Tensor modelling of MIMO communication systems with performance analysis and Kronecker receivers

Michele Nazareth da Costa, Gérard Favier, J M T Romano

Tensor modelling of MIMO communication systems with performance analysis and Kronecker receivers

Michele Nazareth da Costa[a], Gérard Favier[b], J.M.T. Romano[a]

[a] *DSPCom laboratory, University of Campinas, Campinas-SP, Brazil*
[b] *I3S Laboratory, University of Nice Sophia Antipolis, CNRS, Sophia Antipolis, France*

Abstract

The purpose of this paper is manifold. In a first part, we present a new alternating least squares (ALS)-based method for estimating the matrix factors of a Kronecker product, the so-called Kronecker ALS (KALS) method. Four other methods are also briefly described. In a second part, we consider the design of multiple-input multiple-output (MIMO) wireless communication systems using tensor modelling. Eight systems are presented in a unified way, and their theoretical performance is compared in terms of maximal diversity gain. Exploiting a Kronecker product of symbol and channel matrices, and applying the algorithms introduced in the first part, we propose three semi-blind and two supervised receivers, called Kronecker receivers, for jointly estimating the channel and the transmitted symbols. Necessary identifiability conditions are established. Finally, extensive Monte Carlo simulation results are provided to compare the performance of three tensor-based systems, on the one hand, and of the five proposed Kronecker receivers for the tensor space-time-frequency (TSTF) coding system, on the other hand.

Keywords: Channel estimation; Kronecker product; MIMO systems; semi-blind receivers; tensor coding; tensor modelling.

1. Introduction

Kronecker products, also known as tensor products, of matrices are currently used in many signal and image processing applications, like in compressive sensing with Kronecker dictionaries [1] and for image restoration [2]. They are useful in system theory [3] and in numerical linear algebra to write and solve linear matrix equations like Lyapunov and more generally Sylvester equations [4]. They also play an important role to simplify and implement fast transform algorithms like fast Fourier, Walsh-Hadamard, and Haar transforms [5, 6].

Email addresses: nazareth@decom.fee.unicamp.br (Michele Nazareth da Costa), favier@i3s.unice.fr (Gérard Favier), romano@dmo.fee.unicamp.br (J.M.T. Romano)

Recently, Kronecker and Khatri-Rao (column-wise Kronecker) products have been extensively employed in tensor-based system analysis and modelling, since such products naturally appear in matrix unfoldings of basic tensor decompositions, like the parallel factor (PARAFAC) [7] and Tucker [8] ones, and more generally of constrained PARAFAC models [9]. Reviews of the history and applications of the Kronecker product can be found in [3, 10, 5, 6, 11].

In the first part of the paper, we propose a new efficient computational algorithm based on the alternating least-squares (ALS) method for solving the Kronecker product approximation problem, i.e. for determining two matrices $\mathbf{A}$ and $\mathbf{B}$ of predetermined sizes, whose Kronecker product approximates a given matrix $\mathbf{C}$ in the sense of the minimization of the Frobenius norm $\|\mathbf{C}-\mathbf{A}\otimes\mathbf{B}\|_F$. We also briefly describe four other methods for solving this problem.

During the last decade, tensorial approaches have been widely developed to exploit multiple diversities in wireless communication systems. The principle of diversity techniques is to exploit several copies of the information symbols to be recovered at the receiver. This symbol repetition can result from multipath (due to multiantennas at the transmitter and receiver), repeated transmission of same symbols during several time-slots, and also from specific codings like, for instance, space-time (ST), space-frequency (SF) or space-time-frequency (STF) codings, which induce spatial multiplexing and temporal spreading. Tensor-based multiple-input multiple-output (MIMO) systems allow to improve link reliability as well as to jointly and semi-blindly estimate the channel and the transmitted symbols by means of deterministic receivers operating on data blocks. They have the advantage not to require a priori channel knowledge and long training sequences for estimating the channel. Only very few pilot symbols are needed to eliminate scaling ambiguities inherent to each particular tensor model. Moreover, tensor codings lead to natural tensor formulations of transmitted and received signals, and consequently to tensor system modellings. Tensor-based communication systems can be classified according to:

- the type of system (code-division multiple access (CDMA), orthogonal frequency division multiplexing (OFDM), CDMA-OFDM);
- the type of coding (ST, STF; matrices/tensors);
- the presence (in [12, 13, 14, 15, 16]) or not (in [17, 18, 19, 20, 21, 22]) of resource allocation, and their type (matrices/tensors);
- the type of tensor model: PARAFAC [17, 18], block PARAFAC [19, 20], BTD (block term decomposition) [21], CONFAC (constrained PARAFAC) [12], PARAFAC-Tucker2 (PARATUCK2) [13], PARATUCK-(2,4) [14], generalized PARATUCK [15, 16], nested PARAFAC [22].

A brief history of tensor-based systems is now reviewed by beginning with the fundamental work [17], which links direct-sequence CDMA (DS-CDMA) systems with a PARAFAC model. In [18], a space-time (ST) coding based on a Khatri-Rao (KR) product, denoted KRST, was derived by combining

a linear precoding for spatial multiplexing with a linear post-coding for temporal spreading. In [20], the idea of tensor coding was introduced for the first time. A three dimensional tensor allows to combine space-time coding and spatial multiplexing, hence the term space-time multiplexing (STM) coding. The third-order tensor containing the received signals satisfies a block constrained PARAFAC model, with two constraint matrices which depend on the multiplexing parameters.

In [12], a generalized ST spreading scheme was proposed for DS-CDMA systems, using a precoding tensor which allocates the users' data streams and spreading codes to transmit antennas, by means of three resource allocation matrices. The resulting transmission structure led to a third-order tensor model, called CONFAC, for the received signals. In [13], space-time spreading-multiplexing was proposed by combining a matrix precoding with stream- and antenna-to-slot matrix allocations. The third-order tensor of received signals then satisfies a PARATUCK2 model. In [14], the system of [13] was extended by considering a third-order tensor space-time coding, denoted TST, in order to exploit an extra chip diversity. That leads to a fourth-order PARATUCK-(2,4) model for the received signals tensor.

More recently, tensor approaches have been developed for OFDM, and OFDM-CDMA systems. In [22], a double Khatri-Rao STF coding, denoted DKRSTF, was proposed for OFDM systems. This coding constitutes an extension of the KRST coding [18], obtained by combining space-frequency pre-coding with time spreading. The received signals form a fourth-order tensor satisfying a nested PARAFAC model. In [15], a spatial coding matrix is combined with two third-order interaction tensors which control a joint time-frequency allocation of data streams and transmit antennas. In [16], the case of OFDM-CDMA systems is considered. The tensor space-time-frequency (TSTF) coding system was developed with the double objective of increasing the diversity gain by means of a fifth-order coding tensor, which allows to exploit four diversities (space, time, chip, and frequency) at the receiver, and simplifying the resource allocation by using a fourth-order allocation tensor to control the assignment of data streams to transmit antennas in the time-frequency domain. That leads to a generalized PARATUCK model for the fifth-order tensor containing the received signals.

The main contributions of this paper are summarized as follows:

- A new algorithm, called Kronecker-based ALS and denoted KALS, is proposed for solving the Kronecker product approximation problem.
- Eight tensor-based MIMO systems are presented in a unified way, using a generalized PARATUCK model [16].
- A comparative theoretical performance analysis is carried out for the considered tensor-based systems, and the maximal diversity gain is derived under the assumption of flat or frequency-selective fading channels. The transmission rate and the bandwidth of each system are also given.

- Five new receivers exploiting a Kronecker product of the channel and symbol matrices are derived for jointly estimating these matrices, three being semi-blind and two supervised. A necessary identifiability condition is established for each system.

- Extensive Monte Carlo simulation results are shown to compare the performance of three tensor-based systems, with zero-forcing (ZF) receivers in the case of perfect channel knowledge, on the one hand, and with the KALS receivers for joint semi-blind symbol/channel estimation, on the other hand. Then, the performance of the five proposed Kronecker receivers is compared for the TSTF system.

The rest of the paper is organized as follows. In Section 2, we present the KALS method for solving the Kronecker product approximation problem. Four other algorithms are also described. In Section 3, eight tensor-based systems are presented in a unified way using a generalized PARATUCK tensor model. In Section 4, a comparative theoretical performance analysis is carried out for these systems. Section 5 presents five new Kronecker receivers which use the Kronecker product approximation algorithms introduced in Section 2. Simulation results are shown in Section 6 to illustrate and compare the performance of the STF, TST, and TSTF systems, and also of the five proposed Kronecker receivers for the TSTF system. Finally, Section 7 concludes the paper with some perspectives for future work.

Notations and properties: Scalars, column vectors, matrices, and higher-order tensors are written with lower-case, boldface lower-case , boldface upper-case, and calligraphic letters, i.e. (a, $\mathbf{a}$, $\mathbf{A}$, $\mathcal{A}$), respectively. $\mathbf{A}^{\mathrm{T}}$, $\mathbf{A}^{\mathrm{H}}$, $\mathbf{A}^{*}$, and $\mathbf{A}^{\dagger}$ stand for transpose, Hermitian transpose, complex conjugate, and Moore-Penrose pseudo-inverse of $\mathbf{A}$, respectively. $\mathbf{e}_n^{(N)}$ is the n-th canonical basis vector of $\mathbb{R}^N$, $\mathbf{I}_N$ is the identity matrix of order N, $\mathbf{1}_N$ is the $N \times 1$ all-ones column vector, and $\|\cdot\|_F$ is the Frobenius norm. The operator $\mathrm{vec}(\cdot)$ forms a column vector by stacking the columns of its matrix argument, whereas $\mathrm{diag}(\cdot)$ forms a diagonal matrix from its vector argument, and $\mathrm{bdiag}(\mathbf{A}_1, \ldots, \mathbf{A}_K)$ forms a block-diagonal matrix with K diagonal blocks. The inverse of the vectorization operator is denoted unvec, so that $\mathbf{x} = \mathrm{vec}\,(\mathbf{X}) \in \mathbb{C}^{JI} \longleftrightarrow \mathbf{X} = \mathrm{unvec}\,(\mathbf{x}) \in \mathbb{C}^{I \times J}$. By convention, the order of dimensions in a product IJK is linked to the order of variation of the corresponding indices (i, j, k). For instance, given a third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ with entry $x_{i,j,k}$, its tall mode-1 matrix unfolding $\mathbf{X}_{JK \times I} \in \mathbb{C}^{JK \times I}$ corresponds to a combination of its modes (j,k) such that j varies more slowly than k, implying $x_{i,j,k} = [\mathbf{X}_{JK \times I}]_{(j-1)K+k,i}$. The Hadamard, Kronecker, and Khatri-Rao products are denoted by $\odot$, $\otimes$, and $\diamond$, respectively. Given $\mathbf{A} \in \mathbb{C}^{I \times J}, \mathbf{B} \in \mathbb{C}^{K \times L}, \mathbf{C} \in \mathbb{C}^{J \times M}, \mathbf{D} \in \mathbb{C}^{L \times N}$, we have

$$(\mathbf{A} \otimes \mathbf{B})\,(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{A}\,\mathbf{C}) \otimes (\mathbf{B}\,\mathbf{D}) \in \mathbb{C}^{IK \times MN}, \tag{1}$$

$$\mathbf{A} \otimes \mathbf{B} = (\mathbf{A}\,\mathbf{I}_J) \otimes (\mathbf{I}_K\,\mathbf{B}) = (\mathbf{A} \otimes \mathbf{I}_K)\,(\mathbf{I}_J \otimes \mathbf{B})\,, \tag{2}$$

$$\mathbf{A} \otimes \mathbf{B} = (\mathbf{I}_I\,\mathbf{A}) \otimes (\mathbf{B}\,\mathbf{I}_L) = (\mathbf{I}_I \otimes \mathbf{B})\,(\mathbf{A} \otimes \mathbf{I}_L)\,. \tag{3}$$

Given two tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N \times J_{N+1} \times \cdots \times J_{N+P}}$ and $\mathcal{B} \in \mathbb{C}^{I_1 \times \cdots \times I_N \times K_{N+1} \times \cdots \times K_{N+Q}}$, of respective orders $N+P$ and $N+Q$, we define the Hadamard product of $\mathcal{A}$ with $\mathcal{B}$, along their common modes $(i_1, \cdots, i_N)$, as the tensor $\mathcal{C} = \mathcal{A} \underset{\{i_1,\cdots,i_N\}}{\odot} \mathcal{B}$ of order $N+P+Q$ whose entries are given by $c_{i_1,\cdots,i_N,j_{N+1},\cdots,j_{N+P},k_{N+1},\cdots,k_{N+Q}} = a_{i_1,\cdots,i_N,j_{N+1}\cdots,j_{N+P}} b_{i_1,\cdots,i_N,k_{N+1},\cdots,k_{N+Q}}$.

A background with extended bibliography on tensor tools and decompositions, and their applications, is presented in tutorial papers [23, 9, 24]. Concerning the tensor-based systems considered in this paper, a detailed presentation can be found in the references given in Table 1.

2. Kronecker Product Approximation Methods

In this Section, we first present a new ALS-based method for estimating the matrix factors of a Kronecker product, the so-called KALS method. Then, four other methods for solving this problem are briefly described. In Section 5, these methods will be employed to derive three semi-blind and two supervised receivers for seven tensor-based communication systems.

