Received XX Month, XXXX; revised XX Month, XXXX; accepted XX Month, XXXX; Date of publication XX Month, XXXX; date of current version XX Month, XXXX.

Digital Object Identifier 10.1109/XXXX.2022.1234567

Efficient Room Impulse Response Estimation in Wireless Acoustic Sensor Networks using Low-Rank Tensor Decomposition

Matthias Blochberger[1], Graduate Student Member, IEEE, Filip Elvander [2], Member, IEEE, and Toon van Waterschoot[1], Member, IEEE

[1]Department of Electrical Engineering, ESAT-STADIUS, KU Leuven, 3001 Leuven, Belgium
[2]Department of Information and Communications Engineering, Aalto University, 02150 Espoo, Finland

Corresponding author: Matthias Blochberger (email: matthias.blochberger@esat.kuleuven.be).

This paragraph of the first footnote will contain support information, including sponsor and financial support acknowledgment. For example, "This work was supported in part by the U.S. Department of Commerce under Grant 123456."

ABSTRACT We propose an adaptive algorithm that leverages tensor decompositions to identify multi-channel room acoustic systems efficiently. By modelling room impulse responses through low-rank tensor approximations, our method achieves significant dimensionality reduction while preserving essential system characteristics. The algorithm employs an alternating least-squares approach, iterating over various tensor modes and ranks, and incorporates a common factor estimation step that exploits shared structures across channels. Numerical experiments validate that our approach strikes a favourable balance between reducing the number of parameters and maintaining high estimation accuracy, making it a promising tool for complex acoustic system identification tasks.

INDEX TERMS Enter key words or phrases in alphabetical order, separated by commas. Using the *IEEE Thesaurus* can help you find the best standardized keywords to fit your article. Use the thesaurus access request form for free access to the *IEEE Thesaurus*.

I. INTRODUCTION

The task of estimating acoustic room impulse responses (RIRs) from observed input-output data is a commonly required problem in the field of room acoustics. RIRs describe the acoustic path from a source to a microphone and comprise the direct path, if existent, reflections from the room boundaries and objects within the room. The estimation problem can refer to a static system, where the RIRs are assumed time-invariant, or to a time-varying system, where the RIRs change over time due to changes in source or microphone position or changes in the room itself. Estimated RIRs have numerous applications in audio signal processing and room acoustics, including enhancement tasks such as dereverberation [1], source separation [2], acoustic feedback and echo cancellation [3] or for room acoustics analysis, such as the computation of room acoustic parameters, such as the reverberation time RT60 or speech intelligibility predictors C_{50} [4], room geometry estimation [5], [6] and source localisation [7]. A room acoustic system, represented by its RIRs, are commonly modelled as infinite impulse response (IIR) filters [**?**] or as finite impulse response (FIR) filters [**?**]. The IIR approach allows for a more compact representation of the system, with the downside of potential instability [**?**]. The FIR approach is more common in audio signal processing applications, as it is straightforward to implement and estimate, however, in typical room acoustic scenarios, the RIRs can comprise several thousands of taps, leading to issues with dimensionality in estimation and downstream tasks. There are various existing approaches to reduce dimensionality and computational complexity in FIR-based room acoustic system identification, such as frequency-domain processing [8] or low-rank approximations [9]. Here, we focus on the latter, formulating a joint multichannel estimation problem using tensor decompositions [10], [11], specifically the block term decomposition (BTD) [12], [13], [14], a generalization of the canonical polyadic decomposition (CPD) used in [15], [9], [16] and the (generalized) singular value decomposition (SVD) used in [17] and [11].

Wireless sensor networks (WSNs) are becoming more relevant in various technical fields, driven by the increase in network-connected devices and the growing Internet of Things (IoT). They find use in various applications such as environmental monitoring, industrial automation, and smart cities. They are also gaining traction in acoustic and audio signal processing applications. Some of the first proposals for networked audio sensors, such as [18], and [19], started the trend, and now, with the advent of smart homes and smart cities, the potential of using network-connected devices with built-in microphones and loudspeakers, such as phones, tablets and more is ever-increasing. In the context of distributed signal processing tasks such as, e.g., distributed signal estimation [20], [21], [22], noise control and echo cancellation [23], [24], and beamforming [25], [24] have been studied in recent years.

We combine these two research fields and formulate a joint multichannel room acoustic RIR estimation problem in wireless acoustic sensor networks (WASNs). The typical approach to multichannel RIR estimation is to estimate each channel's RIR separately, neglecting potential common structures across channels. In [26] a Kronecker product decomposition, an equivalent formulation of a CPD, is proposed to model a RIR as a time-invariant (shared) part and a time-varying (individual) part. There the time (and therefore space) invariant part is assumed to be known, estimated beforehand, which then reduces the dimensionality and therefore complexity of the individual part in an estimation problem. In [17], [11] it was observed that a multichannel acoustic systems exhibits common structures across channels in their SVD and generalized low-rank approximation of matrices (GLRAM) decompositions, respectively. Relatedly, in [27] a method to estimate RIRs on learned acoustic manifolds is proposed. In previous work, we extended the observation of shared structures to higher-order tensor decompositions and provided theoretical background based on common acoustic poles [**?**]. We showed that exploiting common structures across channels can lead to improved compression performance of RIRs using tensor decompositions. These common structures show up as similarities in subspaces spanned by the factors of the tensor decompositions of RIRs. Here, we apply this observation to the problem of online/adaptive multichannel acoustic RIR estimation. We propose a distributed adaptive algorithm based on the BTD for estimating RIRs in a WASN, where multichannel nodes a) use parameter-efficient multichannel representations and b) nodes collaborate by sharing information about their local estimates. This brings better conditioning, faster convergence and improved estimation stability with the trade-off of reduced steady-state estimation accuracy.

Notation: We use lowercase letters for scalar values (x), lowercase bold letters for vectors ($\mathbf{x} \in \mathbb{R}^{I}$), uppercase bold letters for matrices ($\mathbf{X} \in \mathbb{R}^{I_1 \times I_2}$), and calligraphic bold letters for D-modal tensors ($\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I_1 \times \cdots \times I_D}$). Each of which can be discrete-time-dependent, e.g. $x(n)$ or $\boldsymbol{\mathcal{X}}(n)$. More general linear operators are denoted as uppercase letters (e.g., $A : \mathbb{R} \to \mathbb{R}$). Addressing specific elements of a vector or matrix is done by using brackets, e.g. $\mathbf{x}[i]$, $\mathbf{X}[i_1,i_2]$, and $\boldsymbol{\mathcal{X}}[i_1,\ldots,i_D]$ for the i-th, the (i_1, i_2)-th, and the $(i_1, \ldots, i_D)$-th element of the vector, matrix, and tensor, respectively. The element-wise product, Kronecker product and outer products for vectors/matrices/tensors are denoted with $\odot$, $\otimes$ and $\circ$, respectively. We denote convolution in the time-domain with $*$. Of further importance are tensor-specific operations (see [10] for details): Let the map between a vector $\mathbf{x}$ and tensor $\boldsymbol{\mathcal{X}}$, with $I = \prod_{d=1}^{D} I_d$, be the bijective reshaping operator $\mathcal{R}_{I_1,\ldots,I_D} : \mathbb{R}^{I} \to \mathbb{R}^{I_1 \times \cdots \times I_D}$ such that $\boldsymbol{\mathcal{X}} = \mathcal{R}_{I_1,\ldots,I_D}(\mathbf{x})$. Unfolding the tensor into a matrix $\mathbf{X}_{(d)} \in \mathbb{R}^{I_d \times \prod_{\delta \neq d} I_\delta}$ along mode d is defined as the matrix whose columns are the mode-d fibres of the tensor $\boldsymbol{\mathcal{X}}$. Further, the d-mode product of a tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I_1 \times \cdots \times I_D}$ with a matrix $\mathbf{A} \in \mathbb{R}^{J \times I_d}$ is a matrix-vector product along mode d such that

