Tensor-Based Blind Structured Channel Estimation for Multichannel Systems

Abdulmajid Lawal*, Karim Abed-Meraim†, Azzedine Zerguine*, and Thanh Trung Le‡,
* Electrical Engineering Department and the Center for Communication Systems and Sensing, King Fahd University of Petroleum and Minerals, Dhahran, 31261 Saudi Arabia
† Université d'Orléans, Laboratoire PRISME/IRAuS, Orléans, France
‡ VNU University of Engineering and Technology, Hanoi, Vietnam

***Abstract*—In this work, we propose a blind channel estimation method in multiple-input multiple-output systems using a tensor decomposition approach. Consequently, fundamental link between convolutive channels and block-term decomposition (BTD) is established. The proposed approach leverages the second-order statistics of the received signals to construct a third-order tensor by stacking covariance matrices at different time lags. Then, the channel estimation process consists of two stages. In the first stage, the tensor is decoupled using the type-2 BTD technique to extract the loading factors. In the second stage, a Toeplitz constraint is imposed on the loading factors to obtain the channel matrix. The loading factors are constraint to be identical and have a Toeplitz structure. The numerical simulations show the effectiveness of the proposed method.**

***Index Terms*—Block-term decomposition, Toeplitz structure, Blind, Constraint, Channel estimation.**

I. Introduction

In recent years, tensor decomposition has emerged as a transformative tool for multidimensional data analysis across diverse fields, fueled by advancements in computational resources [1], [2]. This method, which factorizes multidimensional arrays (tensors) into fundamental components, has shown exceptional promise in addressing complex challenges within communication, signal processing, and artificial intelligence, among others [3]–[6]. Of particular relevance is its application in blind communication systems, where numerous tensor decomposition techniques have been tailored for enhanced performance. The canonical polyadic (CP) decomposition, for example, has become foundational in tensor-based blind algorithms [7], while the higher-order singular value decomposition (HOSVD) approach has also proven effective in solving a range of blind communication problems [8], [9]. These methods underscore the critical role of tensor decomposition in advancing robust and scalable solutions for multidimensional data processing.

Current tensor-based decomposition methods are primarily designed for handling instantaneous blind communication systems, often constrained to specific applications, and are less suited for convolutive systems. Consequently, convolutive mixtures frequently need conversion to instantaneous forms via frequency domain representation or transformations before applying existing tensor-based techniques. However, this approach has significant drawbacks, including inconsistencies in filter coefficients across frequencies, as well as scaling and permutation indeterminacies. Addressing these challenges necessitates the development of advanced tensor-based methods capable of directly processing convolutive mixtures, offering robust solutions free from these inherent limitations.

In this work, we propose a novel framework leveraging tensor decomposition to exploit the convolutive characteristics of communication systems for blind channel estimation. The proposed framework consists of two stages. In the first stage, type-2 BTD is applied to the tensor to extract the loading factors, which are constrained to be identical. In the second stage, a Toeplitz structure constraint is imposed on the loading factors to formulate a cost function, which is minimized to obtain the desired channel matrix.

A third-order tensor is constructed by stacking covariance matrices of the received signals, allowing BTD to estimate the block Toeplitz matrix that encapsulates the desired channel information. This approach advances the tensor decomposition literature by introducing an optimized, effective solution for channel estimation in convolutive systems, and eventually demonstrating both efficiency and robustness.

Notations: scalars, vectors, matrices and tensors are represented by lowercase letter x, bold lowercase letter $\mathbf{x}$, bold capital letter $\mathbf{X}$ and bold calligraphic letter $\boldsymbol{\mathcal{X}}$, respectively. $\boldsymbol{\mathcal{X}}_{(n)}$ represents the mode-n matricization of a tensor $\boldsymbol{\mathcal{X}}_{(n)}$. Transpose, conjugate transpose, pseudo-inverse and conjugate operation are, respectively, donated by $(\cdot)^\top$, $(\cdot)^H$, $(\cdot)^\dagger$, and $(\cdot)^*$. The Kronecker product, block Kronecker product, and convolution of matrices $\mathbf{H}$ and $\mathbf{Q}$ are represented by $\mathbf{H} \otimes \mathbf{Q}$, $\mathbf{H} \otimes_b \mathbf{Q}$, and $\mathbf{H} \circledast \mathbf{Q}$, respectively. The function $\mathrm{blkdiag}(\cdot)$ is used to construct a block diagonal matrix. $\|\cdot\|_2$ and $\|\cdot\|_\mathrm{F}$ denote euclidean norm and Frobenius norm, respectively. An identity matrix of size $N \times N$ is represented as $\boldsymbol{I}_N$. The mode-n product between a tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ and a matrix $\mathbf{H} \in \mathbb{C}^{J_n \times I_n}$ produces a new tensor $\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{X}} \times_n \mathbf{H} \in \mathbb{C}^{I_1 \times \cdots \times J_n \times \cdots \times I_N}$. Finally, $\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{X}} \times_n \mathbf{H}$ then $[\boldsymbol{\mathcal{Y}}]_{(n)} = \mathbf{H}[\boldsymbol{\mathcal{X}}]_{(n)}$.