2.1. Kronecker Alternating Least Squares (KALS) method

Consider the Kronecker product of $\mathbf{A} \in \mathbb{C}^{I \times J}$ and $\mathbf{B} \in \mathbb{C}^{K \times L}$

$$\begin{aligned} \mathbf{C} &= \mathbf{A} \otimes \mathbf{B} \in \mathbb{C}^{IK \times JL} \\ &\triangleq \begin{bmatrix} \mathbf{C}^{(1,1)} & \cdots & \mathbf{C}^{(1,J)} \\ \vdots & \ddots & \vdots \\ \mathbf{C}^{(I,1)} & \cdots & \mathbf{C}^{(I,J)} \end{bmatrix} = \begin{bmatrix} a_{1,1}\mathbf{B} & \cdots & a_{1,J}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{I,1}\mathbf{B} & \cdots & a_{I,J}\mathbf{B} \end{bmatrix}, \end{aligned} \tag{4}$$

where $\mathbf{C}^{(i,j)} \triangleq a_{i,j}\mathbf{B} \in \mathbb{C}^{K \times L}$ satisfies the following equation

$$\operatorname{vec}\left(\mathbf{C}^{(i,j)}\right) = a_{i,j} \operatorname{vec}(\mathbf{B}) \in \mathbb{C}^{LK \times 1}. \tag{5}$$

The least-squares (LS) estimate of the coefficient $a_{i,j}$ is given by

$$\hat{a}_{i,j} = \frac{(\operatorname{vec}(\mathbf{B}))^{\mathrm{H}} \operatorname{vec}\left(\mathbf{C}^{(i,j)}\right)}{\|\mathbf{B}\|_F^2}. \tag{6}$$

Reorganize the entries of $\mathbf{C}$ in the Kronecker product $\mathbf{D} = \mathbf{B} \otimes \mathbf{A}$ such as

$$\mathbf{D} \triangleq \mathbf{B} \otimes \mathbf{A} = \mathbf{\Pi}^{(\mathrm{row})}\, \mathbf{C}\, \mathbf{\Pi}^{(\mathrm{col})} \in \mathbb{C}^{KI \times LJ}, \tag{7}$$

$\mathbf{\Pi}^{(\mathrm{row})}$ and $\mathbf{\Pi}^{(\mathrm{col})}$ denoting row and column permutation matrices defined as

$$\begin{aligned} \mathbf{\Pi}^{(\mathrm{row})} &\triangleq \sum_{k=1}^{K} \sum_{i=1}^{I} \mathbf{e}_k^{(K)} \mathbf{e}_i^{(I)^{\mathrm{T}}} \otimes \mathbf{e}_i^{(I)} \mathbf{e}_k^{(K)^{\mathrm{T}}} \in \mathbb{R}^{KI \times IK}, \\ \mathbf{\Pi}^{(\mathrm{col})} &\triangleq \sum_{j=1}^{J} \sum_{l=1}^{L} \mathbf{e}_j^{(J)} \mathbf{e}_l^{(L)^{\mathrm{T}}} \otimes \mathbf{e}_l^{(L)} \mathbf{e}_j^{(J)^{\mathrm{T}}} \in \mathbb{R}^{JL \times LJ}. \end{aligned} \tag{8}$$

For the Kronecker product (7), eqs. (5) and (6) become

$$\text{vec}\Big(\mathbf{D}^{(k,l)}\Big) = b_{k,l}\,\text{vec}(\mathbf{A}) \in \mathbb{C}^{JI\times 1}, \tag{9}$$

$$\hat{b}_{k,l} = \frac{(\text{vec}(\mathbf{A}))^{\text{H}}\,\text{vec}\big(\mathbf{D}^{(k,l)}\big)}{\|\mathbf{A}\|_F^2}. \tag{10}$$

The KALS method consists in iteratively and alternately estimating the matrices $\mathbf{A}$ and $\mathbf{B}$ using (6) and (10), as summarized in Algorithm 1. Note that the Kronecker product is characterized by a scalar ambiguity in the sense that $(\alpha\mathbf{A}, \frac{1}{\alpha}\mathbf{B})$ give the same Kronecker product than $(\mathbf{A}, \mathbf{B})$. This scalar ambiguity can be removed with the knowledge of only one coefficient in $\mathbf{A}$ (or $\mathbf{B}$) as shown in Algorithm 1, where $a_{1,1}$ is assumed a priori known and used for initialization, and $(\hat{\mathbf{A}}(\infty), \hat{\mathbf{B}}(\infty))$ denote the estimates at convergence.

Algorithm 1: KALS

1. Compute the Kronecker product $\mathbf{D}$ from the input matrix $\mathbf{C}$ by using (7)-(8).
2. $it = 0$: Initialize $\hat{\mathbf{B}}(0) = \frac{1}{a_{1,1}}\mathbf{C}^{(1,1)}$.
3. $it = it + 1$.
4. Compute for $i = 1, ..., I; j = 1, ..., J$:
$\hat{a}_{i,j}(it) = \frac{\left(\text{vec}\left(\hat{\mathbf{B}}(it-1)\right)\right)^{\text{H}}\text{vec}\left(\mathbf{C}^{(i,j)}\right)}{\|\hat{\mathbf{B}}(it-1)\|_F^2}$, $\hat{b}_{k,l}(it) = \frac{\left(\text{vec}\left(\hat{\mathbf{A}}(it)\right)\right)^{\text{H}}\text{vec}\left(\mathbf{D}^{(k,l)}\right)}{\|\hat{\mathbf{A}}(it)\|_F^2}$.
5. Repeat steps (3)-(4) until convergence.
6. Eliminate the scaling ambiguity with $\alpha = \frac{a_{1,1}}{\hat{a}_{1,1}(\infty)}$:
$\hat{\mathbf{A}}(\infty) \leftarrow \alpha\,\hat{\mathbf{A}}(\infty)$, $\hat{\mathbf{B}}(\infty) \leftarrow \frac{1}{\alpha}\,\hat{\mathbf{B}}(\infty)$.

When the Kronecker product $\mathbf{C}$ is measured with an additive noise and the first I_p rows of $\mathbf{A}$ are known, i.e. the sub-matrix $\mathbf{A}_{1:I_P} \in \mathbb{C}^{I_P\times J}$, the matrices $(\mathbf{A}_{I_P+1:I}, \mathbf{B})$ can be estimated using the non iterative method, named Kronecker product least-squares (KPLS), proposed in [25, 16] and obtained by combining (6) for estimating $\mathbf{A}_{I_P+1:I} \in \mathbb{C}^{(I-I_P)\times J}$, with the following equation

$$\hat{\mathbf{B}} = \frac{1}{\|\ \mathbf{A}_{1:I_p}\ \|_F^2}\sum_{j=1}^{J}\sum_{i=1}^{I_p} a_{i,j}^{*}\mathbf{C}^{(i,j)}. \tag{11}$$

Another method, called supervised KALS and denoted SKALS, consists in using (11) as the initial value $\hat{\mathbf{B}}(0)$ in step 2 of Algorithm 1. These methods SKALS and KPLS will be exploited in Section 5 for deriving two supervised receivers.

2.2. Kronecker Singular Value Decomposition (KSVD) method

The Kronecker product approximation problem can be solved by computing a rank-one matrix approximation [26], i.e. by computing the singular vectors associated with the largest singular value of a matrix built from the vectorization of the two factors $(\mathbf{A}, \mathbf{B})$. Indeed, we have

$$\|\mathbf{C} - \mathbf{A}\otimes\mathbf{B}\|_F^2 = \|\mathbf{Y} - \text{vec}(\mathbf{B})\,(\text{vec}(\mathbf{A}))^{\text{T}}\,\|_F^2 = \|\mathbf{y} - \text{vec}(\mathbf{A})\diamond\text{vec}(\mathbf{B})\,\|_2^2, \tag{12}$$

where $\mathbf{Y} \in \mathbb{C}^{LK \times JI}$ and $\mathbf{y} = \text{vec}(\mathbf{Y}) \in \mathbb{C}^{JILK \times 1}$ can be obtained by permuting the elements of $\text{vec}(\mathbf{C}) \in \mathbb{C}^{JLIK \times 1}$ as follows

$$\begin{aligned} \mathbf{y} &= \text{vec}(\mathbf{Y}) = \mathbf{\Pi}\, \text{vec}(\mathbf{C}), \\ \mathbf{\Pi} &= \mathbf{I}_J \otimes \left(\sum_{i=1}^{I} \sum_{l=1}^{L} \mathbf{e}_i^{(I)} \mathbf{e}_l^{(L)^{\mathrm{T}}} \otimes \mathbf{e}_l^{(L)} \mathbf{e}_i^{(I)^{\mathrm{T}}} \right) \otimes \mathbf{I}_K \in \mathbb{R}^{JILK \times JLIK}. \end{aligned} \tag{13}$$

Considering the rank-one approximation $\mathbf{Y} \approx \sigma_1 \mathbf{u}\mathbf{v}^{\mathrm{H}}$, where $\mathbf{u}$ and $\mathbf{v}$ are the left and right singular vectors associated with the largest singular value σ_1, one deduces the following estimates $\hat{\mathbf{A}} = \alpha \sqrt{\sigma_1}\, \text{unvec}(\mathbf{v}^*)$, $\hat{\mathbf{B}} = {}^1\!/\!_\alpha \sqrt{\sigma_1}\, \text{unvec}(\mathbf{u})$, where the scalar α is calculated as in Algorithm 1. The *power method* [27] can be applied for computing this rank-one approximation of $\mathbf{Y}$. At each iteration, the right and left singular vectors are calculated as follows

$$\mathbf{v}(it) = \frac{\mathbf{Y}^{\mathrm{H}}\, \mathbf{u}(it-1)}{\left\| \mathbf{Y}^{\mathrm{H}}\, \mathbf{u}(it-1) \right\|_2}, \quad \sigma_1(it) = \left\| \mathbf{Y}\, \mathbf{v}(it) \right\|_2, \quad \mathbf{u}(it) = \frac{\mathbf{Y}\, \mathbf{v}(it)}{\sigma_1(it)}. \tag{14}$$

The KSVD method is summarized in Algorithm 2.

Algorithm 2: KSVD

1. $it = 0$: Initialize $\hat{\mathbf{u}}(0) = \text{vec}\left(\hat{\mathbf{B}}(0)\right) = \frac{1}{a_{1,1}} \text{vec}(\mathbf{C}^{(1,1)})$.
2. Compute the matrix $\mathbf{Y} = \text{unvec}(\mathbf{\Pi}\, \text{vec}(\mathbf{C}))$ with $\mathbf{\Pi}$ defined in (13).
3. $it = it + 1$.
4. Compute the rank-one approximation of $\mathbf{Y} \approx \sigma_1 \mathbf{u}\mathbf{v}^{\mathrm{H}}$ using (14).
5. Repeat steps (3)-(4) until convergence.
6. Compute the estimates of $\mathbf{A}$ and $\mathbf{B}$, with $\alpha = \frac{a_{1,1}}{\hat{a}_{1,1}(\infty)}$:
   $\hat{\mathbf{A}} = \alpha \sqrt{\sigma_1(\infty)}\, \text{unvec}(\mathbf{v}^*(\infty))$, $\hat{\mathbf{B}} = \frac{1}{\alpha} \sqrt{\sigma_1(\infty)}\, \text{unvec}(\mathbf{u}(\infty))$.

2.3. Kronecker Alternating Least Mean Squares (KALMS) method

In [28], the identities (2)-(3) are exploited to iteratively estimate the factors ($\mathbf{A}$,$\mathbf{B}$) by applying a Kronecker-based alternating least mean squares (KALMS) algorithm to the input-output relationship $\mathbf{y}(it) = \mathbf{C}\, \mathbf{x}(it) + \mathbf{e}(it) \in \mathbb{C}^{IK \times 1}$, which defines a MIMO system from the Kronecker product $\mathbf{C} = \mathbf{A} \otimes \mathbf{B}$, with $\mathbf{x}(it) \in \mathbb{C}^{JL \times 1}$ randomly generated. The signal $\mathbf{e}(it)$ representing both measurement noise and modelling error, can be written in the following forms

$$\begin{aligned} \mathbf{e}(it) &= \mathbf{y}(it) - (\mathbf{A} \otimes \mathbf{B})\, \mathbf{x}(it) \\ &= \mathbf{y}(it) - (\mathbf{A} \otimes \mathbf{I}_K)\, (\mathbf{I}_J \otimes \mathbf{B})\, \mathbf{x}(it) \qquad (15) \\ &= \mathbf{y}(it) - (\mathbf{I}_I \otimes \mathbf{B})\, (\mathbf{A} \otimes \mathbf{I}_L)\, \mathbf{x}(it). \qquad (16) \end{aligned}$$

The factors ($\mathbf{A}$,$\mathbf{B}$) are estimated by minimizing the LS cost function $\|\mathbf{e}(it)\|_2^2$. This nonlinear optimization problem is replaced by the alternating minimization

of two quadratic cost functions obtained by fixing one of the factors to its previous estimated value in (15)-(16), i.e.

$$\begin{cases} \mathbf{z}^{\mathbf{A}}(it) \triangleq \left(\mathbf{I}_J \otimes \hat{\mathbf{B}}(it-1)\right)\mathbf{x}(it) \in \mathbb{C}^{JK\times 1} \\ \mathbf{e}^{\mathbf{A}}(it) \triangleq \mathbf{y}(it) - (\mathbf{A}\otimes\mathbf{I}_K)\,\mathbf{z}^{\mathbf{A}}(it) \\ \hat{\mathbf{A}}(it) = \min\limits_{\mathbf{A}} \mathrm{E}\left[\|\mathbf{e}^{\mathbf{A}}(it)\|_2^2\right] \end{cases}, \quad \begin{cases} \mathbf{z}^{\mathbf{B}}(it) \triangleq \left(\hat{\mathbf{A}}(it)\otimes\mathbf{I}_L\right)\mathbf{x}(it) \in \mathbb{C}^{IL\times 1} \\ \mathbf{e}^{\mathbf{B}}(it) \triangleq \mathbf{y}(it) - (\mathbf{I}_I\otimes\mathbf{B})\,\mathbf{z}^{\mathbf{B}}(it) \\ \hat{\mathbf{B}}(it) = \min\limits_{\mathbf{B}} \mathrm{E}\left[\|\mathbf{e}^{\mathbf{B}}(it)\|_2^2\right] \end{cases}.$$

At each iteration, the LMS algorithm is used to update alternately the estimate of $\mathbf{A}$ and $\mathbf{B}$. That results in K and I estimates of $\mathbf{A}$ and $\mathbf{B}$, respectively, due to the presence of the Kronecker products $\hat{\mathbf{A}}(it)\otimes\mathbf{I}_K$ and $\mathbf{I}_I\otimes\hat{\mathbf{B}}(it)$. Taking the mean value of these estimates gives the Algorithm 3.

