Abstract—In this paper, we first introduce two new classes of constrained tensor models that we call generalized PARATUCK-$(N_1,N)$ and Tucker-$(N_1,N)$ models. A new tensor space-time-frequency (TSTF) coding structure is then proposed for MIMO OFDM-CDMA wireless communication systems. Two semi-blind receivers relying on the new generalized PARATUCK model are derived for solving the problem of joint channel and symbol estimation. One is iterative and based on a two-step alternating least squares (ALS) algorithm. The other one is a closed-form and low-complexity solution which consists of the Kronecker product least squares (KPLS) estimation of the symbol matrix and a matrix unfolding of the channel tensor. Uniqueness of the underlying tensor model is discussed and system design requirements are derived for applicability of the ALS and KPLS receivers. We also show that the so-called TSTF system can be viewed as an extension of three existing tensor-based ST/ST/TST coding systems that are described in a unified framework. Computer simulation results illustrate the good performance of the TSTF system which outperforms the considered existing tensor-based systems both in terms of symbol estimation quality and allocation flexibility.

Index Terms—Constrained tensor models, PARATUCK models, Tucker models, Space-time-frequency coding, MIMO communication systems.

I. INTRODUCTION

The use of multiple antennas at both ends of the link in a wireless communication system, commonly known as multiple-input multiple-output (MIMO) system, is considered as one of the key technologies to be deployed in upcoming wireless communications standards [1]. In this context, the integration of multiple-antenna systems with orthogonal frequency division multiplexing (OFDM) and code-division multiple-access (CDMA) transmission has been intensely studied over the past few years [2]–[4]. In MIMO systems, transmit antennas can be employed to achieve high data rates via spatial multiplexing as well as to improve link reliability through space-time/space-frequency or space-time-frequency coding [5], [6].

A common characteristic of all these works is the assumption of perfect channel knowledge at the receiver. When the channel is not known, as it is the case in practice, the receiver design is generally based on suboptimum (linear or nonlinear) equalization/signal separation structures that use training sequences for channel acquisition and tracking, before decoding the transmitted data. However, practical limitations due to the training sequence overhead (which implies a reduction of the information rate) may be prohibitive.

Since the publication of the seminal paper [7], the use of tensor decompositions has gained increased attention in signal processing applications for wireless communications. The practical motivation for tensor modeling comes from the fact that one can simultaneously benefit from multiple forms of diversity to perform multiuser signal separation and channel estimation/equalization under model uniqueness conditions more relaxed than with conventional matrix-based approaches. Aside from their powerful uniqueness properties, tensor models facilitate generalizations that consist of taking more diversities into account while offering more flexibility in terms of spreading and multiplexing across space, time, and frequency, with the possibility of resource allocation. These generalizations lead to higher-order constrained tensor models for the received signals [8], [9].

Additionally, the key characteristics of communication systems modeled using tensor decompositions, not covered by matrix based systems, are the following. Tensor-based receivers do not require the use of long training sequences (i.e. they can operate semi-blindly in the sense that only very few pilot symbols are needed to eliminate scaling ambiguity proper to tensor models), nor the knowledge of channel impulse responses and antenna array responses. Moreover, they do not rely on statistical independence of transmitted signals. Instead, the tensor-based receivers are deterministic, and exploit the multilinear algebraic structure of the received signals, treated as higher order tensors. These receivers operate on data blocks and usually allow a joint channel/symbol estimation by means of an iterative algorithm. The price to pay is associated with the computational complexity of the existing receivers, which is highly depending on the specific algorithm used for channel/symbol estimation and its initialization.

The use of tensor models for designing multiple-antenna space-time transmission schemes has been addressed in several recent works [8]–[15]. In [10], a space-time coding model based on the parallel factors (PARAFAC) decomposition [16], [17] is proposed. The work [11] presents a block tensor model for multiple-access MIMO systems based on spatial multiplexing. It is shown that the received signals follow a
decomposition in blocks of rank-(1, L, L) tensors [18], where the number of blocks corresponds to the number of users. This decomposition was later generalized in [19], [20]. Space-time spreading structures were proposed in [13], [14] by introducing a new third-order constrained factor (CONFAC) tensor model.

A space-time coding structure capitalizing on the PARATUCK2 tensor model [21], [22] was proposed in [15]. The idea was to design different coding structures combining both spreading and multiplexing across space and time, by exploiting PARATUCK2 interaction structure. A generalization of [15] was developed in [8] using a tensor space-time (TST) coding, where a spreading dimension is added into the tensor coding operation. However, all these works are restricted to single-carrier transmissions.

More recently, a generalized fourth-order PARATUCK2 tensor model has been presented for MIMO systems with space-time-frequency (STF) spreading-multiplexing [9]. The core tensor consists of a spatial coding matrix combined with two tridimensional arrays that control the allocation of the data streams and transmit antennas in the joint time-frequency domain, i.e. across time blocks and subcarriers. This transmission scheme generalizes the one in [15] to MIMO communication systems based on multicarrier transmission. The main drawback of this approach, which is also the case in previous works, is related to the computational cost of the receiver processing, based on iterative estimation algorithms, and involving the calculation of matrix inverses at each iteration, which may suffer from a slow convergence.

This paper proposes the use of a new constrained tensor model for the design of MIMO wireless communication systems with space, time and frequency transmit signaling. Its main contributions can be summarized as follows:

- Introduction of new tensor tools like the Hadamard product of two tensors, and general matrix unfoldings of a tensor with possible repetitions of some dimensions. The derivation of such matricized forms is greatly simplified by using mode combinations based on Kronecker products of unit vectors. This original and concise formulation is illustrated with the matricization of the generalized PARATUCK-(2,5) model, Appendix B.

- Introduction of two new classes of tensor models that we call generalized PARATUCK-(N1, N) and Tucker-(N1, N) models, for which the factors of the decompositions are high-order tensors. These models can be viewed as extensions of PARATUCK-(N1, N) and Tucker-(N1, N) models introduced in [23].

- Design of a new tensor space-time-frequency (TST) coding that leads to a generalized PARATUCK-(2,5) model for MIMO OFDM-CDMA systems. The proposed TSTF system combines a fifth-order coding tensor with a fourth-order allocation tensor. It generalizes three existing tensor-based ST/TST/STF coding schemes, offering new performance/complexity tradeoffs and space, time and frequency allocation flexibility.

- Proposition of two semi-blind receivers for joint channel and symbol estimation. The first one is iterative and based on a two-step alternating least squares (ALS) algorithm. The second one is a closed-form and low-complexity solution which consists of the Kronecker product least squares (KPLS) estimation of the symbol matrix and a matrix unfolding of the channel tensor. From the closed-form solution, it is easy to take into account a sequence of pilot symbols to derive a supervised receiver that we call a pilot-assisted KPLS-receiver.

- System design requirements for applicability of the ALS and KPLS receivers, and uniqueness property of the generalized PARATUCK-(2,5) model in the context of our TSTF system.

- Comparison of the proposed TSTF system with three other tensor-based systems through a unified presentation using an equivalent (generalized) Tucker model and by means of simulations.

The rest of this paper is organized as follows. Section II presents the notations and tensor operations that are used throughout the paper. In Section III, we introduce two new tensor models: the generalized PARATUCK-(N1, N) and Tucker-(N1, N) models. Then, in Section IV, constrained tensor models are illustrated in the context of MIMO communication systems with resource allocation. We first present a new tensor-based MIMO transmission system, called tensor space-time-frequency (TSTF) system, arising from a generalized PARATUCK model. In Section V, we propose two semi-blind receivers for jointly estimating the channels and the symbols, and we discuss the uniqueness property of the tensor model of the TSTF system as well as the conditions for applicability of the proposed receivers. We also show that the TSTF system can be viewed as an extension of three recent tensor-based systems that are described in a unified framework. In Section VI, the performance of the TSTF system is evaluated by means of computer simulations and compared with those of existing tensor-based MIMO systems in the semi-blind, pilot-assisted and perfect channel knowledge cases. The paper is concluded in Section VII.

II. NOTATIONS, PROPERTIES, AND TENSOR OPERATIONS

R and C denote the fields of real and complex numbers, respectively. Scalars, column vectors, matrices, and tensors are denoted by lowercase, boldface lowercase, boldface uppercase, and calligraphic uppercase letters, e.g. a, A, and A, respectively. The vector ai (resp. aj) represents the ith row (resp. jth column) of A.

By slicing the third-order tensor $A \in \mathbb{C}^{I \times J \times K}$ along each mode, we get three types of matrix slices, respectively called horizontal, lateral, and frontal slices, denoted as follows

$$A_{i..} \in \mathbb{C}^{I \times K}, \quad A_{j..} \in \mathbb{C}^{K \times J} \quad \text{and} \quad A_{k..} \in \mathbb{C}^{I \times J}$$

with $i = 1, \ldots, I; j = 1, \ldots, J; k = 1, \ldots, K$.

