PARATUCK Semi-Blind Receivers for Relaying Multi-Hop MIMO Systems

Pedro Marinho R. de Oliveira, C. Alexandre Rolim Fernandes, Gérard Favier, Remy Boyer

PARATUCK Semi-Blind Receivers for Relaying Multi-Hop MIMO Systems

Pedro Marinho R. de Oliveira, C. Alexandre Rolim Fernandes, Gérard Favier and Rémy Boyer

***Abstract*—In this paper, two receivers are proposed for a multiple-input multiple-output (MIMO) relaying multi-hop communication system using a Khatri-Rao space-time (KRST) coding at the source and amplify-and-forward (AF) relays. It is shown that the third-order tensor of signals received at the destination satisfies a PARATUCK-(K+1) tensor model, where K is the number of relays. After formulating the system model, the expressions of the Cramér-Rao bound (CRB) for the communication channels are derived for the particular case of a two-hop system, i.e., $K = 1$. The presented tensorial modeling enables a joint semi-blind estimation of the transmitted symbols and the channels. The first proposed estimation algorithm is a non-iterative technique based on a rearrangement of the Kronecker product, while the second proposed receiver is based on the alternating least squares (ALS) algorithm. The uniqueness of the tensor decomposition and the identifiability conditions of the proposed algorithms are discussed. The performance of these receivers is evaluated by means of Monte Carlo simulations.**

***Index Terms*—Relaying Systems, PARATUCK, Multi-Hop, MIMO, Semi-Blind Receivers.**

I. Introduction

Aiming to provide an increase in the coverage area and received signal quality, the concept of relaying communication systems was developed, whith at least one relay node used to assist the communication between the source and the destination [1]. An advantage provided by relaying communications is the spatial diversity gain due to the use of relays. In this work, the amplify-and-forward (AF) protocol is used, due to its easy implementation and good performance.

Multi-hop systems, with several relays connected in a serial way, have the advantage of needing less transmission power than two-hop networks (with only one relay), as the distance between the source and the destination is divided in several shorter links, leading to less severe path losses [1].

Furthermore, multiple-input multiple-output (MIMO) systems provided great advances to the wireless communications, due to the increase in coverage area, capacity and spatial diversity gains. This technology is widely used nowadays, being present in several standards (WIMAX-IEEE 802.16, WLAN-IEEE 802.11N, and many others) [2]. In relaying MIMO schemes, the system benefits from a distributed spatial diversity gain, due to the allocation of relays, and from a concentrated spatial diversity gain, due to the antenna array.

Pedro Marinho R. de Oliveira and Gérard Favier are with the University Côte d'Azur, CNRS, I3S Laboratory, France – {marinho, favier}@i3s.unice.fr.

C. Alexandre Rolim Fernandes is with the Department of Computer Engineering, Federal University of Ceará, Sobral, Brazil – alexandrefernandes@ufc.br.

Rémy Boyer is with the University of Lille-1, CRIStAL Laboratory, France – remy.boyer@univ-lille.fr.

In multi-hop systems, channel estimation is a challenging and relevant problem. Channel state information (CSI) plays an important role for optimizing MIMO relay systems in terms of power allocation, adaptive relaying protocols and space-time coding design.

On the other hand, due to its advantage in exploring the multidimensional nature of signals, tensor decompositions have found applications in several areas, including array signal processing [3] and telecommunications [4], [5], [6], [7], [8]. Tensor decompositions have some advantages over matrix-based methods as, for instance, their uniqueness properties under mild conditions and the fact that the rank of a tensor can exceed its dimensions. Also, tensor analysis has shown to be an efficient approach for channel and/or symbol estimation in relaying MIMO systems [9], [10], [11], [12].

Matrix-based techniques for conventional multi-hop MIMO schemes can be found in [13], [14], [15]. These works show the advantages of multi-hop relaying networks. However, when compared with tensor-based approaches, matrix-based methods generally require stronger constraints in order to guarantee the uniqueness of the estimated parameters, due to the constraint on the matrix rank which is limited by the lowest dimension.

In [7], three semi-blind receivers are proposed for a two-hop MIMO AF relaying system using the Khatri-Rao space-time (KRST) coding [16]. These receivers combine two tensor models (PARAllel FACtor - PARAFAC and PARATUCK-2) that enables the joint estimation of symbols and channels. The tensor modeling of the present paper generalizes the tensor modeling of [7] for the multi-hop case. Moreover, we propose both iterative and non-iterative receivers, contrarily to [7] that only presents iterative methods. In addition, the iterative algorithm proposed in the present paper has some advantages over the ones of [7], as it will be illustrated by the simulation results. In a two-hop MIMO AF relaying system, the work [9] exploited a nested-PARAFAC tensor model and proposed two semi-blind receivers, jointly estimating the symbols and the channels of the communication links. In [10], a tensor space-time coding is used in a MIMO relaying system modeled as a nested Tucker decomposition.

Regarding the use of the PARATUCK decomposition in wireless communications, a blind receiver based on a generalized fourth-order PARATUCK-2 decomposition is proposed in [17] for space-time frequency (STF) MIMO systems. In this tensor model, the core tensor of the PARATUCK-2 decomposition is a spatial coding matrix combined with two third-order tensors. The PARATUCK-2 tensor model is used in [18] to represent a MIMO wireless communication system with space-time spreading-multiplexing. As [17] and [18] used point-to-point systems, [12] used the PARATUCK-3 decomposition to model a three-hop MIMO relaying system and to propose iterative receivers, using channel training sequences. Contrarily to the present paper, the receivers proposed in [12] are restricted to the three-hop case and they consider a supervised scenario. Moreover, in [12], no closed-form algorithm is proposed and the used coding scheme is different from the one of the present paper.

In this work, we propose two semi-blind receivers that jointly estimate the symbols and the channels in a multi-hop AF MIMO relaying system. It is used a simplified KRST coding at the source node, combined with the AF protocol at the relay nodes. The third-order tensor of signals received by the destination node satisfies a PARATUCK-(K+1) decomposition [19], where K is the number of relays. By imposing a simple restriction on the AF relay gains, we derive an alternative representation for the receiver signal tensor, expressing the PARATUCK-(K+1) decomposition as a PARATUCK-2 model. Sufficient uniqueness conditions are provided for the considered tensor decomposition. We also derive the expressions of the expected Cramér-Rao bound (CRB) [20], [21], [22], [23] for the communication channels, considering the particular case of using a single relay, i.e., when the system has only two hops, satisfying the PARATUCK-2 tensor model.

The first proposed receiver is non-iterative and based on a rearrangement of the Kronecker product between the symbols and channel matrices. The second presented receiver is iterative, based on the alternating least squares (ALS) algorithm, to jointly and alternately estimate the matrix factors. Both algorithms estimate the transmitted symbols, the channel matrix of the last hop and a global channel matrix that depends on all the other channel matrices. The identifiability conditions of the proposed algorithms are discussed and their performance is evaluated by means of computational simulations using Monte Carlo runs, showing the good performance of the proposed receivers.

In this paper, we extend the work [11] through the generalization of the system model to the multi-hop case, the proposition of a new receiver and the derivation of uniqueness and identifiability conditions.

The main original contributions of the paper can be summarized as follows:

- We present a MIMO relaying multi-hop communication system using a KRST coding at the source and AF relays.
- We derive a PARATUCK-(K+1) representation for the tensor of received signals. Moreover, we derive an alternative representation for the received signals tensor as a PARATUCK-2 model.
- We derive the expressions of the expected CRB for the communication channels of the PARATUCK-2 model. To the best of our knowledge, this is the first time that a mean square error (MSE) lower bound is derived for such wireless communication systems.
- Two different sets of sufficient uniqueness conditions for the tensor model are provided.
- Two receivers for the considered multi-hop system are proposed.
- The identifiability conditions of the proposed algorithms are derived.
- Simulations are provided to illustrate the performance of the receivers.

The rest of this paper is organized as follows. In Section II, the notation and the tensor prerequisites are introduced. In Section III, the system model is described, while, in Section IV, two semi-blind receivers are proposed. In Section V, we derive the expressions of the expected CRB for the communication channels, considering the two-hop case. Sections VI and VII present the uniqueness properties of the PARATUCK-N tensor model and the identifiability conditions of the proposed algorithms, respectively. In Section VIII the simulation results are presented and, finally, Section IX concludes the paper and presents some perspectives for future work.

II. Notation and Tensor Prerequisites

Scalars, vectors, matrices and tensors are represented, respectively, by lower-case (a, b, c, ...), boldface lower-case (**a**, **b**, **c**, ...), boldface capital (**A**, **B**, **C**, ...) and calligraphic ($\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, ...) letters.

$a_{m,i}$ is the $(m,i)^{th}$ element of the matrix $\mathbf{A} \in \mathbb{C}^{M \times I}$ and $a_{m,i,p}$ is the $(m,i,p)^{th}$ element of the tensor $\mathcal{A} \in \mathbb{C}^{M \times I \times P}$. The transpose, hermitian transpose, Moore-Penrose pseudo-inverse, conjugate, m^{th} row and n^{th} column of the matrix **A**, are respectively represented by $\mathbf{A}^T$, $\mathbf{A}^H$, $\mathbf{A}^\dagger$, $\mathbf{A}^*$, $\mathbf{A}_{m.}$ and $\mathbf{A}_{.n}$. $||\cdot||_F$, $\otimes$ and $\diamond$, denote the Frobenius norm, the Kronecker product and the Khatri-Rao (column-wise Kronecker) product, respectively. The operator diag(**a**) generates a diagonal matrix with the vector argument **a** as the main diagonal, $D_m(\mathbf{A})$ represents the diagonal matrix with the m^{th} row of **A** as the main diagonal, $\text{Bdiag}\{\mathbf{A}_1, ..., \mathbf{A}_P\}$ creates a block diagonal matrix by aligning the input matrices $\mathbf{A}_1, ..., \mathbf{A}_P$ along the diagonal, the operator $\text{vec}(\cdot)$ vectorizes its matrix argument by stacking its columns, while the operator $\text{unvec}(\cdot)$ unvectorizes its vector argument to the original matrix. The operator $SVD(\mathbf{A})$ computes the singular value decomposition (SVD) of the matrix argument **A**. Moreover, $\hat{a}$, $\hat{\mathbf{a}}$ and $\hat{\mathbf{A}}$ represent the estimates of a, **a** and **A**, respectively.

$\Re\{\cdot\}$ and $\Im\{\cdot\}$ represent the real and imaginary parts of the argument, respectively. $\text{Trace}\{\cdot\}$ computes the trace of the matrix argument, $\mathbb{E}$ is the expectation operator and $\mathbf{I}_N$ is the identity matrix of order N.

In this work, the four following properties will be used:

$$\text{vec}(\mathbf{ABC}^T) = (\mathbf{C} \otimes \mathbf{A})\text{vec}(\mathbf{B}), \tag{1}$$

$$D_i(\mathbf{A}) \otimes D_i(\mathbf{D}) = D_i((\mathbf{A}^T \diamond \mathbf{D}^T)^T), \tag{2}$$

$$(\mathbf{A} \otimes \mathbf{D})(\mathbf{B} \otimes \mathbf{E}) = (\mathbf{AB} \otimes \mathbf{DE}), \tag{3}$$

$$\text{vec}(\mathbf{B}\text{diag}(\boldsymbol{\lambda})\mathbf{C}^T) = (\mathbf{C} \diamond \mathbf{B})\boldsymbol{\lambda}, \tag{4}$$

where $\mathbf{A} \in \mathbb{C}^{I \times P}$, $\mathbf{B} \in \mathbb{C}^{P \times M}$, $\mathbf{C} \in \mathbb{C}^{K \times M}$, $\mathbf{D} \in \mathbb{C}^{I \times M}$, $\mathbf{E} \in \mathbb{C}^{M \times K}$ and $\boldsymbol{\lambda} \in \mathbb{C}^{M \times 1}$. Given an arbitrary third-order tensor $\mathcal{A} \in \mathbb{C}^{M \times I \times P}$, its horizontal, lateral and frontal slices are respectively represented by $\mathbf{A}_{m..} \in \mathbb{C}^{I \times P}$, $\mathbf{A}_{.i.} \in \mathbb{C}^{P \times M}$ and $\mathbf{A}_{..p} \in \mathbb{C}^{M \times I}$.

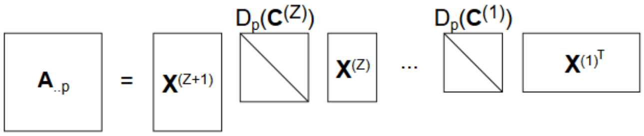

Fig. 1. PARATUCK-Z decomposition in matrix slice terms.

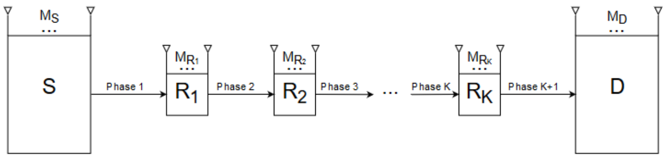

Fig. 2. MIMO cooperative system model with K relays.