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{X}} \times_d \mathbf{A} \in \mathbb{R}^{I_1 \times \cdots \times I_{d-1} \times J \times I_{d+1} \times \cdots \times I_D}$$
$$\iff \mathbf{Y}_{(d)} = \mathbf{A}\mathbf{X}_{(d)} \in \mathbb{R}^{k \times \prod_{\delta \neq d} I_\delta}.$$

We denote the multiple-mode product on a tensor with $D \geq E$ modes as

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{X}} \times_1 \mathbf{A}_1 \cdots \times_E \mathbf{A}_E \equiv \boldsymbol{\mathcal{X}} \times_{d \in [E]} \mathbf{A}_d.$$

We denote the multiple-mode product with all modes except d as

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{X}} \times_{\delta \neq d} \mathbf{A}_\delta \equiv \boldsymbol{\mathcal{X}} \times_{\delta \in [D] \setminus \{d\}} \mathbf{A}_\delta.$$

For tensors $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I_1 \times \cdots \times I_K}$ and $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{J_1 \times \cdots \times J_L}$, where K, L are integers ≥ 1 and index sets $\mathcal{K} \subseteq \{1, \ldots, K\}$ and $\mathcal{L} \subseteq \{1, \ldots, L\}$ with a 1-1 correspondence and matching sizes, write

$$\boldsymbol{\mathcal{Z}} = \langle \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{Y}} \rangle_{\mathcal{K}|\mathcal{L}}$$

for the tensor obtained by contracting $\boldsymbol{\mathcal{X}}$ along modes in $\mathcal{K}$ and $\boldsymbol{\mathcal{Y}}$ along modes in $\mathcal{L}$. The result keeps the remaining modes of $\boldsymbol{\mathcal{X}}$ in order, then the remaining modes of $\boldsymbol{\mathcal{Y}}$ in order. The Frobenius inner product of two tensors $\boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{Y}} \in \mathbb{R}^{I_1 \times \cdots \times I_D}$ is the special case of the contracted tensor product with contraction over all modes $1, \ldots, D$ (denoted $1 : D$)

$$\langle \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{Y}} \rangle := \langle \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{Y}} \rangle_{1:D|1:D} \in \mathbb{R} \tag{1}$$

We use the shorthand $\langle \cdot, \cdot \rangle_{\delta \neq d}$ for partial contractions over all modes except d. More detail on these definitions can be found in [10] and [28].

Todo: Trade-off

Don't say it like that.

There is limited prior work on distributed estimation of estimating room acoustic systems in WASNs such as [29] where a similar concept of exploiting common structures across channels is used for distributed estimation of a common-acoustical-pole room model.

This paper contains the following contributions: First, we introduce the signal model of a convolutive multichannel acoustic system using BTD based on which we formulate the estimation problem as a distributed optimization problem. Second, we propose an adaptive algorithm based on alternating higher-order orthogonal iteration (HOOI) for the BTD [10], [12] and three-operator splitting [30] (forward-backward-backward). We propose the cost function to comprise a data/signal term, a Laplacian regularization term to promote consensus across nodes for shared factors, and a Stiefel manifold constraint to enforce the orthogonality of the factor matrices. The proposed algorithm alternates between updating local estimates at each node (forward) and performing consensus steps (backward) and Stiefel projections (backward). We analyse the computational complexity of the proposed algorithm and show that it is efficient compared to state-of-the-art methods. Finally, we validate the proposed algorithm in numerical experiments using simulated and measured RIRs.

II. SIGNAL MODEL & PROBLEM FORMULATION

Consider a sensor network with K sensors (nodes, vertices) $\mathcal{K} := \{1, \dots, K\}$. Let each node k have M_k microphones, such that the total number of microphones is $M = \sum_{k=1}^{K} M_k$ and each node k have its own BTD filter $F_{k|m}$ for each of its microphones.

A. Signal Model

Suppose we have a single-input-multiple-output (SIMO) M-channel acoustic system that can be modelled by

$$y_{k|m}(n) = x(n) * h_{k|m}(n) + v_{k|m}(n) \tag{2}$$

where for each node $k = 1, \dots, K$ and channel $m = 1, \dots, M_k$, $y_{k|m}(n)$ is the output (microphone) signal, $x(n)$ is the input (source) signal, $h_{k|m}(n)$ is the channel's impulse response, $v_{k|m}(n)$ is the noise (assumed independent). We denote the discrete-time index n. We assume the length of the impulse responses is L. The goal is to estimate all $h_{k|m}(n)$ from the input-output data. It is sufficient and common to treat the M channels separately, i.e. to estimate the impulse responses $h_{k|m}(n)$ independently; however, we want to exploit common structures across channels within the room acoustic system. Therefore, we will formulate a joint estimation problem for all channels.

Let $\boldsymbol{\mathcal{X}}(n) \in \mathbb{R}^{I_1 \times \dots \times I_D}$ be the tensor representation of the input signal at time n, i.e., the tensorization of the last $L = \prod_{d=1}^{D} I_d$ samples $\boldsymbol{\mathcal{X}}(n) = R_{I_1,\dots,I_D}(\mathbf{x}(n))$ where $\mathbf{x}(n) \in \mathbb{R}^{I}$. Analogously, we define $\boldsymbol{\mathcal{H}}_{k|m} = R_{I_1,\dots,I_D}(\mathbf{h}_{k|m})$ as the tensor representation of the impulse response $\mathbf{h}_{k|m}(n) \in \mathbb{R}^{I}$ such that the convolution

$$\begin{aligned} y_{k|m}(n) &= \mathbf{x}^{\top}(n)\mathbf{h}_{k|m} + v_{k|m}(n) \\ &= \langle \boldsymbol{\mathcal{X}}(n), \boldsymbol{\mathcal{H}}_{k|m} \rangle + v_{k|m}(n). \end{aligned} \tag{3}$$

Let us consider a BTD of the tensor $\boldsymbol{\mathcal{H}}_{k|m}$ with R blocks, where each block r has its own core tensor $\boldsymbol{\mathcal{G}}^{(r)} \in \mathbb{R}^{K_1 \times \dots \times K_D}$ and factor matrices $\mathbf{A}^{(r,d)} \in \mathbb{R}^{I_d \times K_d}$ for each mode d. The decomposition is defined as

$$\boldsymbol{\mathcal{H}}_{k|m} \approx \widehat{\boldsymbol{\mathcal{H}}}_{k|m} = \sum_{r=1}^{R} \boldsymbol{\mathcal{G}}_{k|m}^{(r)} \times_{d\in[D]} \mathbf{A}_{k|m}^{(r,d)}. \tag{4}$$

Note that the core tensor dimensions need not be equal across blocks, however for notational convenience we will assume they are. Further, the factors matrices are constrained to have orthonormal columns, i.e. such that $\mathbf{A}_{k|m}^{(r,d)\top}\mathbf{A}_{k|m}^{(r,d)} = \mathbf{I}_{K_d}$, which implies that $K_d \leq I_d$.