$I_N$, $I_N^T$, and $e_{n(N)}$ stand for the identity matrix of order N, the all-ones row vector of dimensions $1 \times N$, and the nth unit vector of the Euclidean space $\mathbb{R}^N$, respectively.

$A^T$, $A^*$, $A^H$, $A^\dagger$, and $r(A)$ denote the transpose, the conjugate, the conjugate (Hermitian) transpose, the Moore-Penrose pseudo-inverse, and the rank of A, respectively. The operators $\text{diag}(\cdot)$ and $\text{bdia}(\cdot)$ form a diagonal and a block-diagonal
matrix from their matrix arguments, respectively, with
\[
\text{bdiag}(A_1, \ldots, A_K) \triangleq \text{bdiag}(A_{i,k}) \in \mathbb{C}^{K_1 \times K_J},
\]
where \( A_{i,k} \) is the \( k^{th} \) block on the diagonal. The operator \( \text{vec}(\cdot) \) transforms a matrix into a column vector by stacking the columns of its matrix argument while the operator \( \text{unvec}(\cdot) \) corresponds to the inverse transformation. The Kronecker, Khatri-Rao (column-wise Kronecker), and Hadamard (element-wise) products are denoted by \( \otimes \), \( \odot \), and \( \cdot \), respectively.

Let us consider a permutation \( S = \{n_1, \ldots, n_N\} \) of the set \( \{1, \ldots, N\} \). For \( A^{(n)} \in \mathbb{C}^{I_n \times R_n}, n = 1, \ldots, N \), we define
\[
\nabla(n) = A^{(n_1)} \oplus A^{(n_2)} \otimes \cdots \otimes A^{(n_N)} = \mathbb{C}^{I_{n_1} \times I_{n_2} \times \cdots \times I_{n_N} \times R_{n_1} \times \cdots \times R_{n_N}}.
\]

By convention, the order of dimensions in (1) is directly related to the order of variation of the associated indices, i.e., the product \( I_{n_1} I_{n_2} \cdots I_{n_N} \) of dimensions means that \( n_1 \) is the index varying the most slowly while \( n_N \) is the index varying the most fastly in the Kronecker products computation. We have the following properties:

For \( A^{(n)} \in \mathbb{C}^{I_n \times R_n}, B^{(n)} \in \mathbb{C}^{R_n \times J_n}, n = 1, \ldots, N \),
\[
\left( \oplus_{n=1}^{N} A^{(n)} \right)^T = \oplus_{n=1}^{N} A^{(n)}^T \in \mathbb{C}^{R_1 \cdots R_N \times I_1 \cdots I_N} \quad (2)
\]
\[
\left( \otimes_{n=1}^{N} B^{(n)} \right) = \otimes_{n=1}^{N} \left( A^{(n)} \right)^T \in \mathbb{C}^{I_1 \cdots I_N \times J_1 \cdots J_N} \quad (3)
\]
For \( A^{(n)} \in \mathbb{C}^{I_n \times J_n}, n = 1, \ldots, N \), and \( B^{(n)} \in \mathbb{C}^{K_n \times I_n}, n = 1, \ldots, p \),
\[
\sum_{n=1}^{N} A^{(n)} \otimes \sum_{p=1}^{P} B^{(p)} = \sum_{n=1}^{N} \sum_{p=1}^{P} \left( A^{(n)} \otimes B^{(p)} \right) \in \mathbb{C}^{I_K \times J_L} \quad (4)
\]

\( C = A \oplus B \leftrightarrow c_{i_1, \ldots, i_N} = a_{i_1, \ldots, i_N} b_{k_1, \ldots, k_N} \)

\( S \cap K \) and \( \mathbb{I} = K \cup K \). For instance, given the two third-order tensors \( A \in \mathbb{C}^{I_1 \times J_2 \times J_3} \) and \( B \in \mathbb{C}^{K_1 \times K_2 \times K_3} \), with \( j_1 = k_1 = i_1, j_2 = k_2 = i_2, j_3 = k_3 = i_3 \), the Hadamard product \( A \odot B \) gives a fourth-order tensor \( C \in \mathbb{C}^{I_1 \times I_2 \times I_3 \times I_4} \) such that
\[
c_{i_1, \ldots, i_8} = a_{i_1, i_2, i_3} b_{i_4, i_5, i_6, i_7}.
\]

Now, we introduce general matrix unfoldings of a tensor, and then we define a mode-n product of two tensors.

A. Matrix unfoldings of a tensor

Let us consider a complex-valued tensor \( \chi \in \mathbb{C}^{I_1 \times \cdots \times I_N} \) of order \( N \) and dimensions \( I_1 \times \cdots \times I_N \), with entries \( (\chi)_{i_1,\ldots,i_N} = x_{i_1,\ldots,i_N} \). We define two subsets \( S_1 \) and \( S_2 \) of the set \( \{1, \ldots, N\} \) such that \( S_1 \cup S_2 = \{1, \ldots, N\} \) with possible repetitions of some dimensions, a general formula for the matricization of \( \chi \in \mathbb{C}^{I_1 \times \cdots \times I_N} \) is given by
\[
\chi_{S_1 \backslash S_2} = \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} x_{i_1,\ldots,i_N} \left( \otimes_{n \in S_1} e_{i_n} \right) \left( \otimes_{n \in S_2} e_{i_n} \right)^T \in \mathbb{C}^{J_1 \times J_2} \quad (5)
\]

with \( J_{n_1} = \prod_{n \in S_1} I_n \), for \( n_1 = 1 \) and 2.

Proposition: The element \( x_{i_1,\ldots,i_N} \) can be expressed as
\[
x_{i_1,\ldots,i_N} = \left( \otimes_{n \in S_1} e_{i_n} \right)^T \chi_{S_1 \backslash S_2} \left( \otimes_{n \in S_2} e_{i_n} \right). \quad (6)
\]

See the proof in Appendix A.

For instance, given a tensor \( \chi \in \mathbb{C}^{I_1 \times I_2 \times I_3} \), with entries \( x_{i_1, i_2, i_3} \), we can define the matrix unfolding \( X_{i_1, I_2 \times I_3} \) with repetition of the second dimension,
\[
X_{i_1, I_2 \times I_3} = \sum_{i_2=1}^{I_2} \sum_{i_3=1}^{I_3} x_{i_1, i_2, i_3} \left( \otimes_{i_2} e_{i_2} \right) \left( \otimes_{i_3} e_{i_3} \right)^T \in \mathbb{C}^{I_1 \times I_2 \times I_3} \quad (7)
\]

B. Mode-n product of two tensors

Let us consider an ordered subset \( S = \{i_{j_1}, \ldots, i_{j_M}\} \) of \( \mathbb{J} = \{j_1, \ldots, j_N\} \), with \( j_m \in \{N_1 + 1, \ldots, N\} \) for \( m = 1, \ldots, M, N_1 \leq N \). The mode-n product of a tensor \( \chi \in \mathbb{C}^{I_1 \times \cdots \times I_N} \) with a tensor \( A \in \mathbb{C}^{I_m \times R_N} \) is defined by \( \chi \times_n A \), with \( n \leq N \) and \( I_N \) being a shortened writing for \( I_1 \times \cdots \times I_N \). This product gives the tensor \( \chi \) of order \( N \) and dimensions \( R_1 \times \cdots \times R_{n-1} \times I_n \times R_{n+1} \times \cdots \times R_N \), such that
\[
x_{r_1, \ldots, r_{n-1}, i_n, r_{n+1}, \ldots, r_N} = \sum_{r_n=1}^{R_n} a_{r_n, i_n} \cdot g_{r_1, \ldots, r_{n-1}, r_{n+1}, \ldots, r_N} \quad (8)
\]

Note that the sum is over the second index of the tensor \( A \) as it is usual for the standard mode-n product of a tensor \( \chi \in \mathbb{C}^{I_1 \times \cdots \times I_N} \) with a matrix \( A \in \mathbb{C}^{I_n \times R_n} \), denoted by
ordered subset \( S \) to \( \mathbb{C}^{I_j \times \cdots \times I_N} \) along the same mode-\( n \), with \( A \in \mathbb{C}^{J_a \times I_j} \) and \( B \in \mathbb{C}^{K_a \times J_a} \), we have the following property [24]

\[
\mathcal{Y} = \mathcal{X} \times_n A \times_n B
\]

or equivalently

\[
\mathcal{Y} = \mathcal{X} \times_n (BA) \in \mathbb{C}^{I_1 \times \cdots \times I_n-1 \times K_n \times I_n+1 \times \cdots \times I_N}.
\]

### III. Constrained Tensor Models

Let us consider an \( N^{th} \)-order tensor \( \mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N} \). We define a generalized PARATUCK-(\( N_1, N \)) model for \( \mathcal{X} \), by means of the following relationship between the input \( (\mathcal{W}) \), output \( (\mathcal{Y}) \), and constraint tensors \((\mathcal{C})\)

\[
x_{i_1, \ldots, i_N} = \sum_{r_1=1}^{R_1} \cdots \sum_{r_N=1}^{R_N} w_{i_1, \ldots, r_N, i_{N+1}, \ldots, i_N} \prod_{n=1}^{N_1} a_{i_{n+1}, r_n} C_{i_{n+1}, r_n, i_N} \mathcal{C}
\]

where the set of indices \( \mathcal{V} = \{i_1, \ldots, r_N, i_{N+1}, \ldots, i_N\} \) is associated with input modes, whereas \( \mathcal{O} = \{i_1, i_N\} \) correspond to output modes.

