

Using Double Contractions to Derive the Structure of Slice-Wise Multiplications of Tensors with Applications to Semi-Blind MIMO OFDM

Kristina Naskovska, André L. F. de Almeida, Martin Haardt

Abstract—The slice-wise multiplication of two tensors is required in a variety of tensor decompositions (including PARAFAC2 and PARATUCK2) and is encountered in many applications, including the analysis of multidimensional biomedical data (EEG, MEG, etc.) or multi-carrier MIMO systems. In this paper, we propose a new tensor representation that is not based on a slice-wise (matrix) description, but can be represented by a double contraction of two tensors. Such a double contraction of two tensors can be efficiently calculated via generalized unfoldings. It leads to new tensor models of the investigated system that do not depend on the chosen unfolding and reveal the tensor structure of the data model (such that all possible unfoldings can be seen at the same time). As an example, we apply this new concept to the design of new receivers for multi-carrier MIMO systems in wireless communications. In particular, we consider MIMO OFDM systems with and without Khatri-Rao coding. The proposed receivers exploit the channel correlation between adjacent subcarriers, require the same amount of training symbols as traditional OFDM techniques, but have an improved performance in terms of the symbol error rate. Furthermore, we show that the spectral efficiency of the Khatri-Rao coded MIMO-OFDM can be increased by introducing "random coding" such that the "coding matrix" also contains useful information symbols. Considering this transmission technique, we derive a tensor model and two types of receivers for randomly coded MIMO-OFDM systems using the double contraction of two tensors.

Index Terms—Tensor modeling, slice-wise multiplications, semi-blind receivers, Khatri-Rao Space-time coding, MIMO-OFDM.

I. INTRODUCTION

In many tensor applications, we only have an element-wise or a slice-wise description of our data/signal model. For instance, there exist only a slice-wise description of the PARATUCK2 decomposition and the PARAFAC2 decomposition corresponding to a certain unfolding of the overall tensor [1], [2]. In the same way, some proposed tensor based models for MIMO-OFDM communication systems have only an element-wise or a slice-wise representation [3]. Further examples include the slice-wise description of MIMO communication systems using two-way relaying [4], [5]. This description of the signal models does not reveal the tensor structure explicitly. Hence, the derivation of all tensor unfoldings is not

always obvious. Therefore, we propose to express the slice-wise multiplication of two tensors in terms of the double contraction operator and use it to derive an explicit tensor structure of the received data tensor.

OFDM is the most widely used multi-carrier technique in current wireless communication systems. It is robust in multipath propagation environments and has a simple and efficient implementation [6], [7]. Using the FFT (Fast Fourier Transform), the complete frequency band is divided into smaller frequency subcarriers. Moreover, the use of the cyclic prefix mitigates the ISI (Inter-Symbol Interference) and the ICI (Inter-Carrier Interference). Typically, the OFDM receiver is implemented in the frequency domain based on a ZF (Zero Forcing) filter. Other more advanced solutions are proposed in [8], as well as optimal training and channel estimation for OFDM systems are proposed in [9], [10].

Tensor based signal processing offers an improved identifiability, uniqueness, and more efficient denoising compared to matrix based techniques. In [3], a MIMO multi-carrier system is modeled using tensor algebra and the PARATUCK2 tensor decomposition resulting in a novel space, time, and frequency coding structure. Similarly in [11], trilinear coding in space, time, and frequency is proposed for MIMO-OFDM systems based on the CP tensor decomposition. By exploiting tensor models, semi-blind receivers are introduced for multi-carrier communications systems in [12] and [13]. All these works use additional spreading that leads to a significantly reduced spectral efficiency to create the tensor structure. Moreover, previous publications on tensor models for multi-carrier communication systems [3], [11], [13], and [12] do not exploit the channel correlation between the adjacent subcarriers. The previously mentioned publications rely on the subcarrier-wise description of the MIMO-OFDM system. This description of the signal models does not reveal the tensor structure explicitly. Hence, the derivation of all tensor unfoldings is not always obvious. Therefore, we propose to express the slice-wise multiplication of two tensors in terms of the double contraction operator. To this end, we summarize important properties of the contraction operator for element-wise and slice-wise multiplications. Using the contraction operator, we derive the tensor structure of the received signal that includes the frequency (subcarrier) mode.

In this paper, we first present the double contraction between an uncoded signal tensor and a channel tensor for OFDM systems, yielding the same spectral efficiency as matrix based approaches (since no additional spreading is used) [14]. By

K. Naskovska and M. Haardt are with the Communications Research Laboratory, Ilmenau University of Technology, P. O. Box 100565, D-98684 Ilmenau, Germany.

A. de Almeida is with the Department of Teleinformatics Engineering, Federal University of Ceará (UFC), Fortaleza, Brazil.

Manuscript received Month Day, Year; revised Month Day, Year.

exploiting this new tensor structure, we can reshape the received signal tensor into the factorization of a sum of Khatri-Rao products. Channel and symbol estimation can be achieved by means of an iterative and recursive least squares procedure originally proposed for blind source separation. Moreover, we propose an application of the double contraction operator to Khatri-Rao coded MIMO-OFDM systems [15]. Due to the Khatri-Rao coding, the signal tensor has more structure, i.e., we can use a CP model to describe it. The Khatri-Rao space-time coding has been introduced in [16]. Later, it has been extended in [17] to Khatri-Rao space-time-frequency coding. In contrast to the state-of-the-art, we exploit the structure of the channel and the contraction properties using the transmit signal tensor and the known coding matrix to propose a receiver based on the LS-KRF. In addition, we reduce the number of required pilot symbols by exploiting the correlation of the channel in the frequency domain. Alternatively, we propose a more spectrally efficient "random coding" model for MIMO-OFDM systems. In this case, we propose to keep the CP structure of the Khatri-Rao coded transmit signal [15] but the "coding matrix" contains useful information symbols. Thus, the proposed randomly coded MIMO-OFDM system has a higher spectral efficiency than a Khatri-Rao coded MIMO-OFDM system. By exploiting the derived tensor structure of the received signal, we also design two types of receivers for the randomly coded MIMO-OFDM systems.

II. TENSOR ALGEBRA AND NOTATION

A. Notation

We use the following notation. Scalars are denoted either as capital or lower-case italic letters, A, a . Vectors and matrices, are denoted as bold-faced lower-case and capital letters, \mathbf{a}, \mathbf{A} , respectively. Tensors are represented by bold-faced calligraphic letters \mathcal{A} . The following superscripts, T , H , $^{-1}$, and $^+$ denote transposition, Hermitian transposition, matrix inversion and Moore-Penrose pseudo matrix inversion, respectively. The outer product, Kronecker product, and Khatri-Rao product are denoted as \circ , \otimes , and \diamond , respectively. Moreover, we denote an inverse Hadamard product (element-wise division) between two matrices of equal dimensions as \oslash . The operators $\|\cdot\|_{\text{F}}$ and $\|\cdot\|_{\text{H}}$ denote the Frobenius norm and the higher order norm, respectively. Similar to matrices, a Kronecker product between two tensors $\mathcal{A} \in \mathbb{C}^{M \times N \times L}$ and $\mathcal{B} \in \mathbb{C}^{P \times Q \times R}$ can be defined as $\mathcal{K} = \mathcal{A} \otimes \mathcal{B} \in \mathbb{C}^{PM \times QN \times RL}$ [18]. Moreover, the n -mode product between a tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ and a matrix $\mathbf{B} \in \mathbb{C}^{J \times I_n}$ is denoted as $\mathcal{A} \times_n \mathbf{B}$, for $n = 1, 2, \dots, N$ [19]. A super-diagonal or identity N -way tensor of dimension $R \times R \times \dots \times R$ is denoted as $\mathcal{I}_{N,R}$. Similarly, an identity matrix of dimension $R \times R$ is denoted as \mathbf{I}_R and we denote a vector of ones of length R as $\mathbf{1}_R$. The n -th 3-mode slice of a tensor $\mathcal{A} \in \mathbb{C}^{I \times J \times N}$ is denoted as $\mathcal{A}_{(\dots, n)}$ and accordingly one element of this tensor is denoted as $\mathcal{A}_{(i,j,n)}$. The operator $\text{diag}(\cdot)$ transforms a vector into a diagonal matrix and the operator $\text{vec}(\cdot)$ transforms a matrix into a vector.

B. The CP Decomposition and Generalized Tensor Unfoldings

The CP tensor decomposition decomposes a given tensor into the minimum number of rank one components. The CP

decomposition of a 4-way, rank R noiseless tensor $\mathcal{A} \in \mathbb{C}^{I \times J \times M \times N}$ is defined as

$$\mathcal{A} = \mathcal{I}_{3,R} \times_1 \mathbf{F}_1 \times_2 \mathbf{F}_2 \times_3 \mathbf{F}_3 \times_4 \mathbf{F}_4, \quad (1)$$

where $\mathbf{F}_1 \in \mathbb{C}^{I \times R}$, $\mathbf{F}_2 \in \mathbb{C}^{J \times R}$, $\mathbf{F}_3 \in \mathbb{C}^{M \times R}$, and $\mathbf{F}_4 \in \mathbb{C}^{N \times R}$ are the factor matrices [19], [20]. In addition to the n -mode unfoldings, generalized matrix unfoldings can be defined by using two subsets of any of the N dimensions [21], [22]. For instance, the set of modes $(1, 2, \dots, N)$ of an N -way tensor \mathcal{A} can be divided into two non-overlapping, P and $N - P$ dimensional subsets, $\alpha^{(1)} = [\alpha_1 \dots \alpha_P]$ and $\alpha^{(2)} = [\alpha_{P+1} \dots \alpha_N]$, respectively. This leads to the generalized unfolding $[\mathcal{A}]_{(\alpha^{(1)}, \alpha^{(2)})}$, where the indices contained in $\alpha^{(1)}$ vary along the rows and the indices contained in $\alpha^{(2)}$ vary along the columns. Here, the index α_1 varies the fastest between the rows, the index α_{P+1} varies the fastest between the columns, P is any number between one and N , and α_n is any of the tensor dimensions. For instance, let us assume the 4-way tensor $\mathcal{A} \in \mathbb{C}^{I \times J \times M \times N}$ defined in equation (1). In the generalized unfolding $[\mathcal{A}]_{([1,2],[3,4])}$ the 1-st mode varies faster than the 2-nd mode along the rows and the 3-rd mode varies faster than the 4-th mode along the columns. Moreover, for a tensor with a CP structure, its unfoldings and generalized unfoldings can be expressed in terms of the factor matrices. For instance, the generalized unfolding $[\mathcal{A}]_{([1,2],[3,4])}$ of the tensor \mathcal{A} satisfies [22], [15]

$$[\mathcal{A}]_{([1,2],[3,4])} = (\mathbf{F}_2 \diamond \mathbf{F}_1) \cdot (\mathbf{F}_4 \diamond \mathbf{F}_3)^{\text{T}}.$$

In a similar way, the rest of the tensor unfoldings and generalized unfoldings can be defined.

C. Tensor Contraction

The contraction $\mathcal{A} \bullet_n^m \mathcal{C}$ between two tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ and $\mathcal{C} \in \mathbb{C}^{J_1 \times J_2 \times \dots \times J_N}$ represents an inner product of the n -th mode of \mathcal{A} with the m -th mode of \mathcal{C} , provided that $I_n = J_m$ [18]. Contraction along several modes of compatible dimensions is also possible and accordingly the contraction along two modes is denoted as $\mathcal{A} \bullet_{n,k}^{m,l} \mathcal{C}$. More specifically, the double contraction between the tensors $\mathcal{A} \in \mathbb{C}^{I \times J \times M \times N}$ and $\mathcal{C} \in \mathbb{C}^{M \times N \times K}$ is defined as [18],

$$(\mathcal{A} \bullet_{3,4}^{1,2} \mathcal{C})_{(i,j,k)} \triangleq \sum_{n=1}^N \sum_{m=1}^M \mathcal{A}_{(i,j,m,n)} \cdot \mathcal{C}_{(m,n,k)} = \mathcal{T}_{(i,j,k)}.$$

This example represents a contraction of the 3-rd and 4-th mode of \mathcal{A} with the 1-st and 2-nd mode of \mathcal{C} , respectively.

Using the concept of the generalized unfoldings, it can be shown that the tensor contraction satisfies

$$[\mathcal{A} \bullet_{3,4}^{1,2} \mathcal{C}]_{([1,2],3)} = [\mathcal{A}]_{([1,2],[3,4])} \cdot [\mathcal{C}]_{([1,2],3)} = \quad (2)$$

$$[\mathcal{A} \bullet_{4,3}^{2,1} \mathcal{C}]_{([1,2],3)} = [\mathcal{A}]_{([1,2],[4,3])} \cdot [\mathcal{C}]_{([2,1],3)}. \quad (3)$$

In the generalized unfolding $[\mathcal{A}]_{([1,2],[3,4])}$ the 1-st mode varies faster than the 2-nd mode between the rows and the 3-rd mode varies faster than the 4-th mode between the columns.