The PARATUCK-Z decomposition of a third-order tensor $\mathcal{A} \in \mathbb{C}^{M \times I \times P}$, in matrix slices terms, is illustrated in Figure 1 and given by:

$$\mathbf{A}_{..p} = \mathbf{X}^{(Z+1)} D_p(\mathbf{C}^{(Z)})\mathbf{X}^{(Z)} ... D_p(\mathbf{C}^{(2)})\mathbf{X}^{(2)} D_p(\mathbf{C}^{(1)}) \times \times \mathbf{X}^{(1)^T} \in \mathbb{C}^{M \times I}, \quad (5)$$

for $p = 1, ..., P$, where $\mathbf{C}^{(z)} \in \mathbb{C}^{P \times R_z}$, for $z = 1, ..., Z$, $\mathbf{X}^{(z)} \in \mathbb{C}^{R_z \times R_{z-1}}$, for $z = 2, ..., Z$, $\mathbf{X}^{(1)} \in \mathbb{C}^{I \times R_1}$ and $\mathbf{X}^{(Z+1)} \in \mathbb{C}^{M \times R_Z}$. In scalar form, the PARATUCK-Z decomposition can be written as:

$$a_{m,i,p} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} ... \sum_{r_Z=1}^{R_Z} x^{(Z+1)}_{m,r_Z} c^{(Z)}_{p,r_Z} x^{(Z)}_{r_Z, r_{Z-1}} ... c^{(2)}_{p,r_2} x^{(2)}_{r_2,r_1} c^{(1)}_{p,r_1} x^{(1)}_{i,r_1}. \quad (6)$$

III. System Model

In the present work, a $(K+1)$-hop one-way MIMO AF relay system with one source (S) node, one destination (D) node and K relays (R_1, R_2, ..., R_K) nodes is considered, as illustrated in Figure 2, where M_X denotes the number of antennas at node X (e.g, M_S denotes the number of antennas at node S). All the channels are assumed to be invariant during the total transmission time and to undergo frequency flat fading. The transmitted symbols are quadrature amplitude modulation (QAM)-modulated or phase shift keying (PSK)-modulated. The relays are half-duplex and the transmission runs in $K+1$ phases. In the first phase, the source transmits the information signals to the first relay, in the second phase, the first relay transmits towards the second relay and so on, until the K^{th} relay (the last one) transmits towards the destination in the $(K+1)^{th}$ phase.

Consider that $\mathbf{H}^{(SR_1)} \in \mathbb{C}^{M_{R_1} \times M_S}$, $\mathbf{H}^{(R_{k-1}R_k)} \in \mathbb{C}^{M_{R_k} \times M_{R_{k-1}}}$ and $\mathbf{H}^{(R_K D)} \in \mathbb{C}^{M_D \times M_{R_K}}$ are MIMO channel matrices of the SR_1, $R_{k-1}R_k$ and $R_K D$ links, respectively, for $k = 2, ...K$. $\mathbf{S} \in \mathbb{C}^{N \times M_S}$ is the matrix with the information symbols multiplexed to the M_S antennas during N consecutive symbol periods. A simplified KRST coding is used at the source to introduce time redundancy:

$$\mathbf{X}_{..p} = D_p(\mathbf{G}_0)\mathbf{S}^T \in \mathbb{C}^{M_S \times N} \quad (7)$$

where $p = 1, ..., P$, $\mathbf{X}_{..p}$ is the p^{th} slice of the transmitted signal tensor $\mathcal{X} \in \mathbb{C}^{M_S \times N \times P}$, $\mathbf{G}_0 \in \mathbb{C}^{P \times M_S}$ is the coding matrix of the source node and P is the number of transmission blocks, each block being composed of N symbol periods. The signals received by the first relay during the p^{th} transmission block are given by the p^{th} matrix slice of the tensor $\tilde{\mathcal{Y}}^{(R_1)} \in \mathbb{C}^{M_{R_1} \times N \times P}$:

$$\tilde{\mathbf{Y}}^{(R_1)}_{..p} = \mathbf{H}^{(SR_1)}\mathbf{X}_{..p} + \mathbf{V}^{(R_1)}_{..p} \in \mathbb{C}^{M_{R_1} \times N}, \quad (8)$$

where $\mathbf{V}^{(R_1)}_{..p} \in \mathbb{C}^{M_{R_1} \times N}$ is the p^{th} matrix slice of the additive white Gaussian noise (AWGN) tensor $\mathcal{V}^{(R_1)} \in \mathbb{C}^{M_{R_1} \times N \times P}$ during the p^{th} transmission block at the relay R_1.

It is considered that all the relays use the AF protocol. Let $\mathbf{G}_k \in \mathbb{C}^{P \times M_{R_k}}$ be the gain matrix of the R_k node, for $k = 1, ..., K$. The amplified signal $D_p(\mathbf{G}_{k-1})\tilde{\mathbf{Y}}^{(R_{k-1})}_{..p} \in \mathbb{C}^{M_{R_{k-1}} \times N}$ is transmitted by R_{k-1} to R_k, during the p^{th} transmission block, for $k = 2, ..., K$. The signals received by R_k can then be written as the matrix slice of the tensor $\tilde{\mathcal{Y}}^{(R_k)} \in \mathbb{C}^{M_{R_k} \times N \times P}$:

$$\tilde{\mathbf{Y}}^{(R_k)}_{..p} = \mathbf{H}^{(R_{k-1}R_k)} D_p(\mathbf{G}_{k-1})\tilde{\mathbf{Y}}^{(R_{k-1})}_{..p} + \mathbf{V}^{(R_k)}_{..p} \in \mathbb{C}^{M_{R_k} \times N}. \quad (9)$$

where $\mathbf{V}^{(R_k)}_{..p} \in \mathbb{C}^{M_{R_k} \times N}$ is the p^{th} matrix slice of the AWGN tensor $\mathcal{V}^{(R_k)} \in \mathbb{C}^{M_{R_k} \times N \times P}$ during the p^{th} transmission block at the relay R_k.

The amplified signal $D_p(\mathbf{G}_K)\tilde{\mathbf{Y}}^{(R_K)}_{..p}$ is then transmitted by R_K to the destination. Finally, the signals during the p^{th} transmission block received by the destination are given by the p^{th} matrix slice of the tensor $\tilde{\mathcal{Y}}^{(D)} \in \mathbb{C}^{M_D \times N \times P}$:

$$\tilde{\mathbf{Y}}^{(D)}_{..p} = \mathbf{H}^{(R_K D)} D_p(\mathbf{G}_K)\tilde{\mathbf{Y}}^{(R_K)}_{..p} + \mathbf{V}^{(D)}_{..p} \in \mathbb{C}^{M_D \times N}, \quad (10)$$

where $\mathbf{V}^{(D)}_{..p} \in \mathbb{C}^{M_D \times N}$ is the p^{th} matrix slice of the AWGN tensor $\mathcal{V}^{(D)} \in \mathbb{C}^{M_D \times N \times P}$ at the destination node.

We can re-express (10) as:

$$\tilde{\mathbf{Y}}^{(D)}_{..p} = \mathbf{Y}^{(D)}_{..p} + \tilde{\mathbf{V}}^{(D)}_{..p} \in \mathbb{C}^{M_D \times N}, \quad (11)$$

where $\mathbf{Y}^{(D)}_{..p}$ is the p^{th} matrix slice of the noiseless signal tensor $\mathcal{Y}^{(D)} \in \mathbb{C}^{M_D \times N \times P}$ given by:

$$\mathbf{Y}^{(D)}_{..p} = \mathbf{H}^{(R_K D)} D_p(\mathbf{G}_K)\mathbf{H}^{(R_{K-1}R_K)} D_p(\mathbf{G}_{K-1}) ... \mathbf{H}^{(SR_1)} D_p(\mathbf{G}_0)\mathbf{S}^T \in \mathbb{C}^{M_D \times N}, \quad (12)$$

or, equivalently:

$$\mathbf{Y}^{(D)}_{..p} = \mathbf{H}^{(R_K D)} D_p(\mathbf{G}_K) \left[\prod_{i=2}^{K} \mathbf{H}^{(R_{i-1}R_i)} D_p(\mathbf{G}_{i-1})\right] \mathbf{H}^{(SR_1)} D_p(\mathbf{G}_0)\mathbf{S}^T \in \mathbb{C}^{M_D \times N}, \quad (13)$$

and $\tilde{\mathbf{V}}^{(D)}_{..p} \in \mathbb{C}^{M_D \times N}$ is the p^{th} matrix slice of the global noise tensor $\tilde{\mathcal{V}}^{(D)}_{..p} \in \mathbb{C}^{M_D \times N \times P}$, given by:

$$\tilde{\mathbf{V}}^{(D)}_{..p} = \mathbf{H}^{(R_K D)} D_p(\mathbf{G}_K)\tilde{\mathbf{V}}^{(R_K)}_{..p} + \mathbf{V}^{(D)}_{..p}, \quad (14)$$

with the global noise matrix at node k defined recursively as:

$$\tilde{\mathbf{V}}^{(R_k)}_{..p} = \mathbf{H}^{(R_{k-1}R_k)} D_p(\mathbf{G}_{k-1}) \tilde{\mathbf{V}}^{(R_{k-1})}_{..p} + \mathbf{V}^{(R_k)}_{..p}, \tag{15}$$

for $k = 2, ..., K$ and $\tilde{\mathbf{V}}^{(R_1)}_{..p} = \mathbf{V}^{(R_1)}_{..p}$.

The received signals in (12) define a third-order tensor which satisfies a PARATUCK-(K+1) decomposition that can be rewritten in scalar form as:

$$y^{(D)}_{m_D,n,p} = \sum_{m_S=1}^{M_S} \sum_{m_{R_1}=1}^{M_{R_1}} \cdots \sum_{m_{R_K}=1}^{M_{R_K}} h^{(R_K D)}_{m_D, m_{R_K}} g^{(K)}_{p, m_{R_K}} \times \\ \times h^{(R_{K-1}R_K)}_{m_{R_K}, m_{R_{K-1}}} g^{(K-1)}_{p, m_{R_{K-1}}} ... h^{(SR_1)}_{m_{R_1}, m_S} g^{(0)}_{p, m_S} s_{n, m_S}. \tag{16}$$

The received signal model described by (12) and (16) is equivalent to the PARATUCK-Z decomposition introduced in (5) and (6), with the following correspondences:

$$(K+1, P, N, M_S, M_{R_1}, ..., M_{R_K}, M_D) \\ \iff (Z, P, I, R_1, R_2, ..., R_Z, M) \tag{17}$$

and

$$(\mathbf{H}^{(R_K D)}, \mathbf{H}^{(R_{K-1}R_K)}, ..., \mathbf{H}^{(SR_1)}, \mathbf{S}, \mathbf{G}_K, ..., \mathbf{G}_0) \\ \iff (\mathbf{X}^{(Z+1)}, \mathbf{X}^{(Z)}, ..., \mathbf{X}^{(2)} \mathbf{X}^{(1)}, \mathbf{C}^{(Z)}, ..., \mathbf{C}^{(1)}). \tag{18}$$

The transmission rate of the source is given by M_S/P and the total transmission rate of the proposed system is $M_S/P(K+1)$.

IV. Symbols and Channels Estimation

This section presents two algorithms that jointly and semi-blindly estimate the symbols and the channels of the multi-hop MIMO relaying system presented in Section III. The first algorithm is called least-squares Kronecker rearrangement-based (LS-KR) and it is based on a rearrangement of the Kronecker product that provides rank-1 matrices. The second algorithm is an iterative ALS-based algorithm called PARATUCK-ALS. In order to simplify the presentation of the proposed algorithms, from now on, we will work with the noiseless part of the received signals.