- \( \mathcal{S}, \mathcal{T}, \) and \( \mathcal{S}_n \) (for \( n = 1, \ldots, N_1 \)) denote ordered subsets of the set \( \mathcal{S} \cup \mathcal{T} = \{i_{N+1}, \ldots, i_N\} \).
- \( a_{i_{n+1}, r_n, s_n} \in \mathbb{C} \), \( i_{n+1}, r_n, i_N, s_n \in \mathcal{C} \), and \( w_{i_1, \ldots, r_N, i_{N+1}, \ldots, i_N} \in \mathbb{C} \) are entries of the tensor factor \( A^{(n)} \), \( n \in \{1, \ldots, N_1\} \), of the constraint tensor \( \mathcal{C} \) which is a Boolean tensor (i.e. with binary entries) of dimensions \( R_1 \times \cdots \times R_N \times I_1 \), and of the input tensor \( \mathcal{W} \in \mathbb{C}^{R_1 \times \cdots \times R_N \times I_1} \), respectively.
- \( I_{S_1}, I_{T} \), and \( I_S \) denote the dimensions associated with the index subsets \( S_1, T \), and \( S \), respectively. For instance, if \( \mathcal{S} = \{i_2, i_3\} \), then \( I_{S} = I_2 \times I_3 \).
- Define the core tensor \( \mathcal{G} \in \mathbb{C}^{R_1 \times \cdots \times R_N \times I_{N+1} \times \cdots \times I_N} \) such that

\[
g_{i_1, \ldots, i_{N+1}, \ldots, i_N} = w_{i_1, \ldots, r_N, i_{N+1}, \ldots, i_N} \prod_{n=1}^{N_1} a_{i_{n+1}, r_n, s_n} C_{i_{n+1}, r_n, i_N} T
\]

It can be written as the Hadamard product of the tensors \( \mathcal{W} \) and \( \mathcal{C} \) along their common modes belonging to \( \mathcal{V} \cup (\mathcal{S} \cap \mathcal{T}) \)

\[
\mathcal{G} = \mathcal{W} \odot_{\mathcal{V} \cup (\mathcal{S} \cap \mathcal{T})} \mathcal{C}.
\]

The generalized PARATUCK-(\( N_1, N \)) model (9) can then be rewritten as follows

\[
x_{i_1, \ldots, i_N} = \sum_{r_1=1}^{R_1} \cdots \sum_{r_N=1}^{R_N} g_{i_1, \ldots, r_N, i_{N+1}, \ldots, i_N} \prod_{n=1}^{N_1} a_{i_{n+1}, r_n} s_n \]

or equivalently

\[
\mathcal{X} = \mathcal{G} \times_{N+1} A^{(n)}.
\]

This model will be called a generalized Tucker-(\( N_1, N \)) model with the core tensor \( \mathcal{G} \), and tensor factors \( A^{(n)}, \) for \( n = 1, \ldots, N_1 \), with entries \( a_{i_{n+1}, r_n, s_n} \). Replacing the tensor factors \( A^{(n)} \) in (10) by matrix factors \( A^{(n)} \), for \( n = 1, \ldots, N_1 \), we obtain a Tucker-(\( N_1, N \)) model defined in [23] as

\[
x_{i_1, \ldots, i_N} = \sum_{r_1=1}^{R_1} \cdots \sum_{r_N=1}^{R_N} g_{i_1, \ldots, r_N, i_{N+1}, \ldots, i_N} \prod_{n=1}^{N_1} a_{i_{n+1}, r_n} s_n
\]

or equivalently

\[
\mathcal{X} = \mathcal{G} \times_{N+1} A^{(n)}
\]

where \( a_{i_{n+1}, r_n} \) is an element of the matrix factor \( A^{(n)} \). \( \in \mathbb{C}^{R_1 \times R_N} \).

The standard \( N^{th} \)-order Tucker model [25] corresponds to a Tucker-(\( N, N \)) model.

Some examples of tensor models that can be directly deduced from the generalized PARATUCK-(\( N_1, N \)) model (9), are described in Table 1.

In telecommunication applications, the tensor \( \mathcal{W} \) corresponds to the code tensor that allows to spread the information symbols in space/time/frequency domains, while the constraint tensor \( \mathcal{C} \) allocates resources like data streams, transmit antennas, and/or subcarriers to transmission time blocks, as illustrated in the next section. The input and output modes correspond to resource and diversity modes, respectively.

### IV. Tensor Space-Time-Frequency (TSTF) Coding

In this section, we consider a MIMO OFDM-CDMA system equipped with \( M \) transmit antennas and \( K \) receive antennas. We denote by \( s_{n,r} \) the \( n^{th} \) symbol of the \( r^{th} \) data stream \( (r = 1, \ldots, R) \), each data stream being composed of \( N \) information symbols. The transmission is decomposed into \( P \) time-slots (data blocks) of \( N \) symbol periods, each one being composed of \( J \) chips. At each symbol period \( n \) of the \( j^{th} \) block, the transceiver transmits a linear combination of the \( n^{th} \) symbols of certain data streams, using a set of transmit antennas and/or sub-carriers. The coding is carried out by means of a fifth-order code tensor \( W \in \mathbb{C}^{M \times R \times F \times P \times J} \) whose dimensions are the numbers of transmit antennas (\( M \)), data streams (\( R \)), sub-carriers (\( F \)), time blocks (\( P \)), and chips (\( J \)). In Table II, the physical meaning of the modes and the main system parameters are summarized.

The transmit antennas and the sub-carriers that are used, and the data streams that are transmitted in each block \( p \) are determined by the allocation tensor \( \mathcal{C} \in \mathbb{R}^{M \times R \times F \times P \times J} \). \( c_{m,r,f,p,j} = 1 \) means that the data stream \( r \) is transmitted using the transmit antenna \( m \), with the sub-carrier \( f \), during the time block \( p \). Each information symbol \( s_{n,r} \) is replicated several times after multiplication by the five-dimensional spreading code \( w_{m,r,f,p,j} \), in such a way that the signal transmitted from the \( m^{th} \) transmit antenna, by the \( f^{th} \) sub-carrier, during the

---

1053-587X (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Table I

<table>
<thead>
<tr>
<th>Models</th>
<th>Scalar writings</th>
<th>mode-r product based writings</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generalized Paratuck-(2,5)</td>
<td>[ x_{i_1,i_2,i_3,i_4} = \frac{R_1}{r_1=1} \sum_{r_2=1}^{R_2} g_{r_1,r_2,i_1,i_3,i_4} a_{(1)}^{r_1} a_{(2)}^{r_2} ]</td>
<td>[ X = G \times_1 A(1) \times_2 A(2) ]</td>
<td>[9]</td>
</tr>
<tr>
<td>Generalized Paratuck-(2,4)</td>
<td>[ x_{i_1,i_2,i_3,i_4} = \frac{R_1}{r_1=1} \sum_{r_2=1}^{R_2} g_{r_1,r_2,i_1,i_3,i_4} a_{(1)}^{r_1} a_{(2)}^{r_2} ]</td>
<td>[ X = G \times_1 A(1) \times_2 A(2) ]</td>
<td>[9]</td>
</tr>
<tr>
<td>Paratuck-(2,4)</td>
<td>[ x_{i_1,i_2,i_3,i_4} = \frac{R_1}{r_1=1} \sum_{r_2=1}^{R_2} g_{r_1,r_2,i_1,i_3,i_4} a_{(1)}^{r_1} a_{(2)}^{r_2} ]</td>
<td>[ X = G \times_1 A(1) \times_2 A(2) ]</td>
<td>[9]</td>
</tr>
<tr>
<td>Paratuck-2</td>
<td>[ x_{i_1,i_2,i_3} = \frac{R_1}{r_1=1} \sum_{r_2=1}^{R_2} g_{r_1,r_2,i_1,i_2,i_3} a_{(1)}^{r_1} a_{(2)}^{r_2} ]</td>
<td>[ X = G \times_1 A(1) \times_2 A(2) ]</td>
<td>[21], [15]</td>
</tr>
<tr>
<td>Tucker-(2,3)</td>
<td>[ x_{i_1,i_2,i_3} = \frac{R_1}{r_1=1} \sum_{r_2=1}^{R_2} g_{r_1,i_1,i_2,i_3} a_{(1)}^{r_1} a_{(2)}^{r_2} ]</td>
<td>[ X = G \times_1 A(1) \times_2 A(2) ]</td>
<td>[23]</td>
</tr>
</tbody>
</table>