D. Contraction Properties for Element-wise and Slice-wise Multiplications

1) *Hadamard product via tensor contraction*: First, let us consider a Hadamard product (element-wise multiplication) between two vectors $\mathbf{a} \in \mathbb{C}^{M \times 1}$ and $\mathbf{b} \in \mathbb{C}^{M \times 1}$, $c_{(m)} = a_{(m)}b_{(m)}$, $\forall m = 1, \dots, M$ ($\mathbf{c} \in \mathbb{C}^{M \times 1}$). The Hadamard product can be expressed via the multiplication of a diagonal matrix and a vector, i.e., $\mathbf{a} \odot \mathbf{b} = \text{diag}(\mathbf{a}) \mathbf{b} = \text{diag}(\mathbf{b}) \mathbf{a}$. Using the fact that a matrix multiplication is equivalent to the contraction $\bullet_{2,1}^{\frac{1}{2}}$, we get

$$\mathbf{a} \odot \mathbf{b} = \text{diag}(\mathbf{a}) \bullet_{2,1}^{\frac{1}{2}} \mathbf{b} = \text{diag}(\mathbf{b}) \bullet_{2,1}^{\frac{1}{2}} \mathbf{a}.$$

Next, for the Hadamard product between two matrices $\mathbf{A} \in \mathbb{C}^{M \times N}$ and $\mathbf{B} \in \mathbb{C}^{M \times N}$, $\mathbf{C}_{(m,n)} = \mathbf{A}_{(m,n)} \mathbf{B}_{(m,n)}$, $\forall m = 1, \dots, M$ and $n = 1, \dots, N$, we can show that $\mathbf{C} = \mathbf{A} \odot \mathbf{B} = \mathcal{D}_A \bullet_{2,4}^{1,2} \mathbf{B} = \mathcal{D}_B \bullet_{2,4}^{1,2} \mathbf{A}$. Here $\mathcal{D}_A \in \mathbb{C}^{M \times M \times N \times N}$ and $\mathcal{D}_B \in \mathbb{C}^{M \times M \times N \times N}$ are diagonal 4-way tensors with non-zero elements $\mathcal{D}_{A(m,m,n,n)} = \mathbf{A}_{(m,n)}$ and $\mathcal{D}_{B(m,m,n,n)} = \mathbf{B}_{(m,n)}$, respectively. As an alternative, we also have

$$\mathbf{C} = \mathbf{A} \odot \mathbf{B} = \mathcal{D}^{(A)} \bullet_{2,3}^{1,3} \mathcal{D}^{(B)},$$

where the diagonal 3-way tensors have the following non-zero elements $\mathcal{D}^{(A)}_{(m,m,n)} = \mathbf{A}_{(m,n)}$ and $\mathcal{D}^{(B)}_{(m,m,n)} = \mathbf{B}_{(m,n)}$. Moreover, these diagonal 3-way tensors can be either defined in terms of slices,

$$\begin{aligned} \mathcal{D}^{(A)}_{(\dots,n)} &= \text{diag}(\mathbf{A}_{(\dots,n)}), \forall n = 1, \dots, N \\ \mathcal{D}^{(B)}_{(m,\dots)} &= \text{diag}(\mathbf{B}_{(m,\dots)}), \forall m = 1, \dots, M \end{aligned}$$

or using tensor notation $\mathcal{D}^{(A)} = \mathcal{I}_{3,M} \times_3 \mathbf{A}^T$ and $\mathcal{D}^{(B)} = \mathcal{I}_{3,N} \times_1 \mathbf{B}$. The diagonal structure of these tensors is visualized in Fig. 1.

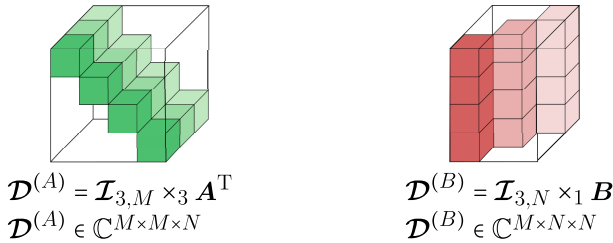


Fig. 1: A visualization of the diagonal structure of the unfoldings for the tensors $\mathcal{D}^{(A)} = \mathcal{I}_{3,M} \times_3 \mathbf{A}^T \in \mathbb{C}^{M \times M \times N}$ and $\mathcal{D}^{(B)} = \mathcal{I}_{3,N} \times_1 \mathbf{B} \in \mathbb{C}^{M \times N \times N}$, for $M = 4$ and $N = 3$.

2) *Slice-wise multiplication via tensor contraction*: A slice-wise multiplication between two tensors $\mathcal{A} \in \mathbb{C}^{M \times N \times K}$ and $\mathcal{B} \in \mathbb{C}^{N \times J \times K}$ is defined as $\mathcal{T}_{1(\dots,k)} = \mathcal{A}_{(\dots,k)} \mathcal{B}_{(\dots,k)}$, $\forall k = 1, \dots, K$. We depict this slice-wise multiplication in Fig. 2. To express this slice-wise multiplication we can diagonalize \mathcal{B} to obtain

$$\mathcal{T}_1 = \mathcal{A} \bullet_{2,3}^{1,4} \mathcal{D}_B \in \mathbb{C}^{M \times J \times K},$$

where $\mathcal{D}_B \in \mathbb{C}^{N \times J \times K \times K}$ has non-zero elements $\mathcal{D}_{B(n,j,k,k)} = \mathcal{B}_{(n,j,k)}$ or $\mathcal{D}_{B(n,j,\dots)} = \text{diag}(\mathcal{B}_{(n,j,\dots)})$, for $n = 1, \dots, N$ and $j = 1, \dots, J$. Further combinations are also possible that lead to the same result, for instance, $\mathcal{T}_2 = \mathcal{D}_B \bullet_{1,4}^{2,3} \mathcal{A} \in \mathbb{C}^{J \times K \times M}$

or $\mathcal{T}_3 = \mathcal{D}_A \bullet_{2,4}^{1,3} \mathcal{B} \in \mathbb{C}^{M \times K \times J}$ with $\mathcal{D}_{A(m,n,k,k)} = \mathcal{A}_{(m,n,k)}$ as diagonal elements (non-zero elements of \mathcal{D}_A). Note that the tensors \mathcal{T}_1 , \mathcal{T}_2 , and \mathcal{T}_3 contain the same elements, but have permuted dimensions. However, the permuted order of the dimensions is not relevant, because we always explicitly declare which dimension is multiplied or unfolded.

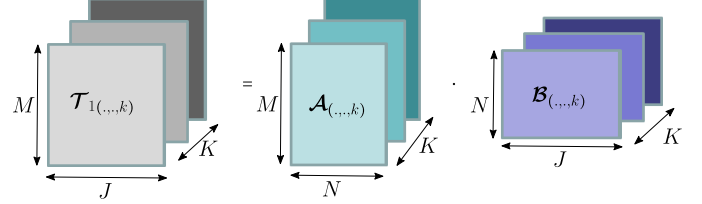


Fig. 2: A slice-wise multiplication between two tensors $\mathcal{A} \in \mathbb{C}^{M \times N \times K}$ and $\mathcal{B} \in \mathbb{C}^{N \times J \times K}$.

3) *Representation of diagonal matrices and diagonal tensors in terms of Khatri-Rao products*: An explicit expression of the diagonalized tensor can be obtained by expressing its generalized unfolding in terms of a Khatri-Rao product with an identity matrix. First, let us consider the column vector $\mathbf{a} \in \mathbb{C}^M$. It can be easily shown that

$$\text{diag}(\mathbf{a}) = \mathbf{I}_M \diamond \mathbf{a}^T.$$

Next, let us consider the reshaping of the matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$ into a diagonal tensor $\mathcal{D}^{(A)} = \mathcal{I}_{3,M} \times_3 \mathbf{A}^T$. By studying the resulting tensor structure, the tensor unfoldings, and the properties of the Khatri-Rao product, we get

$$\left[\mathcal{D}^{(A)} \right]_{([3,2],[1])} = \mathbf{I}_M \diamond \mathbf{A}^T.$$

Likewise, for the tensor $\mathcal{D}^{(B)} = \mathcal{I}_{3,N} \times_1 \mathbf{B} \in \mathbb{C}^{M \times N \times N}$ and the matrix $\mathbf{B} \in \mathbb{C}^{M \times N}$, we have $\left[\mathcal{D}^{(B)} \right]_{([1,3],[2])} = \mathbf{I}_N \diamond \mathbf{B}$.

non-zero elements	generalized unfoldings
$\mathcal{D}_{(m,m)} = \mathbf{a}_{(m)}$	$\mathcal{D} = \mathbf{I}_M \diamond \mathbf{a}^T$
$\mathcal{D}_{(m,n,n)} = \mathbf{A}_{(m,n)}$	$[\mathcal{D}]_{([1,3],[2])} = \mathbf{I}_N \diamond \mathbf{A}$
$\mathcal{D}_{(m,m,n)} = \mathbf{A}_{(m,n)}$	$[\mathcal{D}]_{([3,2],[1])} = \mathbf{I}_M \diamond \mathbf{A}^T$
$\mathcal{D}_{(m,m,n,n)} = \mathbf{A}_{(m,n)}$	$[\mathcal{D}]_{([1,3],[2,4])} = \mathbf{I}_M \diamond \text{vec}(\mathbf{A})^T$
$\mathcal{D}_{(m,n,k,k)} = \mathbf{A}_{(m,n,k)}$	$[\mathcal{D}]_{([1,2,4],[3])} = \mathbf{I}_K \diamond [\mathcal{A}]_{([1,2],[3])}$
$\mathcal{D}_{(m,m,n,k)} = \mathbf{A}_{(m,n,k)}$	$[\mathcal{D}]_{([3,4,2],[1])} = \mathbf{I}_M \diamond [\mathcal{A}]_{([2,3],[1])}$

TABLE I: Link between the diagonalized tensor structures and their generalized unfoldings.

The expression of the diagonalized tensor in terms of its generalized unfoldings and the Khatri-Rao product with an identity matrix can also be obtained for N -way tensors. Hence, there exists a link between the diagonalized tensor structures and their corresponding generalized unfoldings. The corresponding generalized unfolding can always be expressed as a Khatri-Rao product between an identity matrix and a generalized unfolding of the tensor to be diagonalized, where the dimensions that are diagonalized are in the columns of the second matrix. This notation will be used later in this paper and it is given in Table I.

The element-wise or slice-wise multiplication between two arrays (vectors/matrices/tensors) of the same order can be

written in terms of a contraction if the unaffected mode vectors are transformed into a diagonal matrix (by adding an additional array dimension). This diagonalization can be performed using the Khatri-Rao product as shown in Table I.

III. MIMO-OFDM

We assume a MIMO-OFDM system with M_T transmit and M_R receive antennas. One OFDM block consists of N samples, which equals the DFT (Discrete Fourier Transform) length, using the assumption that all N subcarriers are used for data transmission. If guard subcarriers are used, i.e., not all subcarriers are used for data transmission, the number of OFDM samples is smaller than the DFT length. All signals and equations used for the following derivation are in the frequency domain. Moreover, N is the number of subcarriers and K denotes the number of transmitted frames. The received signal in the frequency domain $\tilde{\mathbf{Y}} \in \mathbb{C}^{N \times M_R \times K}$ after the removal of the cyclic prefix is defined by means of the contraction operator

$$\tilde{\mathbf{Y}} = \tilde{\mathcal{H}} \bullet_{2,4}^{1,2} \tilde{\mathcal{S}} + \tilde{\mathcal{N}} = \tilde{\mathbf{Y}}_0 + \tilde{\mathcal{N}}. \quad (4)$$

We use \sim to distinguish the frequency domain from the time domain, i.e., $\tilde{\mathbf{Y}} = \mathbf{Y} \times_1 \mathbf{F}_N$, where $\mathbf{F}_N \in \mathbb{C}^{N \times N}$ is the DFT matrix and \mathbf{Y} is the received signal in the time domain. The transmit signal tensor is denoted as $\tilde{\mathcal{S}} \in \mathbb{C}^{N \times M_T \times K}$ and $\tilde{\mathcal{N}} \in \mathbb{C}^{N \times M_R \times K}$ represents the additive white Gaussian noise in the frequency domain. The tensor $\tilde{\mathbf{Y}}_0 \in \mathbb{C}^{N \times M_R \times K}$ represents the noiseless received signal in the frequency domain after the removal of the cyclic prefix. The frequency-selective propagation channel is represented by a channel tensor $\tilde{\mathcal{H}} \in \mathbb{C}^{N \times N \times M_R \times M_T}$ as we propose in [15] the structure of which is detailed as follows.

A. Channel tensor

We assume that the frequency-selective channel has an impulse response $\mathbf{h}_L^{(m_R, m_T)} \in \mathbb{C}^{L \times 1}$, for each receive-transmit antenna pair, (m_R, m_T) , for $m_R = 1 \dots M_R$ and $m_T = 1 \dots M_T$, and a maximum of L taps. After the removal of the cyclic prefix, the channel matrix in the frequency domain is a diagonal matrix for each receive-transmit antenna pair, $\tilde{\mathbf{H}}^{(m_R, m_T)} = \text{diag}(\mathbf{F}_{N \times L} \cdot \mathbf{h}_L^{(m_R, m_T)}) \in \mathbb{C}^{N \times N}$ [9], [10]. Here, the matrix $\mathbf{F}_{N \times L} \in \mathbb{C}^{N \times L}$ contains the first L columns of the DFT matrix of size $N \times N$. Collecting all the channel matrices in a 4-way channel tensor $\tilde{\mathcal{H}}$, we get

$$\tilde{\mathcal{H}}_{(\dots, m_R, m_T)} = \text{diag}(\mathbf{F}_{N \times L} \cdot \mathbf{h}_L^{(m_R, m_T)}) = \text{diag}(\tilde{\mathbf{h}}^{(m_R, m_T)}). \quad (5)$$

For each receive-transmit antenna pair the channel transfer matrix is a diagonal matrix that is represented by the corresponding slice of the tensor $\tilde{\mathcal{H}}$ as shown in (5). The vector $\tilde{\mathbf{h}}^{(m_R, m_T)} \in \mathbb{C}^{N \times 1}$ contains the frequency domain channel coefficients. An example of a MIMO system with $M_T = 2$ transmit antennas and $M_R = 3$ receive antennas and the corresponding channel vectors is depicted in Fig. 3. We assume that the channel stays constant during the K frames. Note that only in case of cyclic prefix OFDM the channel tensor in the

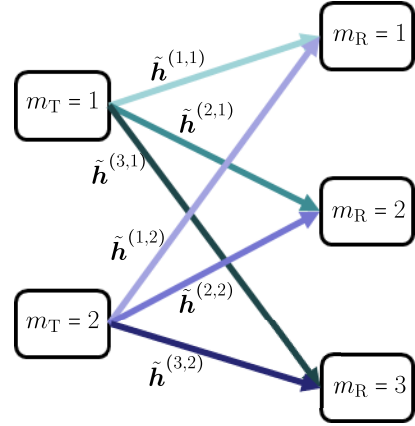


Fig. 3: A MIMO system with $M_T = 2$ transmit antennas and $M_R = 3$ receive antennas.

frequency domain contains diagonal matrices for each receive-transmit antenna pair. In a general multi-carrier system, the frequency domain channel matrix is not necessarily diagonal. However, equation (4) is still satisfied which means that our general model also remains valid for non-orthogonal multi-carrier systems.