Algorithm 3: KALMS

1. Set $\gamma_{\mathbf{A}}$ and $\gamma_{\mathbf{B}}$.
2. $it=0$: Initialize $\hat{\mathbf{B}}(0) = \frac{1}{a_{1,1}}\mathbf{C}^{(1,1)}$ and randomly initialize $\hat{\mathbf{A}}(0)$.
3. $it = it+1$.
4. Randomly generate the input signal $\mathbf{x}(it)$ and compute $\mathbf{y}(it) = \mathbf{C}\,\mathbf{x}(it)$.
5. Update of the estimate $\hat{\mathbf{A}}(it)$:
   $\mathbf{z}^{\mathbf{A}}(it) \triangleq \left(\mathbf{I}_J \otimes \hat{\mathbf{B}}(it-1)\right)\mathbf{x}(it)$, $\mathbf{Z}^{\mathbf{A}}(it) \triangleq \left(\,\mathrm{unvec}\big(\mathbf{z}^{\mathbf{A}}(it)\big)\,\right)^{\mathrm{T}} \in \mathbb{C}^{J\times K}$,
   $\mu^{\mathbf{A}}(it) \triangleq \gamma_{\mathbf{A}}/\left\|\mathbf{z}^{\mathbf{A}}(it)\right\|_2^2$,
   $\mathbf{e}^{\mathbf{A}}(it) \triangleq \mathbf{y}(it) - \left(\hat{\mathbf{A}}(it-1)\otimes\mathbf{I}_K\right)\mathbf{z}^{\mathbf{A}}(it)$, $\mathbf{E}^{\mathbf{A}}(it) \triangleq \left(\,\mathrm{unvec}\big(\mathbf{e}^{\mathbf{A}}(it)\big)\,\right)^{\mathrm{T}} \in \mathbb{C}^{I\times K}$,
   $\hat{\mathbf{A}}(it) = \hat{\mathbf{A}}(it-1) + \frac{\mu^{\mathbf{A}}(it)}{K}\mathbf{E}^{\mathbf{A}}(it)\left(\mathbf{Z}^{\mathbf{A}}(it)\right)^{\mathrm{H}}$.
6. Update of the estimate $\hat{\mathbf{B}}(it)$:
   $\mathbf{z}^{\mathbf{B}}(it) \triangleq \left(\hat{\mathbf{A}}(it)\otimes\mathbf{I}_L\right)\mathbf{x}(it)$, $\mathbf{Z}^{\mathbf{B}}(it) \triangleq \mathrm{unvec}\big(\mathbf{z}^{\mathbf{B}}(it)\big) \in \mathbb{C}^{L\times I}$,
   $\mu^{\mathbf{B}}(it) \triangleq \gamma_{\mathbf{B}}/\left\|\mathbf{z}^{\mathbf{B}}(it)\right\|_2^2$,
   $\mathbf{e}^{\mathbf{B}}(it) \triangleq \mathbf{y}(it) - \left(\mathbf{I}_I \otimes \hat{\mathbf{B}}(it-1)\right)\mathbf{z}^{\mathbf{B}}(it)$, $\mathbf{E}^{\mathbf{B}}(it) \triangleq \mathrm{unvec}\big(\mathbf{e}^{\mathbf{B}}(it)\big) \in \mathbb{C}^{K\times I}$,
   $\hat{\mathbf{B}}(it) = \hat{\mathbf{B}}(it-1) + \frac{\mu^{\mathbf{B}}(it)}{I}\mathbf{E}^{\mathbf{B}}(it)\left(\mathbf{Z}^{\mathbf{B}}(it)\right)^{\mathrm{H}}$.
7. Repeat steps (3)-(6) until convergence.
8. Eliminate the scaling ambiguity with $\alpha = \frac{a_{1,1}}{\hat{a}_{1,1}(\infty)}$:
   $\hat{\mathbf{A}}(\infty) \leftarrow \alpha\,\hat{\mathbf{A}}(\infty)$, $\hat{\mathbf{B}}(\infty) \leftarrow \frac{1}{\alpha}$, $\hat{\mathbf{B}}(\infty)$.

3. Tensor modelling of MIMO communication systems

We first show that the TSTF coding structure, recently proposed in [16] for MIMO OFDM-CDMA systems, allows to deduce seven other tensor-based systems as particular cases.

Consider a MIMO system with M transmit and K receive antennas. The transmission is decomposed into P time blocks of N symbol periods, each one being composed of J chips. During each time block p, the transceiver uses F subcarriers to send R data streams containing N information symbols each, which form the symbol matrix $\mathbf{S} \in \mathbb{C}^{N\times R}$ with entries $s_{n,r}$, $n=1,...,N; r=1,...,R$. The transmission system is characterized by two tensors: a fifth-order

coding tensor $\mathcal{W} \in \mathbb{C}^{M\times R\times F\times P\times J}$ and a fourth-order resource allocation tensor $\mathcal{C} \in \mathbb{R}^{M\times R\times F\times P}$ composed uniquely of 1's and 0's, $c_{m,r,f,p}=1$ meaning that the data stream r is transmitted using the transmit antenna m and the subcarrier f, during the time-block p. At the symbol period n of block p, the transceiver transmits a linear combination of R coded signals according to the equation

$$u_{m,n,f,p,j} = \sum_{r=1}^{R} w_{m,r,f,p,j}\, s_{n,r}\, c_{m,r,f,p}, \tag{17}$$

where the coefficient $c_{m,r,f,p}$ of the allocation tensor $\mathcal{C}$ fixes the space-frequency resource (m, f) used to send the symbol $s_{n,r}$ during the time block p. So, the allocation tensor controls the space-time-frequency spreading-mutiplexing. Eq. (17) shows that the multiplication by the coding tensor $\mathcal{W}$ allows to replicate each symbol $s_{n,r}$ four times, in the space (m), frequency (f), time (p), and chip (j) dimensions. The high-order of the coding tensor is at the origin of a performance improvement over other systems such as the ST, STF, and TST ones. This result will be theoretically established in the next section by a comparative analysis of the diversity gains.

The frequency-selective fading channel coefficients $h_{k,m,f}$ between each pair (m, k) of transmit and receive antennas, at frequency f, are assumed constant during P time-blocks, independent, and circularly symmetric complex Gaussian variables, with zero-mean and unit variance. They form a third-order tensor $\mathcal{H} \in \mathbb{C}^{K\times M\times F}$. In the noiseless case, the received signals define a fifth-order tensor $\mathcal{X} \in \mathbb{C}^{K\times N\times F\times P\times J}$ defined as

$$x_{k,n,f,p,j} = \sum_{m=1}^{M} h_{k,m,f}\, u_{m,n,f,p,j} = \sum_{m=1}^{M}\sum_{r=1}^{R} g_{m,r,f,p,j}\, h_{k,m,f}\, s_{n,r}, \tag{18}$$

$$g_{m,r,f,p,j} = w_{m,r,f,p,j}\, c_{m,r,f,p}.$$

The core tensor $\mathcal{G} \in \mathbb{C}^{M\times R\times F\times P\times J}$ can be interpreted as the Hadamard product of the coding tensor with the allocation tensor, along their common modes (m, r, f, p), i.e. $\mathcal{G} = \mathcal{W} \underset{\{m,r,f,p\}}{\odot} \mathcal{C}$.

The received signal $x_{k,n,f,p,j}$ satisfies the generalized PARATUCK-(2,5) model introduced in [16], and defined as follows

$$x_{i_1,i_2,i_3,i_4,i_5} = \sum_{r_1=1}^{R_1}\sum_{r_2=1}^{R_2} g_{r_1,r_2,i_3,i_4,i_5}\, a^{(1)}_{i_1,r_1,i_3}\, a^{(2)}_{i_2,r_2}, \tag{19}$$

$$g_{r_1,r_2,i_3,i_4,i_5} = w_{r_1,r_2,i_3,i_4,i_5}\, c_{r_1,r_2,i_3,i_4}.$$

Comparing (18) with (19), we deduce the following correspondences

$$\left(I_1, I_2, I_3, I_4, I_5, R_1, R_2, \mathcal{A}^{(1)}, \mathbf{A}^{(2)}\right) \leftrightarrow (K, N, F, P, J, M, R, \mathcal{H}, \mathbf{S})\,. \tag{20}$$

Particular cases

In Tables 1 and 2, we present in a unified way eight tensor-based MIMO systems which can be deduced as particular cases of the TSTF system. In Table 1, the core tensor $\mathcal{G}$ and the received signal tensor $\mathcal{X}$ are given for each system, while Table 2 contains the design parameters for each system rewritten as a generalized PARATUCK-(2,5) model (19).

Table 1: Presentation of eight tensor-based systems.

Systems	**Core tensors**	**Received signals**
TSTF [16]	$g_{m,r,f,p,j} = w_{m,r,f,p,j}\, c_{m,r,f,p}$	$x_{k,n,f,p,j} = \sum_{m=1}^{M} \sum_{r=1}^{R} g_{m,r,f,p,j}\, h_{k,m,f}\, s_{n,r}$
STF [15]	$g_{m,r,f,p} = w_{m,r}\, c^{(\mathcal{H})}_{m,f,p}\, c^{(\mathbf{S})}_{r,f,p}$	$x_{k,n,f,p} = \sum_{m=1}^{M} \sum_{r=1}^{R} g_{m,r,f,p}\, h_{k,m,f}\, s_{n,r}$
TST [14]	$g_{m,r,p,j} = w_{m,r,j}\, c^{(\mathbf{H})}_{m,p}\, c^{(\mathbf{S})}_{r,p}$	$x_{k,n,p,j} = \sum_{m=1}^{M} \sum_{r=1}^{R} g_{m,r,p,j}\, h_{k,m}\, s_{n,r}$
ST [13]	$g_{m,r,p} = w_{m,r}\, c^{(\mathbf{H})}_{m,p}\, c^{(\mathbf{S})}_{r,p}$	$x_{k,n,p} = \sum_{m=1}^{M} \sum_{r=1}^{R} g_{m,r,p}\, h_{k,m}\, s_{n,r}$
STM [20]	$g_{m,r,p} = w_{m,r,p}$	$x_{k,n,p} = \sum_{m=1}^{M} \sum_{r=1}^{R} g_{m,r,p}\, h_{k,m}\, s_{n,r}$
DKRSTF [22]	$g_{m,r,f,p} = w_{m,r,f,p} = \theta_{m,r}\, \omega_{f,r}\, \psi_{p,m}$	$x_{k,n,f,p} = \sum_{m=1}^{M} \sum_{r=1}^{M} g_{m,r,f,p}\, h_{k,m}\, s_{n,r}$
KRST [18]	$g_{m,r,p} = w_{m,r,p} = \theta_{m,r}\, \psi_{p,m}$	$x_{k,n,p} = \sum_{m=1}^{M} \sum_{r=1}^{M} g_{m,r,p}\, h_{k,m}\, s_{n,r}$
DS-CDMA [17]	$g_{j,m} = w_{j,m}$	$x_{k,n,j} = \sum_{m=1}^{M} g_{j,m}\, h_{k,m}\, s_{n,m}$

Table 2: Design parameters for the associated generalized PARATUCK-(2,5) model.

Systems	**Design parameters**										
	I_1	I_2	I_3	I_4	I_5	R_1	R_2	$\mathcal{A}^{(1)}$	$\mathbf{A}^{(2)}$	$\mathcal{C}$	$\mathcal{W}$
TSTF [16]	K	N	F	P	J	M	R	$\mathcal{H}$	$\mathbf{S}$	$\mathcal{C}$	$\mathcal{W}$
STF [15]	K	N	F	P	-	M	R	$\mathcal{H}$	$\mathbf{S}$	$\mathcal{C}^{(\mathcal{H})}, \mathcal{C}^{(\mathbf{S})}$	$\mathbf{W}$
TST [14]	K	N	P	J	-	M	R	$\mathbf{H}$	$\mathbf{S}$	$\mathbf{C}^{(\mathbf{H})}, \mathbf{C}^{(\mathbf{S})}$	$\mathcal{W}$
ST [13]	K	N	P	-	-	M	R	$\mathbf{H}$	$\mathbf{S}$	$\mathbf{C}^{(\mathbf{H})}, \mathbf{C}^{(\mathbf{S})}$	$\mathbf{W}$
STM [20]	K	N	P	-	-	M	R	$\mathbf{H}$	$\mathbf{S}$	-	$\mathcal{W}$
DKRSTF [22]	K	N	F	P	-	M	M	$\mathbf{H}$	$\mathbf{S}$	-	$\mathcal{W}$
KRST [18]	K	N	P	-	-	M	M	$\mathbf{H}$	$\mathbf{S}$	-	$\mathcal{W}$
DS-CDMA [17]	K	N	J	-	-	M	-	$\mathbf{H}$	$\mathbf{S}$	-	$\mathbf{W}$

From these two tables, we can draw the following conclusions:

- Two main features distinguish TSTF from the other systems. The first one concerns the use of a fifth-order tensor ($\mathcal{W}$) for space-time-frequency coding, instead of a fourth-order (or third-order) coding tensor for DKRSTF (or TST, STM, and KRST), respectively, or of a coding matrix in the case of STF, ST, and DS-CDMA. The five dimensional coding tensor allows to increase the diversity gain, which facilitates performance/complexity tradeoffs in all the signaling dimensions. The

second one is linked to the use of a fourth-order allocation tensor ($\mathcal{C}$), while the other systems use either two third-order tensors as with STF, or two matrices as with ST and TST. The use of a single fourth-order allocation tensor provides higher flexibility for allocations.

- The TSTF system can be viewed as an OFDM extension of the TST system with a multicarrier transmission, and a CDMA extension of the STF system. It can also be viewed as an extension of the DKRSTF system which is itself an extension of the KRST one. Indeed, for KRST coding, the coded signals define a third-order tensor $\mathcal{U} \in \mathbb{C}^{M \times N \times P}$ such as

$$u_{m,n,p} = \sum_{r=1}^{M} \theta_{m,r}\, \psi_{p,m}\, s_{n,r} \ , \ \mathbf{U}_{NP \times M} = \mathbf{S}\mathbf{\Theta}^{\mathrm{T}} \diamond \mathbf{\Psi}, \tag{21}$$

  while for the DKRSTF coding, the tensor $\mathcal{U} \in \mathbb{C}^{M \times N \times F \times P}$ is such as

$$u_{m,n,f,p} = \sum_{r=1}^{M} \theta_{m,r}\, \omega_{f,r}\, \psi_{p,m}\, s_{n,r} \ , \ \mathbf{U}_{NFP \times M} = (\mathbf{S} \diamond \mathbf{\Omega})\mathbf{\Theta}^{\mathrm{T}} \diamond \mathbf{\Psi}. \tag{22}$$

  This matrix unfolding highlights the double Khatri-Rao STF coding, the first one corresponding to a space-frequency pre-coding, whereas the second one corresponds to a time post-coding.
  These Eqs. (21) and (22) are to be compared with (17), showing that the KRST and DKRSTF systems exploit third- and fourth-order coded signals tensors, respectively, when TSTF uses a fifth-order coding tensor. Note also the restrictive assumption for DKRSTF which requires the channel $\mathbf{H} \in \mathbb{C}^{K \times M}$ constant across the F subcarriers, while it is a third-order tensor $\mathcal{H} \in \mathbb{C}^{K \times M \times F}$ depending on the F frequencies, in the case of TSTF. Another restriction shared by DKRSTF and KRST concerns the number of transmitted data streams which must be equal to the number of transmit antennas ($R = M$), which is not the case of the other systems.