Table II

| m (k): transmit (receive) antenna index |
| r: data stream index                  |
| p: time slot (data block) index       |
| n: symbol period index                |
| f: chip period index                  |
| k: sub-carrier index                  |
| H ∈ C^{K × M × F}: channel tensor    |
| S ∈ C^{R × K}: symbol matrix         |
| W ∈ C^{R × F × P × J}: code tensor   |
| C ∈ R^{M × R × F × P}: allocation (constraint) tensor |

\[ u_{m,m,f,p} = \sum_{r=1}^{R} w_{m,r,f,p} s_{n,f} c_{m,r,f,p}. \]

In the noiseless case, and assuming flat Rayleigh fading propagation channels, the discrete-time baseband-equivalent model for the signal received at the kth receive antenna during the jth chip period of the nth symbol period of the pth block, and associated with the fth sub-carrier, is given by

\[ x_{k,n,f,p} = \sum_{m=1}^{M} h_{k,m,f} w_{m,n,f,p} \]

The fading coefficients \( h_{k,m,f} \) between the transmit \( m \) and receive \( k \) antennas are random variables. They are assumed to be constant during at least P blocks. The tensor model (11) is a generalized PARATUCK-(2,5) model (9) with \( O = \{ k, n, f, p, j \}, S = \{ f, p \}, T = \{ f, p \}, S_1 = \{ f \}, S_2 = \{ p \}, V = \{ m, r \} \), and the following correspondences

\[ (I_1, I_2, I_3, I_4, I_5, R_1, R_2) \rightarrow (K, N, F, P, J, M, R) \]
\[ (A^{(1)}, A^{(2)}, W, C) \rightarrow (H, S, W, C). \]

Equation (11) can be written as a generalized Tucker-(2,5) model

\[ X = G \times_1 H \times_2 S \]

with the core tensor \( G \in C^{M \times R \times F \times P \times J} \) given by

\[ G = W \odot C, \]

or equivalently

\[ g_{m,r,f,p,j} = u_{m,r,f,p} c_{m,r,f,p}. \]

Figure 1 visualizes the tensor slice \( X_{f,p} \in C^{K \times N \times J} \) of the received signal tensor \( X \in C^{K \times N \times F \times P \times J} \), associated with a particular couple \((f, p)\) of sub-carrier and time-block, defined as

\[ X_{f,p} = G_{f,p} \times_1 H_{f} \times_2 S \]

where \( H_{f} \in C^{K \times M} \), and the tensor slice \( G_{f,p} \in C^{M \times R \times J} \) of the core tensor \( G \) is given by

\[ G_{f,p} = W_{f,p} \odot C_{f,p} \]

\( C_{f,p} \) allocating the data streams and the transmit antennas to the sub-carrier \( f \) and the time-block \( p \).

V. SEMI-BLIND RECEIVERS

Before presenting two semi-blind receivers for jointly estimating the symbol matrix and the channel tensor, we give three matrix representations of the noiseless received signal tensor for the TSF system. These matrix representations are deduced from the matrix unfoldings (28), (29), and (37), demonstrated in Appendix B for the generalized Tucker-(2,5) model. Using the correspondences (12)-(13), and replacing \( A^{(2)}_{14} \) by \( A^{(2)}_{14} F \) and \( I_{14} \otimes A^{(2)} \), respectively, in (28) and (29), we obtain

\[ X_{I_{14}F \times K} = (I_{14} \odot bdiag(H_{f})) G_{I_{14}FM \times R} S^{T}, \]
\[ X_{I_{14}F \times K} = (I_{14} \odot S) G_{I_{14}FM \times R} H_{FM \times K}. \]
After convergence, the estimated symbol matrix \( \hat{\mathbf{S}}_{(t)} \) is obtained as:

\[
\hat{\mathbf{S}}_{(t)} = \underset{\mathbf{S}}{\arg\min} ||\mathbf{X}_{JPFN} - (\mathbf{I}_{JP} \otimes \text{bdiag}(\mathbf{H}_{..f}(t))) \mathbf{G}_{JPFR \times FM} \mathbf{H}_{FM \times K} ||_F^2,
\]

subject to:

\[
\hat{\mathbf{S}}_{(t)} = \underset{\mathbf{s}}{\arg\min} ||\mathbf{X}_{JPFK} - (\mathbf{I}_{JP} \otimes \text{bdiag}(\mathbf{H}_{..f}(t))) \mathbf{G}_{JPFM \times RM} \mathbf{H}_{FM \times K} ||_F^2.
\]

where \( \text{bdiag}(.) \) is defined in Section II, and the core tensor unfoldings are given by:

\[
\mathbf{G}_{JPFM \times R} = \mathbf{W}_{JPFM \times R} \otimes \mathbf{C}_{JPFM \times R},
\]

\[
\mathbf{G}_{JPFR \times FM} = \mathbf{W}_{JPFR \times FM} \otimes \mathbf{C}_{JPFR \times FM},
\]

with:

\[
\mathbf{W}_{JPFR \times FM} = \mathbf{W}_{JPFR \times MF}(\mathbf{I}_F \otimes \mathbf{I}_M),
\]

\[
\mathbf{C}_{JPFM \times R} = (\mathbf{I}_J \otimes \mathbf{I}_{PFM}) \mathbf{C}_{PFM \times R},
\]

\[
\mathbf{C}_{JPFR \times FM} = (\mathbf{I}_J \otimes \mathbf{I}_{PFR}) \mathbf{C}_{PFM \times R}(\mathbf{I}_F \otimes \mathbf{I}_M).
\]

Similarly, for the TSTF system with the correspondences (12)-(13), Eq. (38) becomes:

\[
\mathbf{X}_{NK \times JPF} = (\mathbf{S} \otimes \mathbf{H}_{K \times MF}) \mathbf{G}_{RMF \times JPF}. \quad (19)
\]

### A. ALS-receiver

We assume that the allocation (\( \mathcal{C} \)) and the code (\( \mathcal{W} \)) tensors are known at the receiver, and therefore also the core tensor (\( \mathcal{G} \)). The matrix unfoldings (15) and (16) of the received signal tensor can be used for deriving a two-step ALS algorithm that consists of an alternate minimization of the two conditional LS cost functions (20) and (21) shown on the top of this page, where \( it \) and \( || . ||_F \) denote the iteration number and the Frobenius norm, respectively. \( \mathbf{X}_{JPFK} \) and \( \mathbf{X}_{JPEN} \) represent noisy versions of \( \mathbf{X}_{JPFK} \) and \( \mathbf{X}_{JPEN} \), the noise being assumed to be additive zero-mean complex-valued white Gaussian. After convergence, the estimated symbol matrix \( \hat{\mathbf{S}}_{(\infty)} \) is affected by a scaling ambiguity that must be eliminated before projection onto the alphabet. Although the knowledge of a single pilot symbol is enough for ensuring uniqueness, as shown in Section V.D, in order to improve the performance of this receiver, we assume that the first symbol of each data stream is equal to 1 and known at the receiver, implying that the first row of the symbol matrix is equal to \( \mathbf{I}_F \).

In this case, the final estimate of the symbol matrix is obtained by column normalization, i.e.

\[
\hat{\mathbf{S}}_{(\text{final})} = \hat{\mathbf{S}}_{(\infty)} \left[ \mathbf{D}_1 (\hat{\mathbf{S}}_{(\infty)}) \right]^{-1},
\]

where the operator denoted by \( \mathbf{D}_1(.) \) forms a diagonal matrix from the first row of its matrix argument. The resulting ALS-based semi-blind receiver is described in Table III.