To use this decomposition in the convolutional model, let us define for any input signal x and time n, the projection into the core space,

$$\boldsymbol{\mathcal{W}}_{k|m}^{(r)}(x, n) := \boldsymbol{\mathcal{X}}(n) \times_{d\in[D]} \mathbf{A}_{k|m}^{(r,d)\top}. \tag{5}$$

The channel output $y_{k|m}(n)$ can then be approximated using the BTD filter

$$\hat{y}_{k|m}(n) = F_{k|m}[x](n) + v_{k|m}(n), \tag{6}$$

with

$$F_{k|m}[x](n) := \sum_{r=1}^{R} \langle \boldsymbol{\mathcal{G}}_{k|m}^{(r)}, \boldsymbol{\mathcal{W}}_{k|m}^{(r)}(x, n) \rangle, \tag{7}$$

where $F_{k|m}$ is the BTD filter operator of channel m. It is linear in each of its components, i.e., the cores and the factors, therefore a multilinear operator. We can express the partial derivatives of the operator w.r.t. the cores as

$$\frac{\partial F_{k|m}[x](n)}{\partial \boldsymbol{\mathcal{G}}_{k|m}^{(r)}} = \boldsymbol{\mathcal{W}}_{k|m}^{(r)}(x, n) \tag{8}$$

due to the linearity of the tensor inner product. The partial derivatives w.r.t. the factors are

$$\frac{\partial F_{k|m}[x](n)}{\partial \mathbf{A}_{k|m}^{(r,d)}} = \langle \boldsymbol{\mathcal{Z}}_{k|m}^{(r,d)}(x, n), \boldsymbol{\mathcal{G}}_{k|m}^{(r)} \rangle_{\delta \neq d} \tag{9}$$

where

$$\boldsymbol{\mathcal{Z}}_{k|m}^{(r,d)}(x, n) := \boldsymbol{\mathcal{X}}(n) \times_{\delta \neq d} \mathbf{A}_{k|m}^{(r,\delta)\top} \tag{10}$$

is the mode-d partial projection tensor, i.e., the input projected through all modes except d.

B. Subspace similarities and clustering

Todo: Reference

Can one reference own work that has not even been published yet :/

It has been observed in prior work that room acoustic systems exhibit common structures across channels in their low-rank decompositions [17], [11]. In particular, it was shown in [11] that the factors of the tensor decompositions of the RIRs exhibit common structures across channels, signified by the small principle angles between the subspaces of the right singular vectors of the SVD, a two-dimensional tensor decomposition, of impulse responses within the same room. We further extended this observation to higher-order tensor decompositions in [**?**], showing that these common structures can be partially explained by common acoustic

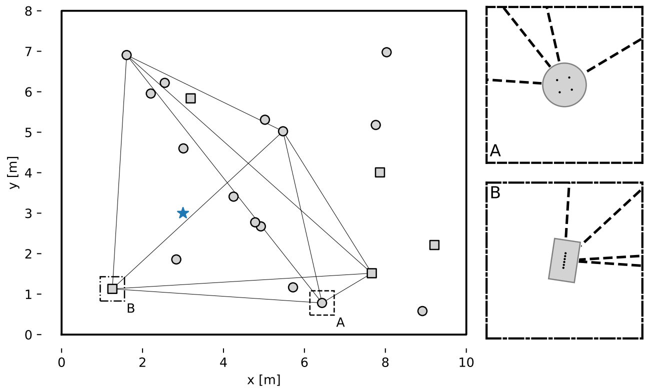

FIGURE 1. **Example of a network topology with $N = 20$ nodes and $M_k \in [5, 10]$ microphones each, randomly placed and arranged in either linear or circular arrays. The nodes are randomly placed in a room of size $10 \times 8 \times 3$ m^3. Indicated are two exemplary clusters of nodes (red, green) forming a (disconnected) graph $\mathcal{V}^{(r,d)}$ for a specific rank-node pair (r, d), while dashed lines indicate edges of other possible clusters in other rank-node pairs. Further, there are two detailed views of nodes showing the microphone arrangements of the randomly placed linear and circular arrays.**

poles across channels and persist also in higher-order decompositions. These common structures show up as similarities in the subspaces spanned by the factor matrices of the tensor decompositions of the RIRs. To exploit the common structures across channels, we propose to estimate common factors across channels, which leads to a reduction in the number of parameters to estimate and improved conditioning of the estimation problem. However, not all channels exhibit the same common structures; therefore, we propose to cluster channels that share similar structures and estimate common factors only within these clusters.

C. Distributed Optimization Problem

[31], [32], [33], [34], [35], [36] Considering the factor's similarities across channels, we propose to estimate shared factors for certain subsets of channels. Note that we consider these clusters to be fixed and known a priori. We differentiate between factors shared across the M_k channels of a single node k, which we consider as shared factors, i.e., the factors are identical across all channels of a node, and factors shared across multiple nodes, which we consider as shared, i.e., the factors are encouraged to be similar but not necessarily identical. For each rank-mode pair (r, d), let us define an undirected graph $\mathcal{V}^{(r,d)} = \{\mathcal{K}, \mathcal{E}^{(r,d)}, \mathbf{W}^{(r,d)}\}$, where $\mathcal{K}$ is the set of nodes, $\mathcal{E}^{(r,d)} \subseteq \mathcal{K} \times \mathcal{K}$ is the set of edges, and $\mathbf{W}^{(r,d)}$ with $w_{ij}^{(r,d)} = w_{ji}^{(r,d)} \geq 0$ is the symmetric weight adjacency matrix defining the similarity between the factors of nodes i and j for the rank-mode pair (r, d). Note that $w_{ii}^{(r,d)} = 0$, i.e., adjacency matrices do not include self-adjacency as opposed to weight matrices in distributed averaging for example [37]. Further, the graph $\mathcal{V}^{(r,d)}$ need not be connected, i.e., it can consist of multiple disconnected subgraphs, which we interpret as clusters of nodes sharing similar factors. The adjacency matrix for each (r, d) defines which nodes share certain factors, so let us define the set of shared factors on node k as

$$\mathcal{C}_k = \{(r, d) \mid \exists\, j \neq k : w_{kj}^{(r,d)} > 0\}. \tag{11}$$