Before presenting the proposed algorithms, we will develop some mathematical expressions for the tensor of received signals. Applying Properties (1) and (3) to (12), we have:

$$\mathbf{y}^{(D)}_p = (\mathbf{S} \otimes \mathbf{H}^{(R_K D)})(D_p(\mathbf{G}_0) \otimes D_p(\mathbf{G}_K)) \times \\ \times \text{vec}(\mathbf{H}^{(R_{K-1}R_K)} D_p(\mathbf{G}_{K-1}) ... D_p(\mathbf{G}_1) \mathbf{H}^{(SR_1)}). \tag{19}$$

Now applying (2) to (19), we get:

$$\mathbf{y}^{(D)}_p = (\mathbf{S} \otimes \mathbf{H}^{(R_K D)}) D_p((\mathbf{G}_0^T \diamond \mathbf{G}_K^T)^T) \times \\ \times \text{vec}(\mathbf{H}^{(R_{K-1}R_K)} D_p(\mathbf{G}_{K-1}) ... D_p(\mathbf{G}_1) \mathbf{H}^{(SR_1)}). \tag{20}$$

Let us define $\mathbf{G}_{0K} = \mathbf{G}_0^T \diamond \mathbf{G}_K^T \in \mathbb{C}^{M_{R_K} M_S \times P}$. Note that $\mathbf{G}_{0K}$ depends on the coding matrices of the first and last hops. (20) can be rewritten as:

$$\mathbf{y}^{(D)}_p = (\mathbf{S} \otimes \mathbf{H}^{(R_K D)}) \times \\ \times \text{diag}(\text{vec}(\mathbf{H}^{(G)}_{..p}))(\mathbf{G}_{0K})_{.p} \in \mathbb{C}^{M_D N \times 1}, \tag{21}$$

with

$$\mathbf{H}^{(G)}_{..p} = \mathbf{H}^{(R_{K-1}R_K)} D_p(\mathbf{G}_{K-1}) ... D_p(\mathbf{G}_1) \mathbf{H}^{(SR_1)} \\ \in \mathbb{C}^{M_{R_K} \times M_S}. \tag{22}$$

where $\mathbf{H}^{(G)}_{..p}$ is the p^{th} matrix slice of the global channel tensor $\mathcal{H}^{(G)} \in \mathbb{C}^{M_{R_K} \times M_S \times P}$ that depends on all the channel matrices, but the last one $\mathbf{H}^{(R_K D)}$, and on the gain matrices $\mathbf{G}_1, ..., \mathbf{G}_{K-1}$. Assuming that the matrices $\mathbf{G}_1$, ..., $\mathbf{G}_{K-1}$ are formed of repeated rows, such that $D_p(\mathbf{G}_k)$ is independent of p, for $k = 1, ..., K-1$, and stacking the column vectors $\mathbf{y}^{(D)}_p$, for $p = 1, ..., P$, side by side into a matrix, we have:

$$\mathbf{Y}^{(D)}_1 = \left[\mathbf{y}^{(D)}_1 \dots \mathbf{y}^{(D)}_P\right] \tag{23}$$

$$= (\mathbf{S} \otimes \mathbf{H}^{(R_K D)}) \text{diag}(\text{vec}(\mathbf{H}^{(G)})) \mathbf{G}_{0K}, \tag{24}$$

with $\mathbf{H}^{(G)} = \mathbf{H}^{(G)}_{..p}$, for $p = 1, ..., P$. Note that the assumption that $\mathbf{G}_1, ..., \mathbf{G}_{K-1}$ are formed of repeated rows means that the same relay gains are used for all transmission blocks. Note that, under this assumption, the PARATUCK-(K+1) received signals tensor can be expressed as a PARATUCK-2 model, with the following matrix factors: $\mathbf{H}^{(R_K D)}$, $\mathbf{G}_K$, $\mathbf{H}^{(G)}$, $\mathbf{G}_0$ and $\mathbf{S}$. Indeed, in this case, (12) can be re-expressed as:

$$\mathbf{Y}^{(D)}_{..p} = \mathbf{H}^{(R_K D)} D_p(\mathbf{G}_K) \mathbf{H}^{(G)} D_p(\mathbf{G}_0) \mathbf{S}^T. \tag{25}$$

We can write from (24):

$$\mathbf{Y}^{(D)}_1 = (\mathbf{S} \otimes \mathbf{H}^{(R_K D)}) \text{diag}(\mathbf{h}^{(G)}) \mathbf{G}_{0K} \in \mathbb{C}^{M_D N \times P}, \tag{26}$$

where $\mathbf{h}^{(G)} = \text{vec}(\mathbf{H}^{(G)})$. Assume that $\mathbf{G}_{0K}$ has a right inverse, i.e. $\mathbf{G}_{0K} \mathbf{G}^{\dagger}_{0K} = \mathbf{I}_{M_{R_K} M_S}$. This means that $\mathbf{G}_{0K}$ is full-row rank, which implies $P \geq M_{R_K} M_S$.

From (26), we may then write:

$$\mathbf{Y}^{(D)}_1 \mathbf{G}^{\dagger}_{0K} = (\mathbf{S} \otimes \mathbf{H}^{(R_K D)}) \text{diag}(\mathbf{h}^{(G)}) \\ \in \mathbb{C}^{M_D N \times M_{R_K} M_S}. \tag{27}$$

A. Least-Squares Kronecker Rearrangement-Based (LS-KR) Algorithm

In [24], a rearrangement of a given Kronecker product matrix as a rank-1 matrix is exploited to estimate the matrix factors of the Kronecker product. Let us define $\mathbf{W}$ as:

$$\mathbf{W} = (\mathbf{S} \otimes \mathbf{H}^{(R_K D)}) \text{diag}(\mathbf{h}^{(G)}) \in \mathbb{C}^{M_D N \times M_{R_K} M_S}, \tag{28}$$

which can be reexpressed as:

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} & \dots & \mathbf{W}_{1M_S} \\ \mathbf{W}_{21} & \mathbf{W}_{22} & \dots & \mathbf{W}_{2M_S} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{W}_{N1} & \mathbf{W}_{N2} & \dots & \mathbf{W}_{NM_S} \end{bmatrix}, \tag{29}$$

where

$$\mathbf{W}_{nm} = s_{nm} \mathbf{H}^{(R_K D)} \text{diag}(\mathbf{H}^{(G)}_{.m}) \in \mathbb{C}^{M_D \times M_{R_K}}, \tag{30}$$

for $n = 1, ..., N$ and $m = 1, ..., M_S$. That leads to:

$$\text{vec}(\mathbf{W}_{nm})^T = s_{nm} \text{vec}(\mathbf{H}^{(R_K D)})^T \text{diag}(\mathbf{H}^{(G)}_{.m} \otimes \mathbf{1}_{M_D}) \\ \in \mathbb{C}^{1 \times M_D M_{R_K}}, \tag{31}$$

where $\mathbf{1}_{M_D}$ is the column vector of length M_D composed of 1s . Let us define:

$$\mathbf{W}_m = \begin{bmatrix} \text{vec}(\mathbf{W}_{1m})^T \\ \vdots \\ \text{vec}(\mathbf{W}_{Nm})^T \end{bmatrix} \in \mathbb{C}^{N \times M_D M_{R_K}}, \quad (32)$$

for $m = 1, ..., M_S$. From (31) and (32), we can write:

$$\mathbf{W}_m = \mathbf{S}_{.m}\text{vec}((\mathbf{H}^{(R_K D)})^T)\text{diag}(\mathbf{H}^{(G)}_{.m} \otimes \mathbf{1}_{M_D}), \quad (33)$$

which is a rank-1 matrix.

The first step of the LS-KR algorithm consists in estimating $\mathbf{W}$ from (27) by means of the LS method. The second step of the proposed algorithm consists in performing the rearrangement described in (29), (31) and (32), and, then, in estimating $\mathbf{S}_{.m}$ as the dominant left singular vector of $\hat{\mathbf{W}}_m$ and vec$(\mathbf{H}^{(R_K D)})$ as the conjugate of the dominant right singular vector of $\hat{\mathbf{W}}_m$, for $m = 1, ..., M_S$. At the end, there will be one estimation of $\mathbf{S}$ and M_S estimations of $\mathbf{H}^{(R_K D)}$. We can then choose any value of m $(1 \leq m \leq M_S)$ for estimating $\mathbf{H}^{(R_K D)}$. In the simulations, we have used $m = 1$.

These estimates have the following ambiguities: $\hat{\mathbf{S}}_{.m} = \mathbf{S}_{.m}\delta_m$ and vec$(\hat{\mathbf{H}}^{(R_K D)})$ = diag$(\mathbf{1}_{M_D} \otimes \mathbf{H}^{(G)}_{.m})vec(\mathbf{H}^{(R_K D)})/\delta_m$, where δ_m is a scalar, for $m = 1, ..., M_S$. Applying the unvec$(\cdot)$ operator to the last equation leads to: $\hat{\mathbf{H}}^{(R_K D)}$ = diag$(\mathbf{H}^{(G)}_{.m})\delta_m \mathbf{H}^{(R_K D)}$, which means that the rows of $\mathbf{H}^{(R_K D)}$ can only be estimated with scalar ambiguities

Moreover, by concatenating the estimations of $\mathbf{S}_{.m}$ for $m = 1, ..., M_S$, we get $\hat{\mathbf{S}} = \mathbf{S}\boldsymbol{\Delta}_s$, where $\boldsymbol{\Delta}_s$ = diag$([\delta_1, ..., \delta_{M_S}]^T) \in \mathbb{C}^{M_S \times M_S}$ is a diagonal ambiguity matrix. This means that the columns of $\mathbf{S}$ can only be estimated with scalar ambiguities. The ambiguity matrix $\boldsymbol{\Delta}_s$ can be estimated and canceled from $\mathbf{S}$ by assuming that the first row of $\mathbf{S}$ is known. This can be done by using one pilot symbol per transmit antenna. Concerning the scaling ambiguity of the matrix $\hat{\mathbf{H}}^{(R_K D)}$, we assumed that the first column of $\mathbf{H}^{(R_K D)}$ is known. In practice, this information can be obtained by a simple LS estimation using a training sequence generated by the relays [9].

The final step of the proposed algorithm is to estimate the global channel matrix $\mathbf{H}^{(G)}$. Applying Property (4) to (24), we have:

$$\text{vec}(\mathbf{Y}_1^{(D)}) = \left[\mathbf{G}_{0K}^T \diamond (\mathbf{S} \otimes \mathbf{H}^{(R_K D)})\right] \text{vec}(\mathbf{H}^{(G)}) \quad (34)$$

$$= \mathbf{C}^{(D)}\mathbf{h}^{(G)} \in \mathbb{C}^{M_D N P \times 1}, \quad (35)$$

where

$$\mathbf{C}^{(D)} = \left[\mathbf{G}_{0K}^T \diamond (\mathbf{S} \otimes \mathbf{H}^{(R_K D)})\right] \in \mathbb{C}^{M_D N P \times M_{R_K} M_S}. \quad (36)$$

The global channel matrix estimation is then given by $\hat{\mathbf{H}}^{(G)}$ = unvec$(\hat{\mathbf{h}}^{(G)})$, where:

$$\hat{\mathbf{h}}^{(G)} = \left(\hat{\mathbf{C}}^{(D)}\right)^{\dagger} \text{vec}(\tilde{\mathbf{Y}}_1^{(D)}), \quad (37)$$

where $\tilde{\mathbf{Y}}_1^{(D)}$ is the noisy version of $\mathbf{Y}_1^{(D)}$ and $\hat{\mathbf{C}}^{(D)}$ = $\left[\mathbf{G}_{0K}^T \diamond (\hat{\mathbf{S}} \otimes \hat{\mathbf{H}}^{(R_K D)})\right]$.

Algorithm 1 - (LS-KR)

1: $\hat{\mathbf{W}} = \tilde{\mathbf{Y}}_1^{(D)} \mathbf{G}_{0K}^{\dagger}$;
2: **for** $m = 1$ to M_S **do**
3: $\quad \hat{\mathbf{W}}_m = \begin{bmatrix} \text{vec}(\hat{\mathbf{W}}_{1m})^T \\ \vdots \\ \text{vec}(\hat{\mathbf{W}}_{Nm})^T \end{bmatrix}$;
4: $\quad \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H = SVD(\hat{\mathbf{W}}_m)$;
5: $\quad \hat{\mathbf{S}}_{.m} = \mathbf{U}_{.1}$;
6: $\quad$ **if** $m = 1$ **then**
7: $\quad\quad \hat{\mathbf{H}}^{(R_K D)}$ = unvec$(\mathbf{V}^*_{.1})$;
8: $\quad$ **end if**
9: **end for**
10: $\hat{\mathbf{h}}^{(G)} = \left[\mathbf{G}_{0K}^T \diamond (\hat{\mathbf{S}} \otimes \hat{\mathbf{H}}^{(R_K D)})\right]^{\dagger} \text{vec}(\tilde{\mathbf{Y}}_1^{(D)})$;
11: **Remove ambiguities**

Note that $\hat{\mathbf{H}}^{(G)}$ will also have ambiguities. Indeed, as $\hat{h}^{(R_K D)}_{m_d, m_r}$ has a scalar ambiguity in the form: $\hat{h}^{(R_K D)}_{m_d, m_r} = h^{(R_K D)}_{m_d, m_r} h^{(G)}_{m_r, m}/\delta_m$, where m corresponds to the index used for estimating $\mathbf{H}^{(R_K D)}$ from $\hat{\mathbf{W}}_m$, $\hat{h}^{(G)}_{m_r, m_s}$ will have the following ambiguity:

$$\hat{h}^{(G)}_{m_r, m_s} = h^{(G)}_{m_r, m_s}\delta_m / h^{(G)}_{m_r, m}, \quad (38)$$

which leads to

$$\hat{\mathbf{H}}^{(G)} = \boldsymbol{\Delta}_h \mathbf{H}^{(G)}, \quad (39)$$

where $\boldsymbol{\Delta}_h = \delta_m \text{diag}([1/h^{(G)}_{1,m}, ..., 1/h^{(G)}_{M_{R_K}, m}]^T) \in \mathbb{C}^{M_{R_K} \times M_{R_K}}$ is a diagonal ambiguity matrix. This means that the rows of $\mathbf{H}^{(G)}$ can only be estimated with scalar ambiguities. Such ambiguities can be eliminated assuming that the first column of $\mathbf{H}^{(R_K D)}$ is known, as already mentioned. This proposed algorithm, denoted by Kronecker Rearrangement (KR), is summarized in Algorithm 1.