---

**TABLE III**

**ALS-BASED RECEIVER**

1. Initialization (\( it=0 \)): randomly draw \( \mathbf{S}_{(0)} \) from the symbol alphabet.
2. \( it=it+1 \).
3. Calculate the LS estimate of the channel tensor \( \hat{\mathbf{H}}_{FM \times K} \) (16).
4. Calculate the LS estimate of the symbol matrix \( \hat{\mathbf{S}}_{(it)}^{(k)} \) (17).
5. Eliminate the scaling ambiguity \( \hat{\mathbf{S}}_{(\text{final})} = \hat{\mathbf{S}}_{(\infty)} \left[ \mathbf{D}_1 (\hat{\mathbf{S}}_{(\infty)}) \right]^{-1} \).
6. Project the estimated symbols onto the symbol alphabet.
7. Return to Step 2 until convergence.
B. KPLS-receiver

Assuming that $G_{RMF \times JPF}$ is right-invertible, which implies $RM \leq JP$, we also propose a non-iterative receiver, named KPLS-receiver. It is based on the LS estimation of the Kronecker product (KP) deduced from (19)

$$\hat{\Gamma} = S \otimes H_{K \times MF} = X_{NK \times JPF} G_{RMF \times JPF}^\dagger \in \mathbb{C}^{NK \times RMF}$$

where $G_{RMF \times JPF}$ has the form (39) in the Appendix, with the correspondences (12)-(13). To simplify the computation of $\hat{\Gamma}$, we design the core tensor $\mathcal{G}$ by choosing its matrix unfolding $G_{RMF \times JPF}$ as a Vandermonde matrix such that

$$g_{m,r,f,p,j} = \frac{1}{\sqrt{N}} \exp\left(\frac{-2\pi i (r-1)M+m-1)(-1)F+(p-1)F+j-1}{N}ight)$$

(22)

where $\omega = \exp\left(\frac{2\pi i}{N}ight)$, with $i^2 = -1$. With this choice of $\mathcal{G}$, we have:

$$G_{RMF \times JPF} = G_{H_{MF \times JPF}}^H.$$ 

Let us first consider the pilot-assisted KPLS-receiver. Denote $\Gamma(n,r)$ the $K \times MF$ sub-matrix of $\Gamma$ containing the symbol $s_{n,r}$. We assume that each data stream contains $N_p$ pilot symbols known at the receiver, which corresponds to a pilot symbol sub-matrix $S_p \in \mathbb{C}^{N_p \times R}$ composed of the first $N_p$ rows of $S \in \mathbb{C}^{N \times R}$, with $N = N_p + N_d$, where $N_d$ is the number of information symbols transmitted per data stream. By definition of the Kronecker product, we have $\Gamma(n,r) = s_{n,r} H_{K \times MF}$. Applying the vec operator to this equation and stacking the resulting vectors, for $r = 1, \ldots, R$ and $n = 1, \ldots, N_p$, we obtain the following equation $\nu = A \text{vec}(H_{K \times MF})$, where

$$\nu = \begin{bmatrix} \text{vec}(\Gamma(1,1)) \\ \vdots \\ \text{vec}(\Gamma(N_p,R)) \end{bmatrix}, \quad A = \begin{bmatrix} s_{1,1} I_{MF K} \\ \vdots \\ s_{N_p, R} I_{MF K} \end{bmatrix}$$

The LS estimate of $\text{vec}(H_{K \times MF})$ that minimizes the LS cost function $\| \nu - A \text{vec}(H_{K \times MF}) \|_2^2$, is then given by

$$\hat{\nu} = (A^H A)^{-1} A^H \nu = \frac{1}{\|S_p\|_F^2} \sum_{r=1}^{N_p} \Gamma(n,r).$$

Applying the unvec operator leads to

$$\hat{H}_{K \times MF} = \frac{1}{\|S_p\|_F^2} \sum_{r=1}^{N_p} s_{n,r} \Gamma(n,r).$$

Once the channel estimated, each information symbol can be separately estimated in minimizing the LS cost function $\| \nu - A \text{vec}(H_{K \times MF}) \|_2^2$, which gives

$$\hat{s}_{n,r} = \frac{\text{vec}(\hat{H}_{K \times MF})^H \text{vec}(\Gamma(n,r))}{\|\hat{H}_{K \times MF}\|_F^2}$$

for $n = N_p + 1, \ldots, N; r = 1, \ldots, R$. The pilot-assisted KPLS-based receiver is summarized in Table IV. As for the ALS-based receiver, the semi-blind KPLS-based receiver used in our simulations is obtained by choosing $N_p = 1$, which corresponds to assuming the knowledge of the first row of the symbol matrix.

C. System design requirements for ALS and KPLS receivers

Applicability of the receivers described in Tables III and IV is conditioned on the left invertibility of $(I_{JP} \otimes \text{blkdiag}(H_{j})) G_{JPF \times Nr}$ and $(I_{JP} \otimes S) G_{JPF \times FM}$ for the ALS-based algorithm, while it depends on the right invertibility of $G_{RMF \times JPF}$ in the case of the KPLS-based algorithm. That induces the following necessary conditions, directly deduced from the dimensions of the matrices to be inverted: $JP \geq \max\{\frac{R}{M}, N_p\}$ and $JP \geq RM$, respectively. In the case of the ALS-based algorithm, applicability means uniqueness of the LS solution of Eq. (15) and (16) for estimating the symbol matrix and the channel tensor. This leads to the theorem below. We have to note that, due to its iterative nature, in the general case the ALS algorithm can converge to a local minimum. However, in our context where the core tensor is known, the two-step ALS receiver resulting from the linearization of the cost functions (20) and (21) around previous estimates, generally converges to the global minimum as we observed in the Monte Carlo simulations whose results are reported in Section VI.

**Theorem:** Let us assume that the symbol matrix $(S)$, the matrix slices $(H_{j}, f = 1, \ldots , F)$ of the channel tensor, and the core tensor unfoldings $G_{JPF \times Nr}$ and $G_{JPF \times FM}$ are full column-rank. Uniqueness of the LS estimates of $\mathcal{H}$ and $S$ is then guaranteed.

**Proof:** Rewrite (15) and (16) as $X_{JPF \times N} = V_1 S^T$ and $X_{JPF \times K} = V_2 H_{FM \times K}$. Uniqueness of the LS solution for estimating $S$ and $H_{FM \times K}$, requires that $V_1$ and $V_2$ be full column-rank. Assuming that $S$ is full-column-rank which implies $N \geq R$, we deduce that $I_{JP} \otimes S$ is also full-column-rank. The assumption $G_{JPF \times FM}$ full-column-rank, and the application of the Sylvester’s inequality $(r(A B)) \geq r(A) + r(B) - n$ for $A \in \mathbb{C}^{N \times M}$ and $B \in \mathbb{C}^{M \times N}$ allow to conclude that $V_2$ is also full-column-rank, which implies the uniqueness of the LS estimate of $H_{FM \times K}$. On the other hand, when $K \geq M$, the assumption of full-column-rank for $H_{j}, f = 1, \ldots , F$, implies that the block-diagonal matrix $I_{JP} \otimes \text{blkdiag}(H_{j})$ is also full-column-rank. The application of the Sylvester inequality combined with the assumption $G_{JPF \times MR}$ full-column-rank implies that $V_1$ is also full-column-rank, which implies the uniqueness of the LS estimate of $S$.

**Remarks:**

- If the propagation channel is assumed to be scattering-rich, it can be modeled with independent
and identically distributed (i.i.d.) entries drawn from a continuous distribution. Then, the matrices $H_j, f = 1, \ldots, F$, are full column-rank with probability one when $K \geq M$. It is important to notice that the random nature of the channel tensor allows to draw the same conclusion when $K \leq M$ implying that the matrices $H_j$ are full row-rank with probability one. Indeed, it is easy to prove that the product $AB$ is full column-rank if $B$ is full column-rank and $A$ is randomly drawn from a continuous distribution, independently of the dimensions of $A$.

- From (14), we can deduce that the rank conditions on the core tensor unfoldings are depending on both the code tensor $\mathcal{W}$ and the allocation tensor $\mathcal{C}$ (see Eq. (17) and (18)). When $\mathcal{G}$ is fixed as the code tensor, e.g. with full allocations, it is much easier to ensure the rank conditions by choosing, for instance, a Vandermonde structure for the code tensor.

\subsection*{D. Uniqueness issue}

Now, we show that when the core tensor has a constrained structure like a Vandermonde structure for instance, then the generalized Tucker-(2,5) model (26) is unique. For this purpose, this model is rewritten as a Tucker-(2,3) model obtained from variable changes $j_1 = (i_3 - 1)I_1 + i_1$ and $j_2 = (i_4 - 1)I_2 + i_2$. As well known, the Tucker models are generally not essentially unique. Indeed, their matrix factors can be only determined up to nonsingular transformations characterized by nonsingular matrices. For the Tucker-(2,3) model (23), let us define nonsingular matrices $U_{i_3} \in \mathbb{C}^{R_3 \times R_1}$ and $V_{i_4} \in \mathbb{C}^{R_4 \times R_2}$ for $i_3 = 1, \ldots, I_3$, and $i_4 = 1, \ldots, I_4$.