With this, we can redefine the factors in the decomposition, such that

$$\mathbf{A}_{k|m}^{(r,d)} = \begin{cases} \mathbf{B}_k^{(r,d)} & \text{if } (r,d) \in \mathcal{C}_k \\ \mathbf{U}_{k|m}^{(r,d)} & \text{otherwise} \end{cases} \quad \forall\, {\scriptstyle k=1,\dots,K, \atop m=1,\dots,M_k} \tag{12}$$

where $\mathbf{B}_k^{(r,d)} \in \mathbb{R}^{I_d \times K_d}$ are the shared factors on node k and $\mathbf{U}_{k|m}^{(r,d)} \in \mathbb{R}^{I_d \times K_d}$ are individual factors. For the remainder of the paper, if not stated otherwise, we will use $\mathbf{A}_{k|m}^{(r,d)}$ if the expression holds for both cases and $\mathbf{B}_k^{(r,d)}$ or $\mathbf{U}_{k|m}^{(r,d)}$ if it only holds for one of the two cases. Let the minimization problem be defined as

$$\min_{\mathbf{\Theta}} \quad f(\mathbf{\Theta}) + g(\mathbf{\Theta}) + h(\mathbf{\Theta}) \tag{13}$$

with the data term

$$f(\mathbf{\Theta}) = \mathbb{E}\left[\sum_{k=1}^{K} f_k(\mathbf{\Theta}, n)\right], \tag{14}$$

being a sum of local cost functions $f_k(\mathbf{\Theta}, n)$ at each node k. It is given by the mean squared error (MSE) between the desired and the estimated output signal, i.e.,

$$f_k(\mathbf{\Theta}, n) = \frac{1}{2}\sum_{m=1}^{M} e_{k|m}^2(\mathbf{\Theta}, n), \tag{15}$$

$$e_{k|m}(\mathbf{\Theta}, n) = y_{k|m}(n) - F_{k|m}(x, n) \quad \forall\, m = 1, \dots, M. \tag{16}$$

We assume that the input signal is sufficiently exciting and that all nodes have access to the input signal without transmission loss or delay. Here, $\mathbb{E}[\cdot]$ denotes the expectation operator and $\mathbf{\Theta} = (\{\boldsymbol{\mathcal{G}}_m^{(r)}\}, \{\mathbf{B}_k^{(r,d)}\}, \{\mathbf{U}_{k|m}^{(r,d)}\})$ is the set of all parameters to be estimated. Then the quadratic penalty $g(\mathbf{\Theta})$ is defined a

$$\begin{aligned} g(\mathbf{\Theta}) = \frac{1}{2}\sum_{(r,d)} \Big[& \lambda^{(r,d)} \sum_{(i,j)}^{\mathcal{E}^{(r,d)}} w_{ij}^{(r,d)} \|\mathbf{B}_i^{(r,d)} - \mathbf{B}_j^{(r,d)}\|_F^2 && \text{(17a)} \\ & + \gamma^{(r,d)} \sum_{k,m} \|\mathbf{A}_{k|m}^{(r,d)} - \mathbf{A}_{k|m,\text{old}}^{(r,d)}\|_F^2 && \text{(17b)} \\ & + \eta^{(r,d)} \sum_{k,m} \|\mathbf{A}_{k|m}^{(r,d)} - \bar{\mathbf{A}}_{k|m}^{(r,d)}\|_F^2 \Big] && \text{(17c)} \end{aligned}$$

where $\lambda^{(r,d)} \geq 0$ is a regularization parameter controlling the strength of the consensus for the rank-mode pair (r, d) and $w_{ij}^{(r,d)}$ is the weight of the edge (i, j) in the graph $\mathcal{V}^{(r,d)}$. The first term (17a) is a consensus regularization term promoting similarity of the shared factors across nodes according to the graph $\mathcal{V}^{(r,d)}$ defined above. The second term (17b) is a recursive regularization term with parameter $\eta^{(r,d)} \geq 0$ to improve convergence by avoiding flat regions in the cost function caused by the scaling and conjugate indeterminacies of the BTD [38], $\mathbf{A}^{(r,d)}_{k|m,\text{old}}$ denotes the previous iteration. The third term (17c) is a prior regularization term with parameter $\eta^{(r,d)} \geq 0$ promoting closeness to a prior estimate $\bar{\mathbf{A}}^{(r,d)}_{k|m}$ if available. The orthonormality penalty term $h(\boldsymbol{\Theta})$ is the sum of indicator functions on the Stiefel manifold,

$$h(\boldsymbol{\Theta}) = \sum_{k,m,r,d} \iota_{\text{St}}(\mathbf{A}^{(r,d)}_{k|m}) \tag{18}$$

where

$$\iota_{\text{St}}(\mathbf{A}) = \begin{cases} 0 & \text{if } \mathbf{A}^\top \mathbf{A} = \mathbf{I} \\ \infty & \text{otherwise} \end{cases}. \tag{19}$$

III. METHOD

To solve the optimization problem in (13), we propose an adaptive algorithm that iterates over all nodes, ranks and modes of the tensor decomposition implementing shared and individual factor updates with forward-backward-backward steps, based on the three-operator splitting strategy [30]. We are going to apply an alternating HOOI-adjacent optimization procedure [10] aggregating updates for shared factors and with subsequent proximal steps for the Laplacian regularization and orthonormality constraints. Let us state the distinct update phases:

Forward on $f(\boldsymbol{\Theta})$: Local parameter sets $\{\boldsymbol{\mathcal{G}}^{(r)}_{k|m}\}$, $\{\mathbf{B}^{(r,d)}_{k}\}$, and $\{\mathbf{U}^{(r,d)}_{k|m}\}$ are updated based on node-local signal data.

Backward on $g(\boldsymbol{\Theta})$: Each node k performs a proximal step on its shared and individual factors promoting consensus based on the exchanged information as well as Tikhonov and prior regularization. This requires neighbourhood exchange: Each node k exchanges its shared factors $\{\mathbf{B}^{(r,d)}_{k}\}$ with its graph neighbours.

Backward on $h(\boldsymbol{\Theta})$: Each factor is projected onto the Stiefel manifold to enforce orthonormality and the cores are realigned accordingly.

Note that in practice the backward steps are not required to be performed at every iteration, but rather can be performed every T_c iterations for some integer $T_c \geq 1$ to reduce communication and computational load. Applying consensus update at intervals will reduce speed of convergence w.r.t. the similarity of factors, however does not affect the feasibility of individual BTDs. The orthonormality of factors is not a hard requirement for a feasible BTD decompositions, however, keeping factors close to orthonormality will benefit problem conditioning and interpretability.

A. Forward: Local Update

We propose to use Levenberg-Marquardt (LM)-type [39] updates using approximate second-order information for the local update phase for better convergence speed and stability. To derive the local factor update steps, let us define the mode-d back-projected regressor,

$$\boldsymbol{\Phi}^{(r,d)}_{m}(x,n) := \left\langle \boldsymbol{\mathcal{Z}}^{(r,d)}_{m}(x,n), \boldsymbol{\mathcal{G}}^{(r)}_{m} \right\rangle_{\delta \neq d} \in \mathbb{R}^{I_d \times K_d}, \tag{20}$$

a core-weighted contraction, representing the sufficient statistic for the factor $\mathbf{A}^{(r,d)}_{m}$. To use approximate second-order information, we will keep a running estimate of the diagonal of the Hessian for each factor. The subsequent explanation holds for all factors, therefore we omit the indexing, sub- and superscripts temporarily. Let

$$\boldsymbol{\Sigma}_\Phi = \mathbb{E}\left[\operatorname{vec}(\boldsymbol{\Phi}) \operatorname{vec}(\boldsymbol{\Phi})^\top\right] \tag{21}$$

be the Hessian matrix for the factor $\mathbf{A}$ in vectorized space, then we can approximate it by its diagonal, which in turn can be approximated by a factorization such that the conditioning can be seen as