In (5), we have defined the channel tensor. However, up to this point, we have not revealed the explicit tensor structure. In order to do so, let us first assume that all channel transfer matrices for the m_T -th transmit and all receive antennas are collected in a diagonal tensor $\tilde{\mathcal{H}}_R^{(m_T)} \in \mathbb{C}^{N \times N \times M_R}$, i.e.,

$$\tilde{\mathcal{H}}_{R(\dots, m_R)}^{(m_T)} = \text{diag}(\tilde{\mathbf{h}}^{(m_R, m_T)}) \quad (6)$$

$$m_R = 1, \dots, M_R, m_T = 1, \dots, M_T$$

Based on this diagonal structure, the tensor $\tilde{\mathcal{H}}_R^{(m_T)}$ can be written as the following CP decomposition

$$\tilde{\mathcal{H}}_R^{(m_T)} = \mathcal{I}_{3, N} \times_1 \mathbf{I}_N \times_2 \mathbf{I}_N \times_3 \tilde{\mathbf{H}}_R^{(m_T)}, \quad (7)$$

where $\tilde{\mathbf{H}}_R^{(m_T)} = [\tilde{\mathbf{h}}^{(1, m_T)} \quad \tilde{\mathbf{h}}^{(2, m_T)} \quad \dots \quad \tilde{\mathbf{h}}^{(M_R, m_T)}]^T \in \mathbb{C}^{M_R \times N}$.

The complete 4-way channel tensor, defined in equation (5) can be obtained by concatenating the $\tilde{\mathcal{H}}_R^{(m_T)}$ tensors along the fourth dimension. Hence, the 4-way channel tensor $\tilde{\mathcal{H}}$ can be expressed as

$$\begin{aligned} \tilde{\mathcal{H}} &= [\tilde{\mathcal{H}}_R^{(1)} \quad \sqcup_4 \quad \tilde{\mathcal{H}}_R^{(2)} \quad \sqcup_4 \quad \dots \quad \tilde{\mathcal{H}}_R^{(M_T)}] \\ &= \sum_{m_T=1}^{M_T} \tilde{\mathcal{H}}_R^{(m_T)} \circ \mathbf{e}_{m_T} \end{aligned} \quad (8)$$

$$= \sum_{m_T=1}^{M_T} \mathcal{D} \times_1 \mathbf{I}_N \times_2 \mathbf{I}_N \times_3 \tilde{\mathcal{H}}_R^{(m_T)} \times_4 \mathbf{e}_{m_T}. \quad (9)$$

Note that $\tilde{\mathcal{H}}$ satisfies a very special BTD, where $\mathcal{D}_{(\dots, 1)} = \mathcal{I}_{3, N} \in \mathbb{R}^{N \times N \times N \times 1}$ ($\mathcal{D} = \mathcal{I}_{4, 1} \otimes \mathcal{I}_{3, N}$) and $\mathbf{e}_{m_T} \in \mathbb{R}^{M_T \times 1}$ is a pinning vector. We prove the BTD structure of the channel tensor \mathcal{H} in Appendix A. In this appendix, we also show that

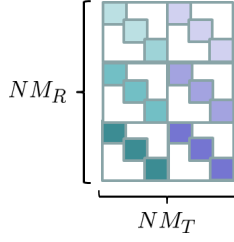


Fig. 4: Visualization of the generalized unfolding $[\tilde{\mathcal{H}}]_{([1,3],[2,4])}$ for a MIMO-OFDM system with parameters $M_T = 2$, $M_R = 3$, and $N = 3$.

the $([1,3],[2,4])$ generalized unfolding of the channel tensor can be expressed as

$$[\tilde{\mathcal{H}}]_{([1,3],[2,4])} = \tilde{\mathbf{H}} \diamond (\mathbf{1}_{M_T}^T \otimes \mathbf{I}_N) \in \mathbb{C}^{NM_R \times NM_T}, \quad (10)$$

where $\tilde{\mathbf{H}} \in \mathbb{C}^{M_R \times NM_T}$ is a matrix containing all non-zero elements of the tensor $\tilde{\mathcal{H}}$ and it is defined as,

$$\begin{aligned} \tilde{\mathbf{H}} &= \begin{bmatrix} \tilde{\mathbf{h}}^{(1,1)T} & \tilde{\mathbf{h}}^{(1,2)T} & \dots & \tilde{\mathbf{h}}^{(1,M_T)T} \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{\mathbf{h}}^{(M_R,1)T} & \tilde{\mathbf{h}}^{(M_R,2)T} & \dots & \tilde{\mathbf{h}}^{(M_R,M_T)T} \end{bmatrix} \\ &= \begin{bmatrix} \tilde{\mathbf{H}}_R^{(1)} & \tilde{\mathbf{H}}_R^{(2)} & \dots & \tilde{\mathbf{H}}_R^{(M_T)} \end{bmatrix} \in \mathbb{C}^{M_R \times NM_T}. \end{aligned} \quad (11)$$

Fig. 4 depicts the structure of the generalized unfolding $[\tilde{\mathcal{H}}]_{([1,3],[2,4])}$ for a MIMO-OFDM system with parameters $M_T = 2$, $M_R = 3$, and $N = 3$.

B. Data transmission

The signal tensor $\tilde{\mathcal{S}}$ in equation (4) contains all data symbols in the frequency domain that are transmitted on N subcarriers, M_T transmit antennas, and K frames. For notational simplicity, we define the following block matrix $\tilde{\mathbf{S}}$ as the transpose of the 3-mode unfolding of $\tilde{\mathcal{S}}$

$$\tilde{\mathbf{S}} = \tilde{\mathcal{S}}_{([1,2],[3])}^T = \begin{bmatrix} \tilde{\mathbf{S}}^{(1)} & \tilde{\mathbf{S}}^{(2)} & \dots & \tilde{\mathbf{S}}^{(M_T)} \end{bmatrix} \in \mathbb{C}^{K \times NM_T}, \quad (12)$$

where $\tilde{\mathbf{S}}^{(m_T)} \in \mathbb{C}^{K \times N}$ contains the symbols transmitted via the m_T -th antenna.

Moreover, we assume that the symbol matrix consists of data and pilot symbols, $\tilde{\mathbf{S}} = \tilde{\mathbf{S}}_d + \tilde{\mathbf{S}}_p$. The matrices $\tilde{\mathbf{S}}_d$ and $\tilde{\mathbf{S}}_p$ represent the data symbols and the pilot symbols, respectively. The matrix $\tilde{\mathbf{S}}_d$ contains zeros at the positions of the pilot symbols. Accordingly, the matrix $\tilde{\mathbf{S}}_p$ contains non-zero elements only at the pilot positions. Typically, there are three ways of arranging the pilot symbol within the OFDM blocks (block, comb, and lattice-type) [6]. We assume a comb-type arrangement of the pilot symbols with equidistant positions in the time and the frequency domains, for each antenna. The spacing in the time domain is denoted by ΔK . Moreover, we assume a subcarrier spacing of ΔF between two pilot symbols. Furthermore, there are positions where neither pilot symbols nor data symbols are allowed to be transmitted. These positions are reserved for the pilot symbols corresponding to the remaining antennas. This results in $M_T \lfloor \frac{N}{\Delta F} \rfloor$ pilot symbols per frame. In comparison, other publications such as [3],

[11], [13], and [12] use NM_T pilot symbols per frame. By exploiting the channel correlation among adjacent subcarriers, a reduced number of pilot symbols can be used for channel estimation.

C. Receiver design

Using the property of the generalized unfoldings in equation (2), the received signal in equation (4) becomes

$$[\tilde{\mathcal{Y}}]_{([1,2],[3])} = [\tilde{\mathcal{H}}]_{([1,3],[2,4])} \tilde{\mathcal{S}}_{([1,2],[3])} + [\tilde{\mathcal{N}}]_{([1,2],[3])}. \quad (13)$$

Next, by substituting the corresponding tensor unfoldings in equation (13), we get

$$[\tilde{\mathcal{Y}}]_{([1,2],[3])} = (\tilde{\mathbf{H}} \diamond (\mathbf{1}_{M_T}^T \otimes \mathbf{I}_N)) \cdot \tilde{\mathbf{S}}^T + [\tilde{\mathcal{N}}]_{([1,2],[3])}. \quad (14)$$

The above equation satisfies an unfolding of a noisy observation of a low-rank tensor with a CP structure. By applying an inverse unfolding for the received signal in the frequency domain after the removal of the cyclic prefix, we get the desired tensor description of the received data tensor

$$\tilde{\mathcal{Y}} = \mathcal{I}_{3, NM_T} \times_1 (\mathbf{1}_{M_T}^T \otimes \mathbf{I}_N) \times_2 \tilde{\mathbf{H}} \times_3 \tilde{\mathbf{S}} + \tilde{\mathcal{N}} \in \mathbb{C}^{N \times M_R \times K}. \quad (15)$$

Our goal is to jointly estimate the channel and the symbols, i.e., $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{S}}$ in equation (15). Note that all factor matrices are flat, resulting in a degenerate CP model in all three modes. Therefore, it is difficult to estimate the channel and the symbols by simply fitting a CP model to the received signal tensor in (15).

Using the prior knowledge of the pilot symbols and their positions, the channel in the frequency domain can be estimated. Naturally, the channel is estimated only at those subcarrier positions where the pilot symbols are located. Afterwards, an interpolation is applied to get the complete channel estimate. Alternatively, as shown in [9], [10] the channel can be first estimated in the time domain and then transformed into the frequency domain. Either way, this leads to a pilot based channel estimate that we denote as $\tilde{\mathbf{H}}_p$, or $\tilde{\mathcal{H}}_p^1$. The pilot based channel estimate is then used to estimate the data symbols. In the remainder of this section, we discuss different ways to estimate the symbols. We use the pilot based channel estimate to initialize the proposed algorithms.

Traditionally, the estimate of the symbols is obtained in the frequency domain with a ZF receiver. In this case, the symbols are calculated by inverting the channel matrix for each subcarrier individually. This ZF receiver using the above defined tensor notation is summarized in Algorithm 1.

Alternatively, if we compute the 1-mode unfolding of the tensor $\tilde{\mathcal{Y}}$ in equation (15), we get

$$[\tilde{\mathcal{Y}}]_{([1],[2,3])} = (\mathbf{1}_{M_T}^T \otimes \mathbf{I}_N) \cdot (\tilde{\mathbf{S}} \diamond \tilde{\mathbf{H}})^T + [\tilde{\mathcal{N}}]_{([1],[3,2])}.$$

¹In our simulations, we use the pilot based channel estimate obtained in the time domain.

Algorithm 1: ZF receiver

```

initialization  $\tilde{\mathcal{H}}_p$ 
for  $n = 1 : N$  do
  |  $\tilde{\mathcal{S}}_{(n,\dots)} \approx \tilde{\mathcal{H}}_{p(n,n,\dots)}^+ \tilde{\mathcal{Y}}_{(n,\dots)}$ 
end
Result:  $\tilde{\mathcal{S}}$ 

```

Taking into account the structure of the matrices $(\mathbf{1}_{M_T}^T \otimes \mathbf{I}_N) \in \mathbb{R}^{N \times N M_T}$, $\tilde{\mathbf{H}}$ in (11), and $\tilde{\mathcal{S}}$ in (12), the 1-mode unfolding becomes

$$[\tilde{\mathcal{Y}}]_{([1],[2,3])} = \sum_{m_T=1}^{M_T} \left(\tilde{\mathcal{S}}^{(m_T)} \diamond \tilde{\mathbf{H}}_R^{(m_T)} \right)^T + [\tilde{\mathcal{N}}]_{([1],[3,2])}.$$

After transposition and omitting the noise term, we get

$$[\tilde{\mathcal{Y}}]_{([2,3],[1])} \approx \sum_{m_T=1}^{M_T} \left(\tilde{\mathcal{S}}^{(m_T)} \diamond \tilde{\mathbf{H}}_R^{(m_T)} \right) \in \mathbb{C}^{M_R K \times N}.$$

This sum of Khatri-Rao products can be resolved in a column-wise fashion. Let $\tilde{\mathbf{y}}_n \in \mathbb{C}^{M_R K \times 1}$ denote the n -th column of $[\tilde{\mathcal{Y}}]_{([2,3],[1])} \in \mathbb{C}^{M_R K \times N}$. After reshaping this vector into the matrix $\tilde{\mathbf{Y}}_n \in \mathbb{C}^{M_R \times K}$, such that $\tilde{\mathbf{y}}_n = \text{vec}(\tilde{\mathbf{Y}}_n)$, it is easy to see that this matrix satisfies

$$\tilde{\mathbf{Y}}_n \approx \tilde{\mathbf{H}}_n \cdot \tilde{\mathcal{S}}_n, \quad (16)$$

where $\tilde{\mathbf{H}}_n$ and $\tilde{\mathcal{S}}_n$ are the n -th slices of $\tilde{\mathcal{H}}_{(n,n,\dots)} \in \mathbb{C}^{M_R \times M_T}$ and $\tilde{\mathcal{S}}_{(n,\dots)} \in \mathbb{C}^{M_T \times K}$, respectively. Note that $\tilde{\mathbf{Y}}_n$ is the n -th slice of $\tilde{\mathcal{Y}}_{(n,\dots)}$. Using the pseudo inverse of the channel, we get the traditional ZF receiver as summarized in Algorithm 1.