The core tensor $\mathcal{C}$ can be replaced by $\mathcal{C} \times [bdiag(U_{i_3})]^{-1} \times [bdiag(V_{i_4})]^{-1}$, and the factor matrices by $A_{bdiag(U_{i_3})}$ and $B_{bdiag(V_{i_4})}$, without modifying the model. Indeed, applying the property (8) of mode-$n$ product gives

$$
\mathcal{C} \times [bdiag(U_{i_3})]^{-1} \times [bdiag(V_{i_4})]^{-1} = \mathcal{C} \times A \times B,
$$

which shows the indeterminacy of the factor matrices $A$ and $B$ up to nonsingular block-diagonal matrices. That corresponds to an indeterminacy of the blocks $A_{i_3}^{(1)}$ and $A_{i_3}^{(2)}$ up to nonsingular matrices $U_{i_3}$ and $V_{i_4}$, respectively, for $i_3 = 1, \ldots, I_3$ and $i_4 = 1, \ldots, I_4$.

In terms of the core tensor of the generalized Tucker-(2,5) model, the indeterminacy can be formulated by the replacement of $\mathcal{G}$ by $\mathcal{G}' = \mathcal{G} \times 1 \times T^{(1)} \times 2 \times T^{(2)}$ where the third-order tensors $T^{(1)} \in \mathbb{C}^{R_3 \times R_1 \times I_3}$ and $T^{(2)} \in \mathbb{C}^{R_4 \times R_2 \times I_4}$ have their $i_3$th and $i_4$th mode-3 slice defined as $(U_{i_3})^{-1}$ and $(V_{i_4})^{-1}$, respectively. On the other hand, the tensor factors $A^{(1)}$ and $A^{(2)}$ are replaced by $A^{(1)}$ and $A^{(2)}$ with their $i_3$th and $i_4$th mode-3 slice given by $A^{(1)}_{i_3} = A^{(1)}_{i_3} U_{i_3}$ and $A^{(2)}_{i_4} = A^{(2)}_{i_4} V_{i_4}$, respectively.

To verify that the triplet $(\mathcal{G}', A^{(1)}, A^{(2)})$ is also a solution for $\mathcal{X}'$, we consider the following unfolding of $\mathcal{X}' = G' \times 1 \times A^{(1)} \times 2 \times A^{(2)}$ deduced from (28)

$$
X'_{I_3 I_4 I_2 I_1} = (I_{I_3} \otimes \text{bdiag}(A_{i_3}^{(1)})) \mathcal{G}'_{I_3 I_4 I_2 I_1} \cdot \text{bdiag}(A_{i_4}^{(2)})^T
$$

with

$$
G'_{I_3 I_4 I_2 I_1} = (I_{I_3} \otimes \text{bdiag}(U_{i_3}^{-1})) \mathcal{G}_{I_3 I_4 I_2 I_1} \cdot \text{bdiag}(V_{i_4}^{-1})^T.
$$

Replacing $G'_{I_3 I_4 I_2 I_1}$ by its expression (25) into (24) and using the associative property (3) give

$$
X'_{I_3 I_4 I_2 I_1} = (I_{I_3} \otimes \text{bdiag}(U_{i_3}^{-1})) \mathcal{G}_{I_3 I_4 I_2 I_1} \cdot (I_{I_4} \otimes \text{bdiag}(V_{i_4}^{-1}))^T
$$

which shows that $\mathcal{X}' = \mathcal{X}$.

Uniqueness can be obtained by imposing some constraints on the core tensor or the matrix factors. For instance, consider the matrix slice $\mathcal{G}_{I_3 I_4 I_2 I_1} \in \mathbb{C}^{R_3 \times R_2}$ of the core tensor, obtained by fixing the modes $i_3, i_4$ and $i_4$. When choosing this matrix slice with a Vandermonde structure, the ambiguity relation (25) gives

$$
G'_{i_3, i_4, i_2, i_1} = U_{i_3}^{-1} G_{i_3, i_4, i_2, i_1} V_{i_4}^{-1}
$$

The only admissible nonsingular matrices $U_{i_3}$ and $V_{i_4}$ are identity matrices multiplied by a scalar, otherwise the Vandermonde structure of $G_{i_3, i_4, i_2, i_1}$ is lost. In our context, where the core tensor representing the TSTF coding operation is assumed to be known at the receiver, the uniqueness of $\mathcal{S}$ and $\mathcal{H}$ up to a scalar factor is guaranteed under mild conditions. This can be seen by considering the matrix unfolding $X_{N K \times P F}$ of the received signal in (19). Let $\mathcal{S}' = \mathcal{S} \otimes H_{K \times MF}$ and $\mathcal{H}_{K \times MF}$ be alternative solutions satisfying the model, where $\mathcal{U} \in \mathbb{C}^{R X \times R}$ and $\mathcal{V} \in \mathbb{C}^{M \times MF}$ are nonsingular transformations. Inserting these matrices into (19) yields $X_{N K \times P F} = (\mathcal{S}' \otimes H_{K \times MF}) G_{RMF \times MF} = (\mathcal{S} \otimes H_{K \times MF})(U \otimes V) G_{RMF \times MF}$. Provided that $\mathcal{S} \otimes H_{K \times MF}$ is full column-rank and $G_{RMF \times MF}$ is full row-rank, we obtain $U \otimes V = I_{RMF}$. Note that such a Kronecker product is equal to the identity matrix if and only if $U$ and $V$ are (scaled) identity matrices that compensate each other, which means that $U = \alpha U$ and $V = \alpha^{-1} V$. This implies that the symbol matrix and channel tensor are
unique up to an unknown scalar factor. This ambiguity can be eliminated through normalization under the knowledge of a single pilot symbol at the receiver. It is worth mentioning, however, that the study of uniqueness of the generalized PARATUCK-(N₁, N) model is an open research topic.

E. Comparison with other tensor-based systems

In Tables V, VI and VII, we present several tensor-based systems in an unified way. The transmission rate (in bits per channel use) for each system is equal to \( \tau \log_2(\mu) \) where \( \mu \) is the cardinality of the information symbol constellation, with \( \tau \) given in Table VI. Due to the time spreading on chips (mode \( j \)), the spectral efficiency of the systems TSTF and TST is divided by the factor \( J \).

All the considered MIMO systems are remodeled by means of an equivalent generalized Tucker-(N₁, N) model. From these Tables, we can conclude that:

- The proposed TSTF system is a MIMO OFDM-CDMA system which generalizes the other three tensor-based systems considered in Tables V-VII. In particular, it combines a fifth-order tensor coding with a fourth-order resource allocation tensor.
- The TSTF system can be viewed as a CDMA extension of the STF system from the coding point of view, the STF system using only a bi-dimensional coding [9].
- The TST system can also be viewed as an OFDM extension of the TST system [8] with a multicarrier transmission and a tensor space-time-frequency (TSTF) coding instead of a third-order tensor coding.
- For each system, the core tensor unfoldings can be expressed in terms of the code matrix (or tensor) and the allocation matrices (or tensors). We have
  i) ST system:
  \[
  G_{PM×R} = \text{diag}(\text{vec}(\Phi^T))(\Psi \circ W),
  \]
  \[
  G_{PR×M} = \text{diag}(\text{vec}(\Psi^T))(\Phi \circ W^T).
  \]

  ii) STF system:
  \[
  G_{PFM×R} = \text{diag}(\text{vec}(\Phi_{M×PF}))(\Psi_{PF×R} \circ W),
  \]
  \[
  G_{PF×FM} = \text{diag}(\text{vec}(\Psi_{RF×PF}))(\Phi_{PF×FM} \circ W_{FR×FM}).
  \]
  with \( W_{FR×FM} = (I_F \otimes I_R)W^T(I_F \otimes I_M) \).

  iii) TST system:
  \[
  G_{JPM×R} = (I_J \otimes \text{diag}(\text{vec}(\Phi^T)))
  \begin{bmatrix}
  \Psi \circ W_{.1} \\
  \vdots \\
  \Psi \circ W_{.J}
  \end{bmatrix},
  \]
  \[
  G_{JPR×M} = (I_J \otimes \text{diag}(\text{vec}(\Psi^T)))
  \begin{bmatrix}
  \Phi \circ W_{.1}^T \\
  \vdots \\
  \Phi \circ W_{.J}^T
  \end{bmatrix}.
  \]

VI. SIMULATION RESULTS

We now present a set of computer simulation results to analyze the behavior of the proposed TSTF MIMO system for various system configurations. In terms of receiver processing, different receiver set-ups are considered. We first show some results concerning a zero forcing (ZF) receiver that operates under perfect channel knowledge. Then, we consider semi-blind and pilot-assisted ALS based receivers. The pilot-assisted ALS receiver consists of using a pilot sequence matrix \( S_p \in \mathbb{C}^{N_x×R} \) to obtain an initial estimate of the MIMO channel. This initial channel estimate is used as the input of an ALS refinement stage that jointly estimates symbols and channel iteratively. On the other hand, the semi-blind receiver avoids the use of pilot sequences, and relies only on the knowledge of the first row of \( S \), for eliminating scaling ambiguities. The goals of the following experiments are multi-fold: i) to evaluate the influence of the different dimensions of the tensor coding on the bit-error-rate (BER) performance; ii) to compare the performance of the TSTF system to that of some existing tensor-based MIMO schemes such as ST [15], TST [8], and STF [9] using semi-blind/pilot-assisted ALS receivers, especially when the number of data streams is much higher than the number of available transmit antennas and iii) to show the benefits of introducing “tensor allocations” of data streams across space, time and frequency domains.

