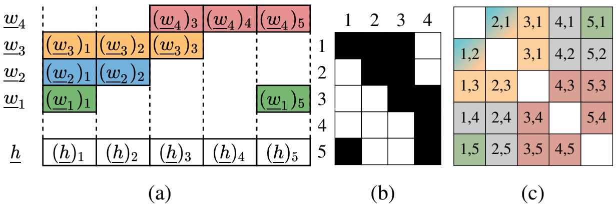


Fig. 1: Problem splitting and parameter mapping for a general case with $M = 5$ channels and $N = 4$ sub-problems. (a) shows the mapping of the local estimate components $(\boldsymbol{w}_i)_j$ to the global consensus components $(\boldsymbol{h})_j$ via $\mathcal{G}(i,j)$, (b) $\mathbf{G}$ with 1 (■) and 0 (□), (c) shows i,j-cross-relations used by sub-problems compared to full M-channel problem.

We replace the cost function minimized in (15) with the separable cost function

$$\tilde{J}(\boldsymbol{h}) = \sum_{i\in\mathcal{N}} \tilde{J}_i(\boldsymbol{w}_i) = \sum_{i\in\mathcal{N}} \boldsymbol{w}_i^{\mathrm{H}} \boldsymbol{P}_i \boldsymbol{w}_i \tag{17}$$

where $\boldsymbol{w}_i$ is defined as the stacked vector of estimated frequency responses analogous to $\boldsymbol{h}$ (cf. Sec. 2.2) and $\boldsymbol{P}$ is constructed as defined in (6), both using the respective subset $\mathcal{C}_i$. If the channel subsets are proper $\mathcal{C}_i \subset \mathcal{M}$, this leads to N sub-problems, each with smaller dimensions than the original centralized problem, reducing complexity. The lowest-dimensionality is attained when each $\mathcal{C}_i$ has two elements, i.e. two channels, with parallel processing, effectively reducing problem size from $ML \times ML$ to $2L \times 2L$. In Fig. 1, we try to visualize which information is used to solve the sub-problems compared to the initial problem for a general case.

3.2. General-Form Consensus ADMM

Minimization problems with separable cost functions can be solved by the well-established method of consensus ADMM [18]. In the previous section we introduced the split cost function (17) which we now set in the context of the general-form consensus problem

$$\text{minimize} \quad \sum_{i\in\mathcal{N}} \tilde{J}_i(\boldsymbol{w}_i) \tag{18}$$

$$\text{subject to} \quad (\boldsymbol{w}_i)_j = \boldsymbol{h}_{\mathcal{G}(i,j)}, \quad i \in \mathcal{N}, j \in \mathcal{C}_i \tag{19}$$

where $\mathcal{G}(i,j) = g$ 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 $(\boldsymbol{h})_g$ (cf. Fig. 1). 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 data of all sub-problems.

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

$$\mathcal{L}_\rho(\boldsymbol{w},\boldsymbol{h},\boldsymbol{u}) = \sum_{i\in\mathcal{N}} \Big(\boldsymbol{w}_i^{\mathrm{H}} \boldsymbol{P}_i \boldsymbol{w}_i + \boldsymbol{u}_i^{\mathrm{H}} \left(\boldsymbol{w}_i - \tilde{\boldsymbol{h}}_i\right) + \left(\boldsymbol{w}_i - \tilde{\boldsymbol{h}}_i\right)^{\mathrm{H}} \boldsymbol{u}_i + \left(\boldsymbol{w}_i - \tilde{\boldsymbol{h}}_i\right)^{\mathrm{H}} \rho \mathbf{I} \left(\boldsymbol{w}_i - \tilde{\boldsymbol{h}}_i\right) \Big). \tag{20}$$

The ADMM then consists of the steps:

$$\boldsymbol{w}_i^{k+1} = \underset{\boldsymbol{w}_i}{\operatorname{argmin}}\, \mathcal{L}_\rho(\boldsymbol{w},\boldsymbol{h}^k,\boldsymbol{u}^k), \tag{21}$$

$$\boldsymbol{h}^{k+1} = \underset{\boldsymbol{h},\|\boldsymbol{h}\|=a}{\operatorname{argmin}}\, \mathcal{L}_\rho(\boldsymbol{w}^{k+1},\boldsymbol{h},\boldsymbol{u}^k) \tag{22}$$

$$\boldsymbol{u}_i^{k+1} = \boldsymbol{u}_i^k + \rho\left(\boldsymbol{w}_i^{k+1} - \tilde{\boldsymbol{h}}_i^{k+1}\right). \tag{23}$$

As this is still denoted as an iterative algorithm, we will introduce the online/adaptive aspect of the proposed one next.