$$\begin{aligned} &\operatorname{vec}^{-1}\left(\boldsymbol{\Sigma}_\Phi^{-1} \operatorname{vec}(\boldsymbol{\Phi})\right) \\ &\qquad \approx \operatorname{diag}(\mathbf{p}_{\text{col}})^{-1} \boldsymbol{\Phi} \operatorname{diag}(\mathbf{p}_{\text{row}})^{-1} =: \tilde{\boldsymbol{\Phi}} \end{aligned} \tag{22}$$

where $\tilde{\boldsymbol{\Phi}}$ is the regressor conditioned by the running estimates $\mathbf{p}_{\text{col}} \in \mathbb{R}^{I_d}$ and $\mathbf{p}_{\text{row}} \in \mathbb{R}^{K_d}$ defined by

$$\mathbf{p}_{\text{col}} \leftarrow \lambda \mathbf{p}_{\text{col}} + (1-\lambda)\hat{\mathbf{p}}_{\text{col}} \tag{23}$$
$$\mathbf{p}_{\text{row}} \leftarrow \lambda \mathbf{p}_{\text{row}} + (1-\lambda)\hat{\mathbf{p}}_{\text{row}} \tag{24}$$

with

$$\hat{\mathbf{p}}_{\text{col}} = \boldsymbol{\Phi}\boldsymbol{\Phi}^\top \mathbf{1}_{I_d} \tag{25}$$
$$\hat{\mathbf{p}}_{\text{row}} = \boldsymbol{\Phi}^\top \boldsymbol{\Phi} \mathbf{1}_{K_d} \tag{26}$$

where λ is a forgetting factor and $\mathbf{1}_{\{I_d, K_d\}}$ is a column vector of ones of appropriate length.

Analogously, we define the conditioned core sufficient statistic from the sufficient statistic in (8). Let there be the Hessian in vectorized core space,

$$\boldsymbol{\Sigma}_{\mathcal{W}} = \mathbb{E}\left[\operatorname{vec}(\boldsymbol{\mathcal{W}}) \operatorname{vec}(\boldsymbol{\mathcal{W}})^\top\right], \tag{27}$$

which we will approximate by its diagonal such that the conditioning can be seen as

$$\operatorname{vec}^{-1}\left(\boldsymbol{\Sigma}_{\mathcal{W}}^{-1} \operatorname{vec}(\boldsymbol{\mathcal{W}})\right) \approx \frac{\boldsymbol{\mathcal{W}}}{\boldsymbol{\mathcal{P}}} =: \tilde{\boldsymbol{\mathcal{W}}}, \tag{28}$$

an element-wise division of the core sufficient statistic $\boldsymbol{\mathcal{W}}$ by the running estimate $\boldsymbol{\mathcal{P}} \in \mathbb{R}^{K_1 \times \cdots \times K_D}$, which is defined by

$$\boldsymbol{\mathcal{P}} \leftarrow \lambda \boldsymbol{\mathcal{P}} + (1-\lambda) \boldsymbol{\mathcal{W}} \odot \boldsymbol{\mathcal{W}}. \tag{29}$$

Note that in practice, we add a small constant $\epsilon > 0$ to the denominators in the definitions of $\tilde{\boldsymbol{\Phi}}$ and $\tilde{\boldsymbol{\mathcal{W}}}$ to avoid numerical issues. Using sub- and superscripts again to index the nodes, ranks, modes, and channels, we can define the factor update step as a dampened LM step, with the update direction at every iteration defined by

$$\delta \mathbf{A}^{(r,d)}_{k|m} = -\frac{e_{k|m}}{\beta_{\text{f}} + \left\langle \boldsymbol{\Phi}^{(r,d)}_{k|m}, \tilde{\boldsymbol{\Phi}}^{(r,d)}_{k|m} \right\rangle} \tilde{\boldsymbol{\Phi}}^{(r,d)}_{k|m} \tag{30}$$

such that then for the individual local factors we have

$$\mathbf{U}^{(r,d)}_{k|m} \leftarrow \mathbf{U}^{(r,d)}_{k|m} - \mu_{\text{f}} \delta \mathbf{A}^{(r,d)}_{k|m} \tag{31}$$

and for the shared factors

$$\mathbf{B}_k^{(r,d)} \leftarrow \mathbf{B}_k^{(r,d)} - \mu_{\text{f}} \sum_{i=1}^{M_k} \delta \mathbf{A}_{k|i}^{(r,d)}. \tag{32}$$

We define the step size for the core update as μ_{c} and the dampening term as β_{c}. The update direction is then

$$\delta \boldsymbol{\mathcal{G}}_{k|m}^{(r)} = -\frac{e_{k|m}(n)}{\beta_{\text{c}} + \langle \boldsymbol{\mathcal{W}}_{k|m}^{(r)}, \tilde{\boldsymbol{\mathcal{W}}}_{k|m}^{(r)} \rangle} \tilde{\boldsymbol{\mathcal{W}}}_{k|m}^{(r)}, \tag{33}$$

and the update step

$$\boldsymbol{\mathcal{G}}_{k|m}^{(r)} \leftarrow \boldsymbol{\mathcal{G}}_{k|m}^{(r)} - \mu_{\text{c}} \delta \boldsymbol{\mathcal{G}}_{k|m}^{(r)}. \tag{34}$$

B. Backwards: Factor Penalty Update

The backwards step at every T_c-th iteration, after each node k exchanged its shared factors $\mathbf{B}_k^{(r,d)}$ with its neighbours according to the cluster edge set $\mathcal{E}^{(r,d)}$ and updates them as

$$\mathbf{A}_{k|m}^{(r,d)} \leftarrow \operatorname{prox}_g\big(\mathbf{A}_{k|m}^{(r,d)}\big) \tag{35}$$

The proximal step size is assumed to be 1, since the penalty weights already act in that manner. We use a first-order approximation based on a single Jacobi iteration [40] of the proximal operator of the Laplacian regularization term yielding

$$\operatorname{prox}_{\tau g}(\mathbf{A}_{k|m}) \approx \frac{\mathbf{A}_{k|m} + \gamma \mathbf{A}_{k|m,\text{old}} + \eta \bar{\mathbf{A}}_{k|m} + \lambda \sum_{l \neq k} w_{kl} \mathbf{B}_l}{1 + \gamma + \eta + \lambda d_k}, \tag{36}$$

where $d_k = \sum_j w_{kj}$ is the degree of node k in the graph $\mathcal{V}$. Note that we omitted the rank-mode superscripts for brevity and kept both shared and individual factors in one expression. For individual factors, the consensus term is zero, i.e., $\lambda = 0$, for shared factors, $\mathbf{A}_{k|m} = \mathbf{B}_k$. See Appendix A for a derivation of this approximation. The implicit update using the Jacobi approximation remains bounded, i.e., non-expansive [41, Ch. 4.5, Ch. 5.2-5.3] [42] ensuring $\|\operatorname{prox}_{\tau g}(\mathbf{A})\|_F \leq \|\mathbf{A}\|_F$ for any matrices $\mathbf{A}$ and penalty parameters chosen such that

$$\frac{\lambda d_i}{1 + \gamma + \eta + \lambda d_i} \leq c_{\max} < 1, \tag{37}$$

where $c_{\max}$ is a safeguard constant less than 1.