Considering that the computational complexity of the SVD of a $M \times N$ matrix is $\mathcal{O}(MN \min\{M, N\})$ and the one of the product of two matrices of dimensions $M \times N$ and $N \times P$ is $\mathcal{O}(MNP)$, the complexity of this algorithm is dominated by the sum of the complexities of lines 1, 4 and 10, given by: $\mathcal{O}(PM_{R_K}M_S[min\{M_{R_K}M_S, P\} + M_D N])$, $\mathcal{O}(M_S N M_D M_{R_K} min\{N, M_D M_{R_K}\})$ and $\mathcal{O}(M_D N P M_{R_K} M_S[1 + min\{M_D N P, M_{R_K} M_S\}])$, respectively. The total complexity is then given by:

$$\mathcal{O}(M_D N M_{R_K}^2 M_S[M_D + P M_S]). \quad (40)$$

If $PM_S > M_D$, then the complexity becomes $\mathcal{O}(M_D N M_{R_K}^2 M_S^2 P)$.

B. PARATUCK-ALS Algorithm

In this subsection we present the second proposed algorithm, based on the ALS method. By stacking $\mathbf{Y}^{(D)^T}_{..p}$ in (12) for $p = 1, ..., P$, we get the mode-2 unfolded matrix given by:

$$\mathbf{Y}_2^{(D)} = \begin{bmatrix} (\mathbf{Y}_{..1}^{(D)})^T \\ \vdots \\ (\mathbf{Y}_{..P}^{(D)})^T \end{bmatrix} \tag{41}$$

$$= \mathbf{B}^{(D)}(\mathbf{H}^{(R_K D)})^T \in \mathbb{C}^{NP \times M_D}, \tag{42}$$

where

$$\mathbf{B}^{(D)} = \begin{bmatrix} \mathbf{S}D_1(\mathbf{G}_0)\mathbf{H}^{(G)^T} D_1(\mathbf{G}_K) \\ \vdots \\ \mathbf{S}D_P(\mathbf{G}_0)\mathbf{H}^{(G)^T} D_P(\mathbf{G}_K) \end{bmatrix} \in \mathbb{C}^{NP \times M_{R_K}}. \tag{43}$$

By stacking $\mathbf{Y}_{..p}^{(D)}$ in (12) for $p = 1, ..., P$, we get the mode-3 unfolded matrix:

$$\mathbf{Y}_3^{(D)} = \begin{bmatrix} \mathbf{Y}_{..1}^{(D)} \\ \vdots \\ \mathbf{Y}_{..P}^{(D)} \end{bmatrix} = \begin{bmatrix} \mathbf{D}_{..1}^{(D)} \\ \vdots \\ \mathbf{D}_{..P}^{(D)} \end{bmatrix} \mathbf{S}^T \tag{44}$$

$$= \mathbf{D}^{(D)}\mathbf{S}^T \in \mathbb{C}^{M_D P \times N}, \tag{45}$$

where

$$\mathbf{D}_{..p}^{(D)} = \mathbf{H}^{(R_K D)} D_p(\mathbf{G}_K)\mathbf{H}^{(G)} D_p(\mathbf{G}_0) \in \mathbb{C}^{M_D \times M_S}. \tag{46}$$

We can then alternately estimate $\mathbf{H}^{(R_K D)}$, $\mathbf{H}^{(G)}$ and $\mathbf{S}$ from (42), (35) and (45), respectively, by means of the ALS method. The PARATUCK-ALS algorithm is summarized in Algorithm 2, where $\hat{\mathbf{A}}_i$ corresponds to the estimate of the matrix $\mathbf{A}$ at the i^{th} iteration, $\tilde{\mathbf{Y}}_2^{(D)}$, $\tilde{\mathbf{Y}}_3^{(D)}$ and vec($\tilde{\mathbf{Y}}_1^{(D)}$) are the noisy versions of $\mathbf{Y}_2^{(D)}$, $\mathbf{Y}_3^{(D)}$ and vec($\mathbf{Y}_1^{(D)}$), respectively. At each iteration, we estimate one of the matrices using the estimates of the other matrices obtained in the previous steps. The iterations will run until the following convergence criterion is satisfied:

$$|\epsilon_i - \epsilon_{i-1}| \leq \delta, \tag{47}$$

where δ is the convergence threshold and

$$\epsilon_i = \frac{||\tilde{\mathbf{Y}}_3^{(D)} - \hat{\mathbf{D}}_i^{(D)}\mathbf{S}_i^T||_F^2}{||\tilde{\mathbf{Y}}_3^{(D)}||_F^2}. \tag{48}$$

At convergence, as it will be shown in the next section, we have the estimation of the symbols and channels matrices with ambiguities, such that:

$$\hat{\mathbf{H}}^{(R_K D)}\mathbf{\Delta}_{R_K D} = \mathbf{H}^{(R_K D)}, \tag{49}$$

$$\hat{\mathbf{S}}\mathbf{\Delta}_S = \mathbf{S}, \tag{50}$$

and

$$\mathbf{\Delta}_{R_K D}^{-1}\hat{\mathbf{H}}^{(G)}\mathbf{\Delta}_S^{-1} = \mathbf{H}^{(G)}, \tag{51}$$

where $\mathbf{\Delta}_{R_K D} \in \mathbb{C}^{M_{R_K} \times M_{R_K}}$ and $\mathbf{\Delta}_S \in \mathbb{C}^{M_S \times M_S}$ are the ambiguity matrices. Assuming that the first row of $\mathbf{S}$ is known by the receiver (through the use of pilot symbols, as mentioned before), $\mathbf{\Delta}_S$ can be estimated and canceled out from $\hat{\mathbf{S}}$ and $\hat{\mathbf{H}}^{(G)}$, leading to $\hat{\mathbf{H}}^{(R_K D)}\mathbf{\Delta}_{R_K D} = \mathbf{H}^{(R_K D)}$, $\hat{\mathbf{S}} = \mathbf{S}$ and $\mathbf{\Delta}_{R_K D}^{-1}\hat{\mathbf{H}}^{(G)} = \mathbf{H}^{(G)}$. The PARATUCK-ALS algorithm is summarized in Algorithm 2.

Algorithm 2 - (PARATUCK-ALS)

1: $i = 0$;
2: $\hat{\mathbf{S}} \leftarrow$ 16-QAM modulated values;
3: $\hat{\mathbf{H}}^{(R_K D)} \leftarrow$ gaussian random values;
4: $\epsilon_1 = 0$;
5: **repeat**
6: $i = i + 1$;
7: $\hat{\mathbf{B}}_i^{(D)} = \begin{bmatrix} \hat{\mathbf{S}}_{i-1}D_1(\mathbf{G}_0)\hat{\mathbf{H}}_{i-1}^{(G)^T} D_1(\mathbf{G}_K) \\ \vdots \\ \hat{\mathbf{S}}_{i-1}D_P(\mathbf{G}_0)\hat{\mathbf{H}}_{i-1}^{(G)^T} D_P(\mathbf{G}_K) \end{bmatrix}$;
8: $\hat{\mathbf{H}}_i^{(R_K D)} = \left(\hat{\mathbf{B}}_i^{(D)^\dagger}\tilde{\mathbf{Y}}_2^{(D)}\right)^T$;
9: $\hat{\mathbf{C}}_i^{(D)} = \mathbf{G}_{0K}^T \diamond (\hat{\mathbf{S}}_{i-1} \otimes \hat{\mathbf{H}}_i^{(R_K D)})$;
10: $\hat{\mathbf{h}}_i^{(G)} = \left(\mathbf{C}^{(D)}\right)^\dagger \text{vec}(\tilde{\mathbf{Y}}_1^{(D)})$;
11: $\hat{\mathbf{H}}_i^{(G)} = \text{unvec}(\hat{\mathbf{h}}_i^{(G)})$;
12: $\hat{\mathbf{D}}_i^{(D)} = \begin{bmatrix} \hat{\mathbf{H}}_i^{(R_K D)} D_1(\mathbf{G}_K)\hat{\mathbf{H}}_i^{(G)} D_1(\mathbf{G}_0) \\ \vdots \\ \hat{\mathbf{H}}_i^{(R_K D)} D_P(\mathbf{G}_K)\hat{\mathbf{H}}_i^{(G)} D_P(\mathbf{G}_0) \end{bmatrix}$;
13: $\hat{\mathbf{S}}_i = \left(\hat{\mathbf{D}}_i^{(D)^\dagger}\tilde{\mathbf{Y}}_3^{(D)}\right)^T$;
14: $\epsilon_i = \frac{||\tilde{\mathbf{Y}}_3^{(D)} - \hat{\mathbf{D}}_i^{(D)}\mathbf{S}_i^T||_F^2}{||\tilde{\mathbf{Y}}_3^{(D)}||_F^2}$;
15: **until** $|\epsilon_i - \epsilon_{i-1}| \leq \delta$
16: **Remove ambiguities**

The complexity per iteration of this algorithm is dominated by the complexity of line 10, which is given by:

$$\mathcal{O}(NPM_D M_S M_{R_K}[1 + 2M_S M_{R_K}]). \tag{52}$$

V. Expected CRB for the Estimated Channels

The CRB is a lower bound for the MSE, or normalized MSE (NMSE), of any unbiased estimator $\hat{\boldsymbol{\theta}}$ of the parameter vector of interest $\boldsymbol{\theta}$, such that:

$$\mathbb{E}||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}||^2 \geq \text{Trace}\{\text{CRB}(\boldsymbol{\theta})\}, \tag{53}$$

where $\text{CRB}(\boldsymbol{\theta})$ is the CRB matrix defined as the inverse of the Fisher Information Matrix (FIM) denoted by $\mathbf{F}(\boldsymbol{\theta})$.

An extension of the above CRB for complex-valued random parameters structured as $\boldsymbol{\theta}_c = [\bar{\boldsymbol{\theta}}^T \tilde{\boldsymbol{\theta}}^T]^T$, where $\bar{\boldsymbol{\theta}} = \Re\{\boldsymbol{\theta}_c\}$ and $\tilde{\boldsymbol{\theta}} = \Im\{\boldsymbol{\theta}_c\}$, with respect to random nuisance parameter $\boldsymbol{\gamma}$ is given by

$$\mathbb{E}||\boldsymbol{\theta}_c - \hat{\boldsymbol{\theta}}_c||^2 \geq \mathbb{E}_{\bar{\boldsymbol{\theta}},\tilde{\boldsymbol{\theta}},\boldsymbol{\gamma}}\Big\{\text{Trace}\{\text{CRB}(\bar{\boldsymbol{\theta}})\} + \text{Trace}\{\text{CRB}(\tilde{\boldsymbol{\theta}})\}\Big\}. \tag{54}$$

We also recall that if the observation follows a complex circular Gaussian distribution such as $\mathbf{y} \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{R})$ then the FIM relatively to a real vector $\boldsymbol{\theta}$ is given by the Slepian-Bangs (SB) formula [21], [22]:

$$[\mathbf{F}(\boldsymbol{\theta})]_{k,j} = 2\Re\left\{\left(\frac{\partial\boldsymbol{\mu}}{\partial[\boldsymbol{\theta}]_k}\right)^H \boldsymbol{R}^{-1}\frac{\partial\boldsymbol{\mu}}{\partial[\boldsymbol{\theta}]_j}\right\} \quad (55)$$

$$+ \operatorname{Trace}\left\{\frac{\partial\boldsymbol{R}}{\partial[\boldsymbol{\theta}]_k}\boldsymbol{R}^{-1}\frac{\partial\boldsymbol{R}}{\partial[\boldsymbol{\theta}]_j}\boldsymbol{R}^{-1}\right\}. \quad (56)$$

In our context, two remarks can be formulated:

1) The symbol matrix $\mathbf{S}$ is discrete-valued. This structure violates the regularity conditions of the CRB. The symbol matrix is then viewed as a random unknown nuisance parameters. This strategy is for instance exploited in [20].
2) The vectorized entries of the channels are be stacked in $\boldsymbol{\theta}_c$ and $\boldsymbol{\gamma} = \mathrm{vec}(\mathbf{S})$. The derivation of the above CRB is mathematically intractable, so we promote the following two alternative strategies.