II. The Type-2 BTD

For comprehension and reader's convenience, the type-2 BTD [10], [11] variant of the block term decomposition (BTD) [12] is briefly presented. The main aim of the type-2 BTD is to factorize a third-order tensor $\boldsymbol{\mathcal{Y}} \in \mathbb{C}^{I \times J \times E}$ into a set of multilinear-rank components $\{\boldsymbol{\mathcal{Y}}_r\}_{r=1}^R$, which can be expressed as follows:

$$\boldsymbol{\mathcal{Y}} = \sum_{r=1}^{R} \boldsymbol{\mathcal{Y}}_r = \sum_{r=1}^{R} \boldsymbol{\mathcal{G}}_r \times_1 \mathbf{H}_r \times_2 \mathbf{Q}_r, \tag{1}$$

where $\boldsymbol{\mathcal{G}}_r \in \mathbb{C}^{L_r \times M_r \times E}$ denotes the core tensor of the r^{th} component of $\boldsymbol{\mathcal{Y}}_r$ and the loading factors $\mathbf{H}_r \in \mathbb{C}^{I \times L_r}$ and $\mathbf{Q}_r \in \mathbb{C}^{J \times M_r}$ are all full column rank matrices. Since (1) is trilinear in $\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_R]$, $\mathbf{Q} = [\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_R]$, and $\boldsymbol{\mathcal{G}} = \mathrm{blkdiag}(\boldsymbol{\mathcal{G}}_1, \boldsymbol{\mathcal{G}}_2, \ldots, \boldsymbol{\mathcal{G}}_R)$, its computation follows the standard alternating least-squares (ALS) method [13]. Additionally, the type-2 BTD is essentially unique under certain mild conditions [12]. To support our algorithm development, we present three mode-n matrix representations of $\boldsymbol{\mathcal{Y}}$.

$$\boldsymbol{\mathcal{Y}}_{(1)} = \mathbf{H}\left[[\boldsymbol{\mathcal{G}}_1 \times_2 \mathbf{Q}_1]_{(1)}^\top, \ldots, [\boldsymbol{\mathcal{G}}_R \times_2 \mathbf{Q}_R]_{(1)}^\top\right]^\top, \tag{2}$$

$$\boldsymbol{\mathcal{Y}}_{(2)} = \mathbf{Q}\left[[\boldsymbol{\mathcal{G}}_1 \times_1 \mathbf{H}_1]_{(2)}^\top, \ldots, [\boldsymbol{\mathcal{G}}_R \times_1 \mathbf{H}_R]_{(2)}^\top\right]^\top, \tag{3}$$

$$\boldsymbol{\mathcal{Y}}_{(3)} = \left[[\boldsymbol{\mathcal{G}}_1]_{(3)}, \ldots, [\boldsymbol{\mathcal{G}}_R]_{(3)}\right](\mathbf{H} \otimes_b \mathbf{Q})^\top. \tag{4}$$

III. The System Model

Let us consider a multiple-input multiple-output communication system equipped with K number of transmitting and M number of receiving antennas. Assuming that at a time instance t, transmitted signal is $\mathbf{x}(t) = [x_1(t), \ldots, x_K(t)]^\top$ through an unknown channel $\bar{\mathbf{H}}(t) \in \mathbb{C}^{M \times K}$, in the presence of noise $\mathbf{n}(t) \in \mathbb{C}^M$. The received signal $\mathbf{y}(t) \in \mathbb{C}^M$ can be expressed as follows:

$$\begin{aligned} \mathbf{y}(t) &= \bar{\mathbf{H}}(t) \circledast \mathbf{x}(t) + \mathbf{n}(t), t = 0, \ldots, T-1 \\ &= \sum_{l=0}^{L} \bar{\mathbf{H}}(l)\mathbf{x}(t-l) + \mathbf{n}(t), \end{aligned} \tag{5}$$

where L represents the channel order, and T denotes the total sample size and

$$\bar{\mathbf{H}}(l) = \begin{bmatrix} h_{11}(l) & \ldots & h_{1K}(l) \\ \vdots & & \vdots \\ h_{M1}(l) & \ldots & h_{MK}(l) \end{bmatrix}.$$