Alternatively, the channel and the symbols on each sub-carrier can be estimated by means of iterative or recursive LS algorithms. Similar algorithms were proposed in [23] and [24] for blind source separation on a single subcarrier. We extend two of the algorithms presented in [24] that are based on projection to our application. We have proposed an extension of these algorithm using enumeration in [14]. In this paper, our focus is on the algorithms using projection since that they are computationally less expensive than the algorithms based on enumeration.

The algorithm ILSP (Iterative Least-Squares with Projection) summarized in Algorithm 2 is an iterative solution based on LS. It is initialized with the pilot based channel estimate, the maximum number of iterations (maxIteration), and the minimum error difference between two consecutive updates (minErr). The ILSP algorithm is essentially an iterative version of the ZF algorithm, where in each iteration the estimated symbols are projected onto the finite alphabet Ω of the transmitted symbols. This finite alphabet depends on the modulation type and the modulation order M_o . Details regarding the convergence for different finite alphabets are discussed in [24]. To estimate the symbols, we compute a pseudo inverse of the channel which leads to the condition $M_R \geq M_T$. The algorithm conditionally updates the channel depending on the rank of the estimated symbol matrix, i.e., the channel is updated if the rank of the symbol matrix $\tilde{\mathcal{S}}_n \in \mathbb{C}^{M_T \times K}$ is M_T , $K \geq M_T$. Note that this is not possible

Algorithm 2: Iterative Least-Squares with Projection (ILSP)

```

initialization  $\tilde{\mathcal{H}}_p$ , maxIteration, minErr
for  $n = 1 : N$  do
  set  $i = 1$ ,  $e = \infty$ 
  while  $i < \text{maxIteration}$  or  $e < \text{minErr}$  do
    |  $\tilde{\mathcal{S}}_n^{(i)} = (\tilde{\mathbf{H}}_n^{(i-1)\text{H}} \tilde{\mathbf{H}}_n^{(i-1)})^{-1} \tilde{\mathbf{H}}_n^{(i-1)\text{H}} \tilde{\mathbf{Y}}_n$ 
    |  $\tilde{\mathcal{S}}_n = \text{proj}(\tilde{\mathcal{S}}_n^{(i)})$ 
    if  $\text{rank}(\tilde{\mathcal{S}}_n^{(i)}) = M_T$  then
      |  $\tilde{\mathbf{H}}_n^{(i)} = \tilde{\mathbf{Y}}_n \tilde{\mathcal{S}}_n^{(i)\text{H}} (\tilde{\mathcal{S}}_n^{(i)} \tilde{\mathcal{S}}_n^{(i)\text{H}})^{-1}$ 
    else
      |  $\tilde{\mathbf{H}}_n^{(i)} = \tilde{\mathbf{H}}_n^{(i-1)}$ 
    end
    |  $i = i + 1$ ,  $e = \|\tilde{\mathbf{H}}_n^{(i-1)} - \tilde{\mathbf{H}}_n^{(i)}\|_F^2$ 
  end
end
Result:  $\tilde{\mathcal{S}}$  and  $\tilde{\mathcal{H}}$ 

```

for all values of M_T , K , and for all patterns of random data symbols from a finite distribution.

Algorithm 3: Recursive Least-Squares with Projection (RLSP)

```

initialization  $\tilde{\mathcal{H}}_p$ ,  $0 \leq \alpha \leq 1$ 
for  $n = 1 : N$  do
  |  $\tilde{\mathcal{S}}_n = (\tilde{\mathbf{H}}_n^{\text{H}} \tilde{\mathbf{H}}_n)^{-1} \tilde{\mathbf{H}}_n^{\text{H}} \tilde{\mathbf{Y}}_n$ 
  |  $\tilde{\mathcal{S}}_n = \text{proj}(\tilde{\mathcal{S}}_n)$ 
  set  $\mathbf{P}^{(0)} = \mathbf{I}_{M_T}$ ,  $\tilde{\mathbf{H}}_n^{(0)} = \tilde{\mathbf{H}}_n$ ,  $\alpha = 1$ 
  for  $k = 1 : K$  do
    |  $\mathbf{s} = \tilde{\mathcal{S}}_n(:,k)$ 
    |  $\tilde{\mathbf{H}}_n^{(k)} = \tilde{\mathbf{H}}_n^{(k-1)} + \frac{(\tilde{\mathbf{Y}}_n(:,k) - \tilde{\mathbf{H}}_n^{(k-1)} \mathbf{s})}{\alpha + \mathbf{s}^{\text{H}} \mathbf{P}^{(k-1)} \mathbf{s}} \mathbf{s}^{\text{H}} \mathbf{P}^{(k-1)}$ 
    |  $\mathbf{P}'^{(k)} = \frac{1}{\alpha} \left( \mathbf{P}'^{(k-1)} - \frac{\mathbf{P}'^{(k-1)} \mathbf{s} \mathbf{s}^{\text{H}} \mathbf{P}'^{(k-1)}}{\alpha + \mathbf{s}^{\text{H}} \mathbf{P}'^{(k-1)} \mathbf{s}} \right)$ 
  end
end
Result:  $\tilde{\mathcal{S}}$  and  $\tilde{\mathcal{H}}$ 

```

The second algorithm, namely RLSP (Recursive Least-Squares with Projections) is a recursive implementation of ILSP. We summarize this algorithm in Algorithm 3. The channel is estimated based on RLS (Recursive Least Squares), where α is the weighting coefficient and \mathbf{P}' denotes the inverse correlation matrix. Due to the computation of the pseudo-inverse of the channel matrix for the algorithm RLSP, $M_R \geq M_T$ should hold.

ILSP has the same computational complexity as traditional ZF receivers with the added complexity of the additional iterations if the symbol matrix has full rank. The RLSP algorithm requires a finite number of iterations that is equal to NK . Taking into account that the multiplication of two complex matrices of sizes $m \times n$ and $n \times l$ requires approximately $2nml$

operations and an inversion of a matrix of size $r \times r$ requires $\frac{2}{3}r^3$ operations, the ZF algorithm requires $\frac{2}{3}M_T^3 + 4M_T^2M_R + 2M_TM_R^2K$ operations per subcarrier. The ILSP algorithm requires $\frac{4}{3}M_T^3 + 6M_T^2M_R + 2M_TM_R^2K + 2M_T^2K + 2M_TM_RK$ operations per subcarrier and iteration. Finally, the RLSP algorithm requires $\frac{2}{3}M_T^3 + 4M_T^2M_R + 2M_TM_R^2K + 2M_TM_RK + 2M_T^2M_RK + 2M_T^2K$ operations per subcarrier and iteration. Therefore, the ILSP algorithm is less complex than the RLSP if the number of antennas is small. However, with the increase of the number of antennas the RLSP algorithm is less complex than the ILSP algorithm.

IV. KHATRI-RAO CODED MIMO-OFDM

In this section, we model a Khatri-Rao coded MIMO-OFDM communication system as a double tensor contraction between a channel and a signal tensor that contains coded symbols. This double tensor contraction is essentially equivalent to the model in (4). However, we assume that the signal tensor contains Khatri-Rao coded symbols.

As in the previously presented MIMO-OFDM model without coding (see Section III), we assume a MIMO-OFDM communication system with M_T transmit and M_R receive antennas. One OFDM block consists of N samples, which equals the DFT length. Moreover, all N subcarriers are used for data transmission. Furthermore, we assume a frequency-selective channel model that stays constant over the transmission of P frames. In contrast to the model presented in Section III, here, we assume that the P frames are divided into K groups of Q blocks (Q corresponds to the spreading factor), $P = K \cdot Q$.

Accordingly, the received signal in the frequency domain is given by

$$\tilde{\mathbf{Y}} = \tilde{\mathcal{H}} \bullet_{2,4}^{1,2} \tilde{\mathcal{X}} + \tilde{\mathcal{N}} = \tilde{\mathbf{Y}}_0 + \tilde{\mathcal{N}} \in \mathbb{C}^{N \times M_R \times K \times Q}, \quad (17)$$

where $\tilde{\mathcal{H}} \in \mathbb{C}^{N \times N \times M_R \times M_T}$ is the channel tensor and $\tilde{\mathcal{X}} \in \mathbb{C}^{N \times M_T \times K \times Q}$ is the signal tensor. The tensor $\tilde{\mathcal{N}} \in \mathbb{C}^{N \times M_R \times K \times Q}$ contains additive white Gaussian noise and $\tilde{\mathbf{Y}}_0 \in \mathbb{C}^{N \times M_R \times K \times Q}$ is the noiseless received signal.

A. Channel tensor

In this section, we use the model of the channel tensor $\tilde{\mathcal{H}}$ defined in equation (9). Moreover, we have defined the generalized unfolding $[\tilde{\mathcal{H}}]_{([1,3],[2,4])}$ in equation (10). Using a permutation matrix, it can be shown that the generalized unfolding $[\tilde{\mathcal{H}}]_{([1,3],[4,2])}$ of the channel is equal to

$$[\tilde{\mathcal{H}}]_{([1,3],[4,2])} = \bar{\mathbf{H}} \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T), \quad (18)$$

where

$$\bar{\mathbf{H}} = \underbrace{\begin{bmatrix} \tilde{\mathbf{H}}_R^{(1)} & \dots & \tilde{\mathbf{H}}_R^{(M_T)} \end{bmatrix}}_{\bar{\mathbf{H}}} \cdot \mathbf{P} \in \mathbb{C}^{M_R \times M_T N}.$$

The permutation matrix $\mathbf{P} \in \mathbb{R}^{N M_T \times M_T N}$ reorders the columns such that the faster increasing index is M_T instead of N and it is defined as $[\tilde{\mathcal{H}}]_{([1,3],[4,2])} = [\tilde{\mathcal{H}}]_{([1,3],[2,4])} \cdot \mathbf{P}$. Recall that the matrices $\tilde{\mathbf{H}} \in \mathbb{C}^{M_R \times N M_T}$ and $\tilde{\mathbf{H}}_R^{(m_T)} \in \mathbb{C}^{M_R \times N}$ are defined in equation (11). The structure of the 4-way channel tensor in the frequency domain $\tilde{\mathcal{H}}$ and its unfoldings are derived in Appendix A.

B. Data transmission

We can impose a CP structure to the transmit signal tensor, if we assume Khatri-Rao coded symbols [16], [17]. The coding is proportional to the number of transmit antennas if we use a spreading factor $Q = M_T$, for each subcarrier $n = 1, 2, \dots, N$. Hence, the generalized unfolding of the signal tensor is

$$[\tilde{\mathcal{X}}]_{([2,1],[4,3])} = [\tilde{\mathbf{S}}_1 \diamond \mathbf{C}_1 \quad \tilde{\mathbf{S}}_2 \diamond \mathbf{C}_2 \quad \dots \quad \tilde{\mathbf{S}}_N \diamond \mathbf{C}_N]^T \\ = \mathbf{I}_{M_T N} (\bar{\mathbf{S}} \diamond \bar{\mathbf{C}})^T, \quad (19)$$

where the matrix $\tilde{\mathbf{S}}_n \in \mathbb{C}^{K \times M_T}$ contains modulated data symbols and $\mathbf{C}_n \in \mathbb{C}^{Q \times M_T}$ is a Vandermonde coding matrix as defined in [16]. The matrices $\bar{\mathbf{S}} = [\tilde{\mathbf{S}}_1 \quad \dots \quad \tilde{\mathbf{S}}_N] \in \mathbb{C}^{K \times M_T N}$ and $\bar{\mathbf{C}} = [\mathbf{C}_1 \quad \dots \quad \mathbf{C}_N] \in \mathbb{C}^{Q \times M_T N}$ contain all symbol and coding matrices for each subcarrier, respectively. Note that $\bar{\mathbf{S}} = \tilde{\mathbf{S}} \cdot \mathbf{P}$, where the matrix $\tilde{\mathbf{S}}$ is defined in equation (12) and $\mathbf{P} \in \mathbb{R}^{N M_T \times M_T N}$ is the above mentioned permutation matrix that reorders the columns such that the faster increasing index is M_T instead of N . Moreover, we assume that $\tilde{\mathbf{S}}$ contains pilot symbols as explained after equation (12). As shown in [16] and as directly follows from (19), the tensor $[\tilde{\mathcal{X}}]_{([2,1],[3,4])}$ satisfies the following CP decomposition

$$[\tilde{\mathcal{X}}]_{([2,1],[3,4])} = \mathcal{I}_{3, M_T N} \times_1 \mathbf{I}_{M_T N} \times_2 \bar{\mathbf{S}} \times_3 \bar{\mathbf{C}}.$$

C. Receiver Design

Using equations (2), (3), and (17) the noiseless received signal can be expressed as

$$[\tilde{\mathbf{Y}}_0]_{([1,2],[4,3])} = [\tilde{\mathcal{H}}]_{([1,3],[4,2])} \cdot [\tilde{\mathcal{X}}]_{([2,1],[4,3])}.$$

Inserting the corresponding unfoldings of the channel and the signal tensor in equation (18) and (19), respectively, the noiseless received signal in the frequency domain is given by

$$[\tilde{\mathbf{Y}}_0]_{([1,2],[4,3])} = (\bar{\mathbf{H}} \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T)) \cdot (\bar{\mathbf{S}} \diamond \bar{\mathbf{C}})^T.$$

The above equation represents an unfolding of a 4-way tensor with a CP structure. Therefore, the noiseless received signal tensor can be expressed as

$$\tilde{\mathbf{Y}}_0 = \mathcal{I}_{4, M_T N} \times_1 (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T) \times_2 \bar{\mathbf{H}} \times_3 \bar{\mathbf{S}} \times_4 \bar{\mathbf{C}} \in \mathbb{C}^{N \times M_R \times K \times Q}. \quad (20)$$

Equation (20) represents the received signal in the frequency domain for all N subcarriers, M_R receive antennas, and P frames after the removal of the cyclic prefix. Depending on the available *a priori* knowledge at the receiver side, channel estimation, symbol estimation, or joint channel and symbol estimation can be performed.