A. Lower Bound for the Channel Between Source and Relay

The channel from R_1 to D is viewed as a random unknown nuisance parameter and the CRB is derived for the channel from S to R_1. In this case, we have

$$\boldsymbol{\theta}_c = [(\bar{\mathbf{h}}^{(SR_1)})^T \ (\tilde{\mathbf{h}}^{(SR_1)})^T]^T, \quad (57)$$

$$\boldsymbol{\gamma} = [(\bar{\mathbf{h}}^{(R_1D)})^T \ (\tilde{\mathbf{h}}^{(R_1D)})^T \ \mathrm{vec}(\mathbf{S})^T]^T. \quad (58)$$

where $\bar{\mathbf{h}}^{(SR_1)} = \mathrm{vec}(\bar{\mathbf{H}}^{(SR_1)})$, $\tilde{\mathbf{h}}^{(SR_1)} = \mathrm{vec}(\tilde{\mathbf{H}}^{(SR_1)})$, $\bar{\mathbf{h}}^{(R_1D)} = \mathrm{vec}(\bar{\mathbf{H}}^{(R_1D)})$ and $\tilde{\mathbf{h}}^{(R_1D)} = \mathrm{vec}(\tilde{\mathbf{H}}^{(R_1D)})$.

For the two-hop case, we define the observation $\tilde{\mathbf{y}}_1^{(D)}$ given by equation (35) corrupted by the following additive noise:

$$\mathbf{b}_1^{(D)} = \begin{bmatrix} (\mathbf{I}_N \otimes [\mathbf{H}^{(R_1D)}D_1(\mathbf{G}_1)])\mathbf{v}_{1,1}^{(R_1)} \\ \vdots \\ (\mathbf{I}_N \otimes [\mathbf{H}^{(R_1D)}D_P(\mathbf{G}_1)])\mathbf{v}_{P,1}^{(R_1)} \end{bmatrix} + \mathbf{v}_1^{(D)}, \quad (59)$$

where $\mathbf{v}_{p,1}^{(R_1)} = \mathrm{vec}(\mathbf{V}_{..p}^{(R_1)}) \in \mathbb{C}^{M_D N\times 1}$ and $\mathbf{v}_1^{(D)} = \mathrm{vec}([\mathrm{vec}(\mathbf{V}_{..1}^{(D)})...\mathrm{vec}(\mathbf{V}_{..P}^{(D)})]) \in \mathbb{C}^{M_D NP\times 1}$.

By (59) conditioned to the observations of the matrices $\mathbf{H}^{(SR_1)}$, $\mathbf{H}^{(R_1D)}$ and $\mathbf{S}$, the statistics of the noise observation is given by:

$$\tilde{\mathbf{y}}_1^{(D)} \sim \mathcal{CN}(\boldsymbol{\mu}_1, \boldsymbol{R}_1), \quad (60)$$

where $\tilde{\mathbf{y}}_1^{(D)} = \mathrm{vec}(\tilde{\mathbf{Y}}_1^{(D)})$ and

$$\boldsymbol{\mu}_1 = \mathbf{C}^{(D)}\mathbf{h}^{(SR_1)}, \quad (61)$$

$$\boldsymbol{R}_1 = \sigma_{R_1}^2 \mathrm{Bdiag}\{\mathbf{I}_N\otimes\boldsymbol{\Lambda}_1, \dots, \mathbf{I}_N\otimes\boldsymbol{\Lambda}_P\} + \sigma_D^2\mathbf{I}_{NPM_D}, \quad (62)$$

with

$$\boldsymbol{\Lambda}_p = \mathbf{H}^{(R_1D)}|D_p(\mathbf{G}_1)|^2(\mathbf{H}^{(R_1D)})^H. \quad (63)$$

Using the SB formula where the second term vanishes, the $(2M_S M_{R_1}) \times (2M_S M_{R_1})$ FIM is given by

$$\mathbf{F}_1(\boldsymbol{\theta}_c) = 2\begin{bmatrix} \bar{\mathbf{M}}_1 & -\tilde{\mathbf{M}}_1 \\ \tilde{\mathbf{M}}_1^T & \bar{\mathbf{M}}_1 \end{bmatrix} \quad (64)$$

where $\bar{\mathbf{M}}_1 = \Re\{(\mathbf{C}^{(D)})^H\boldsymbol{R}_1^{-1}\mathbf{C}^{(D)}\}$ and $\tilde{\mathbf{M}}_1 = \Im\{(\mathbf{C}^{(D)})^H\boldsymbol{R}_1^{-1}\mathbf{C}^{(D)}\}$. The CRB is obtained as the inverse of the FIM. We obtain

$$\mathrm{Trace}\{\mathrm{CRB}(\bar{\mathbf{h}}^{(SR_1)})\} = \frac{1}{2}\mathrm{Trace}\{(\bar{\mathbf{M}}_1 + \tilde{\mathbf{M}}_1\bar{\mathbf{M}}_1^{-1}\tilde{\mathbf{M}}_1)^{-1}\}, \quad (65)$$

$$\mathrm{Trace}\{\mathrm{CRB}(\tilde{\mathbf{h}}^{(SR_1)})\} = \frac{1}{2}\mathrm{Trace}\{\bar{\mathbf{M}}_1^{-1} - \bar{\mathbf{M}}_1^{-1}\tilde{\mathbf{M}}_1(\bar{\mathbf{M}}_1 + \tilde{\mathbf{M}}_1\bar{\mathbf{M}}_1^{-1}\tilde{\mathbf{M}}_1)^{-1}\tilde{\mathbf{M}}_1\bar{\mathbf{M}}_1^{-1}\}. \quad (66)$$

Finally, summing the two above expressions and taking the mathematical expectation with respect to the channel and the symbol matrices, we obtain the lower bound.

B. Lower Bound for the Channel Between Relay and Destination

The channel from S to R_1 is viewed as a random unknown nuisance parameter and the CRB is derived for the channel from R_1 to D. In this case, we get:

$$\boldsymbol{\theta}_c = [(\bar{\mathbf{h}}^{(R_1D)})^T \ (\tilde{\mathbf{h}}^{(R_1D)})^T]^T, \quad (67)$$

$$\boldsymbol{\gamma} = [(\bar{\mathbf{h}}^{(SR_1)})^T \ (\tilde{\mathbf{h}}^{(SR_1)})^T \ \mathrm{vec}(\mathbf{S})^T]^T. \quad (68)$$

Consider the following noisy observation:

$$\tilde{\mathbf{y}}_2^{(D)} = \mathrm{vec}(\tilde{\mathbf{Y}}_2^{(D)}) = (\mathbf{I}_{M_D}\otimes\mathbf{B}^{(D)})\mathbf{h}^{(R_1D)} + \mathbf{b}_2^{(D)}, \quad (69)$$

where the p-th block of matrix $\mathbf{B}^{(D)}$ is given by $\mathbf{S}D_p(\mathbf{G}_0)(\mathbf{H}^{(SR_1)})^T D_p(\mathbf{G}_1)$ and the additive noise is given by

$$\mathbf{b}_2^{(D)} = \begin{bmatrix} ([\mathbf{H}^{(R_1D)}D_1(\mathbf{G}_1)]\otimes\mathbf{I}_N)\mathbf{v}_{1,2}^{(R_1)} \\ \vdots \\ ([\mathbf{H}^{(R_1D)}D_P(\mathbf{G}_1)]\otimes\mathbf{I}_N)\mathbf{v}_{P,2}^{(R_1)} \end{bmatrix} + \mathbf{v}_2^{(D)}, \quad (70)$$

where $\mathbf{v}_{p,2}^{(R_1)} = \mathrm{vec}(\mathbf{V}_{...p}^{(R_1)})^T \in \mathbb{C}^{M_D N\times 1}$ and $\mathbf{v}_2^{(D)} = \mathrm{vec}([\mathrm{vec}(\mathbf{V}_{..1}^{(D)})^T...\mathrm{vec}(\mathbf{V}_{..P}^{(D)})^T]^T) \in \mathbb{C}^{NPM_D\times 1}$.

Conditioned to the observations of the matrices $\mathbf{H}^{(SR_1)}$, $\mathbf{H}^{(R_1D)}$ and $\mathbf{S}$, the statistics of the noise observation is given by:

$$\tilde{\mathbf{y}}_2^{(D)} \sim \mathcal{CN}(\boldsymbol{\mu}_2, \boldsymbol{R}_2), \quad (71)$$

where

$$\boldsymbol{\mu}_2 = (\mathbf{I}_{M_D}\otimes\mathbf{B}^{(D)})\mathbf{h}^{(R_1D)}, \quad (72)$$

$$\boldsymbol{R}_2 = \sigma_{R_1}^2\mathrm{Bdiag}\{\boldsymbol{\Lambda}_1\otimes\mathbf{I}_N, \dots, \boldsymbol{\Lambda}_P\otimes\mathbf{I}_N\} + \sigma_D^2\mathbf{I}_{NPM_D}. \quad (73)$$

As the covariance matrix $\boldsymbol{R}_2$ is dependent of the parameters of interest, the second term in the SB formula cannot vanish and has to be evaluated. We then have

$$\frac{\partial\boldsymbol{R}_2}{\partial[\boldsymbol{\theta}_c]_k} = \sigma_{R_1}^2\mathrm{Bdiag}\Big\{\frac{\partial\mathbf{H}^{(R_1D)}|D_1(\mathbf{G}_1)|^2(\mathbf{H}^{(R_1D)})^H}{\partial[\boldsymbol{\theta}_c]_k}\otimes\mathbf{I}_N, \dots, \frac{\partial\mathbf{H}^{(R_1D)}|D_P(\mathbf{G}_1)|^2(\mathbf{H}^{(R_1D)})^H}{\partial[\boldsymbol{\theta}_c]_k}\otimes\mathbf{I}_N\Big\} \quad (74)$$

where

$$\frac{\partial \mathbf{H}^{(R_1D)}|D_p(\mathbf{G}_1)|^2(\mathbf{H}^{(R_1D)})^H}{\partial[\bar{\mathbf{h}}^{(R_1D)}]_{k=f(i,j)}} = \mathbf{\Gamma}^{i,j}|D_p(\mathbf{G}_1)|^2(\mathbf{H}^{(R_1D)})^H + \mathbf{H}^{(R_1D)}|D_p(\mathbf{G}_1)|^2\mathbf{\Gamma}^{j,i}, \quad (75)$$

$$\frac{\partial \mathbf{H}^{(R_1D)}|D_p(\mathbf{G}_1)|^2(\mathbf{H}^{(R_1D)})^H}{\partial[\tilde{\mathbf{h}}^{(R_1D)}]_{k=f(i,j)}} = i\mathbf{\Gamma}^{i,j}|D_p(\mathbf{G}_1)|^2(\mathbf{H}^{(R_1D)})^H - i\mathbf{H}^{(R_1D)}|D_p(\mathbf{G}_1)|^2\mathbf{\Gamma}^{j,i}, \quad (76)$$

since

$$\frac{\partial \bar{\mathbf{H}}^{(R_1D)}}{\partial[\bar{\mathbf{h}}^{(R_1D)}]_{k=f(i,j)}} = \mathbf{\Gamma}^{i,j}, \; \frac{\partial \tilde{\mathbf{H}}^{(R_1D)}}{\partial[\tilde{\mathbf{h}}^{(R_1D)}]_{k'=f(i',j')}} = \mathbf{\Gamma}^{i',j'} \quad (77)$$

in which $\mathbf{\Gamma}^{i,j}$ is the null matrix except for its (i,j)-th entry, which is equals to one, and $f(.,.)$ is the vectorization function of a matrix. The above expressions allow the computation of the second term of the SB formula. The first term is derived in a similar way as (64) with $\mathbf{C}^{(D)} \to (\mathbf{I}_{M_D} \otimes \mathbf{B}^{(D)})$ and $\boldsymbol{R}_1 \to \boldsymbol{R}_2$. By summing the two terms, we obtain $\mathbf{F}_2(\boldsymbol{\theta}_c)$. Taking the inverse and using expression (54), we obtain the desired lower bound.