If we stack N_w elements of $\mathbf{y}(t)$ into a vector $\mathbf{y}_{N_w}(t)$, the model modifies to

$$\mathbf{y}_{N_w}(t) = \mathbf{H}\mathbf{x}_{N_w}(t) + \mathbf{n}_{N_w}(t), \tag{6}$$

where $\mathbf{y}_{N_w}(t) = [\mathbf{y}^\top(t), \ldots, \mathbf{y}^\top(t - N_w + 1)]^\top \in \mathbb{C}^{MN_w \times 1}$ is the stacked received signal vector, $\mathbf{x}_{N_w}(t) = [\mathbf{x}^\top(t), \ldots, \mathbf{x}^\top(t - L - N_w + 1)]^\top \in \mathbb{C}^{P \times 1}$, $P = KR$ and $R = L + N_w$, and the corresponding noise vector is $\mathbf{n}_{N_w}(t) = [\mathbf{n}^\top(t), \ldots, \mathbf{n}^\top(t - N_w + 1)]^\top \in \mathbb{C}^{MN_w \times 1}$. The channel matrix $\mathbf{H} \in \mathbb{C}^{MN_w \times P}$ is given as

$$\begin{aligned} \mathbf{H} &= \begin{bmatrix} \bar{\mathbf{H}}(0) & \cdots & \bar{\mathbf{H}}(L) & & \mathbf{0} \\ & \ddots & & \ddots & \\ \mathbf{0} & & \bar{\mathbf{H}}(0) & \cdots & \bar{\mathbf{H}}(L) \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{H}_1 & \cdots & \mathbf{H}_{L+1} & \cdots & \mathbf{H}_R \end{bmatrix} \end{aligned} \tag{7}$$

Here, we assumed that the individual input signals are temporary coherent while maintaining their mutual independence [14]. In essence, the correlation between two inputs $x_i(t)$ and $x_j(t)$, for $i \neq j$, follows $\mathbb{E}[x_i(t)x_j(t-\tau)^*] = 0 \ \forall \tau$. Therefore, the correlation matrix that corresponds to (6) can be represented as follows:

$$\begin{aligned} \mathbf{R}_{\mathbf{y}_{N_w}}(t, \tau) &\triangleq \mathbb{E}[\mathbf{y}_{N_w}(t)\mathbf{y}_{N_w}^H(t-\tau)] \\ &= \mathbf{H}\mathbb{E}[\mathbf{x}_{N_w}(t)\mathbf{x}_{N_w}^H(t-\tau)]\mathbf{H}^H \\ &= \mathbf{H}\begin{bmatrix} \mathbf{R}_{\bar{\mathbf{x}}_1}(t,\tau) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{R}_{\bar{\mathbf{x}}_R}(t,\tau) \end{bmatrix}\mathbf{H}^H \\ &\triangleq \mathbf{H}\,\mathrm{blkdiag}(\{\mathbf{R}_{\bar{\mathbf{x}}_r}(t,\tau)\}_{r=1}^R)\mathbf{H}^H, \end{aligned} \tag{8}$$

where $\mathbf{x}_{N_w}(t) = [\mathbf{x}^\top(t), \ldots, \mathbf{x}^\top(t - L - N_w + 1)]^\top \triangleq [\bar{\mathbf{x}}_1^\top(t), \ldots, \bar{\mathbf{x}}_R^\top(t)]^\top$, hence the subscript in (8). Now, if we consider multiple time lags, $\{\tau_j\}_{j=1}^J$, where $J = T - N_w + 1$, we obtain the following:

$$\begin{aligned} \mathbf{R}_{\mathbf{y}_{N_w}}(t, \tau_1) &= \mathbf{H}\,\mathrm{blkdiag}(\{\mathbf{R}_{\bar{\mathbf{x}}_r}(t,\tau_1)\}_{r=1}^R)\mathbf{H}^H \\ &\vdots \\ \mathbf{R}_{\mathbf{y}_{N_w}}(t, \tau_J) &= \mathbf{H}\,\mathrm{blkdiag}(\{\mathbf{R}_{\bar{\mathbf{x}}_r}(t,\tau_J)\}_{r=1}^R)\mathbf{H}^H. \end{aligned} \tag{9}$$

By consecutively stacking $\{\mathbf{R}_{\mathbf{y}_{N_w}}(t, \tau_j)\}_{j=1}^J$ and $\{\mathbf{R}_{\mathbf{x}_{N_w}}(t, \tau_j)\}_{j=1}^J$ along the third mode of $\boldsymbol{\mathcal{R}} \in \mathbb{C}^{MN_w \times MN_w \times J}$ and $\boldsymbol{\mathcal{G}} \in \mathbb{C}^{P \times P \times J}$ we obtain the following representation:

$$\boldsymbol{\mathcal{R}} = \boldsymbol{\mathcal{G}} \times_1 \mathbf{H} \times_2 \mathbf{H} = \sum_{r=1}^{R} \boldsymbol{\mathcal{G}}_{\bar{\mathbf{x}}_r} \times_1 \mathbf{H}_r \times_2 \mathbf{H}_r \tag{10}$$

where $\boldsymbol{\mathcal{G}}_{\bar{\mathbf{x}}_r}(:,:,j) = \mathbf{R}_{\bar{\mathbf{x}}_r}(t, \tau_j)$ and $\mathbf{H}_r$ represents the r^{th} block column of $\mathbf{H}$ as implicitly defined in (7).