- In their original formulation, the received signals tensors satisfy the following tensor models: PARAFAC, block PARAFAC, nested PARAFAC, PARATUCK2, PARATUCK-(2,4), and generalized PARATUCK-(2,5) for the (DS-CDMA, KRST), STM, DKRSTF, ST, TST, and (STF,TSTF) systems, respectively. Moreover, all the systems use iterative ALS-based receivers for jointly estimating the channel and the information symbols.

- The rewriting of the received signals tensors presented in Table 1, by means of generalized PARATUCK models, will allow us to derive closed-form receivers for all systems, in Section 5. These receivers are based on the same Kronecker product between the symbol matrix ($\mathbf{S}$) and the channel matrix ($\mathbf{H}$ or $\mathbf{H}_{K \times FM}$), as summarized in Table 4.

Now, we recall two matrix unfoldings of the tensor $\mathcal{X}$ which will be used in Sections 4 and 5 for deriving the diversity gain and the Kronecker receivers of

the TSTF system (See eqs. (15) and (19) in [16])

$$\mathbf{X}_{N\times JPFK} = \mathbf{S}\,\mathbf{G}_{R\times JPFM}\Big(\mathbf{I}_{JP}\otimes \operatorname{bdiag}\big(\mathbf{H}_{\cdot\cdot 1}^{\mathrm{T}},\cdots,\mathbf{H}_{\cdot\cdot F}^{\mathrm{T}}\big)\Big), \tag{23}$$

$$\mathbf{X}_{NK\times FPJ} = (\mathbf{S}\otimes \mathbf{H}_{K\times FM})\,\mathbf{G}_{RFM\times FPJ}, \tag{24}$$

with

$$\mathbf{G}_{R\times JPFM} \triangleq \Big[\,\mathbf{G}_{\cdot\cdot 1,1,1}^{\mathrm{T}} \;\cdots\; \mathbf{G}_{\cdot\cdot F,P,J}^{\mathrm{T}}\,\Big] \in \mathbb{C}^{R\times JPFM},$$

$$\mathbf{G}_{RFM\times FPJ} \triangleq \begin{bmatrix} \operatorname{bdiag}\left(\begin{bmatrix}\operatorname{vec}\big(\mathbf{G}_{1,1,1\cdot\cdot}^{\mathrm{T}}\big)^{\mathrm{T}}\\ \vdots \\ \operatorname{vec}\big(\mathbf{G}_{M,1,1\cdot\cdot}^{\mathrm{T}}\big)^{\mathrm{T}}\end{bmatrix},\cdots,\begin{bmatrix}\operatorname{vec}\big(\mathbf{G}_{1,1,F\cdot\cdot}^{\mathrm{T}}\big)^{\mathrm{T}}\\ \vdots \\ \operatorname{vec}\big(\mathbf{G}_{M,1,F\cdot\cdot}^{\mathrm{T}}\big)^{\mathrm{T}}\end{bmatrix}\right) \\ \vdots \\ \operatorname{bdiag}\left(\begin{bmatrix}\operatorname{vec}\big(\mathbf{G}_{1,R,1\cdot\cdot}^{\mathrm{T}}\big)^{\mathrm{T}}\\ \vdots \\ \operatorname{vec}\big(\mathbf{G}_{M,R,1\cdot\cdot}^{\mathrm{T}}\big)^{\mathrm{T}}\end{bmatrix},\cdots,\begin{bmatrix}\operatorname{vec}\big(\mathbf{G}_{1,R,F\cdot\cdot}^{\mathrm{T}}\big)^{\mathrm{T}}\\ \vdots \\ \operatorname{vec}\big(\mathbf{G}_{M,R,F\cdot\cdot}^{\mathrm{T}}\big)^{\mathrm{T}}\end{bmatrix}\right) \end{bmatrix},$$

$$\mathbf{G}_{\cdot\cdot f,p,j} = \mathbf{W}_{\cdot\cdot f,p,j} \underset{\{m,r\}}{\odot} \mathbf{C}_{\cdot\cdot f,p}, \qquad \mathbf{G}_{m,r,f\cdot\cdot} = \mathbf{W}_{m,r,f\cdot\cdot} \underset{\{p\}}{\odot} \mathbf{c}_{m,r,f\cdot} \tag{25}$$

where $\mathbf{H}_{K\times FM}$ is a matrix unfolding of the channel tensor $\mathcal{H}\in\mathbb{C}^{K\times M\times F}$, and $\mathbf{H}_{\cdot\cdot f}\in\mathbb{C}^{K\times M}$ is a matrix slice obtained by fixing the index f of $\mathcal{H}$, which corresponds to the channel matrix associated with the subcarrier f. Similarly, $\{\mathbf{G}_{\cdot\cdot f,p,j},\mathbf{W}_{\cdot\cdot f,p,j},\mathbf{C}_{\cdot\cdot f,p}\}\in\mathbb{C}^{M\times R}$ denote matrix slices of $\{\mathcal{G},\mathcal{W},\mathcal{C}\}$, obtained by fixing (f,p,j), whereas $\{\mathbf{G}_{m,r,f\cdot\cdot},\mathbf{W}_{m,r,f\cdot\cdot}\}\in\mathbb{C}^{P\times J}$ and $\mathbf{c}_{m,r,f\cdot}\in\mathbb{C}^{P\times 1}$ denote matrix slices of $\{\mathcal{G},\mathcal{W}\}$, and a vector slice of $\mathcal{C}$, respectively, obtained by fixing (m,r,f).

4. Performance analysis

We first analyze the theoretical performance of the TSTF system by deriving its diversity gain. For this system, it is possible to jointly estimate the information symbols and the channel, without decoding of a codeword as required by standard ST and STF codings. The performance analysis is based on the pairwise error probability (PEP) of the maximum likelihood (ML) estimator of the symbol matrix $\mathbf{S}$. The diversity gain d is defined as the negative of the asymptotic slope of the plot PEP(ρ) on a log-log scale, where ρ denotes the received signal-to-noise ratio (SNR). We assume that the receiver has perfect knowledge of the channel, allocation, and coding tensors.

Let us recall that the average PEP between $\mathbf{S}$ and $\hat{\mathbf{S}}$, conditioned on channel realizations, is given by [29, 30]

$$P\Big(\mathbf{S}\to\hat{\mathbf{S}}\Big) = Q\left(\frac{1}{2\sqrt{N_0/2}}\|\mathcal{X}-\hat{\mathcal{X}}\|_F\right) = \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \exp\left(-\frac{\|\mathcal{X}-\hat{\mathcal{X}}\|_F^2}{4N_0\sin^2(\beta)}\right)\mathrm{d}\beta, \tag{26}$$

where $N_0/2$ is the noise variance per (real and imaginary) dimension and $Q(\cdot)$ is the complementary cumulative distribution function of a Gaussian variable, written using the Craig's formula [30]. Fixing the indices (f, p, j) in (18) gives

$$\mathbf{X}_{\cdot\cdot f,p,j} = \mathbf{H}_{\cdot\cdot f}\, \mathbf{U}_{\cdot\cdot f,p,j} = \mathbf{H}_{\cdot\cdot f}\, \mathbf{G}_{\cdot\cdot f,p,j}\, \mathbf{S}^{\mathrm{T}} \in \mathbb{C}^{K\times N}. \tag{27}$$

Defining the estimation error of codeword matrix slices

$$\mathbf{E}^{(f,p,j)} \triangleq \mathbf{U}_{\cdot\cdot f,p,j} - \hat{\mathbf{U}}_{\cdot\cdot f,p,j} = \mathbf{G}_{\cdot\cdot f,p,j}\Big(\mathbf{S} - \hat{\mathbf{S}}\Big)^{\mathrm{T}} \in \mathbb{C}^{M\times N} \tag{28}$$

and using (27), we have

$$\begin{aligned}\|\mathcal{X} - \hat{\mathcal{X}}\|_F^2 &= \sum_{f=1}^{F}\sum_{p=1}^{P}\sum_{j=1}^{J} \|\mathbf{X}_{\cdot\cdot f,p,j} - \hat{\mathbf{X}}_{\cdot\cdot f,p,j}\|_F^2 = \sum_{f=1}^{F}\sum_{p=1}^{P}\sum_{j=1}^{J} \|\mathbf{H}_{\cdot\cdot f}\, \mathbf{E}^{(f,p,j)}\|_F^2 \\ &= \sum_{f=1}^{F}\sum_{p=1}^{P}\sum_{j=1}^{J} \mathrm{tr}\Big(\mathbf{H}_{\cdot\cdot f}\, \mathbf{A}^{(f,p,j)}\, \mathbf{H}_{\cdot\cdot f}^{\mathrm{H}}\Big) = \sum_{f=1}^{F}\sum_{p=1}^{P}\sum_{j=1}^{J} y^{(f,p,j)},\end{aligned} \tag{29}$$

where $\mathbf{A}^{(f,p,j)} \triangleq \mathbf{E}^{(f,p,j)}\big(\mathbf{E}^{(f,p,j)}\big)^{\mathrm{H}}$ is Hermitian nonnegative definite, and

$$y^{(f,p,j)} \triangleq \mathrm{tr}\Big(\mathbf{H}_{\cdot\cdot f}\, \mathbf{A}^{(f,p,j)}\, \mathbf{H}_{\cdot\cdot f}^{\mathrm{H}}\Big) = \big[\mathrm{vec}\big(\mathbf{H}_{\cdot\cdot f}^{\mathrm{T}}\big)\big]^{\mathrm{T}}\Big(\mathbf{I}_K \otimes \mathbf{A}^{(f,p,j)}\Big)\mathrm{vec}\big(\mathbf{H}_{\cdot\cdot f}^{\mathrm{H}}\big).$$

The channel coefficients $h_{k,m,f}$ being assumed i.i.d and drawn from a circular symmetric complex Gaussian distribution with zero-mean and unit variance, application of the theorem E.1 on page 418 of [31] leads to

$$P\Big(\mathbf{S} \to \hat{\mathbf{S}}\Big) = \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \prod_{f=1}^{F}\prod_{p=1}^{P}\prod_{j=1}^{J} \left[\det\left(\mathbf{I}_M + \frac{1}{4N_0\sin^2(\beta)}\mathbf{A}^{(f,p,j)}\right)\right]^{-K} \mathrm{d}\beta.$$

In order to simplify the calculation of the integral, we use the Chernoff bound [30, 31], obtained by taking $\sin^2(\beta) = 1$, which gives the following upper bound

$$P\Big(\mathbf{S} \to \hat{\mathbf{S}}\Big) \leq \prod_{f=1}^{F}\prod_{p=1}^{P}\prod_{j=1}^{J} \left[\det\left(\mathbf{I}_M + \frac{1}{4N_0}\mathbf{A}^{(f,p,j)}\right)\right]^{-K}. \tag{30}$$

Since $\det(\mathbf{I} + \alpha\mathbf{A}) = \prod_{i=1}^{\mathrm{rank}(\mathbf{A})}(1 + \alpha\lambda_i(\mathbf{A}))$, where $\lambda_i(\mathbf{A})$ denotes an eigenvalue of $\mathbf{A}$, we can rewrite (30) as

$$P\Big(\mathbf{S} \to \hat{\mathbf{S}}\Big) \leq \prod_{f=1}^{F}\prod_{p=1}^{P}\prod_{j=1}^{J}\prod_{i=1}^{r^{(f,p,j)}} \left(1 + \frac{1}{4N_0}\lambda_i^{(f,p,j)}\right)^{-K}, \tag{31}$$

where $\lambda_i^{(f,p,j)}$ denotes the non-zero eigenvalues of $\mathbf{A}^{(f,p,j)}$, and $r^{(f,p,j)} \triangleq \text{rank}\left(\mathbf{A}^{(f,p,j)}\right) = \text{rank}\left(\mathbf{E}^{(f,p,j)}\right)$. At high SNR, i.e. for small values of N_0, the upper bound on the PEP becomes

$$P\left(\mathbf{S} \to \hat{\mathbf{S}}\right) \leq \prod_{f=1}^{F} \prod_{p=1}^{P} \prod_{j=1}^{J} \prod_{i=1}^{r^{(f,p,f)}} \left(\lambda_i^{(f,p,j)}\right)^{-K} \left(\frac{1}{4N_0}\right)^{-K \sum\limits_{f,p,j} r^{(f,p,j)}}, \tag{32}$$

which gives the following diversity gain

$$d^{\text{TSTF}} = K \sum_{f=1}^{F} \sum_{p=1}^{P} \sum_{j=1}^{J} r^{(f,p,j)}. \tag{33}$$

Assuming $\mathbf{S}$ is full column rank, which implies $N \geq R$ (or $N \geq M$ for KRST and DKRSTF), and applying the property $\text{rank}(\mathbf{AB}) \leq \min(\text{rank}(\mathbf{A})\,\text{rank}(\mathbf{B}))$, we deduce from (28) that $r^{(f,p,j)} = \text{rank}\left(\mathbf{E}^{(f,p,j)}\right) \leq \min(M, R)$, $\forall f, p, j$. It is interesting to note from (25) and (28) that the maximum rank of $\mathbf{G}_{..f,p,j}$, and consequently of $\mathbf{E}^{(f,p,j)}$, depends on the allocation tensor $\mathcal{C}$.

Define $\alpha^{(f,p)}$ and $\beta^{(f,p)}$ as the numbers of transmit antennas used and of data streams transmitted with the subcarrier f, during the time block p. Noting that $\mathbf{C}_{..f,p}$, and consequently $\mathbf{G}_{..f,p,j}$ for all j, have $M - \alpha^{(f,p)}$ zero rows and $R - \beta^{(f,p)}$ zero columns, we deduce that $r^{(f,p,j)} = \text{rank}(\mathbf{G}_{..f,p,j}) \leq \min\left(\alpha^{(f,p)}, \beta^{(f,p)}\right)$ for all j. So, a maximal diversity gain is given by:

$$d_{\max}^{\text{TSTF}} = KJ \sum_{f=1}^{F} \sum_{p=1}^{P} \min\left(\alpha^{(f,p)}, \beta^{(f,p)}\right). \tag{34}$$

The expression (34) can be upper bounded by $d_{\max}^{\text{TSTF}} = KFPJ \min(M, R)$ when choosing $\alpha^{(f,p)} = M$, $\beta^{(f,p)} = R$, for all (f, p), which includes a full allocation strategy corresponding to the case where all data streams are transmitted by all antennas, using all subcarriers, during each time block p. Therefore, the TSTF coding provides higher diversity gain than standard matrix ST coding schemes that ensure a maximal diversity gain equal to KM. Moreover, for fixed numbers (K and M) of receive and transmit antennas, the maximal diversity gain $d_{\max}^{\text{TSTF}}$ can be increased by increasing the design parameters F, P, and J.