VI. Uniqueness Properties

A. Uniqueness Property

Assuming that the matrices $\mathbf{G}_1$, ..., $\mathbf{G}_{K-1}$ are identical, we have $\mathbf{H}^{(G)} = \mathbf{H}^{(G)}_{..p}$, for $p = 1, ..., P$. In this case, the PARATUCK-(K+1) model (12) leads to:

$$\mathbf{Y}^{(D)}_{..p} = \mathbf{H}^{(R_KD)}D_p(\mathbf{G}_K)\mathbf{H}^{(G)}D_p(\mathbf{G}_0)\mathbf{S}^T. \quad (78)$$

The uniqueness of the PARATUCK-2 tensor model was established in [19]. It was demonstrated that, if the following conditions are satisfied:

- $\mathbf{H}^{(R_KD)}$, $\mathbf{G}_K$, $\mathbf{H}^{(G)}$, $\mathbf{G}_0$ and $\mathbf{S}$ are full rank matrices;
- $\mathbf{H}^{(G)}$ has entries different from zero;
- $\mathbf{G}_K$ and $\mathbf{G}_0$ have the same number of columns, i.e. $M_S = M_{R_K}$;

then any alternate representation of the tensor $\mathcal{Y}$ of the form:

$$\mathbf{Y}^{(D)}_{..p} = \tilde{\mathbf{H}}^{(R_KD)}D_p(\tilde{\mathbf{G}}_K)\tilde{\mathbf{H}}^{(G)}D_p(\tilde{\mathbf{G}}_0)\tilde{\mathbf{S}}^T, \quad (79)$$

where $\tilde{\mathbf{H}}^{(R_KD)}$, $\tilde{\mathbf{G}}_K$, $\tilde{\mathbf{H}}^{(G)}$, $\tilde{\mathbf{G}}_0$ and $\tilde{\mathbf{S}}$ have the same dimensions of their counterparts in (78), satisfies the following relationships:

$$\tilde{\mathbf{H}}^{(R_KD)}(\mathbf{\Pi}_{R_KD}\mathbf{\Delta}_{R_KD}) = \mathbf{H}^{(R_KD)}, \quad (80)$$

$$\tilde{\mathbf{S}}(\mathbf{\Pi}_S\mathbf{\Delta}_S) = \mathbf{S}, \quad (81)$$

$$(\mathbf{\Delta}_{G_K}\mathbf{\Delta}^{-1}_{R_KD}\mathbf{\Pi}^T_{R_KD})\tilde{\mathbf{H}}^{(G)}(\mathbf{\Pi}_S\mathbf{\Delta}^{-1}_S\mathbf{\Delta}_{G_0}) = \mathbf{H}^{(G)}, \quad (82)$$

$$(z_p\mathbf{\Pi}^T_{R_KD})D_p(\tilde{\mathbf{G}}_K)(\mathbf{\Pi}_{R_KD}\mathbf{\Delta}^{-1}_{G_K}) = D_p(\mathbf{G}_K), \quad (83)$$

$$(z_p^{-1}\mathbf{\Pi}^T_S)D_p(\tilde{\mathbf{G}}_0)(\mathbf{\Pi}_S\mathbf{\Delta}^{-1}_{G_0}) = D_p(\mathbf{G}_0), \quad (84)$$

for $p = 1, ..., P$, where $\mathbf{\Delta}_{R_KD} \in \mathbb{C}^{M_{R_K}\times M_{R_K}}$, $\mathbf{\Delta}_S \in \mathbb{C}^{M_S\times M_S}$, $\mathbf{\Delta}_{G_K} \in \mathbb{C}^{M_{R_K}\times M_{R_K}}$ and $\mathbf{\Delta}_{G_0} \in \mathbb{C}^{M_S\times M_S}$ are diagonal matrices, $\mathbf{\Pi}_{R_KD} \in \mathbb{C}^{M_{R_K}\times M_{R_K}}$ and $\mathbf{\Pi}_S \in \mathbb{C}^{M_S\times M_S}$ are permutation matrices, and z_p are nonzero scalars.

Assuming that $\mathbf{G}_K$ and $\mathbf{G}_0$ are known at the receiver, we get $\mathbf{\Pi}_{R_KD} = \mathbf{I}_{M_{R_K}}$, $\mathbf{\Pi}_S = \mathbf{I}_{M_S}$ $z_p\mathbf{\Delta}^{-1}_{G_K} = \mathbf{I}_{M_{R_K}}$ and $z_p^{-1}\mathbf{\Delta}^{-1}_{G_0} = \mathbf{I}_{M_S}$, for $p = 1, ..., P$, where $\mathbf{I}_M$ is the identity matrix of dimension M. This leads to $z_1 = z_2 = ... = z_P = z$ and $\mathbf{\Delta}_{G_K} = z\mathbf{I}_{M_{R_K}}$ and $\mathbf{\Delta}_{G_0} = z^{-1}\mathbf{I}_{M_S}$.

(80), (81) and (82) can then be respectively rewritten as $\tilde{\mathbf{H}}^{(R_KD)}\mathbf{\Delta}_{R_KD} = \mathbf{H}^{(R_KD)}$, $\tilde{\mathbf{S}}\mathbf{\Delta}_S = \mathbf{S}$ and $\mathbf{\Delta}^{-1}_{R_KD}\tilde{\mathbf{H}}^{(G)}\mathbf{\Delta}^{-1}_S = \mathbf{H}^{(G)}$.

B. Alternative Uniqueness Condition

The following theorem establishes sufficient alternative uniqueness conditions for the PARATUCK-2 tensor model, based on the fact that $\mathbf{G}_K$, $\mathbf{G}_0$ and the first row of $\mathbf{S}$ are known.

Theorem: Assuming that $\mathbf{G}_K$, $\mathbf{G}_0$ and the first row of $\mathbf{S}$ are known, if $r(\mathbf{G}_{0K}) = M_{R_K}M_S$, where $r(\mathbf{A})$ denotes the rank of the matrix $\mathbf{A}$ and $\mathbf{G}_{0K} = \mathbf{G}_0^T \diamond \mathbf{G}_K^T \in \mathbb{C}^{M_{R_K}M_S\times P}$, the PARATUK-2 model (78) is unique up the following ambiguities: $\tilde{\mathbf{H}}^{(R_KD)} = \mathbf{H}^{(R_KD)}\mathbf{\Delta}_{R_KD}$ and $\tilde{\mathbf{H}^{(G)}} = \mathbf{\Delta}^{-1}_{R_KD}\mathbf{H}^{(G)}$, where $\mathbf{\Delta}_{R_KD} \in \mathbb{C}^{M_{R_K}\times M_{R_K}}$ is a diagonal matrix.

Proof: Let us consider an alternate representation of the tensor $\mathcal{Y}$ of the form:

$$\mathbf{Y}^{(D)}_1 = (\tilde{\mathbf{S}} \otimes \tilde{\mathbf{H}}^{(R_KD)})\text{diag}(\tilde{\mathbf{h}}^{(G)})\mathbf{G}_{0K}. \quad (85)$$

where $\tilde{\mathbf{S}}$, $\tilde{\mathbf{H}}^{(R_KD)}$ and $\tilde{\mathbf{h}}^{(G)}$ have the same dimensions of their counterparts in (24). If $\mathbf{G}_{0K}$ is full row rank, then:

$$(\mathbf{S} \otimes \mathbf{H}^{(R_KD)})\text{diag}(\mathbf{h}^{(G)}) = (\tilde{\mathbf{S}} \otimes \tilde{\mathbf{H}}^{(R_KD)})\text{diag}(\tilde{\mathbf{h}}^{(G)}). \quad (86)$$

We can deduce from (86):

$$\mathbf{H}^{(R_KD)}D_m\left(\mathbf{H}^{(G)^T}\right)s_{n,m} = \tilde{\mathbf{H}}^{(R_KD)}D_m\left(\tilde{\mathbf{H}^{(G)}}^T\right)\tilde{s}_{n,m}, \quad (87)$$

for $n = 1, ..., N$ and $m = 1, ..., M_S$. From (87), we can conclude that $\mathbf{H}^{(R_KD)}$ is unique up to column scaling ambiguities, i.e., $\tilde{\mathbf{H}}^{(R_KD)} = \mathbf{H}^{(R_KD)}\mathbf{\Delta}_{R_KD}$, with $\mathbf{\Delta}_{R_KD} \in \mathbb{C}^{M_{R_K}\times M_{R_K}}$ being a diagonal matrix.

Let us denote by $\mathbf{w}_{m_s,m_r} \in \mathbb{C}^{N\times 1}$ the vector containing the elements $1, (M_D+1), (2M_D+1), ..., [(N-1)M_D+1]$ of the $[(m_s-1)M_{R_K}+m_r]^{th}$ column of the matrix $(\mathbf{S} \otimes \mathbf{H}^{(R_KD)})\text{diag}(\mathbf{h}^{(G)})$, for $1 \le m_s \le M_S$ and $1 \le m_r \le M_{R_K}$. We can write from (87):

$$\mathbf{w}_{m_s,m_r} = \mathbf{s}_{m_s}h^{(R_KD)}_{1,m_r}h^{(G)}_{m_r,m_s} \quad (88)$$

where $\mathbf{s}_{m_s} \in \mathbb{C}^{N\times 1}$ is the m_s^{th} row of $\mathbf{S}$. We can then conclude from (86):

$$\mathbf{s}_{m_s}h^{(R_KD)}_{1,m_r}h^{(G)}_{m_r,m_s} = \tilde{\mathbf{s}}_{m_s}\tilde{h}^{(R_KD)}_{1,m_r}\tilde{h}^{(G)}_{m_r,m_s}, \quad (89)$$

which means that $\mathbf{S}$ is unique up to column scaling ambiguities, i.e., $\tilde{\mathbf{S}} = \mathbf{S}\mathbf{\Delta}_S$, with $\mathbf{\Delta}_S \in \mathbb{C}^{M_S\times M_S}$ being a diagonal matrix. Moreover, assuming that the first row of $\mathbf{S}$ is known, $\mathbf{\Delta}_S$ can be estimated and canceled out from $\tilde{\mathbf{S}}$ leading to $\tilde{\mathbf{S}} = \mathbf{S}$.

From (89), we can then write:

$$h_{1,m_r}^{(R_K D)} \mathbf{h}_{.m_r}^{(G)} = \tilde{h}_{1,m_r}^{(R_K D)} \tilde{\mathbf{h}}_{.m_r}^{(G)} \in \mathbb{C}^{1\times M_S}, \tag{90}$$

where $\mathbf{h}_{.m_r}^{(G)}$ is the m_r^{th} row of $\mathbf{H}^{(G)}$, for $1 \le m_r \le M_{R_K}$, which means that $\mathbf{H}^{(G)}$ is unique up to row scaling ambiguities, i.e., $\tilde{\mathbf{H}^{(G)}} = \mathbf{\Delta}_G \mathbf{H}^{(G)}$, with $\mathbf{\Delta}_G \in \mathbb{C}^{M_{R_K}\times M_{R_K}}$ being a diagonal matrix. In addition, it can deduced from (90) that $\mathbf{\Delta}_G \mathbf{\Delta}_{R_K D} = \mathbf{I}_{M_{R_K}}$.

□

Note that the above uniqueness conditions imply $P \ge M_{R_K} M_S$. It should be remarked that the uniqueness conditions of Subsections VI.A and VI.B are sufficient, but not necessary, which means that the conditions of these two subsections do not have to be satisfied simultaneously.

VII. Identifiability Conditions

A. LS-KR Algorithm

The identifiability conditions for the LS-KR algorithm are:

i. $r(\mathbf{G}_{0K}) = M_{R_K} M_S$,
ii. $r\left(\mathbf{C}^{(D)}\right) = M_{R_K} M_S$.

Condition i implies $M_{R_K} M_S \le P$. In this work, we consider that the coding matrices $\mathbf{G}_0 \in \mathbb{C}^{P\times M_S}$ and $\mathbf{G}_K \in \mathbb{C}^{P\times M_{R_K}}$ are such that $\mathbf{G}_{0K} = \mathbf{F}_P^{(M_{R_K} M_S)} \in \mathbb{C}^{M_{R_K} M_S \times P}$, where $\mathbf{F}_P^{(M_{R_K} M_S)}$ is a truncated discrete Fourier transform (DFT) matrix, i.e., $\mathbf{F}_P^{(M_{R_K} M_S)}$ is a matrix that contains the first $M_{R_K} M_S$ rows of a DFT matrix of order P, with $M_{R_K} M_S \le P$, which leads to $r(\mathbf{G}_{0K}) = M_{R_K} M_S$. The $(p+1)^{th}$ column of $\mathbf{F}_P^{(M_{R_K} M_S)}$ can be represented by the following Kronecker product:

$$\begin{bmatrix} 1 \\ \omega^p \\ \vdots \\ \omega^{(M_{R_K} M_S - 1)p} \end{bmatrix} = \begin{bmatrix} 1 \\ \omega^{pM_{R_K}} \\ \vdots \\ \omega^{(M_S-1)pM_{R_K}} \end{bmatrix} \otimes \begin{bmatrix} 1 \\ \omega^p \\ \vdots \\ \omega^{(M_{R_K}-1)p} \end{bmatrix}. \tag{91}$$

for $p = 0, ..., P-1$ and $\omega = e^{-\frac{2\pi j}{P}}$. We then set the coding matrices $\mathbf{G}_0$ and $\mathbf{G}_K$ to be equal to [25]:

$$\mathbf{G}_0 = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & \omega^{M_{R_K}} & \dots & \omega^{(P-1)M_{R_K}} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{(M_S-1)M_{R_K}} & \dots & \omega^{(M_S-1)(P-1)M_{R_K}} \end{bmatrix}^T. \tag{92}$$

and

$$\mathbf{G}_K = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & \omega & \dots & \omega^{(P-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{(M_{R_K}-1)} & \dots & \omega^{(M_{R_K}-1)(P-1)} \end{bmatrix}^T. \tag{93}$$

The matrix $\mathbf{G}_{0K}$ used in the proposed receivers is then orthogonal, avoiding noise amplification.