Equation (10) illustrates a modified form of type-2 Block Term Decomposition as presented in (1), where the factor matrices $\mathbf{H}$ and $\mathbf{Q}$ are enforced to be identical. Notably, under certain mild conditions, (10) is effectively unique, if $MN_w \geq P$ and $J \geq 3$. Specifically, the estimate $\tilde{\mathbf{H}}$ of $\mathbf{H}$ is unique up to basic ambiguities, meaning $\tilde{\mathbf{H}} = \mathbf{H}\boldsymbol{\Pi}\boldsymbol{\Lambda}$ where $\boldsymbol{\Pi}$ represents a block permutation matrix, and $\boldsymbol{\Lambda}$ is a square nonsingular block-diagonal matrix.

IV. Proposed Optimization Framework

This section introduces the proposed two-stage channel estimation framework. First, the type-2 BTD optimization is applied to extract loading factors, followed by a Toeplitz constraint cost function for blind channel estimation as follows:

$$\arg\min_{\boldsymbol{\mathcal{G}},\mathbf{H},\mathbf{Q}} \left\| \boldsymbol{\mathcal{R}} - \sum_{r=1}^{R} \boldsymbol{\mathcal{G}}_{\bar{\mathbf{x}}_r} \times_1 \mathbf{H}_r \times_2 \mathbf{Q}_r \right\|_{\mathrm{F}}^2 \ s.t.\ \mathbf{H} = \mathbf{Q}, \tag{11}$$

where $\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_R]$ and $\mathbf{Q} = [\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_R]$. To estimate the channel, the Lagrangian function is utilized to enforce minimization while also imposing a Toeplitz structure constraint, as follows:

$$\boldsymbol{\mathcal{L}}(\boldsymbol{\mathcal{G}},\mathbf{H},\mathbf{Q},\mathbf{U}) = \min_{\boldsymbol{\mathcal{G}},\mathbf{H},\mathbf{Q}} \left\|\boldsymbol{\mathcal{R}} - \sum_{r=1}^{R} \boldsymbol{\mathcal{G}}_{\bar{\mathbf{x}}_r} \times_1 \mathbf{H}_r \times_2 \mathbf{Q}_r\right\|_{\mathrm{F}}^2 + \frac{\rho}{2}\|\mathbf{H}-\mathbf{Q}+\mathbf{Q}\|_{\mathrm{F}}^2 - \frac{\rho}{2}\|\mathbf{U}\|_{\mathrm{F}}^2, \tag{12}$$

where $\mathbf{U}$ represents the scaled dual variable and $\rho > 0$ is a regularized parameter. The primary objective is to minimize the Lagrangian function with respect to $\mathbf{H}$, $\mathbf{Q}$, and $\boldsymbol{\mathcal{G}}$ while maintaining the dual variable $\mathbf{U}$. The closed form solution of the problem is presented as follows:

$$\mathbf{H} = \left(\boldsymbol{\mathcal{R}}_{(1)}\mathbf{Z}_H^H + \rho(\mathbf{Q}+\mathbf{U})\right)\left(\mathbf{Z}_H\mathbf{Z}_H^H + \rho\mathbf{I}_P\right)^{-1} \tag{13}$$

$$\mathbf{Q} = \left(\boldsymbol{\mathcal{R}}_{(2)}\mathbf{Z}_Q^H + \rho(\mathbf{Q}-\mathbf{U})\right)\left(\mathbf{Z}_Q\mathbf{Z}_Q^H + \rho\mathbf{I}_P\right)^{-1} \tag{14}$$

and

$$\left[[\boldsymbol{\mathcal{G}}_{\bar{\mathbf{x}}_1}]_{(3)},\ldots,[\boldsymbol{\mathcal{G}}_{\bar{\mathbf{x}}_R}]_{(3)}\right] = \boldsymbol{\mathcal{R}}_{(3)}\left((\mathbf{B}\otimes_b\mathbf{A})^\top\right)^\dagger \tag{15}$$

where $\mathbf{Z}_H$ and $\mathbf{Z}_Q$ are defined as

$$\mathbf{Z}_H = \left[[\boldsymbol{\mathcal{G}}_{\bar{\mathbf{x}}_1}\times_2\mathbf{Q}_1]_{(1)}^\top,\ldots,[\boldsymbol{\mathcal{G}}_{\bar{\mathbf{x}}_R}\times_2\mathbf{Q}_R]_{(1)}^\top\right]^\top \tag{16}$$

$$\mathbf{Z}_Q = \left[[\boldsymbol{\mathcal{G}}_{\bar{\mathbf{x}}_1}\times_1\mathbf{H}_1]_{(2)}^\top,\ldots,[\boldsymbol{\mathcal{G}}_{\bar{\mathbf{x}}_R}\times_1\mathbf{H}_R]_{(2)}^\top\right]^\top \tag{17}$$

The stationary point of the optimization function $\boldsymbol{\mathcal{L}}(\cdot)$ is found using an iterative procedure.