C. Backwards: Orthonormalization & Core Realignment

At regular intervals, each T_c-th iteration, the factors are re-orthogonalized by the proximal operator of the Stiefel manifold constraint, see Appendix B. The factors are re-orthogonalized as

$$\mathbf{A}_{k|m}^{(r,d)} \leftarrow \operatorname{prox}_{\tau h}\big(\mathbf{A}_{k|m}^{(r,d)}\big) \tag{38}$$

and let the right normalizer be such that

$$\mathbf{A}_{k|m}^{(r,d)} = \operatorname{prox}_{\tau h}\big(\mathbf{A}_{k|m}^{(r,d)}\big) \mathbf{R}_{k|m}^{(r,d)} \tag{39}$$

which can be computed by an orthogonalization procedure, i.e., using a Polar decomposition, or an approximation [40].

Finally, the cores are realigned to the re-orthogonalized factors using the right-normalizer

$$\boldsymbol{\mathcal{G}}_{k|m}^{(r)} \leftarrow \boldsymbol{\mathcal{G}}_{k|m}^{(r)} \times_d \mathbf{R}_{k|m}^{(r,d)\top}. \tag{40}$$

Algorithm 1 explains the order of operations in this procedure.

Algorithm 1 Online BTD with Core Pushback

1: **Input:** Data tensor stream $\boldsymbol{\mathcal{X}}(n)$, decomposition parameters $R, K_1, \ldots, K_D$, common factors $(r,d) \in \mathcal{C}$, forgetting factor λ, step size μ

2: **Initialize:**

IV. COMPUTATIONAL COMPLEXITY

We count real floating point operations (FLOPs) per hop, i.e., for each time iteration. Let us repeat definitions, so let there be N nodes, each with M_k channels, and therefore a total of $M = \sum_{k=1}^N M_k$ channels. The filter length is $L = \prod_{d=1}^D I_d$ and the core dimensions are K_d for each mode $d = 1, \ldots, D$. Let us define $K_{\text{sum}} = \sum_{d=1}^D K_d$ and $K_{\text{prod}} = \prod_{d=1}^D K_d$. A multi-mode product has an approximate cost of $\approx LK_{\text{sum}}$, due to the sequential reduction of the tensor dimensions. A single core inner product costs K_{prod}. A thin QR decomposition (QR) of an $I_d \times K_d$ matrix has a cost of $\approx 4 I_d K_d^2$. Here, $R_{\text{sh}}^{(d)}$ and $R_{\text{ind}}^{(d)}$ are the number of shared and individual factors for mode d, respectively, such that $R = R_{\text{sh}}^{(d)} + R_{\text{ind}}^{(d)}$. Further, $\Delta^{(d)}$ is the average node degree of the cluster graph for mode d. The constants $c_{\mathbf{A}}^{(d)} \in [5, 10]$ and $c_{\boldsymbol{\mathcal{G}}} \in [3, 6]$ are small constants. One iteration of the proposed algorithm consists of: (i) Forward projection and contractions, (ii) Factor LM updates, (iii) Adapt-then-combine (ACT) mixing, (iv) Stiefel projection and core realignment, (v) Core LM update. The concrete number of FLOPs per step is summarized in Table 1:

Step	Cost [FLOPs]
(i)	$MR\left(LDK_{\text{sum}} + \left(\sum_d I_d\right) K_{\text{prod}}\right)$
(ii)	$MR \sum_d c_{\mathbf{A}}^{(d)} I_d K_d$
(iii)	$\sum_d R_{\text{sh}}^{(d)} \Delta^{(d)} I_d K_d$
(iv)	$\sum_d \left(MR_{\text{ind}}^{(d)} + NR_{\text{sh}}^{(d)}\right) I_d K_d^2 + MRK_{\text{prod}} K_{\text{sum}}$
(v)	$MRc_{\boldsymbol{\mathcal{G}}} K_{\text{prod}}$

TABLE 1. Computational Complexity per Iteration

Memory consumption is $\mathcal{O}(MRK_{\text{prod}} + NR \sum_d R_{\text{sh}}^{(d)} I_d K_d + M \sum_d R_{\text{ind}}^{(d)} I_d K_d)$, where the first term is for the cores, the second for the shared factors, and the last for the individual factors. Communication per ACT step is $\sum_d R_{\text{sh}}^{(d)} \Delta^{(d)} I_d K_d$ scalars. The overall dominant scaling can be stated as $\mathcal{O}(MRLDK_{\text{sum}})$, which is linear in the number of channels, filter length, and decomposition rank. Further, the complexity numbers are upper bounds, as in practice, intermediate variables can be reused such as certain projections used in prediction and update steps as well as for shared factors.

Further, the diffusion step as well as the orthogonalization and core realignment steps are only performed at regular intervals, which reduces the average complexity per iteration.

Taking a full-length FIR LM/recursive least squares (RLS) algorithm, which in general has a cost of $\mathcal{O}(ML^2)$ per iteration, we can see the scaling advantage of the proposed method. In the comparison with a typical frequency-domain normalized least mean squares (NLMS) algorithm with complexity of $\mathcal{O}(ML \log L)$, the computational advantage depends on the chosen rank and number of modes, generally however, frequency-domain adaptive filtering has lower per-hop complexity. The advantage of the proposed method is the significantly reduced number of parameters to estimate, which can lead to faster convergence to useful estimates and lower memory requirements. **??** shows the complexity comparison for different filter lengths and decomposition ranks.

V. NUMERICAL EXPERIMENTS

To validate the proposed method, we perform numerical experiments. We simulate networks of nodes in a reverberant room environment. Each node has one or multiple microphones, and the graph connectivity $\mathcal{V}^{(r,d)}$ is defined a priori based on clustering of subspace similarity of BTDs estimated from true RIRs. This spectral clustering approach was introduced in [**?**]. We compare different network sizes, number of microphones per node, decomposition parameters, and choices of penalty parameters.

VI. CONCLUSION

We proposed a novel adaptive algorithm for efficient estimation of multichannel RIRs based on the BTD. We first show how to formulate the convolution model using BTD and then derive an adaptive algorithm based on three-operator splitting to solve the resulting optimization problem in a distributed manner. The algorithm is further extended to incorporate consensus regularization terms to promote agreement among nodes in a WASN.