Let us compare the MIMO-OFDM tensor model and the Khatri-Rao coded MIMO-OFDM tensor model in equations (15) and (20), respectively. First, the factor matrices in these equations have different index orderings. In equation (15) the faster increasing index in N , whereas in equation (20) the faster increasing index in M_T along the columns of the factor matrices. We use \sim and $-$ to distinguish the different index orderings of the factor matrices. Recall that we have defined a permutation matrix \mathbf{P} that considers the reordering of the columns of the factor matrices. Moreover, equation (20) has an additional tensor dimension (the 4-mode) corresponding to

the coding technique and the spreading factor Q . Furthermore, taking into account the permutation matrix \mathbf{P} , we get equation (15) from equation (20) for $Q = 1$ and $\bar{\mathbf{C}} = \mathbf{1}_{M_T N}^T$ (i.e., no coding and the spreading factor equals one).

Using equation (20), the channel and the data symbols can be jointly estimated from the $([1,4],[3,2])$ generalized unfolding of the noise corrupted received signal

$$[\tilde{\mathcal{Y}}]_{([1,4],[3,2])} \approx (\bar{\mathbf{C}} \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T)) \cdot (\bar{\mathbf{H}} \diamond \bar{\mathbf{S}})^T.$$

Under the assumption that $Q = M_T$, $(\bar{\mathbf{C}} \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T)) \in \mathbb{C}^{N \times M_T N}$ is a block diagonal, left invertible matrix and known at the receiver. Using the properties of the coding matrices defined in [16], i.e., $\mathbf{C}_n^H \mathbf{C}_n = M_T \mathbf{I}_{M_T}$, we have

$$\bar{\mathbf{Y}} \triangleq \frac{1}{M_T} (\bar{\mathbf{C}} \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T))^H \cdot [\tilde{\mathcal{Y}}]_{([1,4],[3,2])} \approx (\bar{\mathbf{H}} \diamond \bar{\mathbf{S}})^T.$$

After transposition, $\bar{\mathbf{Y}}^T \approx \bar{\mathbf{H}} \diamond \bar{\mathbf{S}}$ can be approximated by the Khatri-Rao product between the channel and the data symbols. Therefore, the channel and the data symbols can be jointly estimated based on the LSKRF as in [25].

Using the LSKRF, the matrices $\bar{\mathbf{H}}$ and $\bar{\mathbf{S}}$ can be identified up to one complex scaling factor ambiguity per column. Hence, the estimated matrices satisfy the following relations

$$\hat{\bar{\mathbf{H}}} = \bar{\mathbf{H}} \cdot \mathbf{\Lambda}, \quad (21)$$

$$\hat{\bar{\mathbf{S}}} = \bar{\mathbf{S}} \cdot \mathbf{\Lambda}^{-1}, \quad (22)$$

where $\mathbf{\Lambda} \in \mathbb{C}^{M_T N \times M_T N}$ is a diagonal matrix with diagonal elements equal to the $M_T N$ complex scaling ambiguities. The simplest way to resolve the scaling ambiguity is by assuming the knowledge of one row of the matrix $\bar{\mathbf{S}} \in \mathbb{C}^{K \times M_T N}$. This corresponds to $M_T N$ pilot symbols, i.e., one pilot symbol per transmit antenna and subcarrier. Since traditional MIMO-OFDM communication systems use less pilot symbols than $M_T N$, we propose to use the same amount of pilot symbols and exploit the channel correlation between adjacent subcarriers in order to estimate the scaling matrix. We transmit pilot symbols on positions with equidistant spacing in the frequency and the time domain. With the prior knowledge of the pilot symbols and their positions, we can obtain an initial channel estimate as in traditional MIMO-OFDM systems (see Section III). We denote this pilot based channel estimate by $\hat{\mathcal{H}}_p$ ($\bar{\mathbf{H}}_p$). The pilot based channel estimate is then used to estimate the scaling ambiguity $\mathbf{\Lambda}$ in equation (21) as

$$\hat{\mathbf{\Lambda}} = \text{diag} \left(\frac{1}{M_R} \sum_{m_R=1}^{M_R} \hat{\bar{\mathbf{H}}}_{(m_R, \cdot)} \oslash \bar{\mathbf{H}}_{p(m_R, \cdot)} \right).$$

By multiplying the solution of the LSKRF with the diagonal matrix $\hat{\mathbf{\Lambda}}$, the scaling ambiguity in equation (22) is resolved and the data symbols can be demodulated. Note that the proposed Khatri-Rao receiver estimates the channel and the symbols in a semi-blind fashion. First, the channel and the symbols are jointly estimated without any *a priori* information. The pilot based channel estimate is then used to resolve the scaling ambiguity affecting the columns of $\hat{\bar{\mathbf{H}}}$ and $\hat{\bar{\mathbf{S}}}$. Therefore, the optimal length and repetition of the piloting sequences are identical as for the traditional OFDM systems.

Algorithm 4: Khatri-Rao (KR) receiver

Initialization: $\hat{\mathcal{H}}_p$ and $\bar{\mathbf{C}}$;

1. Compute $\bar{\mathbf{Y}} = \frac{1}{M_T} (\bar{\mathbf{C}} \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T))^H \cdot [\tilde{\mathcal{Y}}]_{([1,4],[3,2])}$.

2. Compute the LSKRF of $\bar{\mathbf{Y}}^T$ using the algorithm proposed in [25] which gives $\hat{\bar{\mathbf{H}}}$ and $\hat{\bar{\mathbf{S}}}$;

3. Compute the scaling matrix

$\hat{\mathbf{\Lambda}} = \text{diag} \left(\frac{1}{M_R} \sum_{m_R=1}^{M_R} \hat{\bar{\mathbf{H}}}_{(m_R, \cdot)} \oslash \bar{\mathbf{H}}_{p(m_R, \cdot)} \right)$. The matrix $\bar{\mathbf{H}}_{p(m_R, \cdot)}$ is defined as in equation (18) using the estimated channel tensor $\hat{\mathcal{H}}_p$;

4. Resolve the scaling ambiguity $\bar{\mathbf{H}} = \hat{\bar{\mathbf{H}}} \cdot \hat{\mathbf{\Lambda}}^{-1}$ and $\bar{\mathbf{S}} = \hat{\bar{\mathbf{S}}} \cdot \hat{\mathbf{\Lambda}}$.

Result: $\bar{\mathbf{S}}$ and $\bar{\mathbf{H}}$

We summarize the steps of the proposed Khatri-Rao (KR) receiver in Algorithm 4.

Furthermore, the channel estimate resulting from the KR receiver can be used for channel tracking in future transmission frames if the channel has not changed drastically. If the channel estimate is used for tracking, it could be improved by means of an additional LS estimate from $[\tilde{\mathcal{Y}}]_{([2,4,1],[3])}$ with the knowledge of the estimated and projected symbols onto the finite alphabet Ω , i.e., $Q(\bar{\mathbf{S}}) = \text{proj}(\bar{\mathbf{S}})$. The finite alphabet Ω depends on the modulation type and the modulation order M_O .

$$\hat{\mathbf{H}}_{LS}^T = ((\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T) \diamond \bar{\mathbf{C}} \diamond Q(\bar{\mathbf{S}}))^+ \cdot [\tilde{\mathcal{Y}}]_{([2,4,1],[3])}$$

However, we can also use this improved channel estimation to further improve the performance of the KR receiver. Using this updated channel estimate an improved estimate of the diagonal scaling matrix $\hat{\mathbf{\Lambda}}$ can be calculated and with that an enhanced estimate of the symbols, $\hat{\bar{\mathbf{S}}}_{LS}$, using equation (22). Note that, instead of just one LS estimate of the channel and the symbols the performance can be further enhanced with additional iterations leading to an iterative receiver. Note that the symbol matrix $\hat{\bar{\mathbf{S}}}_{LS}$ can be estimated in the least squares sense from the 3-mode unfolding of equation (20), but the estimation of $\hat{\mathbf{\Lambda}}$ is computationally cheaper. The KR receiver with its enhancement via LS is summarized in Algorithm 5.

Algorithm 5: Khatri-Rao receiver and its enhancement via Least-Squares (KR+LS)

1. Apply Algorithm 4

2. Project the symbols onto the finite alphabet Ω , i.e., $Q(\bar{\mathbf{S}}) = \text{proj}(\bar{\mathbf{S}})$;

3. Compute an enhanced channel estimate

$\hat{\mathbf{H}}_{LS}^T = ((\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T) \diamond \bar{\mathbf{C}} \diamond Q(\bar{\mathbf{S}}))^+ \cdot [\tilde{\mathcal{Y}}]_{([2,4,1],[3])}$;

4. Improve the estimate of the diagonal scaling matrix

$\hat{\mathbf{\Lambda}}_{LS} = \text{diag} \left(\frac{1}{M_R} \sum_{m_R=1}^{M_R} \hat{\bar{\mathbf{H}}}_{(m_R, \cdot)} \oslash \hat{\mathbf{H}}_{LS(m_R, \cdot)} \right)$;

5. Compute an enhanced estimate of the symbol matrix

$\hat{\bar{\mathbf{S}}}_{LS} = \bar{\mathbf{S}} \cdot \hat{\mathbf{\Lambda}}_{LS}$.

Result: $\hat{\bar{\mathbf{S}}}_{LS}$ and $\hat{\mathbf{H}}_{LS}$

Due to the additional LS based estimates the KR+LS algorithm has higher computational complexity than the KR

algorithm.

V. RANDOMLY CODED MIMO-OFDM

In Section IV, we have proposed a tensor model for KR coded MIMO-OFDM systems that introduces an additional CP structure to the signal tensor. The additional CP structure of the signal tensor is achieved by means of Khatri-Rao coding. However, using the Khatri-Rao coding, we add additional spreading that reduces the spectral efficiency of the system. Therefore, in this section we propose to keep the CP structure of the signal tensor proposed in Section IV, but to introduce random coding. We introduce the random coding, which means that the "coding matrix" also contains data symbols.

As in Section IV, the received signal in the frequency domain after the removal of the cyclic prefix is given by

$$\tilde{\mathcal{Y}} = \tilde{\mathcal{H}} \bullet_{4,2}^{2,1} \tilde{\mathcal{X}} + \tilde{\mathcal{N}} = \tilde{\mathcal{Y}}_0 + \tilde{\mathcal{N}} \in \mathbb{C}^{N \times M_R \times K \times Q}, \quad (23)$$

where $\tilde{\mathcal{H}} \in \mathbb{C}^{N \times N \times M_R \times M_T}$ is the channel tensor and $\tilde{\mathcal{X}} \in \mathbb{C}^{N \times M_T \times K \times Q}$ is the signal tensor. The tensor $\tilde{\mathcal{N}}$ contains additive white Gaussian noise and $\tilde{\mathcal{Y}}_0$ is the noiseless received signal. As for the KR coded MIMO-OFDM system, we transmit $P = KQ$ frames that are divided into K groups of Q blocks ("spreading factor"). The number of subcarriers is N , and M_R and M_T denote the number of receive and transmit antennas, respectively.

A. Channel tensor

We model the channel tensor $\tilde{\mathcal{H}}$ according to equation (9). Details regarding this model are also provided in Appendix A. In this section, we use the generalized unfolding $[\tilde{\mathcal{H}}]_{([1,3],[4,2])} = \bar{\mathbf{H}} \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T)$ that is defined in (18).