The diversity gain and the maximal diversity gain for the other systems can be easily deduced from (33)-(34), using the unified presentation in Table 1, which leads to the results presented in Table 3, where $\alpha^{(p)}$ and $\beta^{(p)}$ ($\alpha^{(f,p)}$ and $\beta^{(f,p)}$) denote the number of non-zero elements of $\mathbf{c}_{\cdot p}^{(\mathbf{H})}$ and $\mathbf{c}_{\cdot p}^{(\mathbf{S})}$ ($\mathbf{c}_{f,p}^{(\mathcal{H})}$ and $\mathbf{c}_{f,p}^{(\mathbf{S})}$), for TST and ST (STF, respectively).

Remark that, in the case of DKRSTF, KRST and DS-CDMA, the maximal diversity gain is proportional to M since the number of data streams is equal to the number of transmit antennas ($R = M$) for these systems.

Note also that for a full allocation strategy ($\alpha^{(f,p)} = \alpha^{(p)} = M$, $\beta^{(f,p)} = \beta^{(p)} = R$, for all $p \in \{1, ..., P\}$ and $f \in \{1, ..., F\}$), i.e. in choosing all the entries of

the allocation matrix/tensor equal to 1, the maximal diversity gains are upper bounded by $d_{\max}^{\mathrm{STF}} = KFP\min(M,R)$, $d_{\max}^{\mathrm{TST}} = KJP\min(M,R)$, and $d_{\max}^{\mathrm{ST}} = KP\min(M,R)$, showing that the TSTF coding provides the highest diversity gain. Comparing TST and STF, with $F = J$, we conclude that $d_{\max}^{\mathrm{STF}} = d_{\max}^{\mathrm{TST}}$. However, when all the subcarriers are not used, we have $d^{\mathrm{STF}} < d^{\mathrm{TST}}$, explaining why the TST system offers better performance than STF when full frequency allocation is not considered.

Table 3: Diversity gains

Systems	**Diversity gains**	**Maximal diversity gains**	τ
TSTF [16]	$K\sum_{f=1}^{F}\sum_{p=1}^{P}\sum_{j=1}^{J}\operatorname{rank}\left(\mathbf{G}_{..f,p,j}\left(\mathbf{S}-\hat{\mathbf{S}}\right)^{\mathrm{T}}\right)$ $\mathbf{G}_{..f,p,j} = \mathbf{W}_{..f,p,j} \underset{\{m,r\}}{\odot} \mathbf{C}_{..f,p}$	$KJ\sum_{f=1}^{F}\sum_{p=1}^{P}\min\left(\alpha^{(f,p)},\beta^{(f,p)}\right)$	$\frac{R}{FP}$
STF [15]	$K\sum_{f=1}^{F}\sum_{p=1}^{P}\operatorname{rank}\left(\mathbf{G}_{..f,p}\left(\mathbf{S}-\hat{\mathbf{S}}\right)^{\mathrm{T}}\right)$ $\mathbf{G}_{..f,p} = \operatorname{diag}\left(\mathbf{c}_{\cdot f,p}^{(\mathcal{H})}\right)\mathbf{W}\operatorname{diag}\left(\mathbf{c}_{\cdot f,p}^{(\mathbf{S})}\right)$	$K\sum_{f=1}^{F}\sum_{p=1}^{P}\min\left(\alpha^{(f,p)},\beta^{(f,p)}\right)$	$\frac{R}{FP}$
TST [14]	$K\sum_{j=1}^{J}\sum_{p=1}^{P}\operatorname{rank}\left(\mathbf{G}_{..p,j}\left(\mathbf{S}-\hat{\mathbf{S}}\right)^{\mathrm{T}}\right)$ $\mathbf{G}_{..p,j} = \operatorname{diag}\left(\mathbf{c}_{\cdot p}^{(\mathbf{H})}\right)\mathbf{W}_{..j}\operatorname{diag}\left(\mathbf{c}_{\cdot p}^{(\mathbf{S})}\right)$	$KJ\sum_{p=1}^{P}\min\left(\alpha^{(p)},\beta^{(p)}\right)$	$\frac{R}{P}$
ST [13]	$K\sum_{p=1}^{P}\operatorname{rank}\left(\mathbf{G}_{..p}\left(\mathbf{S}-\hat{\mathbf{S}}\right)^{\mathrm{T}}\right)$ $\mathbf{G}_{..p} = \operatorname{diag}\left(\mathbf{c}_{\cdot p}^{(\mathbf{H})}\right)\mathbf{W}\operatorname{diag}\left(\mathbf{c}_{\cdot p}^{(\mathbf{S})}\right)$	$K\sum_{p=1}^{P}\min\left(\alpha^{(p)},\beta^{(p)}\right)$	$\frac{R}{P}$
STM [20]	$K\sum_{p=1}^{P}\operatorname{rank}\left(\mathbf{G}_{..p}\left(\mathbf{S}-\hat{\mathbf{S}}\right)^{\mathrm{T}}\right)$ $\mathbf{G}_{..p} = \mathbf{W}_{..p}$	$KP\min(M,R)$	$\frac{R}{P}$
DKRSTF [22]	$K\sum_{f=1}^{F}\sum_{p=1}^{P}\operatorname{rank}\left(\mathbf{G}_{..f,p}\left(\mathbf{S}-\hat{\mathbf{S}}\right)^{\mathrm{T}}\right)$ $\mathbf{G}_{..f,p} = \operatorname{diag}(\psi_{p\cdot})\boldsymbol{\Theta}\operatorname{diag}(\omega_{f\cdot})$	$KFPM$	$\frac{M}{FP}$
KRST [18]	$K\sum_{p=1}^{P}\operatorname{rank}\left(\mathbf{G}_{..p}\left(\mathbf{S}-\hat{\mathbf{S}}\right)^{\mathrm{T}}\right)$ $\mathbf{G}_{..p} = \operatorname{diag}(\psi_{p\cdot})\boldsymbol{\Theta}$	KPM	$\frac{M}{P}$
DS-CDMA [17]	$K\sum_{j=1}^{J}\operatorname{rank}\left(\operatorname{diag}(\mathbf{g}_{j\cdot})\left(\mathbf{S}-\hat{\mathbf{S}}\right)^{\mathrm{T}}\right)$	KJM	M

The transmission rate, in bits per channel use, is given by $R_b = \tau\log_2(\mu)$, where μ denotes the cardinality of the symbol alphabet set, i.e. the number of constellation points, and the ratio τ is given in Table 3 for all systems. As expected, for STF and TSTF, an increase of P and/or F decreases the transmission rate, while an increase of R increases it. For TSTF, TST, STF and ST, the bandwidth is given by, $\frac{FJ}{T}$, $\frac{J}{T}$, $\frac{F}{T}$, and $\frac{1}{T}$, respectively, where T is the symbol period.

5. Kronecker semi-blind receivers

Assuming a perfect knowledge of the coding and allocation matrices/tensors at the receiver, we propose five receivers for all the systems presented in

Section 3. These receivers based on a Kronecker product approximation use the algorithms described in Section 2. So, they are called KALS, SKALS, KPLS, KSVD, and KALMS. Note that SKALS and KPLS correspond to supervised receivers since they use a pilot sequence constituted by the first N_p rows of $\mathbf{S}$. In the sequel, we detail the Kronecker receivers for the TSTF system, the matrix unfolding (24) being exploited to jointly estimate the symbol matrix $\mathbf{S}$ and the unfolding $\mathbf{H}_{K\times FM}$ of the channel tensor. Similar Kronecker receivers can be easily derived for the other systems using the matrix unfoldings given in Table 4. Note that, for TST, ST, STM, KRST, and DKRSTF, the unfolding $\mathbf{H}_{K\times FM}$ is replaced by the channel matrix $\mathbf{H}$.

Table 4: Matrix unfoldings of the received signals tensor $\mathcal{X}$.

Systems	**Unfoldings with a Kronecker Product**	**Identifiability conditions**
TSTF	$\mathbf{X}_{NK\times FPJ} = (\mathbf{S}\otimes\mathbf{H}_{K\times FM})\,\mathbf{G}_{RFM\times FPJ}$ $\mathbf{G}_{RFM\times FPJ}\in\mathbb{C}^{RFM\times FPJ}$ defined in (25)	$PJ\geq MR$
STF†	$\mathbf{X}_{NK\times FP} = (\mathbf{S}\otimes\mathbf{H}_{K\times FM})\,\mathbf{G}_{RFM\times FP}$ $\mathbf{G}_{RFM\times FP} = \mathbf{\Pi}\ \mathrm{bdiag}\big(\mathbf{G}^{(1)},\ldots,\mathbf{G}^{(F)}\big)\in\mathbb{C}^{RFM\times FP}$ $\mathbf{G}^{(f)}\triangleq\left(\left(\mathbf{C}^{(\mathbf{S})}_{f\cdot}\diamond\mathbf{C}^{(\mathcal{H})}_{f\cdot}\right)^{\mathrm{T}}\diamond\mathrm{vec}^{\mathrm{T}}(\mathbf{W})\right)^{\mathrm{T}}\in\mathbb{C}^{RM\times P},\ \forall f$	$P\geq MR$
TST	$\mathbf{X}_{NK\times PJ} = (\mathbf{S}\otimes\mathbf{H})\,\mathbf{G}_{RM\times PJ}$ $\mathbf{G}_{RM\times PJ}\triangleq\left(\left(\mathbf{C}^{(\mathbf{S})}\diamond\mathbf{C}^{(\mathbf{H})}\right)^{\mathrm{T}}\diamond\mathbf{W}_{J\times RM}\right)^{\mathrm{T}}\in\mathbb{C}^{RM\times PJ}$	$PJ\geq MR$
ST	$\mathbf{X}_{NK\times P} = (\mathbf{S}\otimes\mathbf{H})\,\mathbf{G}_{RM\times P}$ $\mathbf{G}_{RM\times P}\triangleq\left(\left(\mathbf{C}^{(\mathbf{S})}\diamond\mathbf{C}^{(\mathbf{H})}\right)^{\mathrm{T}}\diamond\mathrm{vec}^{\mathrm{T}}(\mathbf{W})\right)^{\mathrm{T}}\in\mathbb{C}^{RM\times P}$	$P\geq MR$
STM	$\mathbf{X}_{NK\times P} = (\mathbf{S}\otimes\mathbf{H})\,\mathbf{W}_{RM\times P}$	$P\geq MR$
KRST	$\mathbf{X}_{NK\times P} = (\mathbf{S}\otimes\mathbf{H})\,\mathbf{G}_{RM\times P}$ $\mathbf{G}_{RM\times P}\triangleq\left(\left(\mathbf{1}_R^{\mathrm{T}}\otimes\mathbf{\Psi}\right)\diamond\mathrm{vec}^{\mathrm{T}}(\mathbf{\Theta})\right)^{\mathrm{T}}\in\mathbb{C}^{RM\times P}$	$P\geq MR$
DKRSTF	$\mathbf{X}_{NK\times FP} = (\mathbf{S}\otimes\mathbf{H})\,\mathbf{G}_{RM\times FP}$ $\mathbf{G}_{RM\times FP}\triangleq\left(\left(\mathbf{\Omega}\otimes\mathbf{\Psi}\right)\diamond\mathrm{vec}^{\mathrm{T}}(\mathbf{\Theta})\right)^{\mathrm{T}}\in\mathbb{C}^{RM\times FP}$	$FP\geq MR$

† $\mathbf{\Pi}$ denotes a $(RFM\times FRM)$ permutation matrix which can be easily deduced from (8).

5.1. Kronecker receivers for the TSTF system

In the case of TSTF, assuming that $\mathbf{G}_{RFM\times FPJ}$ is full row-rank to be right invertible, the LS estimate of the Kronecker product in (24) is given by

$$\begin{aligned}\mathbf{Y}_{NK\times RFM}\triangleq\mathbf{S}\otimes\mathbf{H}_{K\times FM} &= \mathbf{X}_{NK\times FPJ}\,\mathbf{G}^{\dagger}_{RFM\times FPJ}\in\mathbb{C}^{NK\times RFM} \\ &= \begin{bmatrix}\mathbf{Y}^{(1,1)}_{K\times FM} & \cdots & \mathbf{Y}^{(1,R)}_{K\times FM}\\ \vdots & \ddots & \vdots\\ \mathbf{Y}^{(N,1)}_{K\times FM} & \cdots & \mathbf{Y}^{(N,R)}_{K\times FM}\end{bmatrix},\ \text{with}\ \mathbf{Y}^{(n,r)}_{K\times FM}\triangleq s_{n,r}\,\mathbf{H}_{K\times FM}.\end{aligned}\tag{35}$$

The matrix factors $(\mathbf{S},\mathbf{H}_{K\times FM})$ of the Kronecker product $\mathbf{Y}_{NK\times RFM}$ can be estimated using the algorithms described in Section 2.

For applying the KALS method described in Algorithm 1, we have to compute the Kronecker product $\mathbf{D} = \mathbf{Y}_{KN\times FMR} \triangleq \mathbf{H}_{K\times FM} \otimes \mathbf{S}$ given by

$$\mathbf{Y}_{KN\times FMR} = \mathbf{\Pi}^{(\text{row})}\, \mathbf{Y}_{NK\times RFM}\, \mathbf{\Pi}^{(\text{col})} = \begin{bmatrix} \mathbf{Y}_{N\times R}^{(1,1)} & \cdots & \mathbf{Y}_{N\times R}^{(1,FM)} \\ \vdots & \ddots & \vdots \\ \mathbf{Y}_{N\times R}^{(K,1)} & \cdots & \mathbf{Y}_{N\times R}^{(K,FM)} \end{bmatrix} \tag{36}$$

$$\mathbf{Y}_{N\times R}^{(k,m_f)} \triangleq h_{k,m,f}\, \mathbf{S}, \text{ with } m_f \triangleq m + (f-1)M \in [1,\ldots,FM],$$

where the permutation matrices $\mathbf{\Pi}^{(\text{row})}$ and $\mathbf{\Pi}^{(\text{col})}$ can be easily deduced from (8) with the following correspondences $\{I,J,K,L\} \longleftrightarrow \{N,R,K,FM\}$. Note that the scalar ambiguity inherent to the Kronecker product can be removed by the knowledge of only one symbol at the receiver. Algorithms 4, 5, and 6 describe the semi-blind KALS, KSVD, and KALMS receivers for the TSTF system. The matrix $\tilde{\mathbf{Y}}_{NK\times RFM}$ denotes a noisy version of $\mathbf{Y}_{NK\times RFM}$ defined in (35), with $\mathbf{X}_{NK\times FPJ}$ replaced by $\tilde{\mathbf{X}}_{NK\times FPJ}$, an unfolding of the noisy received signal tensor $\tilde{\mathcal{X}} = \mathcal{X} + \sigma\mathcal{V}$, where $\mathcal{V}$ is an additive noise tensor, and σ is adjusted according to the desired SNR.