3.3. Online ADMM-BSI

We introduced the original problem (7) with time-dependent data matrix. Following from this, also the data term in (20) is time-dependent, which from here on out will be denoted with the additional superscript time index m as $\boldsymbol{w}_i^{\mathrm{H}} \boldsymbol{P}_i^m \boldsymbol{w}_i$. A thorough look at time-varying data terms in ADMM can be found in [19, 20] where it is referred to as “Online-ADMM”.

We transform the iterative batch processing method (21)-(23) 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 $\boldsymbol{w}_i$ (21) can be solved by various algorithms, in this case however we perform a Newton update step

$$\boldsymbol{w}_i^{m+1} = \boldsymbol{w}_i^m - \mu \boldsymbol{V}_i^m \left(\boldsymbol{P}_i^m \boldsymbol{w}_i^m + \boldsymbol{u}_i^m + \rho\left(\boldsymbol{w}_i^m - \tilde{\boldsymbol{h}}_i^m\right)\right) \tag{24}$$

where $0 < \mu \leq 1$ is a step size and $\boldsymbol{V}_i^m = (\boldsymbol{P}_i^m + \rho\mathbf{I})^{-1}$ is the inverse Hessian of the problem. As this inverse is costly to compute, it is approximated by a diagonalized matrix

$$\tilde{\boldsymbol{V}}_i^m = \operatorname{diag}\left\{ \left(\operatorname{diag}\{\boldsymbol{P}_i^m\} + \rho\mathbf{1}\right)^{-1} \right\}, \tag{25}$$

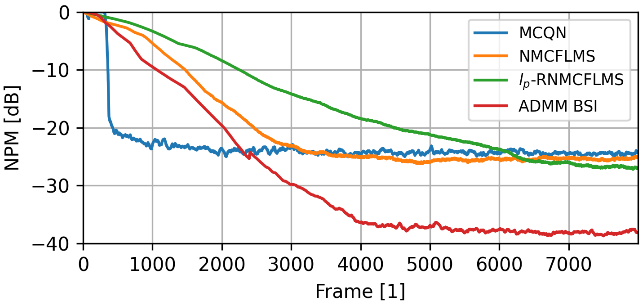
similar to NMCFLMS [13], which is straightforward to compute.

For the update step of the consensus variable $\boldsymbol{h}$, it may be readily verified (cf. [18]) that the solution to (22) is given by

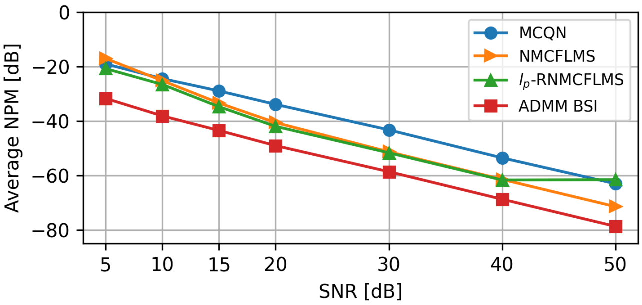
$$\boldsymbol{h}^{m+1} = a \frac{\bar{\boldsymbol{w}}^{m+1} + \frac{1}{\rho}\bar{\boldsymbol{u}}^m}{\left\| \bar{\boldsymbol{w}}^{m+1} + \frac{1}{\rho}\bar{\boldsymbol{u}}^m \right\|}. \tag{26}$$

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

$$(\bar{\boldsymbol{w}}^{m+1})_g = \frac{1}{N_g} \sum_{\mathcal{G}(i,j)=g} (\boldsymbol{w}_i^{m+1})_j, \quad g,i,j \in \mathcal{M}, \tag{27}$$



(a)



(b)

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

and

$$(\bar{\boldsymbol{u}}^m)_g = \frac{1}{N_g} \sum_{\mathcal{G}(i,j)=g} (\boldsymbol{u}_i^m)_j, \quad g, i, j \in \mathcal{M}. \tag{28}$$