B. Data Transmission

As previously mentioned, we impose a CP structure on the signal tensor $\tilde{\mathcal{X}}$ similar to the Khatri-Rao coding proposed in Section IV. For the generalized unfolding $([2,1],[4,3])$ of the signal tensor, we have

$$[\tilde{\mathcal{X}}]_{([2,1],[4,3])} = [\bar{\mathbf{S}}_1 \diamond \mathbf{C}'_1 \quad \bar{\mathbf{S}}_2 \diamond \mathbf{C}'_2 \quad \dots \quad \bar{\mathbf{S}}_N \diamond \mathbf{C}'_N]^T \\ = \mathbf{I}_{M_T N} (\bar{\mathbf{S}} \diamond \bar{\mathbf{C}}')^T, \quad (24)$$

where the matrix $\bar{\mathbf{S}}_n \in \mathbb{C}^{K \times M_T}$ contains modulated data symbols. In contrast to the Khatri-Rao coding in Section IV, here, we assume that the first row of the matrix $\mathbf{C}'_n \in \mathbb{C}^{Q \times M_T}$ contains only ones, whereas the remaining $Q-1$ rows contain modulated data symbols. Hence, the "coding matrix" (the matrix \mathbf{C}'_n in (19) represents the coding matrix) contains also random entries. We refer to this transmission technique as random coding. Moreover, the matrices $\bar{\mathbf{S}} = [\bar{\mathbf{S}}_1 \quad \dots \quad \bar{\mathbf{S}}_N] \in \mathbb{C}^{K \times M_T N}$ and $\bar{\mathbf{C}}' = [\mathbf{C}'_1 \quad \dots \quad \mathbf{C}'_N] \in \mathbb{C}^{Q \times M_T N}$ contain symbol and random coding matrices for each subcarrier, respectively. Note that $\bar{\mathbf{S}}$ is defined as in Section IV, i.e., $\bar{\mathbf{S}} = \tilde{\mathbf{S}} \cdot \mathbf{P}$, where the matrix $\tilde{\mathbf{S}}$ is defined in equation (12) and $\mathbf{P} \in \mathbb{R}^{N M_T \times M_T N}$ is the permutation matrix that reorders the columns such that the faster increasing index is M_T instead of N . Moreover, we assume that $\bar{\mathbf{S}}$ contains pilot symbols

as explained after equation (12). As shown in [16] and as directly follows from (24), the tensor $[\tilde{\mathcal{X}}]_{([2,1],[3,4])}$ satisfies the following CP decomposition

$$[\tilde{\mathcal{X}}]_{([2,1],[3,4])} = \mathcal{I}_{3, M_T N} \times_1 \mathbf{I}_{M_T N} \times_2 \bar{\mathbf{S}} \times_3 \bar{\mathbf{C}}'.$$

C. Receiver Design

Using equations (2) and (23), for the noiseless received signal, we get

$$[\tilde{\mathcal{Y}}_0]_{([1,2],[4,3])} = [\tilde{\mathcal{H}}]_{([1,3],[4,2])} \cdot [\tilde{\mathcal{X}}]_{([2,1],[4,3])}. \quad (25)$$

Inserting the corresponding unfoldings of the channel tensor and the signal tensor, i.e., inserting (18) and (24) into (25), we obtain

$$[\tilde{\mathcal{Y}}_0]_{([1,2],[4,3])} = (\bar{\mathbf{H}} \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T)) \cdot (\bar{\mathbf{S}} \diamond \bar{\mathbf{C}}')^T.$$

The above equation represents an unfolding of a 4-way tensor with CP structure. Therefore, it can be expressed as

$$\tilde{\mathcal{Y}}_0 = \mathcal{I}_{4, M_T N} \times_1 (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T) \times_2 \bar{\mathbf{H}} \times_3 \bar{\mathbf{S}} \times_4 \bar{\mathbf{C}}' \in \mathbb{C}^{N \times M_R \times K \times Q}. \quad (26)$$

Equation (26) represents the noiseless received signal in the frequency domain for all N subcarriers, M_R receive antennas, and P frames after the removal of the cyclic prefix for MIMO-OFDM system with RC (Random Coding). Note that the CP decomposition in (26) is degenerate in all four modes.

Depending on the available *a priori* knowledge at the receiver side, channel estimation, symbol estimation, or joint channel and symbol estimation can be performed. For instance, from the 3-mode unfolding of the tensor \mathcal{Y}_0 in (26), we can obtain

$$\bar{\mathbf{S}} = [\mathcal{Y}_0]_{(3)} \cdot \left[(\bar{\mathbf{C}}' \diamond \bar{\mathbf{H}} \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T))^T \right]^+, \quad (27)$$

provided that $M_R Q \geq M_T$. Moreover, from the 4-mode unfolding and 2-mode unfolding of tensor \mathcal{Y}_0 in (26), we can obtain $\bar{\mathbf{C}}'$ and $\bar{\mathbf{H}}$, respectively.

$$\bar{\mathbf{C}}' = [\mathcal{Y}_0]_{(4)} \cdot \left[(\bar{\mathbf{S}} \diamond \bar{\mathbf{H}} \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T))^T \right]^+ \quad (28)$$

$$\bar{\mathbf{H}} = [\mathcal{Y}_0]_{(2)} \cdot \left[(\bar{\mathbf{C}}' \diamond \bar{\mathbf{S}} \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T))^T \right]^+ \quad (29)$$

Note that we can compute $\bar{\mathbf{C}}'$ via a pseudo-inverse if $M_R K \geq M_T$. This condition is a design requirement for the proposed MIMO-OFDM system with random coding.

For noisy observations such as (23), the equations (27)-(29) hold approximately. In this case, we can use the equations (27)-(29) to estimate the symbols and the channel in an ALS fashion. However, there is no guarantee of convergence if we initialize the ALS algorithm randomly. Therefore, we propose to use the pilot based channel estimate $\bar{\mathbf{H}}_p$ to obtain initial estimates of the matrices $\bar{\mathbf{S}}$ and $\bar{\mathbf{C}}'$ based on LSKRF. This pilot based channel estimate is obtained from the pilot symbols in $\bar{\mathbf{S}}$ and the first row of $\bar{\mathbf{C}}'$ that has entries equal to one. From the $([3,4],[1,2])$ generalized unfolding of the noisy observation \mathcal{Y} , we get

$$[\tilde{\mathcal{Y}}]_{([3,4],[1,2])} \approx [\bar{\mathbf{C}}' \diamond \bar{\mathbf{S}}] \cdot [\bar{\mathbf{H}}_p \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T)]^T.$$

Algorithm 6: Random Coding-Khatri-Rao (RC-KR) receiver

 Initialization: $\bar{\mathcal{H}}_p$;

1. Compute

$$\bar{\mathbf{Y}} = [\mathcal{Y}]_{([3,4],[1,2])} \cdot \left[\left(\bar{\mathbf{H}}_p \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T) \right)^T \right]^+;$$

 2. Compute the LSKRF of $\bar{\mathbf{Y}}$ using the algorithm proposed in [25] which gives $\hat{\mathbf{C}}'$ and $\hat{\mathbf{S}}$;

 3. Compute the scaling matrix $\hat{\mathbf{A}} = \text{diag} \left(\hat{\mathbf{C}}'_{(1,\cdot)} \odot \bar{\mathbf{C}}'_{(1,\cdot)} \right)$.
 (The first row of the matrix $\bar{\mathbf{C}}'$ contains only ones);

 4. Resolve the scaling ambiguity $\bar{\mathbf{C}}' = \hat{\mathbf{C}}' \cdot \hat{\mathbf{A}}^{-1}$ and
 $\bar{\mathbf{S}} = \hat{\mathbf{S}} \cdot \hat{\mathbf{A}}$;

Result: $\bar{\mathbf{S}}$ and $\bar{\mathbf{C}}'$

Algorithm 7: Random Coding-Khatri-Rao + ALS (RC-KR+ALS) receiver

Apply Algorithm 6;

while does not exceed the maximum number of iterations, does not reach a predefined minimum, or the error of the cost function has not changed within two consecutive iterations **do**
if $\text{rank} \left(\left[\bar{\mathbf{C}}' \diamond \bar{\mathbf{S}} \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T) \right]^T \right) = M_T N$ **then**

Update

$$\hat{\mathbf{H}} = [\mathcal{Y}_0]_{(2)} \cdot \left[\left(\hat{\mathbf{C}}' \diamond \hat{\mathbf{S}} \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T) \right)^T \right]^+$$

else

 keep the previous estimate of $\hat{\mathbf{H}}$;

end

 Update $\hat{\mathbf{C}}' = [\mathcal{Y}_0]_{(4)} \cdot \left[\left(\hat{\mathbf{S}} \diamond \hat{\mathbf{H}} \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T) \right)^T \right]^+$;

 Project $\hat{\mathbf{C}}' = \text{proj} \left(\hat{\mathbf{C}}' \right)$ onto the finite alphabet Ω ;

 Update $\hat{\mathbf{S}} = [\mathcal{Y}_0]_{(3)} \cdot \left[\left(\hat{\mathbf{C}}' \diamond \hat{\mathbf{H}} \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T) \right)^T \right]^+$;

 Project $Q(\hat{\mathbf{S}}) = \text{proj} \left(\hat{\mathbf{S}} \right)$ onto the finite alphabet Ω .

end
Result: $\hat{\mathbf{S}}$, $\hat{\mathbf{C}}'$, and $\hat{\mathbf{H}}$

Given $\bar{\mathbf{H}}_p$ and $M_R \geq M_T$, from $[\tilde{\mathbf{Y}}]_{([3,4],[1,2])} \cdot \left[\left(\bar{\mathbf{H}}_p \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T) \right)^T \right]^+ \approx [\bar{\mathbf{C}}' \diamond \bar{\mathbf{S}}]$ based on LSKRF, we obtain $\hat{\mathbf{S}}$ and $\hat{\mathbf{C}}'$. However, the matrices $\hat{\mathbf{S}}$ and $\hat{\mathbf{C}}'$ are estimated up to one complex scaling ambiguity per column. We exploit the first row of the matrix $\bar{\mathbf{C}}'$ to estimate this ambiguity (recall that the elements of the first row of the matrix $\bar{\mathbf{C}}'$ are set to one). After resolving the scaling ambiguity, we propose to iterate between the equations (27)-(29) to enhance the accuracy of the receiver. Hence, we propose two receivers RC-KR (Random Coding-Khatri-Rao) and RC-KR+ALS (Random Coding-Khatri-Rao+Alternating Least-Squares) for randomly coded MIMO-OFDM systems. These two algorithms are summarized in Algorithm 6 and Algorithm 7, respectively. The RC-KR receiver exploits the LSKRF to compute an estimate of the symbol matrices $\bar{\mathbf{S}}$ and $\bar{\mathbf{C}}'$, assuming that $M_R \geq M_T$, the

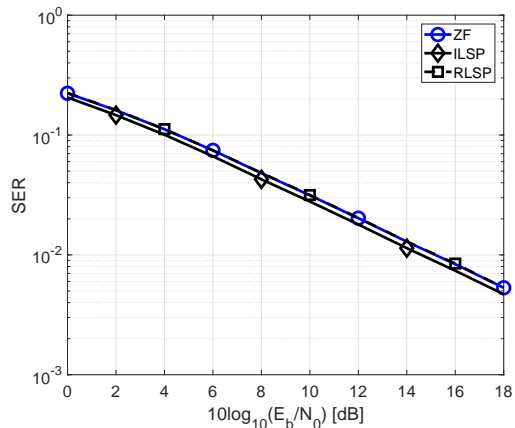


Fig. 5: SER for a 2×2 OFDM system with parameters $\Delta K = 8$, $\Delta F = 3$, $N = 128$, $K = 8$.

first row on the matrix $\bar{\mathbf{C}}'$ contains only ones, and a pilot based channel estimate $\bar{\mathbf{H}}_p$ is already available. The initial steps of the RC-KR receiver are equivalent to the RC-KR receiver. In a following steps using the RC-KR+ALS receiver, we estimate the channel matrix and both symbol matrices in an ALS fashion. Therefore, the RC-KR+ALS receiver exploits the LSKRF to initialize the ALS algorithm. The ALS algorithm is stopped if it exceeds the maximum number of iterations that is set to 5, reaches a predefined minimum of the cost function $\|\tilde{\mathbf{Y}} - \mathcal{I}_{4, M_T N} \times_1 (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T) \times_2 \hat{\mathbf{H}} \times_3 \hat{\mathbf{S}} \times_4 \hat{\mathbf{C}}'\|_{\mathbf{H}}^2 / \|\tilde{\mathbf{Y}}\|_{\mathbf{H}}^2$, or if the error of the cost function has not changed within two consecutive iterations. The RC-KR algorithm has a higher computational complexity than the RC-KR+ALS algorithm due to the additional ALS iterations, as shown in Algorithm 7.

VI. SIMULATION RESULTS

In this section, we evaluate the performance of the proposed receivers for MIMO-OFDM systems using Monte-Carlo simulations. First, we compare the performance of ZF, ILSP, and RLSP, (i.e., Algorithms 1-3) using 5000 realizations. We consider a 2×2 OFDM system, with K frames, and $N = 128$ subcarriers. The pilot symbols are transmitted on every third subcarrier such that $\Delta F = 3$ and only during the first frame, i.e, $\Delta K = K$. Using these pilots, we obtain a pilot based channel estimate with which we initialize all of the algorithms. The transmitted data symbols are independent and modulated using 4-QAM (Quadrature Amplitude Modulation). The frequency selective propagation channel is modeled according to the 3GPP (3rd Generation Partnership Project) Pedestrian A channel (Ped A) [26]. The duration of the cyclic prefix is 32 samples and the weighting factor $\alpha = 1$, for the recursive LS. The maximum number of iterations for the iterative algorithm is set to 7.

In Fig. 5 we depict the SER (Symbol Error Rate) as a function of the E_b/N_0 (energy per bit/ noise power spectral density) in dB for $K = 8$. The performance of ILSP and RLSP is similar to the ZF performance and it depends on the number of frames. As shown in Fig. 5, increasing the number of frames leads to a slightly better SER than using

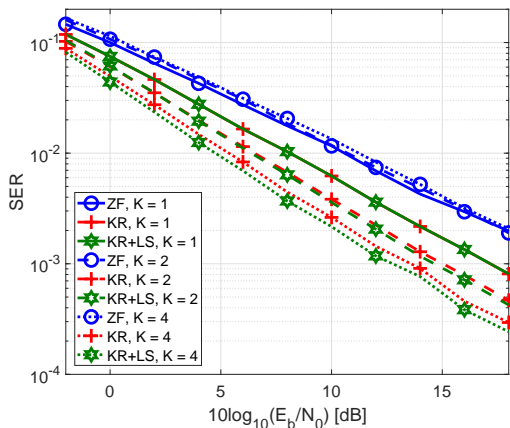


Fig. 6: SER comparison for different numbers of transmitted blocks.

a ZF receiver. Note that the transmitted data symbols are independent and randomly drawn with no guarantee that the matrices $\tilde{\mathbf{S}}_n$ are of rank M_T . Therefore, in many cases the number of iterations is equal to one. In all of the simulated cases the iterative algorithm converges after 3 iterations. As in [24], we also observe that the iterative algorithms have a better performance than the recursive one for an increased number of frames. However, the recursive algorithm, RSLP requires less computational complexity than the iterative one, ILSP.