Regarding Condition ii, assuming that all the elements of $\mathbf{H}^{(R_K D)}$ are independent and drawn from a continuous distribution, this channel matrix have full rank. Moreover, if the modulation cardinality and the number N of symbols are not small, then the matrix $\mathbf{S}$ has full rank with a high probability. Using (36) and the following relationships: $r(\mathbf{A} \diamond \mathbf{B}) \ge \max(r(\mathbf{A}), r(\mathbf{B}))$ and $r(\mathbf{A} \otimes \mathbf{B}) = r(\mathbf{A})\, r(\mathbf{B})$, we get:

$$\begin{aligned} r\left(\mathbf{C}^{(D)}\right) &\ge \max\left(r\left(\mathbf{G}_{0K}^T\right), r(\mathbf{S})\, r\left(\mathbf{H}^{(R_K D)}\right)\right) \\ &\ge \max\left(M_{R_K} M_S, \min(N, M_S) \min(M_D, M_{R_K})\right) \\ &\ge M_{R_K} M_S. \end{aligned} \tag{94}$$

On the other hand, the rank of $\mathbf{C}^{(D)}$ should be smaller than or equal to its dimensions, i.e. $\min(M_D NP, M_{R_K} M_S) \ge r\left(\mathbf{C}^{(D)}\right)$. Thus, if $M_D NP \ge M_{R_K} M_S$, we get $r(\mathbf{C}^{(D)}) = M_{R_K} M_S$.

$$\mathbf{C}^{(D)} = \left[\mathbf{G}_{0K}^T \diamond \left(\mathbf{S} \otimes \mathbf{H}^{(R_K D)}\right)\right] \in \mathbb{C}^{M_D NP \times M_{R_K} M_S}. \tag{95}$$

B. PARATUCK-ALS Algorithm

The ALS algorithm requires the following identifiability conditions:

i. $r\left(\mathbf{B}^{(D)}\right) = M_{R_K}$,
ii. $r\left(\mathbf{C}^{(D)}\right) = M_{R_K} M_S$,
iii. $r\left(\mathbf{D}^{(D)}\right) = M_S$.

where $\mathbf{B}^{(D)}$, $\mathbf{C}^{(D)}$ and $\mathbf{D}^{(D)}$ are respectively defined in (43), (36) and (46). These conditions requires, respectively: $M_{R_K} \le NP$, $M_D NP \ge M_{R_K} M_S$ and $M_S \le M_D P$. Note that Condition ii is the same for the LS-KR algorithm.

Assuming that all the elements of the channel matrices $\mathbf{H}^{(R_K D)}$, $\mathbf{H}^{(SR_1)}$ and $\mathbf{H}^{(R_{k-1} R_k)}$, for $k = 2, ..., K$, are independent and drawn from a continuous distribution, all these channel matrices have full rank. Moreover, as above explained, we can consider that the matrix $\mathbf{S}$ has full rank with a high probability. Assuming that the matrices $\mathbf{G}_0, ..., \mathbf{G}_K$ have no zeros, we can then conclude:

$$\begin{aligned} r\left(\mathbf{S} D_p(\mathbf{G}_0) \mathbf{H}^{(SR_1)^T} ... \mathbf{H}^{(R_{K-1} R_K)^T} D_p(\mathbf{G}_K)\right) = \\ \min(M_S, M_{R_1}, ..., M_{R_K}, N), \end{aligned} \tag{96}$$

for $p = 1, ..., P$. Then, from (43), a sufficient but not necessary relationship for assuring the above stated Condition i is $M_{R_K} \le M_S, M_{R_1}, ..., M_{R_{K-1}} N$.

Finally, from (46), we can then conclude that:

$$r\left(\mathbf{D}_p^{(D)}\right) = \min(M_S, M_{R_1}, ..., M_{R_K}, M_D), \tag{97}$$

which means that a sufficient but not necessary relationship for assuring the Condition iii is $M_S \le M_{R_1}, ..., M_{R_K}, M_D$.

VIII. Simulations Results

In this section, simulation results that evaluate the performance of the proposed receivers by means of Monte Carlo simulations are presented. The criteria of performance used are the symbol error rate (SER), the block error rate (BLER) and the NMSE of the channels, given by:

$$NMSE = \frac{1}{S_{MC}} \left(\sum_{s=1}^{S_{MC}} \frac{||\mathbf{H}_s - \hat{\mathbf{H}}_s||_F^2}{||\mathbf{H}_s||_F^2} \right), \quad (98)$$

where S_{MC} is the number of Monte Carlo samples, $\mathbf{H}_s$ is channel matrix ($\mathbf{H}^{(R_K D)}$ or $\mathbf{H}^{(G)}$) and $\hat{\mathbf{H}}_s$ is the estimate of $\mathbf{H}_s$ at the s^{th} Monte Carlo run. The SER and NMSE are shown as a function of the SNR in dB at each communication link, that is given by:

$$SNR = 10 \log_{10} \left(\frac{P_t}{\sigma_V^2 (K+1)} \right), \quad (99)$$

where P_t is the total transmission power and σ_V^2 is the noise variance. The gain matrix $\mathbf{G}_k$, for $1 \leq k \leq K-1$ has equal rows, with the elements of the row being independent and identically distributed (i.i.d) and drawn form a zero-mean complex Gaussian distribution. When not stated otherwise it is used 16-QAM, $P = 8, N = 100, M_S = M_{R_k} = M_D = 2$, for $k = 1, ..., K$. The wireless channels undergo frequency-flat Rayleigh fading, with the elements of the channel matrices being i.i.d with variance following a large scale exponential path-loss model with exponent equal to 4. The source, relays and destination nodes are placed in a straight line, with the relays being equally spaced between the source and destination. The total transmission power is equal to 1 and it is equally divided by the source and the relays. The convergence threshold δ of the PARATUCK-ALS algorithm is 10^{-6}.

The proposed receivers were compared to the iterative receiver proposed in [7], also based on the ALS algorithm. The proposed receivers were also compared to a PARATUCK2 Zero Forcing (PARATUCK2-ZF) receiver that estimates the $\mathbf{S}$ matrix assuming that all the channels are known, using the following estimator:

$$\hat{\mathbf{S}}_{ZF} = [\mathbf{D}^{(D)^\dagger} \tilde{\mathbf{Y}}_3^{(D)}]^T, \quad (100)$$

Figure 3 shows the SER versus SNR, comparing the proposed receivers with the receiver proposed in [7], using its original coding matrix and the DFT coding matrix, and with the PARATUCK2-ZF receiver. In this simulation it was considered a single relay ($K = 1$) and a 4-QAM modulation, as the receiver of [7] only works for two-hop systems. We can see that both proposed receivers have a similar performance. One should expect the PARATUCK-ALS to provide a small SER due to the fact the estimations are refined at each iteration for this algorithm. However, the DFT-based coding used leads to an orthogonal matrix $\mathbf{G}_{0K}$, which considerably enhances the estimations of the LS-KR method, since it does not change the noise norm. Moreover, as expected, the PARATUCK2-ZF method outperformed all the other techniques. However, the proposed methods provided SERs close to the ZF receiver. Note also that the proposed techniques provided much better SERs than the receiver of [7], using both types of coding, which stands out the effectiveness of the proposed receivers. We should point out that, contrarily to the proposed LS-KR method, the receiver of [7] does not exploit the orthogonality of the matrix $\mathbf{G}_{0K}$.

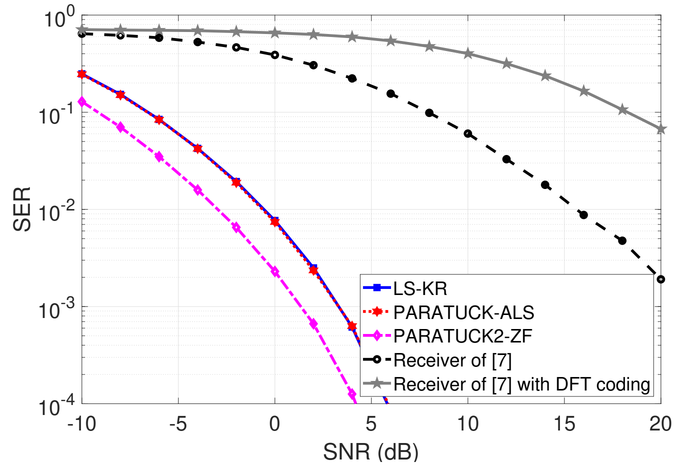

Fig. 3. SER *versus* SNR for the LS-KR, PARATUCK-ALS, PARATUCK2-ZF receivers and the receiver of [7] using its original Vandermonde matrix coding and the DFT coding used in the proposed receivers.

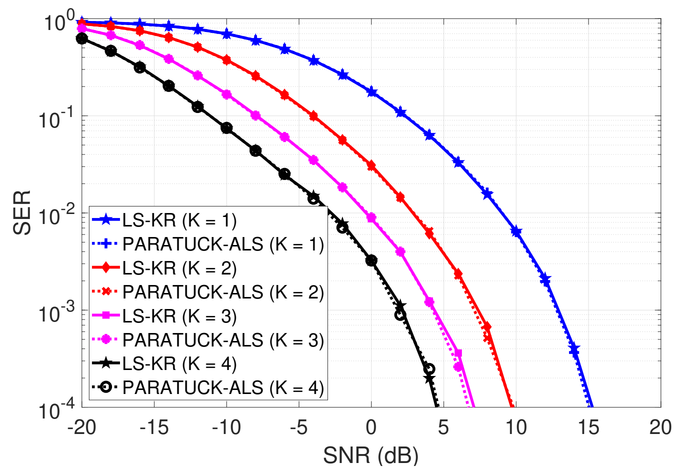

Fig. 4. SER *versus* SNR for the LS-KR and PARATUCK-ALS receivers, for several values of K.

Figure 4 shows the SER versus SNR, varying the number of relays, for the two proposed receivers. Again, we can see that both receivers have a similar performance. It can also be viewed that, as the number of relays increases, the receivers provide a smaller SER. This is due to the fact that, when more relays are used to assist the communication, we will have shorter distances between the nodes, leading to less severe path-losses. This result stands out the advantage of multi-hop systems.

Figure 5 shows the SER versus SNR, varying the number of antennas at the destination, for the two proposed receivers, using $K = 4$. As expected, the values of SER decrease as the number of antennas is increased. This is due to the spatial diversity at the destination node. It can also be seen that both proposed receivers present similar performance.

Figure 6 shows the BLER and SER versus SNR, varying the number of antennas at the source, for $K = 3$. As

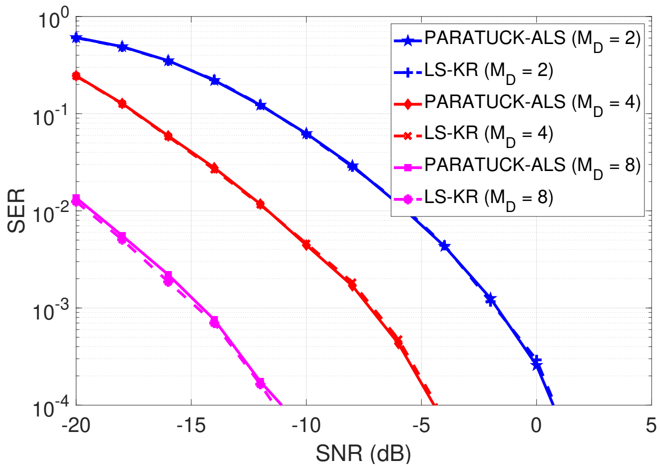

Fig. 5. SER *versus* SNR varying the number of antennas at the destination for the LS-KR and PARATUCK-ALS receivers.

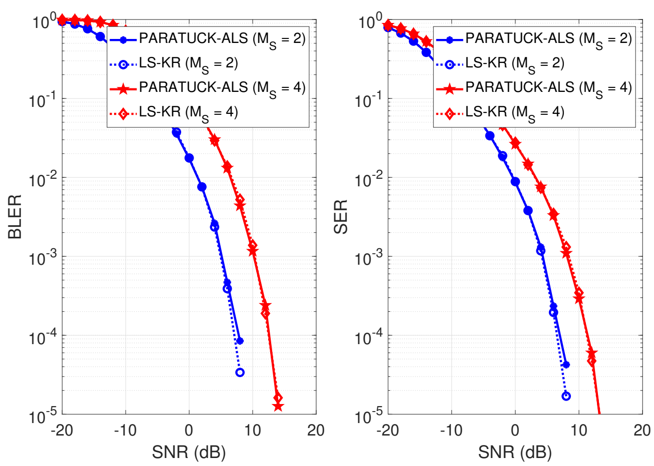

Fig. 6. BLER and SER *versus* SNR for the LS-KR and PARATUCK-ALS varying the number of antennas at the source, for $K = 3$.

expected, when the number of antennas at the source increases, maintaining fixed the number of antennas in the other nodes, the SER and BLER tend to reach higher values, as more information is being transmitted without increasing the number of receive antennas. It can also be observed that BLER and SER present similar behaviors.