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

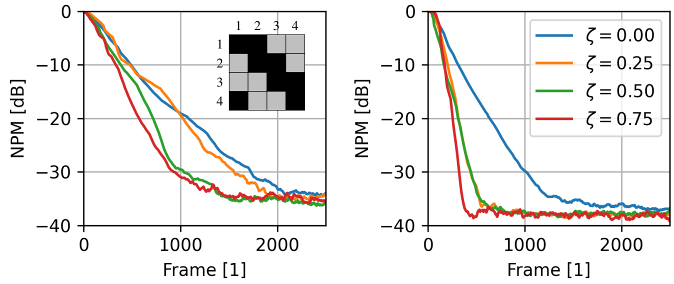
4. NUMERICAL EVALUATION

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

$$\mathrm{NPM}(m) = 20\,\log_{10}\left(\frac{\left\|\mathbf{h}(m) - \frac{\mathbf{h}^{\mathrm{T}}(m)\mathbf{h}_{\mathrm{t}}}{\mathbf{h}_{\mathrm{t}}^{\mathrm{T}}\mathbf{h}_{\mathrm{t}}}\mathbf{h}(m)\right\|_2}{\|\mathbf{h}(m)\|_2}\right) \tag{29}$$

where $\mathbf{h}$ is the stacked vector of impulse response estimates (cf. Sec. 2.2), which are the inversely Fourier-transformed estimated frequency responses stacked in $\boldsymbol{h}$ and $\mathbf{h}_{\mathrm{t}}$ is the ground truth. Further, the signal-to-noise ratio (SNR) for the experiments is defined as

$$\mathrm{SNR} = 10\log_{10}\left(\frac{\sigma_s^2\|\mathbf{h}_t\|}{M\sigma_v^2}\right) \tag{30}$$



(a) $M = 4, N = 4$ (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.

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 experiment 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 100 seconds ($f_s = 8000$) of WGN to ensure convergence of all algorithms. The step sizes are hand tuned for stable convergence: $\mu_{\mathrm{MCQN}} = 0.5$, $\mu_{\mathrm{NMCFLMS}} = 0.4$, $\mu_{l_p-\mathrm{NMCFLMS}} = 0.3$, $\mu_{\mathrm{ADMM}} = 0.6$, $\rho = 1$, $\eta = 0.98$. Fig. 2 shows the median of 30 Monte-Carlo runs where the averaged NPM of the last 100 frames is taken as measurement. 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 experiment 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 in 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.

5. CONCLUSIONS

In this paper, a novel adaptive ADMM algorithm for blind system identification was developed. The algorithm separates the BSI problem into lower-dimensional sub-problems to reduce complexity and allow parallel processing while maintaining steady-state error performance and convergence speed. Comparison to state-of-the-art algorithms in numerical simulations shows improved steady-state error measures for randomly generated impulse responses. A further experiment to assess the influence of sub-problem overlap shows that steady-state performance is not affected by the separation 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] JM Mecidel, "Tutorial on higher-order statistics in signal processing and system theory," *Proc. IEEE*, vol. 79, no. 3, pp. 278–305, 1991.

[5] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Transactions on information Theory*, vol. 40, no. 2, pp. 340–349, 1994.

[6] G. Xu, H. Liu, L. 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] S. Gannot and M. Moonen, "Subspace Methods for Multimicrophone Speech Dereverberation," *EURASIP J. Adv. Signal Process.*, vol. 2003, no. 11, pp. 769285, Dec. 2003.

[9] K. I. Diamantaras and T. 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] Q. Mayyala, K. Abed-Meraim, and A. Zerguine, "Structure-Based Subspace Method for Multichannel Blind System Identification," *IEEE Signal Processing Letters*, vol. 24, no. 8, pp. 1183–1187, Aug. 2017.

[11] Y. 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] Y. Huang and J. Benesty, "Adaptive multi-channel least mean square and Newton algorithms for blind channel identification," *Signal Processing*, p. 12, 2002.

[13] Y. 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] H. He, J. Chen, J. Benesty, and T. 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] B. Jo and P. 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] E. A.P. Habets and P. 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] S. Boyd, N. Parikh, E. Chu, 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.

[19] H. Wang and A. Banerjee, "Online Alternating Direction Method," *International Conference on Machine Learning*, June 2012.

[20] 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.