Here, $\mathbf{W}^{(i)}$ and β_i are auxiliary variables designed to enhance the convergence speed of the iterative procedure. The incorporation of Equations (18d) to (18f) accelerates the augmented Lagrangian approach, improving both computational efficiency and estimation accuracy. The stopping criteria is implemented based on the maximum iteration number, I_{max}, or the following criteria:

$$\|\mathbf{H}^{(i)}-\mathbf{Q}^{(i)}\|_{\mathrm{F}}^2 \le \alpha_{pre},\ \|\rho(\mathbf{H}^{(i)}-\mathbf{H}^{(i-1)})\|_{\mathrm{F}}^2 \le \alpha_{dual}, \tag{18}$$

where

$$\alpha_{pre} = \sigma_{abs}\sqrt{MKN_wR} + \sigma_{rel}\max\{\|\mathbf{H}^{(i)}\|_2, \|\mathbf{Q}^{(i)}\|_2\},$$

and

$$\alpha_{dual} = \sigma_{abs}\sqrt{MKN_wR} + \sigma_{rel}\max\{\|\rho\mathbf{U}^{(i)}\|_2\},$$

where $\sigma_{abs} > 0$ and $\sigma_{rel} > 0$ denotes the absolute and relative tolerance, respectively. In practice, the value of ρ can be set to unity, $\sigma_{abs} = 10^{-4}$ and $\sigma_{rel} = 10^{-2}$ for good performance. After obtaining the matrix $\mathbf{H}$ from the tensor decomposition, the channel estimate is obtained as $\hat{\mathbf{H}} = \mathbf{HW}$, where $\mathbf{W}$ is a matrix that enforces the Toeplitz structure. A cost function that exploit the Toeplitz structure to estimate the channel coefficient is formulated as given below:

$$\begin{aligned} J &= \|\mathbf{J}_1\hat{\mathbf{H}}\mathbf{J}_2 - \mathbf{J}_3\hat{\mathbf{H}}\mathbf{J}_4\|_{\mathrm{F}}^2 + \|\mathbf{J}_5\hat{\mathbf{H}}\mathbf{J}_6\|_{\mathrm{F}}^2 + \|\mathbf{J}_7\hat{\mathbf{H}}\mathbf{J}_8\|_{\mathrm{F}}^2 \\ &= \|\mathbf{J}_1\mathbf{HW}\mathbf{J}_2 - \mathbf{J}_3\mathbf{HW}\mathbf{J}_4\|_{\mathrm{F}}^2 + \|\mathbf{J}_5\mathbf{HW}\mathbf{J}_6\|_{\mathrm{F}}^2 \\ &+ \|\mathbf{J}_7\mathbf{HW}\mathbf{J}_8\|_{\mathrm{F}}^2 \end{aligned} \tag{19}$$

where $\mathbf{J}_1 = [\mathbf{I}_{MN_w-M}, \mathbf{0}_{MN_w-M,M}]$, $\mathbf{J}_2 = [\mathbf{I}_{P-K}, \mathbf{0}_{P-K,K}]^\top$, $\mathbf{J}_3 = [\mathbf{0}_{MN_w-M,M}, \mathbf{I}_{MN_w-M}]$,

Algorithm 1: The proposed optimization framework.

Initialize $\mathbf{H}^{(0)}$, $\mathbf{Q}^{(0)}$, $\mathbf{U}^{(0)}$, and $\boldsymbol{\mathcal{G}}^{(0)}$ at $i=0$

while *the current iteration is less than* I_{max} **do**

$$\mathbf{H}^{(i)} = \arg\min_{\mathbf{H}} \|\boldsymbol{\mathcal{R}} - \sum_{r=1}^{R}\boldsymbol{\mathcal{G}}_{\bar{\mathbf{x}}_r}^{(i-1)} \times_1 \mathbf{H}_r \times_2 \mathbf{Q}_r^{(i-1)}\|_{\mathrm{F}}^2 + \frac{\rho}{2}\|\mathbf{Q}^{(i-1)} - \mathbf{H} + \mathbf{U}^{(i-1)}\|_{\mathrm{F}}^2 \tag{18a}$$

$$\mathbf{Q}^{(i)} = \arg\min_{\mathbf{Q}} \|\boldsymbol{\mathcal{R}} - \sum_{r=1}^{R}\boldsymbol{\mathcal{G}}_{\bar{\mathbf{x}}_r}^{(i-1)} \times_1 \mathbf{H}_r^{(i)} \times_2 \mathbf{Q}_r\|_{\mathrm{F}}^2 + \frac{\rho}{2}\|\mathbf{Q}^{(i-1)} - \mathbf{H} + \mathbf{U}^{(i-1)}\|_{\mathrm{F}}^2 \tag{18b}$$