When a training sequence $\mathbf{S}_{1:N_p} \in \mathbb{C}^{N_p\times R}$ composed of the first N_p rows of $\mathbf{S}$, is used, the supervised receiver SKALS is obtained in replacing the initialization by $\hat{\mathbf{H}}_{K\times FM}(0) = \frac{1}{\|\mathbf{S}_{1:N_p}\|_F^2} \sum_{n=1}^{N_p} \sum_{r=1}^{R} s_{n,r}^{*} \tilde{Y}_{K\times FM}^{(n,r)}$ in step 2 of Algorithm 4. The supervised receiver KPLS is described in Algorithm 7.

From Algorithms 4-7, we remark that:

- The KALS, SKALS and KALMS receivers which are based on the ALS and ALMS algorithms, respectively, are iterative, whereas the KPLS and KSVD ones are closed-form solutions. However, due to the application of the power method for computing the rank-one approximation, KSVD is also iterative. Its convergence speed depends on the ratios $\frac{\sigma_i}{\sigma_1}$, and more particularly the ratio $\frac{\sigma_2}{\sigma_1}$, the singular values σ_i for $i > 1$ being introduced by the additive noise tensor $\mathcal{V}$.
- The KALS, SKALS and KSVD receivers need to apply permutation matrices, which is not the case of KPLS and KALMS.

Algorithm 4: Semi-blind KALS receiver

1. Compute the LS estimate $\tilde{\mathbf{Y}}_{NK\times RFM} = \tilde{\mathbf{X}}_{NK\times FPJ}\,\mathbf{G}_{RFM\times FPJ}{}^{\dagger}$ and $\tilde{\mathbf{Y}}_{KN\times FMR}$ using (36), with $\mathbf{Y}_{NK\times RFM}$ replaced by $\tilde{\mathbf{Y}}_{NK\times RFM}$.
2. $it = 0$: Initialize $\hat{\mathbf{H}}_{K\times FM}(0) = \frac{1}{s_{1,1}}\tilde{\mathbf{Y}}^{(1,1)}_{K\times FM}$.
3. $it = it + 1$.
4. Compute $\hat{\mathbf{S}}(it)$:
   $\hat{s}_{n,r}(it) = \dfrac{\mathrm{vec}\big(\hat{\mathbf{H}}_{K\times FM}(it-1)\big)^{\mathrm{H}}\,\mathrm{vec}\big(\tilde{\mathbf{Y}}^{(n,r)}_{K\times FM}\big)}{\big\|\hat{\mathbf{H}}_{K\times FM}(it-1)\big\|_F^2}$.
5. Compute $\hat{\mathbf{H}}_{K\times FM}(it)$:
   $\hat{h}_{k,m,f}(it) = \dfrac{\mathrm{vec}\big(\hat{\mathbf{S}}(it)\big)^{\mathrm{H}}\,\mathrm{vec}\Big(\tilde{\mathbf{Y}}^{(k,m_f)}_{N\times R}\Big)}{\big\|\hat{\mathbf{S}}(it)\big\|_F^2}$, $m_f = m + (f-1)M$.
6. Repeat steps (3)-(5) until convergence.
7. Eliminate the scaling ambiguity with $\alpha = {s_{1,1}}/{\hat{s}_{1,1}(\infty)}$:
   $\hat{\mathbf{S}}(\infty) \leftarrow \alpha\,\hat{\mathbf{S}}(\infty)$, $\quad \hat{\mathbf{H}}_{K\times FM}(\infty) \leftarrow \frac{1}{\alpha}\,\hat{\mathbf{H}}_{K\times FM}(\infty)$.
8. Project the estimated symbols onto the symbol alphabet.

Algorithm 5: Semi-blind KSVD receiver

1. Compute the LS estimate $\tilde{\mathbf{Y}}_{NK\times RFM}$ as in step (1) of Alg. 4, and $\tilde{\mathbf{Y}}_{FMK\times RN} = \mathrm{unvec}\big(\mathbf{\Pi}\,\mathrm{vec}(\tilde{\mathbf{Y}}_{NK\times RFM})\big)$, with (13) and (20).
2. $it = 0$: Initialize $\hat{\mathbf{u}}(0) = \mathrm{vec}\Big(\hat{\mathbf{H}}_{K\times FM}(0)\Big) = \frac{1}{s_{1,1}}\,\mathrm{vec}(\tilde{\mathbf{Y}}^{(1,1)}_{K\times FM})$.
3. Compute the rank-one approximation of $\tilde{\mathbf{Y}}_{FMK\times RN}$ using steps (3)-(5) of Alg. 2.
4. Compute the estimates of $\mathbf{S}$ and $\mathbf{H}_{K\times FM}$, with $\alpha = {s_{1,1}}/{\hat{s}_{1,1}(\infty)}$:
   $\hat{\mathbf{S}} = \alpha\,\sqrt{\sigma_1(\infty)}\,\mathrm{unvec}(\mathbf{v}^*(\infty))$, $\hat{\mathbf{H}}_{K\times FM} = \frac{1}{\alpha}\,\sqrt{\sigma_1(\infty)}\,\mathrm{unvec}(\mathbf{u}(\infty))$.
5. Project the estimated symbols onto the symbol alphabet.

Algorithm 6: Semi-blind KALMS receiver

1. Set $\gamma_{\mathbf{S}}$ and $\gamma_{\mathbf{H}}$.
2. Compute the LS estimate of $\mathbf{C} = \tilde{\mathbf{Y}}_{NK\times RFM}$ as in step (1) of Alg. 4.
3. $it = 0$: Initialize $\hat{\mathbf{H}}_{K\times FM}(0) = \frac{1}{s_{1,1}}\tilde{\mathbf{Y}}^{(1,1)}_{K\times FM}$ and randomly initialize $\hat{\mathbf{S}}(0)$.
4. $it = it + 1$.
5. Randomly generate the input signal $\mathbf{x}(it)$ and compute $\mathbf{y}(it) = \mathbf{C}\,\mathbf{x}(it)$.
6. Compute $\hat{\mathbf{S}}(it)$ and $\hat{\mathbf{H}}_{K\times FM}(it)$ as in steps (5)-(6) of Alg. 3.
7. Repeat steps (4)-(6) until convergence.
8. Eliminate the scaling ambiguity as in step (7) of Alg. 4.
9. Project the estimated symbols onto the symbol alphabet.

Algorithm 7: Supervised KPLS receiver

1. Compute the LS estimate $\tilde{\mathbf{Y}}_{NK\times RFM}$ as in step (1) of Alg. 4.
2. Compute the LS estimate of $\mathbf{H}_{K\times FM}$:
   $\hat{\mathbf{H}}_{K\times FM} = \frac{1}{\|\mathbf{S}_{1:N_p}\|_F^2}\sum_{n=1}^{N_p}\sum_{r=1}^{R} s^*_{n,r}\tilde{Y}^{(n,r)}_{K\times FM}$.
3. Compute the LS estimate of $\mathbf{S}_{N_p+1:N}$:
   $\hat{s}_{n,r} = \dfrac{\mathrm{vec}\big(\hat{\mathbf{H}}_{K\times FM}\big)^{\mathrm{H}}\,\mathrm{vec}\big(\tilde{\mathbf{Y}}^{(n,r)}_{K\times FM}\big)}{\big\|\hat{\mathbf{H}}_{K\times FM}\big\|_F^2}$, $n = N_p+1,\cdots,N$; $r = 1,\cdots,R$.
4. Project the estimated symbols onto the symbol alphabet.

5.2. Identifiability conditions

For TSTF, the identifiability condition is linked with the calculation of the LS estimate (35) of the Kronecker product, that is the full row-rank property of $\mathbf{G}_{RFM\times FPJ}$. Premultiplying this expression by the permutation matrix

$$\mathbf{\Pi} \triangleq \sum_{l=1}^{FM}\sum_{r=1}^{R} \mathbf{e}_l^{(FM)}\mathbf{e}_r^{(R)\mathrm{T}} \otimes \mathbf{e}_r^{(R)}\mathbf{e}_l^{(FM)\mathrm{T}} \in \mathbb{R}^{FMR\times RFM}, \tag{37}$$

$\mathbf{G}_{RFM\times FPJ}$ can be rewritten as the following block-diagonal matrix:

$$\mathbf{\Pi}\,\mathbf{G}_{RFM\times FPJ} = \mathrm{bdiag}\Big(\mathbf{G}^{(1)}_{MR\times PJ}, \cdots, \mathbf{G}^{(F)}_{MR\times PJ}\Big) \in \mathbb{C}^{FMR\times FPJ}, \tag{38}$$

where $\mathbf{G}^{(f)}_{MR\times PJ}$ corresponds to a matrix unfolding of the core tensor $\mathcal{G} \in \mathbb{C}^{M\times R\times F\times P\times J}$ obtained by fixing f, and by combining its first two modes on the rows and its last two modes on the columns. As the premultiplication by the permutation matrix $\mathbf{\Pi}$ does not modify the row rank, one concludes that $\mathbf{G}_{RFM\times FPJ}$ is full row-rank if and only if all diagonal blocks $\mathbf{G}^{(f)}_{MR\times PJ}$ are themselves full row-rank, which implies the necessary condition $PJ \geq MR$. From the unfoldings given in Table 4, it is easy to deduce the identifiability condition for the other systems, as summarized in Table 4. From these results, one can conclude that TST and TSTF are more flexible than ST ant STF in terms of identifiability.

All the results previously established for TSTF imply that the desired transmission rate T_r and bandwidth B determine the design parameters (F, P, J, R) and the modulation μ-phase shift-keying (PSK) through $T_r = \frac{R}{FP}\log_2\mu$ and $B = \frac{FJ}{T}$, under the constraint $PJ \geq MR$ representing the identifiability condition, with an achievable maximal diversity gain equal to $KFPJ\min(M, R)$.

Recall that diversity techniques are employed to combat the impact of channel fading on the bit error rate (BER), by providing to the receiver several versions of each transmitted symbol. For TSTF, four diversities (frequency, time, chip, space) are simultaneously exploited owing the tensor coding, combined with the use of multiple antennas and multiple carriers. That explains the expression of the maximal diversity gain as a function of (F, P, J, K, M). Note that the diversity gain, directly linked with the BER, and the multiplexing gain, proportional to the data rate, cannot be simultaneously maximized. That corresponds to the well-known diversity-multiplexing tradeoff. In our simulations, the diversity gain will be exploited to determine the design parameters, with a fixed transmission rate identical for all the systems to be compared. In practice, the adjustment of the design parameters must take into account some system constraints such as an a priori fixed number of transmit and receive antennas, an available bandwidth, or a desired transmission rate. So, the expressions of the diversity gains, of the transmission rates, and of the bandwidths, derived in Section 4 for all the considered systems, as functions of (F, P, J, M, R, K), can help the user to choose these design parameters.

The impact of an increase of these parameters on the transmission rate, the bandwidth, and the BER, is summarized in Table 5, where the signs + and - mean an increase and a decrease of transmission rate and bandwidth, and a performance improvement and degradation, respectively.

Table 5: Impact of the design parameters (F, P, J, M, R, K) on data rate/bandwidth/BER.

Increase of design parameters	Data rate	Bandwidth	BER
Number of subcarriers (F)	$-$	$+$	$+$
Number of data blocks (P)	$-$		$+$
Length of spreading code (J)		$+$	$+$
Number of transmit antennas (M)			$+$
Number of data streams (R)	$+$		$-$
Number of receive antennas (K)			$+$

5.3. Comparative complexity analysis

For TSTF, the Kronecker receivers involve the computation of one pseudo-inverse for the LS estimation of the Kronecker product $\mathbf{S} \otimes \mathbf{H}_{K\times FM}$, which represents the most costly operation. Using the full allocation strategy ($\mathcal{C} = \mathbf{1}_{M\times R\times F\times P}$, a fourth-order tensor composed of ones) and a Vandermonde structure for $\mathbf{G}_{RFM\times FPJ}$ allows to simplify the calculation of its pseudo-inverse, with a complexity in $\mathcal{O}\big(F^2JKMNPR\big)$ complex multiplications, common to all the receivers. Besides this computation, the non-iterative KPLS receiver requires $\mathcal{O}(FKMNR)$ multiplications to estimate the symbols and the channel, while the iterative (KALS, SKALS, KSVD), and KALMS receivers require $\mathcal{O}(FKMNR)$, and $\mathcal{O}(FKMN + KRN)$ multiplications at each iteration, respectively.

In Table 6, we present the computational cost of the receivers, taking into account the iterations number N_{it} for convergence of the iterative methods. From this table, one concludes that the supervised non-iterative KPLS receiver is the least costly, whereas the three non-supervised receivers have approximately the same computational cost per iteration.

Table 6: Computational complexity of TSTF

Receivers	**Computational cost**
KPLS	$\mathcal{O}(F^2JKMNPR + FKMNR)$
KALS	$\mathcal{O}(F^2JKMNPR + N_{it}\,(FKMNR))$
KSVD	$\mathcal{O}(F^2JKMNPR + N_{it}\,(FKMNR))$
KALMS	$\mathcal{O}(F^2JKMNPR + N_{it}\,(FKMN + KNR))$

6. Simulation results

The simulations have two main goals. First, we compare the performance of three tensor-based systems (TSTF, TST, STF), with two issues:

(i) Perfect channel knowledge at the receiver, i.e. using the zero-forcing (ZF) receivers given in Table 7, deduced from (23) and the tensor models in Table 1, with full allocation. Fig. 1 illustrates the impact of the design parameters (F, P, J) on the BER of each system, versus SNR, for the same bandwidth and transmission rate. See Subsection 6.1.

(ii) Joint channel/symbols estimation using the KALS receivers, with random allocations, except a full frequency allocation and $c_{m,r,f,1}=c^{(\mathcal{H})}_{m,f,1}=c^{(\mathbf{S})}_{r,f,1}=c^{(\mathbf{H})}_{m,1}=c^{(\mathbf{S})}_{r,1}=1, \forall(m,r,f)$, for ensuring each data stream is sent at least once (during the first time block p=1). Fig. 2 shows the impact of the design parameters on system identifiability and the BER, with the same transmission rate, inducing different modulation formats, and two different bandwidths and numbers of data streams ($B=4/T, R\in\{2,4\}$ in the left figure, and $B=8/T, R\in\{4,8\}$ in the right figure). See Subsection 6.2.