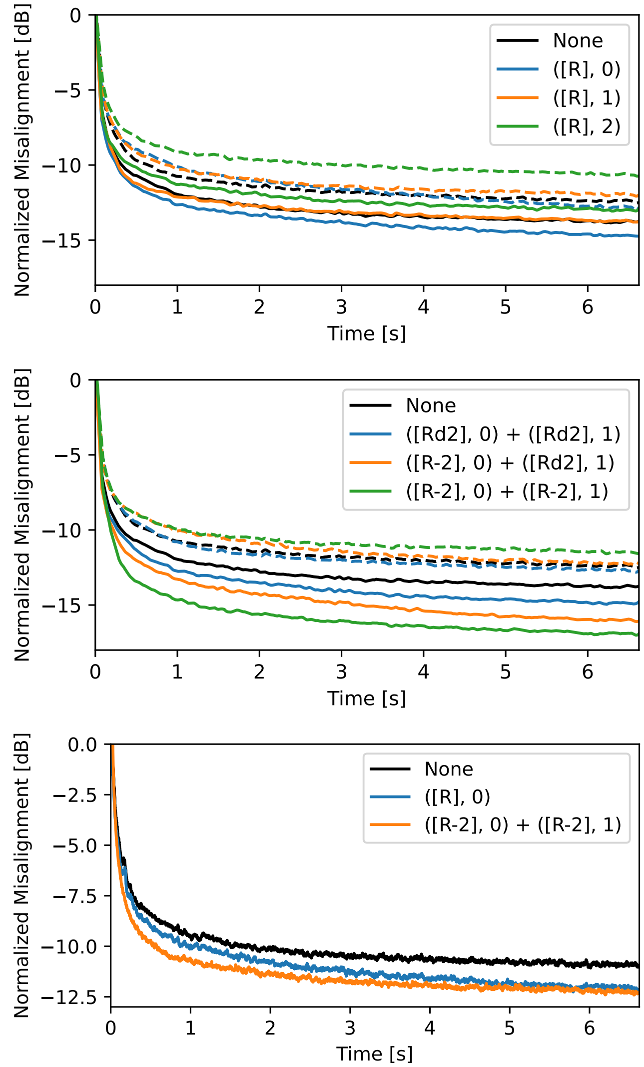

FIGURE 2. Preliminary results for different choices of common factors. $D = 3$, $l_h = 512$. It is chosen to have the same compression ratio for all choices of common factors. (Top) All factors of a mode are considered common. (Middle) Combination of common factors across modes. (Bottom) Comparison of the best choice of single-mode common factors and the best combination of modes.

APPENDIX

A. Proximal Operator of g

The proximal operator of the penalty term is defined as [32]

$$\mathrm{prox}_g\big(\mathbf{Y}\big) = \arg\min_{\mathbf{X}} \left\{ g(\mathbf{X}) + \frac{1}{2}\|\mathbf{X} - \mathbf{Y}\|_F^2 \right\}, \quad (41)$$

with $g(\boldsymbol{\Theta})$ as defined in (17). Here we assume that $\mathbf{X}, \mathbf{Y}$ are stacked versions of the relevant factors $\mathbf{X}_i, \mathbf{Y}_i$, either indexed by node or channel. Let us drop the superscript (r, d) for readability, the following holds for all. The exact proximal operator solves

$$\big[(1 + \gamma + \eta)\mathbf{I} + \lambda\mathbf{L}\big]\mathbf{X} = \mathbf{Y} + \gamma\mathbf{X}(n-1) + \eta\bar{\mathbf{X}}, \quad (42)$$

with

$$\mathbf{L} = (\mathbf{D} - \mathbf{W}) \otimes \mathbf{I}_{I_d}, \quad (43)$$

being the Laplacian, $\mathbf{D}$ is the degree matrix with diagonal elements $d_i = \sum_j w_{ij}$ and $\mathbf{W}$ is the (weighted) adjacency matrix for the graph $\mathcal{V}$ with elements $w_{ij} \geq 0$. For each mode d, the Kronecker product with the identity matrix expands the Laplacian to the size of the factors. The proximal

$$\mathbf{X} = \mathrm{prox}_g\big(\mathbf{Y}\big) \quad (44)$$

can be solved using conjugate gradient (CG), or approximated using a single Jacobi iteration [40]. For an individual factor $\mathbf{X}_i$, this gives the update

$$\mathbf{X}_i \leftarrow \frac{\mathbf{Y}_i + \gamma\mathbf{X}_i(n-1) + \eta\bar{\mathbf{X}}_i + \lambda\sum_{j\neq i} w_{ij}\mathbf{X}_j}{1 + \gamma + \eta + \lambda d_i}. \quad (45)$$

The exact proximal operator is firmly non-expansive since the smallest eigenvalue of $(1+\gamma+\eta)\mathbf{I} \leq 1+\gamma\eta$. The one-step Jacobi map is a row-stochastic averaging and a constant shift, therefore it is non-expansive under the following safeguard on the step size: With $\gamma, \eta < 1$, choose λ such that for every node i

$$\frac{\lambda d_i}{1+\gamma+\eta+\lambda d_i} < c_{\max} < 1, \tag{46}$$

which represents the mixing coefficient, with $c_{\max}$ being a constant smaller than 1.

B. Proximal Operator of the Stiefel Manifold Indicator

The proximal operator of the indicator function ι_{St} on the Stiefel set $\mathrm{St}(m,n) = \{\mathbf{X} \in \mathbb{R}^{m\times n} \mid \mathbf{X}^\top \mathbf{X} = \mathbf{I}\}$ is defined as [32]

$$\operatorname{prox}_{\iota_{\mathrm{St}}}(\mathbf{Y}) = \arg\min_{\mathbf{X}} \left\{ \iota_{\mathrm{St}}(\mathbf{X}) + \frac{1}{2}\|\mathbf{X}-\mathbf{Y}\|_F^2 \right\}, \tag{47}$$

where $\iota_{\mathrm{St}}(\mathbf{X}) = 0$ if $\mathbf{X}^\top\mathbf{X} = \mathbf{I}$ and ∞ else. With $\mathbf{Y} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$, the optimal solution is

$$\operatorname{prox}_{\iota_{\mathrm{St}}}(\mathbf{Y}) = \mathbf{U}\mathbf{V}^\top = \mathbf{Y}(\mathbf{Y}^\top\mathbf{Y})^{-1/2} \tag{48}$$

which is the left factors of the polar decomposition of $\mathbf{Y}$ [43], [40] or also done using the SVD. Using the QR decomposition instead of the Polar decomposition or SVD yields a retraction onto the Stiefel manifold as well, however, it is not the closest-point projection and therefore not the exact proximal operator.

REFERENCES

[1] P. A. Naylor and N. D. Gaubitch, Eds., *Speech Dereverberation*, ser. Signals and Communication Technology. Springer Science & Business Media, 2010.

[2] S. Gannot, E. Vincent, S. Markovich-Golan, and A. Ozerov, "A Consolidated Perspective on Multimicrophone Speech Enhancement and Source Separation," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 25, no. 4, pp. 692–730, Apr. 2017.

[3] T. van Waterschoot, "Design and evaluation of digital signal processing algorithms for acoustic feedback and echo cancellation," *J. Acoust. Soc. Am.*, vol. 126, no. 6, pp. 3373–3373, Dec. 2009.

[4] J. S. Bradley, "Predictors of speech intelligibility in rooms," *J. Acoust. Soc. Am.*, vol. 80, no. 3, pp. 837–845, Sep. 1986.

[5] Y. E. Baba, A. Walther, and E. A. P. Habets, "3D Room Geometry Inference Based on Room Impulse Response Stacks," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 26, no. 5, pp. 857–872, May 2018.

[6] K. MacWilliam, F. Elvander, and T. van Waterschoot, "Simultaneous Acoustic Echo Sorting and 3-D Room Geometry Inference," in *Proc. 2023 IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP '23)*. Rhodes Island, Greece: IEEE, 2023, pp. 1–5.

[7] C. Evers, H. W. Löllmann, H. Mellmann, A. Schmidt, H. Barfuss, P. A. Naylor, and W. Kellermann, "The LOCATA Challenge: Acoustic Source Localization and Tracking," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 28, pp. 1620–1643, 2020.

[8] Y. Avargel and I. Cohen, "System Identification in the Short-Time Fourier Transform Domain With Crossband Filtering," *IEEE Trans. Audio Speech Lang. Process.*, vol. 15, no. 4, pp. 1305–1319, May 2007.