$$\boldsymbol{\mathcal{G}}^{(i)} = \arg\min_{\boldsymbol{\mathcal{G}}} \|\boldsymbol{\mathcal{R}} - \sum_{r=1}^{R}\boldsymbol{\mathcal{G}}_{\bar{\mathbf{x}}_r} \times_1 \mathbf{H}_r^{(i)} \times_2 \mathbf{Q}_r^{(i)}\|_{\mathrm{F}}^2 \tag{18c}$$

$$\mathbf{U}^{(i)} = \mathbf{W}^{(i-1)} + \mathbf{Q}^{(i)} - \mathbf{H}^{(i)} \tag{18d}$$

$$\beta_i = \frac{1+\sqrt{1+4\beta_{i-1}^2}}{2} \tag{18e}$$

$$\mathbf{W}^{(i)} = \mathbf{U}^{(i)} + \frac{\beta_{i-1}-1}{\beta_i}(\mathbf{U}^{(i)} - \mathbf{U}^{(i-1)}) \tag{18f}$$

$$i = i+1$$

end

$\mathbf{J}_4 = [\mathbf{0}_{P-K,K}, \mathbf{I}_{P-K}]^\top$, $\mathbf{J}_5 = [\mathbf{I}_M, \mathbf{0}_{M,MN_w-M}]$, $\mathbf{J}_6 = [\mathbf{0}_{P-n_a,n_a}\mathbf{I}_{P-n_a}]^\top$, $\mathbf{J}_7 = [\mathbf{0}_{MN_w-M}, \mathbf{I}_{MN_w-M}]$ $\mathbf{J}_8 = [\mathbf{I}_K\ \mathbf{0}_{K,P-K}]^\top$ are all selection matices used to select a specific region of the channel matrix.

It is important to note that the proposed cost function comprises three main components. The first part enforces the Toeplitz structure on the block diagonal matrix elements, the second part minimizes the zeros along the columns, and the third part minimizes the zeros along the rows of the channel matrix.

Using the Kronecker product property, the cost function J can be vectorized as follow.

$$\begin{aligned} J &= \|(\mathbf{J}_2^\top\otimes(\mathbf{J}_1\mathbf{H}) - \mathbf{J}_4^\top\otimes(\mathbf{J}_3\mathbf{H}))\mathbf{w}\|_2^2 \\ &+ \|(\mathbf{J}_6^\top\otimes(\mathbf{J}_5\mathbf{H}))\mathbf{w}\|_2^2 + \|(\mathbf{J}_8^\top\otimes(\mathbf{J}_7\mathbf{H}))\mathbf{w}\|_2^2 \\ &= \mathbf{w}^H(\mathbf{J}_2^\top\otimes(\mathbf{J}_1\mathbf{H}) - \mathbf{J}_4^\top\otimes(\mathbf{J}_3\mathbf{H}))^H \\ &\times (\mathbf{J}_2^\top\otimes(\mathbf{J}_1\mathbf{H}) - \mathbf{J}_4^\top\otimes(\mathbf{J}_3\mathbf{H}))\mathbf{w} \\ &+ \mathbf{w}^H(\mathbf{J}_6^\top\otimes(\mathbf{J}_5\mathbf{H}))^H(\mathbf{J}_6^\top\otimes(\mathbf{J}_5\mathbf{H}))\mathbf{w} \\ &+ \mathbf{w}^H(\mathbf{J}_8^\top\otimes(\mathbf{J}_7\mathbf{H}))^H(\mathbf{J}_8^\top\otimes(\mathbf{J}_7\mathbf{H}))\mathbf{w} \\ &= \mathbf{w}^H\mathbf{K}_1^H\mathbf{K}_1\mathbf{w} + \mathbf{w}^H\mathbf{K}_2^H\mathbf{K}_2\mathbf{w} + \mathbf{w}^H\mathbf{K}_3^H\mathbf{K}_3\mathbf{w} \\ &= \mathbf{w}^H[\mathbf{K}_1^H\mathbf{K}_1 + \mathbf{K}_2^H\mathbf{K}_2 + \mathbf{K}_3^H\mathbf{K}_3]\mathbf{w} \\ &= \mathbf{w}^H[\mathbf{K}^H\mathbf{K}]\mathbf{w}, \end{aligned} \tag{20}$$

where $\mathbf{K}_1 = (\mathbf{J}_2^\top \otimes (\mathbf{J}_1\mathbf{H}) - \mathbf{J}_4^\top \otimes (\mathbf{J}_3\mathbf{H}))$, $\mathbf{K}_2 = (\mathbf{J}_6^\top \otimes (\mathbf{J}_5\mathbf{H}))$, $\mathbf{K}_3 = (\mathbf{J}_8^\top \otimes (\mathbf{J}_7\mathbf{H}))$ and $\mathbf{K} = [\mathbf{K}_1^\top \ \mathbf{K}_2^\top \ \mathbf{K}_3^\top]^\top$ are implicitly defined in (21). The optimal $\hat{\mathbf{w}}$ is determined by performing the singular value decomposition (SVD) of $\mathbf{K}^H\mathbf{K}$ and selecting the eigenvector corresponding to the smallest eigenvalue. The obtained vector $\hat{\mathbf{w}}$ is reshaped into a matrix $\hat{\mathbf{W}} \in \mathbb{C}^{MN_w \times MN_w}$ and used to obtain the channel Toeplitz matrix estimate $\hat{\mathbf{H}} = \mathbf{H}\hat{\mathbf{W}}$. The channel taps are obtained by averaging the diagonal elements of the $\hat{\mathbf{H}}$ matrix.