Table 7: **ZF Receivers**

Systems	**Estimation of the symbol matrix S**
TSTF	$\hat{\mathbf{S}}=\tilde{\mathbf{X}}_{N\times JPFK}\left(\mathbf{G}_{R\times JPFM}\left(\mathbf{I}_{JP}\otimes \text{bdiag}\left(\mathbf{H}^{\mathrm{T}}_{\cdot\cdot 1},\cdots,\mathbf{H}^{\mathrm{T}}_{\cdot\cdot F}\right)\right)\right)^{\dagger}$ $\mathbf{G}_{R\times JPFM}\triangleq\left[\begin{array}{ccc}\mathbf{G}^{\mathrm{T}}_{\cdot\cdot 1,1,1} & \cdots & \mathbf{G}^{\mathrm{T}}_{\cdot\cdot F,P,J}\end{array}\right]$ $\quad \mathbf{G}_{\cdot\cdot f,p,j}=\mathbf{W}_{\cdot\cdot f,p,j}\underset{\{m,r\}}{\odot}\mathbf{C}_{\cdot\cdot f,p}$
STF	$\hat{\mathbf{S}}=\tilde{\mathbf{X}}_{N\times PFK}\left(\mathbf{G}_{R\times PFM}\left(\mathbf{I}_{P}\otimes \text{bdiag}\left(\mathbf{H}^{\mathrm{T}}_{\cdot\cdot 1},\cdots,\mathbf{H}^{\mathrm{T}}_{\cdot\cdot F}\right)\right)\right)^{\dagger}$ $\mathbf{G}_{R\times PFM}\triangleq\left[\begin{array}{ccc}\mathbf{G}^{\mathrm{T}}_{\cdot\cdot 1,1} & \cdots & \mathbf{G}^{\mathrm{T}}_{\cdot\cdot F,P}\end{array}\right]$ $\quad \mathbf{G}_{\cdot\cdot f,p}=\mathbf{W}\underset{\{m,r\}}{\odot}\left(\mathbf{c}^{(\mathcal{H})}_{\cdot f,p}\mathbf{c}^{(\mathbf{S})^{\mathrm{T}}}_{\cdot f,p}\right)$
TST	$\hat{\mathbf{S}}=\tilde{\mathbf{X}}_{N\times JPK}\left(\mathbf{G}_{R\times JPM}\left(\mathbf{I}_{JP}\otimes \mathbf{H}^{\mathrm{T}}\right)\right)^{\dagger}$ $\mathbf{G}_{R\times JPM}\triangleq\left[\begin{array}{ccc}\mathbf{G}^{\mathrm{T}}_{\cdot\cdot 1,1} & \cdots & \mathbf{G}^{\mathrm{T}}_{\cdot\cdot P,J}\end{array}\right]$ $\quad \mathbf{G}_{\cdot\cdot p,j}=\mathbf{W}_{\cdot\cdot j}\underset{\{m,r\}}{\odot}\left(\mathbf{c}^{(\mathbf{H})}_{\cdot p}\mathbf{c}^{(\mathbf{S})^{\mathrm{T}}}_{\cdot p}\right)$

Second, several points are analyzed for the TSTF system considering the full allocation strategy (except for the first item (i)):

(i) Impact of allocations on the BER of each transmitted data stream taken separately, using the ZF receiver. See Subsection 6.3;

(ii) Impact of the pilot sequence length on the BER and channel NMSE obtained with the SKALS and KPLS receivers. See Subsection 6.4;

(iii) Comparison of the five proposed Kronecker receivers in terms of BER, channel and reconstruction NMSEs, and convergence speed. See Subsection 6.5.

The reconstruction NMSE is evaluated in terms of the Kronecker product estimate $\hat{\mathbf{Y}}_{NK\times RFM}=\hat{\mathbf{S}}\otimes\hat{\mathbf{H}}_{K\times FM}$ averaged over 2000 Monte Carlo runs. For all the simulations with the Kronecker receivers, a maximum number of 100 iterations was fixed, which ensures the convergence as shown in Figure 6. The transmitted symbols are randomly drawn from a 16-PSK alphabet, except in Subsection 6.2. The spreading code ($\mathbf{W}$ or $\mathcal{W}$) is constructed with

a Vandermonde structure ensuring the existence of the pseudo-inverse required by the ZF and Kronecker receivers. Flat Rayleigh fading channels for the TST system, and frequency-selective Rayleigh fading channels for STF and TSTF were simulated. As mentioned in Subsection 5.1, the noisy received signals tensor is given by $\tilde{\mathcal{X}} = \mathcal{X} + \sigma\mathcal{V}$, the elements of $\mathcal{V}$ being zero-mean complex valued Gaussian variables with unit variance, and σ being adjusted according to the desired SNR defined as $\text{SNR}_{\text{dB}} = 10\log_{10}(\|\mathcal{X}\|_F^2/\|\sigma\mathcal{V}\|_F^2)$.

6.1. Impact of the design parameters (F, P, J) with ZF receivers

Fig. 1 plots the BER versus SNR, obtained with the ZF receivers, in the case of full allocation, with the same transmission rate (1 bit/channel use) and bandwidth $(4/T)$, for all systems. From this figure, one can conclude that:

- As expected, TSTF with $(F, P, J) = (4, 2, 1)$ and STF with $(F, P) = (4, 2)$ give nearly the same BER since both systems are characterized by the same diversity gain with FPJ=8. Similarly, TSTF with $(F, P, J) = (1, 8, 4)$ and TST with $(P, J) = (8, 4)$ provide close BERs, with $FPJ = 32$, which explains the performance improvement.
- To illustrate the impact of the product FPJ on the BER performance, i.e. the diversity gain, we simulated TSTF with three sets of parameters $(F, P, J) = \{(4, 2, 1), (2, 4, 2), (1, 8, 4)\}$, corresponding to three different values of $FPJ = \{8, 16, 32\}$, with the same transmission rate and bandwidth. We observe a SNR gap of around 7.5dB for a BER value of 10^{-3}, when the value of FPJ is multiplied by 4 (from 8 to 32).
- The TSTF system allows more flexibility for choosing the design parameters, depending on available bandwidth, desired transmission rate, and fixed numbers of antennas. This flexibility is brought by the fifth-order coding tensor which exploits four spreading dimensions (space, frequency, time, chip) for each data stream, whereas the TST and STF systems exploit only three spreading dimensions.

6.2. Impact of the design parameters (F, P, J) with KALS receivers

Fig. 2 shows the BER versus SNR obtained with the KALS receivers of the three considered systems, with the same transmission rate (1 bit/channel use), two different bandwidths (B=$4/T$ and $8/T$), and two different numbers of data streams ($R \in \{2, 4\}$ and $R \in \{4, 8\}$), for the left and right figures, respectively. From these simulation results, one can draw the following conclusions:

- The STF design parameters ($(P, F) = (2, 4)$ for $R = 2$, and $(P, F) = (2, 8)$ for $R = 4$, with a 16-PSK modulation), do not satisfy the identifiability condition ($P \geq MR$) which explains the bad BER performance. An alternative solution consists in applying the Levenberg-Marquardt-based receiver proposed in [15], which is much more complex to implement than the KALS receiver. Note that an increase of P allows to satisfy the identifiability condition at the cost of a data rate decrease.

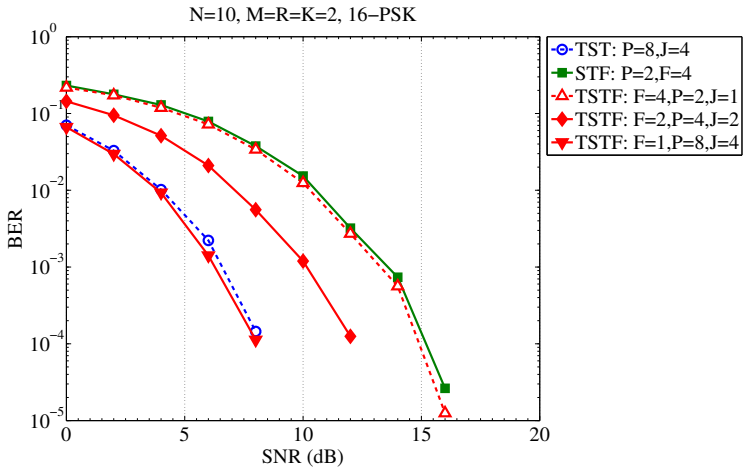

Figure 1: ZF receivers: Impact of the design parameters (F, P, J) on the BER.

- In contrast, TST and TSTF give good BER performances, with $PJ \in \{4, 8, 16, 32\}$ such that $PJ \geq MR$, which shows these two systems are more flexible for satisfying the identifiability condition owing the chip diversity J. On the left figure, we consider a bandwidth $B = 4/T$, and two different values of R. For $R = 2$, TST with $(P, J) = (2, 4)$ and TSTF with $(F, P, J) = (2, 2, 2)$ are characterized by the same diversity gain, which induces close BER performance. The same behaviour occurs for $R = 4$, TST with $(P, J) = (4, 4)$ and TSTF with $(F, P, J) = (2, 4, 2)$. Note that TST is used with a BPSK modulation, while TSTF uses a QPSK modulation for ensuring the same data rate. On the right figure, corresponding to $B = 8/T$, TST and TSTF use the same modulation (BPSK). TSTF provides a better BER performance with a SNR gap of 2dB and 3dB, for a BER value of 10^{-3}, in the cases $R = 8$ and $R = 4$, respectively. This improvement is due to the full frequency allocation. As expected, the BER increases when the number of data streams is increased from 2 to 4 (left figure), and from 4 to 8 (right figure).

6.3. TSTF-ZF receiver: Impact of the allocations

Consider three different allocation tensors $\mathcal{C} \in \mathbb{C}^{M \times R \times F \times P}$, for $f \in \{1, 2, 3\}$:

Case 1: $\mathbf{C}^{(f)}_{MR \times P} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, Case 2: $\mathbf{C}^{(f)}_{MR \times P} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}$, Case 3: $\mathbf{C}^{(f)}_{MR \times P} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$,

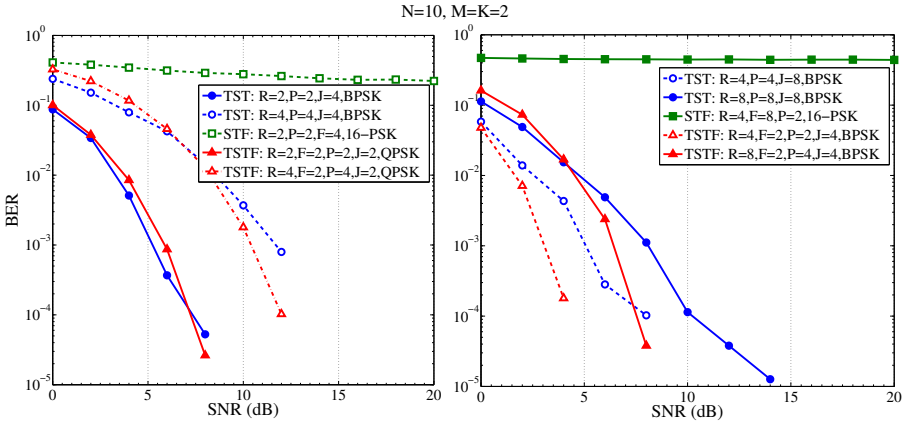

Figure 2: KALS receivers: Impact of the design parameters and the modulation on the BER.

Fig. 3 depicts the BER per data stream versus SNR, with the system parameters $K=M=J=2, R=P=F=3$, $N=10$. For the three configurations, the three data streams are sent using the three subcarriers, which ensures a full frequency diversity. Case 3 corresponds to a full allocation, meaning that all the data streams are sent by all the transmit antennas, during the three time blocks. In Case 1 (Case 2), only the first (first two, respectively) data stream(s) is (are) transmitted by both antennas, during the three time blocks. These three choices of the allocation tensor illustrate three different levels of redundancy in the space domain, with full spreading in the time and frequency domains. More generally, the allocation tensor can be used to exploit different levels of space, time and/or frequency diversities.

As expected, in Case 3, the BERs of the three data streams are very close since all three benefit from full allocation. In Case 1, the first data stream which is fully spread (in space, time, and frequency), has the smallest BER, the BERs of the other two data streams being nearly the same. In Case 2, the first two data streams have the best performance, with very close BERs, because they benefit from full space-time-frequency diversities, the third data stream being partially spread in space due to its transmission with both antennas, only during the first time block.

6.4. TSTF-KPLS/SKALS/KALS receivers: Impact of the pilot sequence length

In this section, we illustrate the impact of the pilot sequence length ($N_p \in \{1, 2, 3\}$) on the BER and channel NMSE obtained with the supervised SKALS and KPLS receivers, in the case of full allocation, with $M=R=F=J=2$, $K=P=4$ and $N=10$. A comparison is also made with the semi-blind KALS receiver. From Fig. 4, one can conclude that the SKALS and KPLS receivers provide very close BERs, and a BER improvement with respect to the KALS receiver. This BER

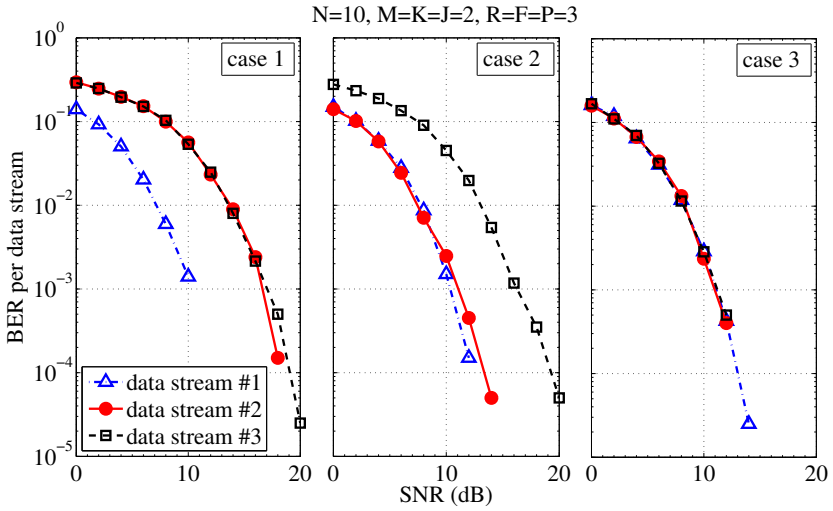

Figure 3: TSTF-ZF receiver: Impact of data stream allocations.

improvement saturates from N_p=2, with a SNR gap of about 3dB for a BER of 10^{-3}. In terms of channel NMSE, the best performance is obtained with the SKALS algorithm due to its iterative nature, contrarily to KPLS. Note that an increase of N_p beyond 2 allows to improve the channel estimation with KPLS, which is much less pronounced with SKALS. However, the channel estimation with KPLS remains worse than with KALS and SKALS.