[9] M. Jalmby, F. Elvander, and T. van Waterschoot, "Low-Rank Room Impulse Response Estimation," *IEEE/ACM Trans. Audio Speech Lang. Process.*, pp. 1–13, 2023.

[10] T. G. Kolda and B. W. Bader, "Tensor Decompositions and Applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, Aug. 2009.

[11] M. Jälmby, F. Elvander, and T. van Waterschoot, "Multi-Channel Low-Rank Convolution of Jointly Compressed Room Impulse Responses," *IEEE Open J. Signal Process.*, vol. 5, pp. 850–857, 2024.

[12] L. De Lathauwer and D. Nion, "Decompositions of a Higher-Order Tensor in Block Terms—Part III: Alternating Least Squares Algorithms," *SIAM J. Matrix Anal. & Appl.*, vol. 30, no. 3, pp. 1067–1083, Jan. 2008.

[13] L. De Lathauwer, "Decompositions of a Higher-Order Tensor in Block Terms—Part II: Definitions and Uniqueness," *SIAM J. Matrix Anal. & Appl.*, vol. 30, no. 3, pp. 1033–1066, Jan. 2008.

[14] ——, "Decompositions of a Higher-Order Tensor in Block Terms—Part I: Lemmas for Partitioned Matrices," *SIAM J. Matrix Anal. & Appl.*, vol. 30, no. 3, pp. 1022–1032, Jan. 2008.

[15] M. Jälmby, F. Elvander, and T. van Waterschoot, "Low-Rank Tensor Modeling of Room Impulse Responses," in *Proc. 29th European Signal Process. Conf. (EUSIPCO '21)*, Aug. 2021, pp. 111–115.

[16] ——, "Compression of room impulse responses for compact storage and fast low-latency convolution," *EURASIP J. Audio, Speech, Music Process.*, vol. 2024, no. 1, p. 45, Sep. 2024.

[17] G. Huang, J. Benesty, and J. Chen, "Dimensionality Reduction of Room Acoustic Impulse Responses and Applications to System Identification," *IEEE Signal Process. Lett.*, vol. 30, pp. 1107–1111, 2023.

[18] M. Maroti, G. Simon, A. Ledeczi, and J. Sztipanovits, "Shooter localization in urban terrain," *Computer*, vol. 37, no. 8, pp. 60–61, 2004.

[19] W.-P. Chen, J. Hou, and L. Sha, "Dynamic clustering for acoustic target tracking in wireless sensor networks," *IEEE Trans. Mob. Comput.*, vol. 3, no. 3, pp. 258–271, 2004.

[20] A. Bertrand, "Applications and trends in wireless acoustic sensor networks: A signal processing perspective," in *Proc. 18th IEEE Symp. Commun. Vehicular Technol. Benelux (SCVT '11)*, Nov. 2011, pp. 1–6.

[21] A. Bertrand and M. Moonen, "Distributed Adaptive Node-Specific Signal Estimation in Fully Connected Sensor Networks—Part II: Simultaneous and Asynchronous Node Updating," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5292–5306, Oct. 2010.

[22] ——, "Distributed Adaptive Node-Specific Signal Estimation in Fully Connected Sensor Networks—Part I: Sequential Node Updating," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5277–5291, Oct. 2010.

[23] M. Ferrer, M. de Diego, G. Piñero, and A. Gonzalez, "Affine projection algorithm over acoustic sensor networks for active noise control," *IEEE Trans. Audio Speech Lang. Process.*, vol. 29, pp. 448–461, 2021.

[24] S. Ruiz, T. van Waterschoot, and M. Moonen, "Distributed combined acoustic echo cancellation and noise reduction using GEVD-based distributed adaptive node specific signal estimation with prior knowledge," in *Proc. 28th European Signal Process. Conf. (EUSIPCO '20)*. Amsterdam, Netherlands: IEEE, Jan. 2021, pp. 206–210.

[25] Y. Zeng and R. C. Hendriks, "Distributed Delay and Sum Beamformer for Speech Enhancement via Randomized Gossip," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 22, no. 1, pp. 260–273, Jan. 2014.

[26] G. Huang, J. Benesty, J. Chen, C. Paleologu, S. Ciochină, W. Kellermann, and I. Cohen, "Acoustic System Identification with Partially Time-Varying Models Based on Tensor Decompositions," in *Proc. Intl. Workshop Acoust. Echo Noise Control (IWAENC '22)*, 2022, pp. 1–5.

[27] T. Hardenbicker and P. Jax, "Online System Identification on Learned Acoustic Manifolds Using an Extended Kalman Filter," in *Proc, 2024 Int, Workshop Acoust. Signal Enhanc. (IWAENC '24)*, Sep. 2024, pp. 339–343.

[28] D. Pandey, A. Venugopal, and H. Leib, "Linear to multi-linear algebra and systems using tensors," Dec. 2023.

[29] T. van Waterschoot and M. Moonen, "Distributed estimation and equalization of room acoustics in a wireless acoustic sensor network," in *Proc. 20th European Signal Process. Conf. (EUSIPCO '12)*, Aug. 2012, pp. 2709–2713.

[30] D. Davis and W. Yin, "A three-operator splitting scheme and its optimization applications," *Set-valued var. anal.*, vol. 25, no. 4, pp. 829–858, 2017.

[31] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK ; New York: Cambridge University Press, 2004.

[32] N. Parikh and S. Boyd, "Proximal Algorithms," *Found. Trends Optim*, vol. 1, no. 3, pp. 127–239, Jan. 2014.

[33] R. T. Rockafellar, "Monotone Operators and the Proximal Point Algorithm," *SIAM J. Control Optim.*, vol. 14, no. 5, pp. 877–898, Aug. 1976.

[34] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction

method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.
[35] S. Shalev-Shwartz, "Online Learning and Online Convex Optimization," *FNT in Machine Learning*, vol. 4, no. 2, pp. 107–194, 2011.
[36] J. C. Butcher, *Numerical Methods for Ordinary Differential Equations*, 1st ed. Wiley, Mar. 2008.
[37] L. Xiao, S. Boyd, and S. Lall, "Distributed Average Consensus with Time-Varying Metropolis Weights," *Automatica*, vol. 1, pp. 1–4, 2006.
[38] C. Navasca, L. De Lathauwer, and S. Kindermann, "Swamp reducing technique for tensor decomposition," in *Proc. 16th European Signal Process. Conf. (EUSIPCO '08)*, Aug. 2008, pp. 1–5.
[39] D. W. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *J. Soc. Ind. Appl. Math.*, vol. 11, no. 2, pp. 431–441, Jun. 1963.
[40] G. H. Golub and C. F. Van Loan, *Matrix Computations*, fourth edition ed., ser. Johns Hopkins Studies in the Mathematical Sciences. Baltimore: The Johns Hopkins University Press, 2013.
[41] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. New York: Springer, 2011.
[42] P. L. Combettes, "Solving monotone inclusions via compositions of nonexpansive averaged operators," *Optimization*, 2004.
[43] N. J. Higham, *Functions of Matrices: Theory and Computation*. Society for Industrial and Applied Mathematics, Jan. 2008.

SECOND B. AUTHOR, photograph and biography not available at the time of publication.

THIRD C. AUTHOR JR. (Member, IEEE), photograph and biography not available at the time of publication.