V. Computational Complexity

In our proposed method, the overall computational complexity is determined by the sum of two key components: the tensor decomposition required to obtain the factor matrix, and the subsequent channel estimation procedure. The updates for both $\mathbf{H}$ in (13) and $\mathbf{Q}$ in (14) have the same computational steps which requires $\mathcal{O}(M^2N_wRKJ + MN_wR^2K^2J + N_w^3K^3)$ flops. In the case of $\boldsymbol{\mathcal{G}}$ presented in (15), the computational cost is $\mathcal{O}(M^2N_w^2R^2KJ + M^2N_w^2KR^2\min(M^2N_w^2, KR^2))$ flops. For $\mathbf{U}$, $\mathbf{Z}$ and α the computation cost required to update them is inexpensive and is $\mathcal{O}(MN_wKR)$, $\mathcal{O}(MN_wKR)$, and $\mathcal{O}(1)$ respectively. Hence, the total complexity of the type-2 BTD is $\mathcal{O}(M^2N_w^2R^2K\max(J, R^2K) + R^3K^3)$. The computational complexity of the channel estimation part is $\mathcal{O}(K^3R^3)$. Therefor, the overall complexity is $\mathcal{O}(M^2N_w^2R^2K\max(J, R^2K) + R^3K^3 + K^3R^3)$.

VI. Simulations

In this section, we validate the efficacy of the proposed type-2 BTD-based channel estimation through comprehensive numerical simulations and performance comparisons with the widely adopted ALS method [15] and its variants, including ALS-LSH [16] and ALS-ELSC [17]. In all scenarios, we first apply tensor decomposition, followed by the imposition of a Toeplitz constraint on the loading factors to estimate the channel. The performance evaluation is conducted using the mean squared error (MSE), defined as

$$\mathrm{MSE} = \sqrt{\frac{1}{N_m}\sum_{i=1}^{N_m} \|\hat{\mathbf{H}}_i - \mathbf{H}_i\|^2}, \tag{21}$$

where $N_m = 100$ denotes the number of Monte Carlo runs. Throughout the simulations, we adopt the following parameter settings: $K = 2$, $M = 4$, $L = 2$, $N_w = 5$, and $T = 100$, unless stated otherwise.

Figure. 1 depicts the MSE performance as a function of the signal-to-noise ratio (SNR) for the proposed approach and the benchmark methods. The results demonstrate that the proposed approach outperforms competing methods throughout the SNR range.

Figure 2 illustrates the variation in the data size (T) against the MSE. It is evident that increasing the size of the data leads to a corresponding improvement in the performance of all the considered methods. In particular, the proposed method consistently achieves a sustainable performance gain across all data sizes.

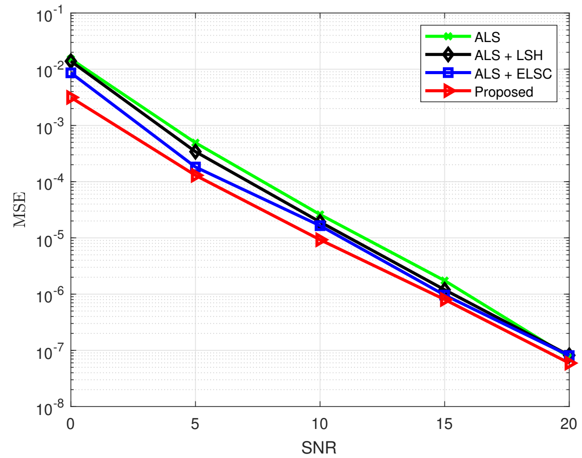

Fig. 1. Channel estimation MSE versus SNR comparison.

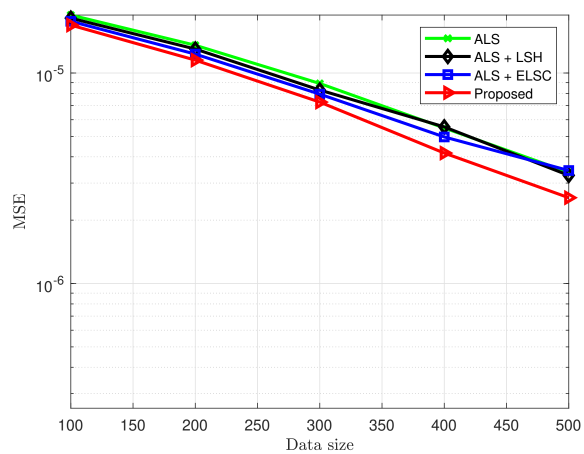

Fig. 2. MSE versus data size comparison at SNR = 10 dB.