The results represent an average of \( L = 10^4 \) Monte Carlo runs. At each run, the space-frequency MIMO channel coefficients are drawn from a complex-valued Gaussian distribution with zero-mean and unit variance. The transmitted symbols are randomly drawn from 4-PSK or 16-QAM alphabets. The BER curves are plotted as a function of the signal-to-noise ratio (SNR). At each run, the additive noise power is fixed according to the desired SNR given by

\[
\text{SNR}_{\text{dB}} = 10\log_{10} \frac{||X||^2_F}{||N||^2_F},
\]

where \( X \in \mathbb{C}^{K×N×F×P×J} \) is an additive noise tensor whose entries are circularly-symmetric complex-valued Gaussian random variables. Unless otherwise stated, our simulations assume that the space-stream-frequency-time allocation tensor \( C \in \mathbb{C}^{M×R×F×P} \) is equal to the “all-ones tensor”, i.e., full allocations, meaning that data streams are fully spread across all transmit antennas, sub-carriers and time-slots. The core tensor (\( \Phi \)) is then equal to the coding tensor (\( \Psi \)). Such an allocation structure which greatly simplifies the system design, corresponds to the case where all the entries of the allocation matrices/tensors, in Table V, are chosen equal to 1. It is worth noting that under this particular allocation configuration, the TSTF system operates with its maximum diversity gain. That is not the case for the STF, TST and ST systems. Indeed, for these systems, such an allocation configuration induces a simplification of the core tensor \((g_{m,r,f,p} = \omega_{m,r,p,j} = \omega_{m,r,j} \text{ and } g_{m,r,p} = \omega_{m,r}, \text{ respectively, from Table VI})\), which makes the received signals independent on \( p \) and corresponds to a loss of time diversity besides the absence of some diversities with respect to the TSTF system. This remark points...
out the crucial influence of the allocation structure on the performance of these three systems, contrarily to the TSTF one for which the choice of an all-ones allocation tensor gives the best performance.

A. ZF receiver performance under perfect channel knowledge

We first evaluate the performance of the TSTF coding system under perfect channel knowledge and with an all-ones allocation tensor. In this case, we use a ZF receiver for estimating the transmitted signals, given by:

\[
\hat{S}_{ZF}^T = \left[(I_{JP} \otimes \text{bdiag}(H_{J})) G_{JPFM \times R}\right]^\top \tilde{X}_{JPFK \times N}
\]

We consider a TSTF MIMO system with \( K = M = 2 \), \( R = 8 \), and \( N = 50 \). Figure 2 depicts the performance of different system configurations corresponding to different combinations of \( F, P \) and \( J \) values. Each \((F,P,J)\) set-up is related with a different special case of the TSTF system. For \((F,P,J) = (1,8,1)\), the TSTF coding system reduces to a particular case of the STF system presented in [9] with a full fourth-order coding tensor instead of a coding matrix with two allocation tensors.

It can be seen from Fig. 2, that the TSTF system with \((F,P,J) = (2,2,2)\) offers an improved performance over the other system configurations using the same overall spreading degree given by the product \( F^4 P J = 8 \) of coding dimensions. Such a gain comes from a more efficient tensor coding operation, which is done across more signal dimensions. It is worth noting that the TSTF system allows to exploit five coding dimensions (space, stream, frequency, time, chip) by means of a fifth-order coding tensor. That is not the case for the ST, TST and STF systems which exploit at the most three dimensions. Moreover, the use of a single allocation tensor of fourth-order offers more flexibility for allocations while facilitating the design of the coding tensor. Note also that the transmission rate of TST and TSTF is twice that of ST and STF.

B. Performance of pilot-assisted and semi-blind ALS receivers under different settings

In this Sub-section, we assess the BER performance by considering semi-blind and pilot-assisted ALS receivers for a joint channel and symbol estimation. We consider \( N_p = 10 \) and \( N_q = 50 \). The results are shown in Figure 3 for \( K = 2 \), \( M = 2 \), and \( R = 8 \). Again, it can be noted that the TSTF system offers an improved performance over the extensions of the ST, TST and STF ones. The performance gains of TSTF over the other systems are more pronounced for higher SNRs. These results corroborate the efficient use of both spatial and
For the chosen configurations, the four tensor-based systems have similar computational complexities since the product performance is obtained when the semi-blind ALS receiver for the TSTF system. Improved antennas and sub-carriers on the BER performance offered by KF Note the symmetrical roles played by KF (approximately 1dB) in favour of configuration with KF diversity gain (slope of the BER curves). Both configurations imply the same number of channel coefficients (2 = (2, 4, 1) − (2, 2, 2)). Although both configurations imply the same number of channel coefficients to be estimated, the improved performance of configuration (2, 4) is due to an additional coding gain brought by the coding across more sub-carriers.

Figure 4 displays the influence of the number of receive antennas and sub-carriers on the BER performance of the semi-blind ALS receiver for the TSTF system.

C. Impact of tensor allocations

As shown in (14) (or (17)-(18)), the core tensor G can be decomposed as the Hadamard product of the code tensor W with the allocation tensor C composed of 1’s and 0’s. The allocation tensor controls the assignment of data streams to transmit antennas, time-slots and sub-carriers. Our previous experiments assumed an allocation tensor composed of all-ones entries. Herein, the goal is to show that the allocation structure can indeed be exploited to cover situations where data streams have different levels of coding redundancy over space, time and/or frequency. The fixed system parameters are K = 2, R = 3, M = 4, F = P = J = 2, and N = 10. The semi-blind ALS algorithm is used at the receiver. We consider the following structure for the allocation tensor:

- data stream 1: \(c_{m,1,f,p} = 1\), for \(m \in \{1, 2, 3\}\) and \(f, p \in \{1, 2\}\);
- data stream 2: \(c_{m,2,f,p} = 1\), for \(m \in \{1, 2, 3\}, f = 1\) and \(p \in \{1, 2\}\);
- data stream 3: \(c_{m,3,f,p} = 1\), for \(m = 4\) and \(f, p \in \{1, 2\}\).

The remaining elements are equal to 0. From this choice, it can be deduced that i) the first and second data streams are partially spread in space (i.e. across three antennas), while the third data stream is transmitted by a single antenna only; ii) the first and third data streams are fully spread in time and frequency, while the second data stream is associated only with the first frequency but fully spread in time. Such an allocation scheme is useful, for instance, when we want to provide the data streams with different levels of protection against fading, which translates into different levels of redundancy in space, time and frequency. Figure 5 depicts the individual BER of each data stream. Note that the first data stream, which is fully spread in time and frequency domains, has the best performance followed by the second data stream, which is only partially spread in space and frequency. The third data stream that is not spread in space, presents the worst performance, compared to the other ones.
D. KPLS vs. ALS receivers

In the next experiment, we evaluate the performance of the KPLS receiver. The results are depicted in Figure 6 for a system with $K = M = 2$, $R = 4$, $P = 4$, $J = 2$, $N = 50$, $N_p = 10$, and 16-QAM. A comparison with the semi-blind ALS is shown for $F = 1$ (dashed lines) and $F = 2$ (solid lines). In both cases, KPLS outperforms ALS when a semi-blind setting is considered for these receivers. These gains come from the fact that KPLS is a closed-form solution that exploits the Kronecker product structure of the received data, while ALS is an iterative algorithm whose performance is sensitive to initialization. Additionally, KPLS is less computationally complex than ALS. When pilot sequences are used, an extra performance gain is obtained, as can be seen in this figure.

E. KPLS vs. ZF receivers

Now, we consider the semi-blind KPLS receiver and evaluate the impact of channel estimation on the BER performance of the TSTF system in comparison with the performance of the ZF receiver that assumes perfect channel knowledge. The results are shown in Figure 7 for a system with $K = 2$, $M = 2$, $R = 4$, $P = 4$, $J = 2$, $N = 50$ and 16-QAM. For a fixed BER value (e.g. $10^{-2}$), it can be observed that the SNR gap is around 5dB for $F = 1$, and increases for higher values of $F$. This is expected due to the higher number of parameters to be estimated. Note that such a gap can be reduced by resorting to pilot sequences.