Next, we compare the performance of the traditional frequency domain ZF receiver, the proposed Khatri-Rao (KR) receiver (see Algorithm 4) and the proposed Khatri-Rao receiver with one additional LS iteration (see Algorithm 5). In these simulations, we average the results over 5000 realizations and Ped A channel [26] are assumed. In Fig. 6, we depict the SER is as a function of E_b/N_0 for different numbers of transmitted blocks. In this case, we consider a MIMO system with the following parameters $N = 128$, $Q = 2$, $M_T = 2$, $M_R = 2$, $\Delta K = K$, $\Delta F = 4$ and different numbers of blocks K (the number of blocks is indicated in the legend). Note that the KR and the KR+LS receivers benefit from the increased number of frames as the channel has been kept constant during the $P = Q \cdot K$ frames. Moreover, as the number of frames increases, the advantages of the enhancement via LS become more pronounced.

In another experiment, we compare the receivers proposed in Section III (ILSP and RLSP) for a MIMO-OFDM system with the receivers proposed in Section IV for a Khatri-Rao coded MIMO-OFDM system. We assume that both systems have $N = 128$ subcarriers, $M_T = 2$ transmit antennas, and $M_R = 2$ receive antennas. Moreover, for both systems we assume that $\Delta F = 10$ is the subcarrier spacing between two pilot symbols in the frequency domain and $\Delta K = K$ is the spacing between two pilot symbols in the time domain. The OFDM system has the following parameters $K = 10$, $\Delta K = 10$, and the symbols are modulated using 4-QAM. The KR coded OFDM system has the following parameters $K = 5$, $\Delta K = 5$, $Q = 2$, $P = KQ = 10$ and the symbols are modulated using 16-QAM. Hence, we transmit 2 bits/symbol with both systems. In Fig. 7, we depict the SERs for these two systems.

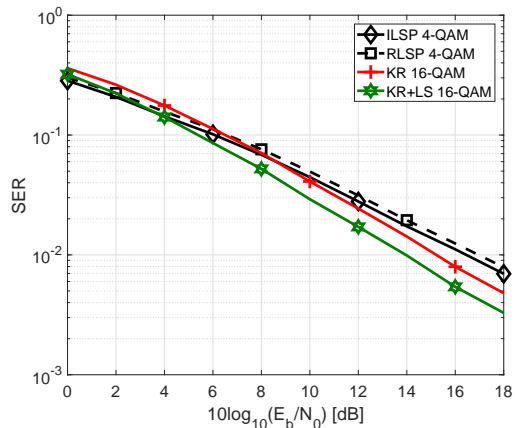


Fig. 7: SER for 2×2 OFDM and KR coded OFDM systems, $N = 128$ and $\Delta F = 10$. The OFDM system has the following parameters $K = 10$, $\Delta K = 10$, and the symbols are modulated using 4-QAM. The KR coded OFDM system has the following parameters $K = 5$, $\Delta K = 5$, $Q = 2$, $P = KQ = 10$ and the symbols are modulated using 16-QAM. Hence, both systems transmit 2 bits/symbol.

The KR receiver has similar accuracy to the ILSP and the RLSP algorithms that improves with the increased SNR. The KR+LS receiver outperforms the ILSP algorithm and the KR algorithm in terms of SER. Recall that the KR coded OFDM model in equation (20) has a richer tensor structure than the OFDM model in equation (15) due to the coding. The KR algorithm and the KR-LS algorithm effectively exploit this structure to estimate the channel and the symbols. Note that the KR-LS algorithm computes an improved estimate of the scaling matrix. Therefore, KR-LS leads to lower SER levels than the ILSP and KR algorithms.

In our next experiment, we evaluate the performance of the proposed RC-KR and RC-KR+ALS receivers for randomly coded MIMO-OFDM systems using Monte-Carlo simulations. We consider 2×2 systems, with $N = 128$ subcarriers, and $P = KQ$ frames. Moreover, the spacing between two pilot symbols in the time domain and in the frequency domain are denoted by ΔK and ΔF respectively. The frequency selective propagation channel is modeled according to the 3GPP Pedestrian A channel [26]. The duration of the cyclic prefix is 32 samples. In these simulations we consider 5000 realizations. In Fig. 8, we provide an SER comparison for two scenarios. For both scenarios, we assume 2×2 randomly coded OFDM system, $Q = 2$, and the symbols are modulated using 4-QAM modulation. Moreover, $K = 5$, $\Delta F = 10$, and $\Delta K = 5$, for the first scenario, whereas for the second scenario $K = 3$, $\Delta F = 5$, and $\Delta K = 3$. Hence, in the first scenario we estimate more symbols than in the second scenario, using less pilot symbols. As expected, we achieve a lower SER if more pilot symbols are used because they lead to a more accurate initial pilot based channel estimate. Moreover, in Fig. 8 we see that the RC-KR+ALS receiver outperforms the RC-KR receiver. Thus, we benefit from the additional iterations and from exploiting the complete tensor structure. In

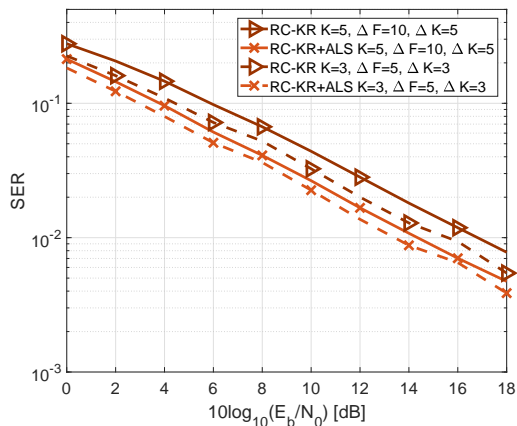


Fig. 8: SER for a 2×2 randomly coded OFDM system with parameters $N = 128$, $Q = 2$, K , ΔK , ΔF , and the symbols are modulated using 4-QAM. The parameters K , ΔK and ΔF are indicated in the legend.

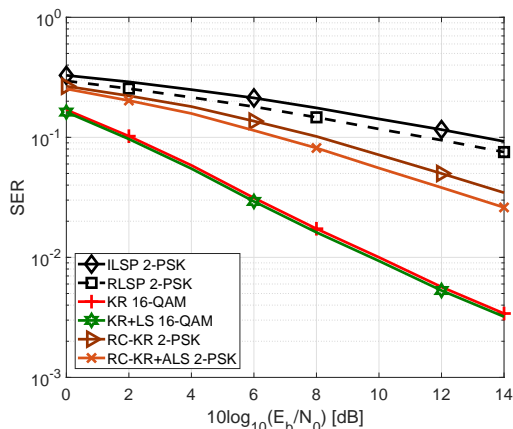


Fig. 9: SER for 4×4 KR coded OFDM, randomly coded OFDM, and traditional OFDM systems for $N = 128$ and $\Delta F = 10$. The KR coded OFDM system assumes $K = 2$, $\Delta K = 2$, $Q = 4$, $P = KQ = 8$ and the symbols are modulated using 16-QAM. The randomly coded OFDM system assumes $K = 2$, $\Delta K = 2$, $Q = 4$, $P = KQ = 8$ and the symbols are modulated using 2-PSK. The OFDM system has the following parameters $K = 8$, $\Delta K = 8$, and the symbols are modulated using 2-PSK.

contrast to RC-KR, RC-KR+ALS also estimates the channel matrix. Furthermore, the accuracy gain of the RC-KR+ALS receiver is more pronounced if we initialize the RC-KR+ALS with a less accurate pilot based channel estimate (the gain is more pronounced for the solid lines than for the dashed lines in Fig. 8).

Finally, in Fig. 9, we depict the SER performance for a 4×4 MIMO system, considering the following receivers: (i) ILSP receiver (Algorithm 2), (ii) RLSP receiver (Algorithm 3), (iii) KR receiver (Algorithm 4), (iv) KR-LS receiver (Algorithm 5), (v) RC-KR receiver (Algorithm 6), and (vi) RC-KR+ALS (Algorithm 7). To ensure a fair comparison in terms

of spectral efficiency, the following parameters were chosen for the different receivers: The KR coded OFDM system assumes $N = 128$, $\Delta F = 10$, $K = 2$, $\Delta K = 2$, $Q = 4$, $P = KQ = 8$ and the symbols are modulated using 16-QAM. For the RC coded OFDM system we assume $N = 128$, $\Delta F = 10$, $K = 2$, $\Delta K = 2$, $Q = 4$, $P = KQ = 8$ and the symbols are modulated using BPSK symbols. The OFDM system assumes $N = 128$, $\Delta F = 10$, $K = 8$, $\Delta K = 8$, and BPSK symbols. We see that the RC-KR receiver outperforms ILSP and RLSP algorithms. In addition, the KR and KR-LS receivers for KR-coded OFDM have different slopes than the uncoded OFDM and the randomly coded OFDM, exhibiting a better performance, as expected.

VII. CONCLUSION AND DISCUSSION

In this paper, we have first presented a tensor model for MIMO-OFDM systems using the generalized tensor contraction operator between a channel tensor and a transmit signal tensor and the new properties of this contraction operator. The proposed model is a very general and flexible way of describing the received signal in MIMO-OFDM systems for all subcarriers jointly. We have also proposed a new representation of the channel tensor using a 4-way tensor with a special BTD structure. The resulting tensor model of the received signal enables the design of the traditional ZF receiver and facilitates the design of an iterative LS receiver based on projections and a recursive version of the iterative receiver ILSP denoted by RLSP. These algorithms based on projections have a better performance than the ZF receiver if the number of transmitted frames is large enough. Moreover, the accuracy of the ILSP algorithm depends on the rank of the transmitted symbol matrices. Therefore, its performance in terms of the SER depends on the chosen modulation order and the modulation scheme. Hence, the system can be modified such that only specific code words are used. Moreover, the recursive algorithm can be modified such that it exploits the channel correlation in time varying scenarios. Note that we already exploit the correlation of the channel among adjacent subcarriers that leads to a reduced number of pilot symbols as compared to other tensor models.

Next, we use Khatri-Rao coding for the transmission of the OFDM symbols leading to Khatri-Rao coded MIMO-OFDM systems. The generalized tensor model using the contraction operator has been extended to the Khatri-Rao coded MIMO-OFDM system in a straightforward way. In this case, the transmit signal tensor has a CP structure. By exploiting the overall tensor model, we propose a receiver based on the LSKRF. This receiver requires the same amount of training symbols as traditional OFDM techniques, but it has an improved performance in terms of the SER. Hence, we benefit from the additional tensor structure of the transmitted signal to achieve a tensor gain. In addition, we have proposed to improve the performance of this receiver further by means of an additional LS iteration. Note that the Khatri-Rao coding strategy has a reduced spectral efficiency than the uncoded MIMO-OFDM system. Therefore, we propose an alternative transmission technique where the "coding matrices" contain

random data symbols. Thereby, this transmission technique also imposes a CP structure on the transmit signal tensor. Using the resulting received signal tensor, we have proposed two receivers for randomly coded MIMO-OFDM systems. The first proposed receiver RC-KR estimates the symbol matrices based on the LSKRF. The second proposed RC-KR+ALS receiver is an ALS algorithm initialized with the estimates of the symbol matrices using the RC-KR receiver. The proposed RC-KR+ALS algorithm outperforms the iterative receivers for MIMO-OFDM because it exploits the additional tensor structure of the signal tensor. Unlike the receivers for Khatri-Rao coded MIMO-OFDM, both receivers for the randomly coded MIMO-OFDM assume that $M_R \geq M_T$. However, the randomly coded system has a higher spectral efficiency than the Khatri-Rao coded system.

In the future, for the Khatri-Rao coded MIMO-OFDM system we can consider not just one additional LS iteration, but several iterations leading to an ALS based receiver initialized using the LSKRF. We should also consider the design of optimal orthogonal pilot sequences specific to the KR receiver. Moreover, we can consider recursive LS instead of LS in order to relax this condition. The randomly coded MIMO-OFDM system can be modified such that both symbol matrices contain symbols from different constellations and/or different modulation orders. This will lead to a resulting transmit signal tensor with diverse entries and potentially improved performance for the receivers in terms of the SER. We should investigate which combinations of modulation orders and modulation types are suitable for different SNRs. Furthermore, the general tensor model for multi-carrier systems proposed in this paper can be extended to other multi-carrier techniques such as UPMC and FBMC. Even more, this model can be extended to relay-assisted systems and multi-user systems, which would imply modeling the received signal tensor as a train of tensors linked by a sequence of double contractions, each hop representing one double contraction. Another promising perspective is to assume a low-rank structure of the channel tensor, which will be a more realistic assumption for millimeter wave MIMO-OFDM systems. Exploiting the double contraction model and the additional structure of the low-rank channel tensor would lead to new blind receivers.