In the final experiment, the SNR value is set to 4 dB to compare the convergence rate of the proposed method with other benchmark methods. The results indicate that the proposed method achieves faster convergence and higher estimation accuracy compared to ALS, ALS + LSH, and ALS + ELSC. Figure 3 illustrates this performance very well.

VII. Conclusion

In this work, a two-stage approach for channel estimation using tensor decomposition is introduced. Our method used type-2 BTD to obtain the loading factor of a covariance matrix tensor. A Toeplitz criterion is then enforced on the obtained loading factor which is then minimized subsequently resulting in the estimation of the channel coefficients. This combined strategy markedly improves channel estimation performance. Moreover, the core principle of our approach is versatile and can be applied to other estimation problems involving tensors with inherent structural properties.

References

[1] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.

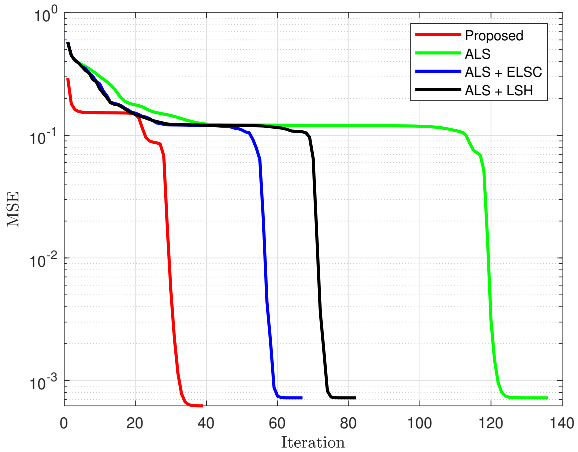

Fig. 3. MSE vs the number of iterations.

[2] K. Abed-Meraim, N. L. Trung, A. Hafiane, *et al.*, "A contemporary and comprehensive survey on streaming tensor decomposition," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 11, pp. 10897–10921, 2022.

[3] A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H. A. Phan, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE signal processing magazine*, vol. 32, no. 2, pp. 145–163, 2015.

[4] J. Du, M. Han, Y. Chen, L. Jin, H. Wu, and F. Gao, "Tensor decompositions for integrated sensing and communications," *IEEE Communications Magazine*, vol. 62, no. 9, pp. 128–134, 2024.

[5] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Transactions on signal processing*, vol. 65, no. 13, pp. 3551–3582, 2017.

[6] M. Moinuddin and A. Zerguine, "Tensor least mean fourth algorithm," in *2023 57th Asilomar Conference on Signals, Systems, and Computers*, pp. 1006–1009, 2023.

[7] M. Sørensen and L. De Lathauwer, "Blind signal separation via tensor decomposition with vandermonde factor: Canonical polyadic decomposition," *IEEE Transactions on Signal Processing*, vol. 61, no. 22, pp. 5507–5519, 2013.

[8] L. De Lathauwer, B. De Moor, and J. Vandewalle, "Blind source separation by higher-order singular value decomposition," in *Proc. EUSIPCO*, vol. 1, pp. 175–178, 1994.

[9] V. D. Vrabie, N. Le Bihan, and J. I. Mars, "Multicomponent wave separation using hosvd/unimodal-ica subspace method," *Geophysics*, vol. 71, no. 5, pp. V133–V143, 2006.

[10] T. T. Le, K. Abed-Meraim, P. Ravier, O. Buttelli, and A. Holobar, "Tensor decomposition meets blind source separation," *Signal Processing*, vol. 221, p. 109483, 2024.

[11] T. T. Le, K. Abed-Meraim, P. Ravier, O. Buttelli, and A. Holobar, "Tensorial convolutive blind source separation," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5575–5579, IEEE, 2024.

[12] L. De Lathauwer, "Decompositions of a higher-order tensor in block terms—part ii: Definitions and uniqueness," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1033–1066, 2008.

[13] L. De Lathauwer and D. Nion, "Decompositions of a higher-order tensor in block terms—part iii: Alternating least squares algorithms," *SIAM journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1067–1083, 2008.

[14] H. Bousbia-Salah, A. Belouchrani, and K. Abed-Meraim, "Jacobi-like algorithm for blind signal separation of convolutive mixtures," *Electronics Letters*, vol. 37, no. 16, p. 1, 2001.

[15] L. De Lathauwer and D. Nion, "Decompositions of a higher-order tensor in block terms—part iii: Alternating least squares algorithms," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1067–1083, 2008.

[16] M. Rajih, P. Comon, and R. A. Harshman, "Enhanced line search: A novel method to accelerate parafac," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1128–1147, 2008.

[17] D. Nion and L. De Lathauwer, "An enhanced line search scheme for complex-valued tensor decompositions. application in ds-cdma," *Signal Processing*, vol. 88, no. 3, pp. 749–755, 2008.