F. KPLS receiver: Channel estimation performance

This section shows the performance of the semi-blind and pilot-assisted KPLS receivers in terms of channel estimation accuracy. The NMSE of the estimated channel, averaged over
L = 5000 Monte Carlo runs, is given by

\[ \text{NMSE}_{\text{DB}} = 10 \log_{10} \left( \frac{1}{L} \sum_{\ell=1}^{L} \frac{||\mathcal{H}(\ell) - \hat{\mathcal{H}}(\ell)||_F^2}{||\mathcal{H}(\ell)||_F^2} \right), \]

where \( \hat{\mathcal{H}}(\ell) \) denotes the \( \ell \)-th realization of the channel tensor, and \( \mathcal{H}(\ell) \) is the estimated channel tensor at the \( \ell \)-th run. The system parameters are the same as those of the previous section, with \( N = 100 \). Figure 8 shows the NMSE results for KPLS with different numbers of pilots. In all cases, the NMSE decreases linearly as a function of the SNR. It can be seen that KPLS yields satisfactory results in both semi-blind and pilot-assisted configurations. As expected, using a few pilots yields performance gains in terms of channel estimation accuracy, at the expense of some loss of bandwidth efficiency.

VII. CONCLUSION

A tensor-based space-time-frequency (TSTF) MIMO OFDM-CDMA transmission system with a flexible tensorial allocation structure has been developed with two semi-blind receivers, the so-called ALS- and KPLS-receivers. The original closed-form solution associated with the KPLS receiver presents the advantages to be non-iterative, of low computational complexity, and easily applicable in a supervised context, i.e. with a pilot sequence. Rewriting the generalized PARATUCK-(2,5) model as a generalized Tucker-(2,5) model, with different choices for the core tensor, has allowed to show that the proposed TSTF system can be viewed as an extension of three existing tensor-based ST/TST/STF systems, as summarized in Tables V-VII. It turns out that our TSTF system involves interesting performance/complexity tradeoffs in all the signaling dimensions, offers improved performance, and enjoys a higher allocation flexibility compared to previous tensor-based solutions. Finally, it is worth noting that the constrained tensor models introduced in this paper (i.e. generalized PARATUCK-(\( N_1, N \)) and Tucker-(\( N_1, N \)) models) open new perspectives of applications in other areas where constraints are needed to model different types of interactions between factors and new linear dependency structures. Among some perspectives of this work, we can mention the uniqueness issue of the generalized PARATUCK-(\( N_1, N \)) model, as already noted, and also the development of a procedure both for designing a good code and for optimizing the allocation tensor.

APPENDIX A

Proof of Proposition (6): Using the expression (5) of \( X_{S_1 S_2} \) and the properties (2) and (3) of the Kronecker product, we have

\[ \left( \otimes_{n \in S_1} e_{(1)n}^{(I)} \right)^T X_{S_1 S_2} \left( \otimes_{n \in S_2} e_{(2)n}^{(I)} \right) = \sum_{j_1=1}^{I_1} \cdots \sum_{j_N=1}^{I_N} x_{j_1,\ldots,j_N} \left( \otimes_{n \in S_1} e_{(1)n}^{(I)} \right)^T \left( \otimes_{n \in S_2} e_{(2)n}^{(I)} \right) = \sum_{j_1=1}^{I_1} \cdots \sum_{j_N=1}^{I_N} x_{j_1,\ldots,j_N} \prod_{n \in S_1} \delta_{n,j_n} \prod_{n \in S_2} \delta_{n,j_n} = x_{1,\ldots,N}. \]

APPENDIX B

Unfoldings of the generalized Tucker-(2,5) model: Let us consider the generalized Tucker-(2,5) model

\[ x_{1,2,3,4,5} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} g_{r_1,r_2,3,4,5} a_{(1)}^{(r_1)} a_{(2)}^{(r_2)} \]

\[ \mathcal{X} = \mathcal{G} \times_1 \mathcal{A}^{(1)} \times_2 \mathcal{A}^{(2)}. \]

APPENDIX C

Proof: From the formulae (6) and (7), we deduce the following expressions for \( a_{1,1,1,3}^{(1)} \), \( a_{1,2,2,4}^{(2)} \), and \( g_{1,2,3,4,5}^{(r_1,r_2,3,4,5)} \):

\[ a_{1,1,1,3}^{(1)} = (e_{13}^{(1)} \otimes e_{13}^{(1)})^T A_{1,1,1,3}^{(1)} e_{13}^{(1)}, \]

\[ a_{1,2,2,4}^{(2)} = (e_{14}^{(2)} \otimes e_{14}^{(2)})^T A_{1,2,2,4}^{(2)} e_{14}^{(2)}, \]

\[ g_{1,2,3,4,5}^{(r_1,r_2,3,4,5)} \cdot (e_{14}^{(2)} \otimes e_{14}^{(2)})^T = (e_{13}^{(1)} \otimes e_{13}^{(1)}) \cdot (e_{14}^{(2)} \otimes e_{14}^{(2)})^T \]

\[ \cdot G_{1,1,1,3}^{(r_1)} e_{13}^{(1)} e_{14}^{(2)} e_{13}^{(1)} e_{14}^{(2)} \]

\[ = (e_{13}^{(1)} \otimes e_{13}^{(1)}) \cdot (e_{14}^{(2)} \otimes e_{14}^{(2)})^T \]

\[ \cdot G_{1,1,1,3}^{(r_1)} e_{13}^{(1)} e_{14}^{(2)} e_{13}^{(1)} e_{14}^{(2)} \].
The matrix representation $X_{I_5 I_4 I_3 I_2}$ of $X$ can be deduced from (5)

$$X_{I_5 I_4 I_3 I_2} = \sum_{i_5=1}^{I_5} \sum_{i_4=1}^{I_4} \sum_{i_3=1}^{I_3} \sum_{i_2=1}^{I_2} x_{i_1 i_2 i_3 i_4 i_5} \cdot (e_{i_5} \otimes e_{i_4} \otimes e_{i_3} \otimes e_{i_2}) e_{i_2}^T,$$

Replacing $x_{i_1 i_2 i_3 i_4 i_5}$ and $g_{r_1 r_2 r_3 i_4 i_5}$ by their expressions (26) and (31) into (33) gives

$$X_{I_5 I_4 I_3 I_2} = \sum_{i_5=1}^{I_5} \sum_{i_4=1}^{I_4} \sum_{i_3=1}^{I_3} \sum_{i_2=1}^{I_2} \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} a_{i_1 r_1 i_2 r_2 i_3 i_4 i_5} \cdot (e_{i_5} \otimes e_{i_4} \otimes e_{i_3} \otimes e_{r_1 i_2}) (e_{r_1 i_2})^T.$$

Applying properties (2)-(4) of the Kronecker product yields

$$\sum_{i_5=1}^{I_5} \sum_{i_4=1}^{I_4} \sum_{i_3=1}^{I_3} \sum_{i_2=1}^{I_2} \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} a_{i_1 r_1 i_2 r_2 i_3 i_4 i_5} \cdot (e_{i_5} \otimes e_{i_4} \otimes e_{i_3} \otimes e_{r_1 i_2}) (e_{r_1 i_2})^T = e_{i_5} \otimes e_{i_4} \otimes e_{i_3} \otimes e_{r_1 i_2} (e_{r_1 i_2})^T.$$

Replacing (35) and (36) into (34) gives

$$X_{I_5 I_4 I_3 I_2} = \sum_{i_4=1}^{I_4} (I_{I_5} \otimes e_{i_4} \otimes e_{i_4} \otimes e_{i_4} \otimes e_{i_3}) e_{i_3}^T.$$

Taking into account the relation $I_{I_5} \otimes I_{I_4} = I_{I_5 I_4}$ and the
block partitioned form of \( \mathbf{G}_{I_2I_1I_1I_3R_1} \), we obtain (28) with

\[
\mathbf{A}_{I_2R_2I_1I_2}^{(2)} = \begin{bmatrix}
\mathbf{A}_{I_1I_1}^{(2)T} & \\
\vdots & \\
\mathbf{A}_{I_1I_1}^{(2)T} & \\
\mathbf{A}_{I_1I_1}^{(2)T}
\end{bmatrix} \in \mathbb{C}^{I_2 \times R_2 \times I_1}.
\]

Proceeding in the same way, with \( g_{r_1r_2I_1I_3} \) replaced by (32) instead of (31) into (33), it is easy to prove (29).

Another useful unfolding of \( \chi \) is \( X_{I_2I_1I_1I_3I_3} \), given by

\[
X_{I_2I_1I_1I_3I_3} = (A_{I_2I_2}^{(2)} \otimes A_{I_1I_1}^{(1)}) G_{R_2I_3R_1I_3I_1I_1I_1I_3}.
\]

In the particular case where the second factor \( A^{(2)} \) is a matrix, as encountered with the TSTF system, Eq (37) becomes

\[
X_{I_2I_1I_1I_3I_3} = (A^{(2)} \otimes A_{I_1I_1}^{(1)}) G_{R_2R_1I_1I_3I_1I_1I_3I_3}.
\]

where \( G_{R_2R_1I_1I_3I_1I_1I_3I_3} \) is defined in (39) on the top of the previous page.

REFERENCES