APPENDIX

DERIVATION OF THE 4-WAY CHANNEL TENSOR IN THE FREQUENCY DOMAIN AND ITS UNFOLDINGS

Let us assume a MIMO-OFDM system with M_T transmit antennas and M_R receive antennas. Such a system is depicted in Fig. 3, for $M_T = 2$ and $M_R = 3$. As shown in Section III, we can define a 4-way channel tensor $\tilde{\mathcal{H}} \in \mathbb{C}^{N \times N \times M_R \times M_T}$ in equation (5) by concatenating the channel tensors for each transmit antenna, i.e., $\tilde{\mathcal{H}}_R^{(m_T)} \in \mathbb{C}^{N \times N \times M_R}$ along the 4-mode. The tensors $\tilde{\mathcal{H}}_R^{(m_T)} \in \mathbb{C}^{N \times N \times M_R}$ contain the channel vectors for the m_T -th transmit antenna and all receive antennas as defined in equation (6), for $m_T = 1, \dots, M_T$. Recall that these tensors have a CP structure, i.e. $\tilde{\mathcal{H}}_R^{(m_T)} = \mathcal{I}_{3,N} \times_3 \tilde{\mathbf{H}}_R^{(m_T)}$,

for $m_T = 1, \dots, M_T$. The matrices $\tilde{\mathbf{H}}_R^{(m_T)}$ ($m_T = 1, \dots, M_T$) are defined in equation (7). Hence, the 4-way channel tensor is

$$\tilde{\mathcal{H}} = \left[\tilde{\mathcal{H}}_R^{(1)} \quad \sqcup_4 \quad \tilde{\mathcal{H}}_R^{(2)} \quad \sqcup_4 \quad \dots \quad \tilde{\mathcal{H}}_R^{(M_T)} \right]$$

We can rewrite this concatenation by means of an outer product with a pinning vector e_{m_T} . Moreover, if we substitute the CP structure of the tensor $\tilde{\mathcal{H}}_R^{(m_T)}$, we get

$$\begin{aligned} \tilde{\mathcal{H}} &= \sum_{m_T=1}^{M_T} \tilde{\mathcal{H}}_R^{(m_T)} \circ e_{m_T} \\ &= \sum_{m_T=1}^{M_T} \left(\mathcal{I}_{3,N} \times_1 \mathbf{I}_N \times_2 \mathbf{I}_N \times_3 \tilde{\mathbf{H}}_R^{(m_T)} \right) \circ e_{m_T}. \end{aligned}$$

Replacing the outer product by an n -mode product, we have

$$\tilde{\mathcal{H}} = \sum_{m_T=1}^{M_T} \mathcal{D} \times_1 \mathbf{I}_N \times_2 \mathbf{I}_N \times_3 \tilde{\mathbf{H}}_R^{(m_T)} \times_4 e_{m_T}, \quad (30)$$

where $\mathcal{D}_{(\dots,1)} = \mathcal{I}_{3,N}$. Note that the tensor $\mathcal{D} \in \mathbb{R}^{N \times N \times N \times 1}$ is a 4-way tensor, but its 4-mode is a singleton dimension. We can define this tensor in terms of a Kronecker product, which yields $\mathcal{D} = \mathcal{I}_{4,1} \otimes \mathcal{I}_{3,N}$. Equation (30) represents a very special BTD where the block terms are equivalent in all modes, but the 3-mode and the 4-mode. Next, we can replace the sum in (30) with a block diagonal core tensor and factor matrices partitioned accordingly.

$$\begin{aligned} \tilde{\mathcal{H}} &= \text{blkdiag}(\mathcal{I}_{4,1} \otimes \mathcal{I}_{3,N})_{m_T=1}^{M_T} \times_1 \left[\mathbf{I}_N \quad \dots \quad \mathbf{I}_N \right] \\ &\quad \times_2 \left[\mathbf{I}_N \quad \dots \quad \mathbf{I}_N \right] \times_3 \underbrace{\left[\tilde{\mathbf{H}}_R^{(1)} \quad \dots \quad \tilde{\mathbf{H}}_R^{(M_T)} \right]}_{\tilde{\mathbf{H}}} \\ &\quad \times_4 \left[e_1 \quad \dots \quad e_{M_T} \right] \end{aligned}$$

Further, we rewrite the block diagonal structure and the partitioned factor matrices using Kronecker products

$$\begin{aligned} \tilde{\mathcal{H}} &= (\mathcal{I}_{4,M_T} \otimes \mathcal{I}_{3,N}) \times_1 (\mathbf{1}_{M_T}^T \otimes \mathbf{I}_N) \times_2 (\mathbf{1}_{M_T}^T \otimes \mathbf{I}_N) \quad (31) \\ &\quad \times_3 \tilde{\mathbf{H}} \times_4 \mathbf{I}_{M_T}. \quad (32) \end{aligned}$$

This last equation explicitly reveals the structure of the channel tensor $\tilde{\mathcal{H}}$. Exploiting this structure, we can define any of the tensor unfoldings. For the generalized unfolding $[\tilde{\mathcal{H}}]_{([1,3],[2,4])}$, from equation (32), we get

$$\begin{aligned} [\tilde{\mathcal{H}}]_{([1,3],[2,4])} &= [\tilde{\mathbf{H}} \otimes (\mathbf{1}_{M_T}^T \otimes \mathbf{I}_N)] \\ &\quad [\mathcal{I}_{4,M_T} \otimes \mathcal{I}_{3,N}]_{([1,3],[2,4])} \left[\mathbf{I}_{M_T} \otimes \mathbf{1}_{M_T}^T \otimes \mathbf{I}_N \right] \quad (33) \end{aligned}$$

Next, we have

$$\begin{aligned} [\mathcal{I}_{4,M_T} \otimes \mathcal{I}_{3,N}]_{([1,3],[2,4])} \left[\mathbf{I}_{M_T} \otimes \mathbf{1}_{M_T}^T \otimes \mathbf{I}_N \right] &= \mathbf{I}_{NM_T} \diamond \mathbf{I}_{NM_T} \\ \text{for the second part in (33). Recognize that } \mathbf{I}_{NM_T} \diamond \mathbf{I}_{NM_T} &= \mathbf{J}_{NM_T} \\ \text{is a selection matrix that converts a Kronecker product} & \\ \text{into a Khatri-Rao. Using this property, (33) becomes} & \end{aligned}$$

$$[\tilde{\mathcal{H}}]_{([1,3],[2,4])} = \tilde{\mathbf{H}} \diamond (\mathbf{1}_{M_T}^T \otimes \mathbf{I}_N). \quad (34)$$

Moreover, the generalized unfolding $[\tilde{\mathcal{H}}]_{([1,3],[4,2])}$ can also be derived directly from equation (32). However, to simplify the final result is not straightforward because N

is the faster rising index along the columns of the factor matrix $\tilde{\mathbf{H}}$ in equation (32). On the other hand, M_T varies faster than N along the columns in the generalized unfolding $[\tilde{\mathcal{H}}]_{([1,3],[4,2])}$. Therefore, we derive this generalized unfolding by means of a permutation matrix $\mathbf{P} \in \mathbb{R}^{N M_T \times M_T N}$. The permutation matrix \mathbf{P} reorders the columns such that the faster increasing index is M_T instead of N and is defined as $[\tilde{\mathcal{H}}]_{([1,3],[4,2])} = [\tilde{\mathcal{H}}]_{([1,3],[2,4])} \cdot \mathbf{P}$. Hence,

$$[\tilde{\mathcal{H}}]_{([1,3],[4,2])} = [\tilde{\mathbf{H}} \diamond (\mathbf{1}_{M_T}^T \otimes \mathbf{I}_N)] \cdot \mathbf{P}. \quad (35)$$

Considering that the permutation matrix \mathbf{P} reorders the columns in equation (35) and the Khatri-Rao product is a column-wise operator (Khatri-Rao product is column-wise Kronecker product), the following equality holds

$$\begin{aligned} [\tilde{\mathcal{H}}]_{([1,3],[4,2])} &= [\tilde{\mathbf{H}} \diamond (\mathbf{1}_{M_T}^T \otimes \mathbf{I}_N)] \cdot \mathbf{P} \\ &= [\tilde{\mathbf{H}} \cdot \mathbf{P}] \diamond [(\mathbf{1}_{M_T}^T \otimes \mathbf{I}_N) \cdot \mathbf{P}]. \end{aligned}$$

The permutation matrix for $M_T = 2$ and $N = 3$ is given by

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Finally, using $(\mathbf{1}_{M_T}^T \otimes \mathbf{I}_N) \cdot \mathbf{P} = (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T)$ and defining $\bar{\mathbf{H}} = \tilde{\mathbf{H}} \cdot \mathbf{P}$, we get

$$[\tilde{\mathcal{H}}]_{([1,3],[4,2])} = \bar{\mathbf{H}} \diamond (\mathbf{I}_N \otimes \mathbf{1}_{M_T}^T).$$

ACKNOWLEDGMENT

This work has been supported by CAPES/PROBRAL Proc. no. 88887.144009/2017-00 and CAPES/PRINT Proc. no. 88887.311965/2018-00.

REFERENCES

- [1] R. A. Harshman and M. E. Lundy, "Uniqueness proof for a family of models sharing features of Tucker's three-mode factor analysis and PARAFAC/CANDECOMP," *Psychometrika*, vol. 61, no. 1, pp. 133–154, 1996.
- [2] R. A. Harshman, "PARAFAC2: Mathematical and Technical Notes," *CLA Working Papers in Phonetics (University Microfilms, Ann Arbor, Michigan, No. 10,085)*, pp. 30–44, 1972.
- [3] A. L. F. de Almeida, G. Favier, and L. R. Ximenes, "Space-time-frequency (STF) MIMO communication systems with blind receiver based on a generalized PARATUCK2 model," *IEEE Transactions on Signal Processing*, vol. 61, no. 8, pp. 1895–1909, 2013.
- [4] J. Zhang, A. Nimr, K. Naskovska, and M. Haardt, "Enhanced tensor based semi-blind estimation algorithm for relay-assisted MIMO systems," *Latent Variable Analysis and Signal Separation, Volume 9237 of the series Lecture Notes in Computer Science, in Proceedings of the 12-th International Conference on Latent Variable Analysis and Signal Separation (LVA/ICA) (E. Vincent, A. Yeredor, Z. Koldovsky, and P. Tichavsky, eds.)*, pp. 64 – 72, 2015.
- [5] L. R. Ximenes, G. Favier, A. L. F. de Almeida, and Y. C. B. Silva, "PARAFAC-PARATUCK semi-blind receivers for two-hop cooperative MIMO relay systems," *IEEE Transactions on Signal Processing*, vol. 62, pp. 3604 – 3615, 2014.
- [6] T. Hwang, C. Yang, G. Wu, S. Li, and G. Ye Li, "OFDM and Its Wireless Applications: A Survey," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 1673–1694, 2009.
- [7] B. Farhang-Boroujeny, "OFDM versus filter bank multicarrier," *IEEE Signal Processing Magazine*, vol. 28, no. 3, pp. 92 – 112, 2011.
- [8] M. Speth, S. A. Fechtel, G. Fock, and H. Meyr, "Optimum receiver design for wireless broad-band systems using OFDM. I," *IEEE Transactions on Communications*, vol. 47, no. 11, pp. 1668–1677, 1999.
- [9] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Transactions Signal Processing*, vol. 5, pp. 1615–1624, 2003.
- [10] D. Hu, L. Yang, Y. Shi, and L. He, "Optimal pilot sequence design for channel estimation in MIMO OFDM systems," *IEEE Communications Letters*, vol. 10, pp. 1–3, 2006.
- [11] A. L. F. de Almeida and G. Favier, "Unified tensor model for space-frequency spreading-multiplexing (SFSM) MIMO communication systems," *EURASIP Journal on Advances in Signal Processing*, vol. 48, 2013.
- [12] G. Favier and A. L. F. de Almeida, "Tensor space-time-frequency coding with semi-blind receivers for MIMO wireless communication systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 22, pp. 5987–6002, 2014.
- [13] K. Liu, J. P. C. L. da Costa, H. C. So, and A. L. F. de Almeida, "Semi-blind receivers for joint symbol and channel estimation in space-time-frequency MIMO-OFDM systems," *IEEE Transactions Signal Processing*, vol. 61, no. 21, pp. 5444–5457, 2013.
- [14] K. Naskovska, M. Haardt, and A. L. F. de Almeida, "Generalized Tensor Contractions for an Improved Receiver design in MIMO-OFDM Systems," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2018, pp. 3186–3190.
- [15] —, "Generalized tensor contraction with application to Khatri-Rao Coded MIMO OFDM systems," in *Proc. IEEE 7th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2017, pp. 286–290.
- [16] N. D. Sidiropoulos and R. S. Budampati, "Khatri-Rao Space-Time codes," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2396–2407, 2002.
- [17] A. L. F. de Almeida and G. Favier, "Double Khatri-Rao Space-Time-Frequency coding using semi-blind PARAFAC based receiver," *IEEE Signal Processing Letters*, vol. 20, pp. 471–474, 2013.
- [18] P. Comon, "Tensors: A brief introduction," *IEEE Signal Processing Magazine*, vol. 31, no. 3, pp. 44–53, 2014.
- [19] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, pp. 455–500, 2009.
- [20] A. Cichocki, D. Mandic, A. Phan, C. Caiafa, G. Zhou, Q. Zhao, and L. De Lathauwer, "Tensor decompositions for signal processing applications: From two-way to multiway," *IEEE Signal Processing Magazine*, vol. 32, pp. 145–163, 2015.
- [21] X. Luciani and L. Albera, "Semi-algebraic canonical decomposition of multi-way arrays and joint eigenvalue decomposition," in *Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2011, pp. 4104–4107.
- [22] F. Roemer, C. Schroeter, and M. Haardt, "A semi-algebraic framework for approximate CP decompositions via joint matrix diagonalization and generalized unfoldings," in *Proc. of the 46th Asilomar Conference on Signals, Systems, and Computers*, November 2012, pp. 2023–2027.
- [23] S. Talwar, M. Viberg, and A. Paulraj, "Blind estimation of multiple co-channel digital signals using an antenna array," *IEEE Signal Processing Letters*, vol. 1, no. 2, pp. 29–31, 1994.
- [24] —, "Blind separation of synchronous co-channel digital signals using an antenna array. I. Algorithms," *IEEE Transactions on Signal Processing*, vol. 44, no. 5, pp. 1184–1197, 1996.
- [25] F. Roemer and M. Haardt, "Tensor-Based channel estimation (TENCE) for Two-Way relaying with multiple antennas and spatial reuse," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Taipei, Taiwan*, April 2009, pp. 3641–3644.
- [26] ITU-R Recommendation M.1225, "Guidelines for evaluation of radio transmission technologies for IMT-2000," Recommendation, 1997.