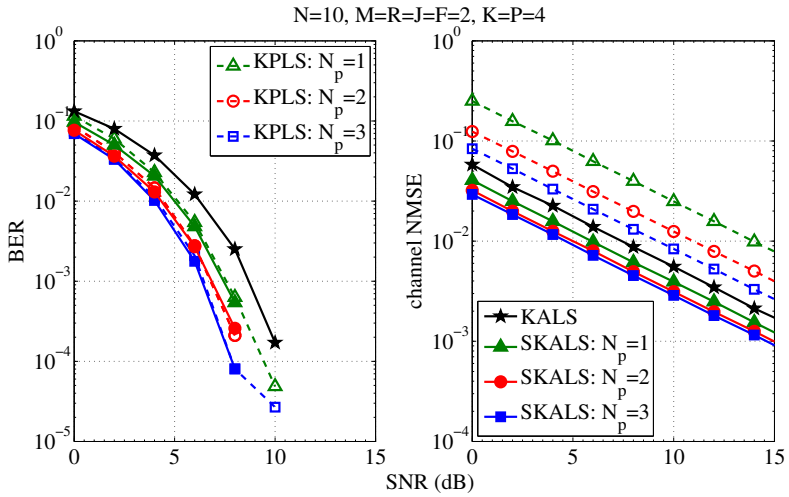

Figure 4: TSTF-KPLS/SKALS/KALS receivers: Impact of pilot length.

6.5. Comparison of the Kronecker receivers for TSTF

We compare four Kronecker receivers (KALS, SKALS, KALMS, KSVD) of the TSTF system, in terms of BER, channel NMSE versus SNR (Fig. 5), and reconstruction NMSE versus iterations number (Fig. 6). For this last figure, the SNR was set to 20dB, and the maximum number of iterations was fixed to 80 for the semi-blind receivers. The parameters are M=R=F=J=2, K=P=4 and N=10. For KALMS, the step sizes $\gamma_{\mathbf{S}}$=0.5 and $\gamma_{\mathbf{H}}$=2 were determined from experiments, so as to obtain a good performance-convergence speed trade-off. From the simulation results, one can draw the following conclusions:

- The supervised SKALS receiver allows to improve significantly the BER and channel NMSE in comparison with the semi-blind receivers.
- The KALS and KSVD receivers provide very close performances in terms of BER, channel NMSE, and reconstruction NMSE, requiring only three iterations for convergence, with nearly the same computational cost. Moreover, we can note the remarkable performance of these semi-blind receivers which allow to nullify the BER for a SNR greater than 10dB.
- The KALMS receiver is the least successful among the three semi-blind receivers, in terms of BER, of channel estimation, and particularly of convergence speed. Due to a much greater average number of iterations for convergence (around N_{it}=60 instead of N_{it}=3 for KALS and KSVD), the KALMS receiver has the highest computational cost.

In summary, one can conclude that the KALS and KSVD receivers provide the best performance-computational cost trade-off among the three semi-blind receivers which use the knowledge of only one symbol. A short pilot sequence exploited by the SKALS receiver allows to improve notably the performance.

7. Conclusion

In the first part of the paper, we have presented five methods for solving the Kronecker product approximation problem: KALS, SKALS, KPLS, KSVD, and KALMS, the first two ones being new. In the second part, eight tensor-based MIMO communication systems have been presented in a unified way, using a generalized PARATUCK model. This presentation has allowed to highlight the evolution of MIMO system design since the pioneering work [17], with the introduction of coding and resource allocation tensors, and an extension to OFDM and CDMA-OFDM systems. A comparative theoretical performance analysis has been carried out, establishing the maximal diversity gain for all considered systems, as a function of the design parameters (K, M, F, P, J), and the allocation matrices/tensors. This study clearly shows the benefits of tensor codings, the more recently proposed TSTF system providing the highest maximal diversity gain, as corroborated by computer simulations. Exploiting the algorithms presented in the first part, five Kronecker receivers have been

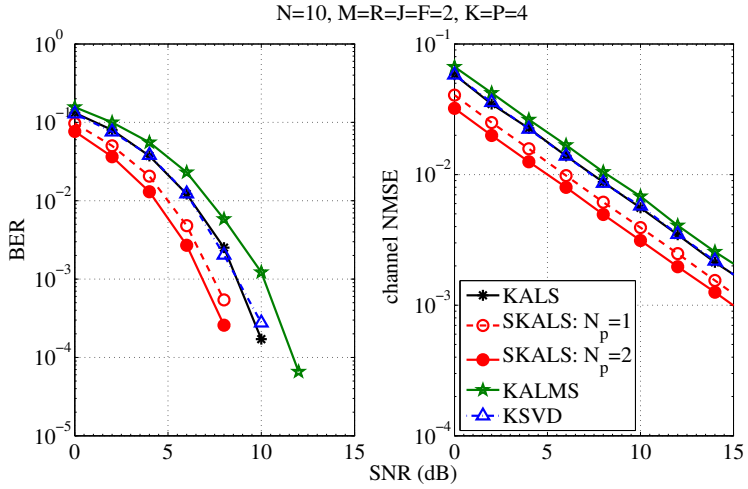

Figure 5: TSTF-Kronecker receivers: Performance comparison.

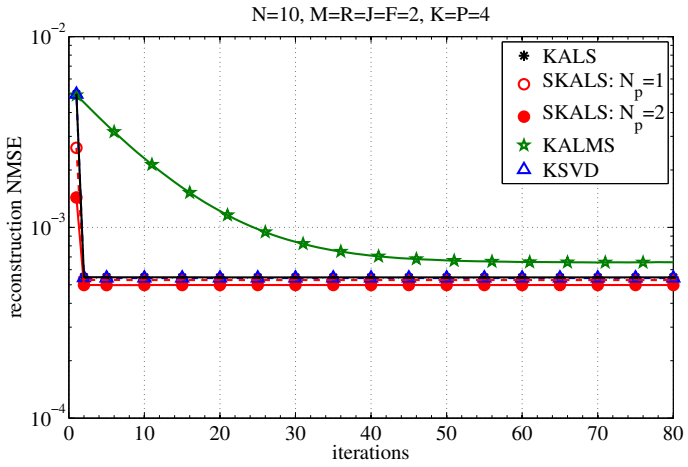

Figure 6: TSTF-Kronecker receivers: Convergence speed comparison.

proposed using Kronecker products between symbol and channel matrices. This new formulation of receivers shows that the systems differentiate by the matrix to be pseudo-inverted for computing the LS estimate of the Kronecker product. This matrix depends both on coding and allocation matrices/tensors. Necessary conditions have been derived for joint channel/symbols estimation, showing that TSTF and TST provide more flexibility than STF and ST. Based on computer simulations, we have compared the performance of TSTF, STF, and TST, with the purpose of illustrating the diversity gain of each system. Finally, the five proposed Kronecker receivers have been compared for the

TSTF system, in terms of BER, channel and reconstruction NMSEs, speed of convergence, and computational cost, showing that the KALS and KSVD receivers provide the best trade-off between computational cost and quality of symbol and channel estimation. The role played by the allocation tensor has also been illustrated.

Perspectives of this work include the optimization of the coding tensor, and extensions of recently developed tensor-based relay systems [32]-[33] by incorporating, at the source and relay nodes, coding and allocation tensors, with Kronecker receivers as proposed in this paper. Allocation tensors can be useful for affecting different levels of coding redundancy at each data stream, or for allocating a subchannel, i.e. an OFDM subcarrier, to each user of a multiuser orthogonal frequency-division multiple access (OFDMA) system. It should be interesting to combine such allocations with a resource allocation optimization [34], [35]. Finally, the multiuser case, with multipath propagation for each user, constitutes also an interesting perspective of this work.

Acknowledgements

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Brazil, under Grant 2014/23936-4.

[1] M. F. Duarte, R. Baraniuk, Kronecker compressive sensing, IEEE Transactions on Image Processing 21 (2) (2012) 494–504.

[2] J. G. Nagy, M. K. Ng, L. Perrone, Kronecker product approximation for image restoration with reflexive boundary conditions, SIAM J. Matrix Anal. Appl. 25 (3) (2004) 829–841.

[3] J. Brewer, Kronecker products and matrix calculus in system theory, IEEE Transactions on Circuits and Systems 25 (9) (1978) 772–781.

[4] R. Bartels, G. Stewart, Solution of the matrix equation AX+XB=C, Comm of the ACM 15 (9) (1972) 820–826.

[5] P. A. Regalia, S. K. Mitra, Kronecker products, unitary matrices and signal processing applications, SIAM Review 31 (4) (1989) 586–613.

[6] N. P. Pitsianis, The Kronecker product in approximation and fast transform generation. Ph. D. Thesis, Cornell University, USA, 1997.

[7] R. A. Harshman, Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-modal factor analysis, UCLA Working Papers in Phonetics 16 (1970) 1–84.

[8] L. Tucker, Some mathematical notes of three-mode factor analysis, Psychometrika 31 (3) (1966) 279–311.

[9] G. Favier, A. L. F. de Almeida, Overview of constrained PARAFAC models, Eurasip J. Adv. Signal Process. 5 (May) (2014) 1–41.

[10] H. Henderson, F. Pukelsheim, S. R. Searle, On the history of the Kronecker product, Linear and Multilinear Algebra 14 (1983) 113–120.

[11] C. Van Loan, The Kronecker product. A product of the times, in: SIAM Conf. on Applied Linear Algebra, Monterey, California, USA, 2009.

[12] A. L. F. de Almeida, G. Favier, J. C. M. Mota, A constrained factor decomposition with application to MIMO antenna systems, IEEE Trans. Signal Process. 56 (6) (2008) 2429–2442.

[13] A. L. F. de Almeida, G. Favier, J. C. M. Mota, Space-time spreading multiplexing for MIMO wireless communication systems using the PARATUCK-2 tensor model, Signal Process. 89 (11) (2009) 2103–2116.

[14] G. Favier, M. N. da Costa, A. L. F. de Almeida, J. M. T. Romano, Tensor space-time (TST) coding for MIMO wireless communication systems, Signal Processing 92 (4) (2012) 1079–1092.

[15] A. L. F. de Almeida, G. Favier, L. R. Ximenes, Space-time-frequency (STF) MIMO communication systems with blind receiver based on a generalized PARATUCK2 model, IEEE Trans. Signal Process. 61 (8) (2013) 1895–1909.

[16] G. Favier, A. L. F. de Almeida, Tensor space-time-frequency coding with semi-blind receivers for MIMO wireless communication systems, IEEE Trans. Signal Process. 62 (22) (2014) 5987–6002.

[17] N. D. Sidiropoulos, G. B. Giannakis, R. Bro, Blind PARAFAC receivers for DS-CDMA systems, IEEE Trans. Signal Process. 48 (3) (2000) 810–823.

[18] N. D. Sidiropoulos, R. S. Budampati, Khatri-Rao space-time codes, IEEE Trans. Signal Process. 50 (10) (2002) 2396–2407.

[19] A. de Baynast, L. de Lathauwer, B. Aazhang, Blind PARAFAC receivers for multiple access-multiple antenna systems, in: Proc. IEEE 58th Fall Vehicular Technology Conf. (VTC 2003), Vol. 2, 2003, pp. 1128–1132.

[20] A. L. F. de Almeida, G. Favier, J. C. M. Mota, Space-time multiplexing codes: A tensor modeling approach, in: Proc. IEEE 7th Workshop SPAWC '06, Cannes, France, 2006.

[21] L. de Lathauwer, A. de Baynast, Blind deconvolution of DS-CDMA signals by means of decomposition in rank-(1,L,L) terms, IEEE Transactions on Signal Processing 56 (4) (2008) 1562–1571.

[22] A. L. F. de Almeida, G. Favier, Double Khatri-Rao space-time -frequency coding using semi-blind PARAFAC based receiver, IEEE Signal Process. Lett. 20 (5) (2013) 471–474.

[23] T. G. Kolda, B. W. Bader, Tensor decompositions and applications, SIAM Review 51 (3) (2009) 455–500.

[24] A. Cichocki, D. Mandic, L. D. Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, H. A. Phan, Tensor decompositions for signal processing applications: From two-way to multiway component analysis, IEEE Signal Process. Magazine 32 (2) (2015) 145–163.

[25] M. N. da Costa, Tensor Space-Time coding for MIMO wireless communication systems, Ph.D. thesis, University of Campinas (UNICAMP) & University of Nice-Sophia Antipolis (March 2014).

[26] C. F. Van Loan, N. P. Pitsianis, Linear algebra for large scale and real-time applications, Kluwer Academic Pub., Dordrecht, Netherlands, 1997, Ch. Approximation with Kronecker products, pp. 293–314.

[27] G. H. Golub, C. F. V. Loan, Matrix computations, Johns Hopkins University Press, 1996.

[28] M. Rupp, S. Schwarz, Gradient-based approaches to learn tensor products, in: Proc. of European Signal Processing Conference (EUSIPCO), Nice, France, 2015.

[29] V. Tarokh, N. Seshadri, A. R. Calderbank, Space-time codes for high data rate wireless communication: performance criterion and code construction, IEEE Trans. Inform. Theory 44 (2) (1998) 744–765.

[30] E. A. Lee, D. G. Messerschmitt, Digital communications, Kluwer Academic, 1993.

[31] B. Clerckx, C. Oestges, MIMO wireless networks: Channels, techniques and standards for multi-antenna, multi-user and multi-cell systems, Academic Press, 2013.

[32] L. R. Ximenes, G. Favier, A. L. F. de Almeida, Semi-blind receivers for non-regenerative cooperative MIMO communications based on nested PARAFAC modeling, IEEE Trans. Signal Process. 63 (18) (2015) 4985–4998.

[33] G. Favier, C. A. R. Fernandes, A. L. F. de Almeida, Nested Tucker tensor decomposition with application to MIMO relay systems using tensor space time coding (TSTC), Signal Processing 128 (2016) 318–331.

[34] Z. Shen, J. G. Andrews, B. L. Evans, Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints, IEEE Transactions on Wireless Communications 4 (6) (2005) 2726–2737.

[35] J. Perez, J. Via, A. Nazabal, Optimal resource allocation in OFDMA broadcast channels using dynamic programming, In Tech, 2011, Ch. 6 in Recent advances in wireless communications and networks, pp. 117–138.