Figure 7 compares the NMSE of the channels $\mathbf{H}^{(R_K D)}$ and $\mathbf{H}^{(G)}$ versus SNR, for the two proposed receivers, using $K = 3$. We can see that the estimation of $\mathbf{H}^{(R_K D)}$ is better than the one of $\mathbf{H}^{(G)}$, for both receivers. This is due to the fact that we estimate and cancel the ambiguity matrix of $\mathbf{H}^{(R_K D)}$ assuming that its first row is known, while the ambiguity matrix of $\mathbf{H}^{(G)}$ is estimated from the ambiguity matrix of $\mathbf{H}^{(R_K D)}$, without assuming any knowledge of $\mathbf{H}^{(G)}$. Moreover, the PARATUCK-ALS provides better estimations of $\mathbf{H}^{(R_K D)}$ than the LS-KR receiver, as the PARATUCK-ALS is an iterative algorithm that refines the estimations at each iteration. Regarding the estimation of $\mathbf{H}^{(G)}$, the PARATUCK-

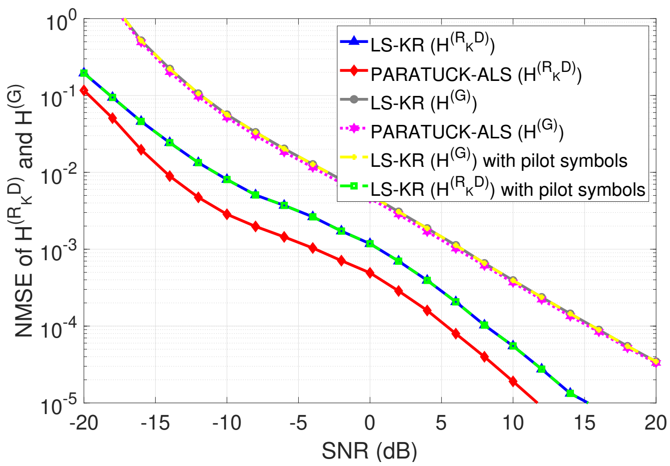

Fig. 7. NMSE *versus* SNR for $\mathbf{H}^{(R_K D)}$ and $\mathbf{H}^{(G)}$ for the LS-KR and PARATUCK-ALS receivers, for $K = 3$.

ALS provides a small performance gain with respect to the LS-KR receiver. This is due to the fact that the ambiguity matrix of this channel is estimated in an indirect way, decreasing the performance of both algorithms.

Figure 7 also shows the NMSE of $\mathbf{H}^{(R_K D)}$ and $\mathbf{H}^{(G)}$ provided by the LS-KR receiver considering that the ambiguities are removed using pilot symbols sent by the last relay to the destination. We assumed that, at each symbol period, one pilot symbol is sent by one of the M_{R_K} antennas of the last relay, while the other antennas remain silent. The training period occurs during M_{R_K} symbol periods, without coding. These pilot symbols are used to estimate the ambiguity matrix $\mathbf{\Delta}_{R_K D}$. It can be viewed that the NMSE of both $\mathbf{H}^{(R_K D)}$ and $\mathbf{H}^{(G)}$ using pilot symbols is quite close to the NMSE obtained assuming the knowledge of the first row of $\mathbf{H}^{(R_K D)}$. This results shows that a priori knowledge of the first row of $\mathbf{H}^{(R_K D)}$ is deduced from a short training step inducing no performance loss.

Figures 8 and 9 show the NMSE of $\mathbf{H}^{(G)}$ and $\mathbf{H}^{(R_K D)}$, respectively, versus SNR, by varying the number of relays, for the two proposed receivers. As in Figure 4, we can see that the provided NMSEs are better when the number of relays increases, due to the less severe path-loss. Moreover, the PARATUCK-ALS always provides better estimations of $\mathbf{H}^{(R_K D)}$ than the LS-KR receiver, due to the reason previously explained for Figure 7. For the $\mathbf{H}^{(G)}$ channel, the proposed techniques provide roughly the same performance, as also explained for Figure 7.

In order to compare the channel estimates of the proposed techniques with the one of [7], we use the NMSE of the effective channel $\mathbf{H}^{(eff)} = \mathbf{H}^{(R_K D)}\mathbf{H}^{(G)}$ as the figure of merit, as the technique of [7] does not estimate the channel matrices individually, but only the product $\mathbf{H}^{(R_1 D)}\mathbf{H}^{(G)}$. Moreover, as the receiver of [7] only works for two-hop systems, the parameters of this simulation are the same as in the simulation of Figure 3. Note that, for $K = 1$, $\mathbf{H}^{(G)} = \mathbf{H}^{(R_1 D)}$. Figure 10 compares NMSE of $\mathbf{H}^{(eff)}$ obtained with the proposed receivers and

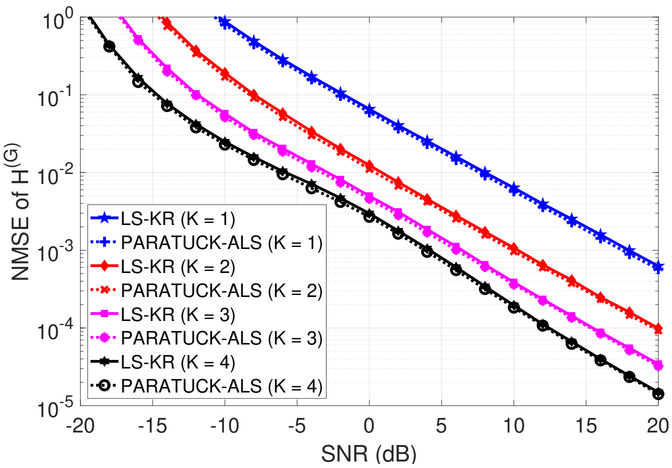

Fig. 8. NMSE of $\mathbf{H}^{(G)}$ *versus* SNR for the LS-KR and PARATUCK-ALS receivers, for several values of K.

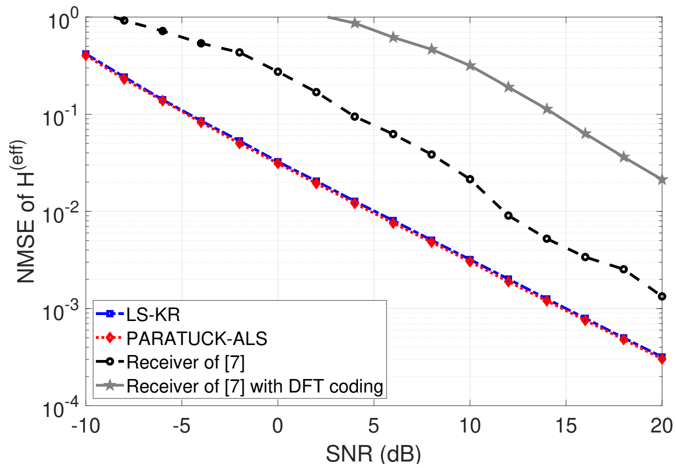

Fig. 10. NMSE *versus* SNR for the effective channel $\mathbf{H}^{(eff)}$ for the LS-KR, PARATUCK-ALS receivers and the receiver of [7] using its original Vandermonde matrix coding and the DFT coding used in the proposed receivers.

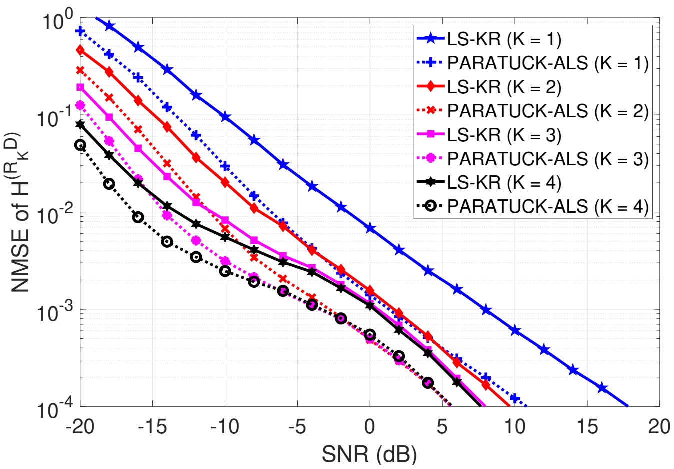

Fig. 9. NMSE of $\mathbf{H}^{(R_K D)}$ *versus* SNR for the LS-KR and PARATUCK-ALS receivers, for several values of K.

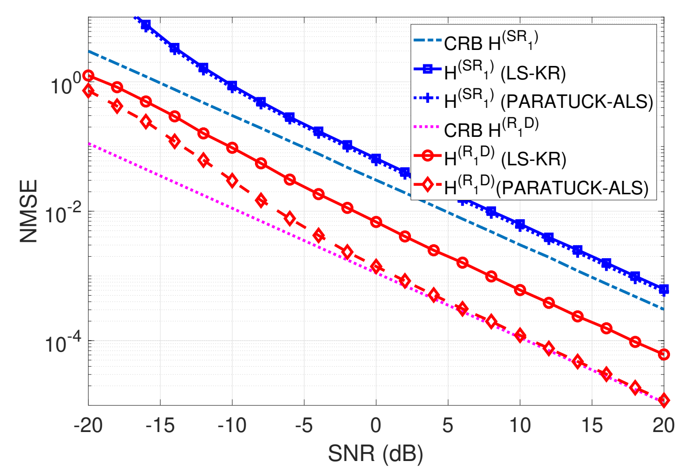

Fig. 11. CRBs for the channels $\mathbf{H}^{(SR_1)}$ and $\mathbf{H}^{(R_1 D)}$ and their respective NMSEs generated by LS-KR and PARATUCK-ALS, for $K = 1$.

with the one of [7]. It can be seen that both proposed receivers provided much better NMSEs than the method of [7] using both coding matrices. As previously explained, this is due to the fact that both of the proposed receivers exploit efficiently the orthogonal matrix $\mathbf{G}_{0K}$, contrarily to the receiver of [7].

Figure 11 shows the CRBs of the channels for the PARATUCK-2 model, i.e., when only a single relay is used, as well as the NMSE of these channels provided by the two proposed receivers. It can be observed that, for channel $\mathbf{H}^{(SR_1)}$, the NMSE provided by both receivers is very close to the lower bound, as the ambiguities of this channel estimate are eliminated in function of the ambiguity matrices of $\mathbf{H}^{(R_1 D)}$ and $\mathbf{S}$. For the channel $\mathbf{H}^{(R_1 D)}$, the NMSE provided by the LS-KR receiver is very close to the lower bound, showing that this receiver is close to optimality, while the one provided by the PARATUCK-ALS receiver touches the bound, reaching the optimal case.

IX. Conclusion

The main contribution of this work is the proposal of two semi-blind receivers for one-way multi-hop cooperative MIMO AF relaying systems. The considered communication system uses the KRST coding at the source and at the relays. It was shown that the received signals in this transmission scheme satisfy a PARATUCK-N tensor model. One of the proposed receivers is non-iterative, based on a rearrangement of the Kronecker product, while the other is iterative, based on the ALS algorithm. Uniqueness conditions were derived, as well as the identifiabiliy conditions of the proposed algorithms. Also, the expressions of the expected CRB for the two-hop case were derived and simulated.

Simulation results have shown a better performance of the proposed receivers in the considered scenarios, when compared to the receiver of [7]. Simulation results also showed that both proposed receivers are close to the optimal case, the PARATUCK2-ZF receiver. The results corroborated the fact that increasing the number of hops improves the SER and NMSE performances, highlighting the advantage of multi-hop networks.

In future works, we aim to develop more efficient algorithms and generalize the proposed ones for other system models (e.g., two-way relaying, OFDM, GFDM, etc.). The use of the channel state information (CSI) provided by the proposed techniques in space-time coding setting is also a perspective, as well as the development of the CRB for the general case of K relays. Finally, a deeper analysis of the performance of the proposed receivers using forward error correction (FEC) coding will be considered in future work.

Acknowledgements

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (PQ/Process no 304053/2016-3 and Universal 01/2016) and Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP) (PRONEM 01/2016). Pedro Marinho R. de Oliveira was funded by a Master scholarship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). He is now funded by a PhD scholarship from the IT Doctoral School (ED STIC) of the Université Côte d'Azur.

References

[1] K. J. R. Liu, A. K. Sadek, W. Su, and A. Kwasinski, *Cooperative Communications and Networking*. Cambridge University Press, 2009.

[2] S. Plevel, S. Tomazic, T. Javornik, and G. Kandus, "MIMO: Wireless Communications", Encyclopedia of Wireless and Mobile Communications, DOI: 10.1081/E-EWMC-120043484. 2008.

[3] N. D. Sidiropoulos, R. Bro and G. B. Giannakis, "Parallel Factor Analysis in Sensor Array Processing", *IEEE Transactions on Signal Processing*, vol. 48, no. 8, pp. 2377-2388, Aug. 2000.

[4] A. L. F. de Almeida, G. Favier and J. C. M. Mota, "PARAFAC-Based Unified Tensor Modeling of Wireless Communication Systems with Application to Blind Multiuser Equalization", *Signal Processing - Special Issue on Tensor Signal Processing*, vol. 87, no. 2, pp. 337-351, Feb. 2007.

[5] C. A. R. Fernandes, A. L. F. de Almeida and D. B. da Costa, "Unified Tensor Modeling for Blind Receivers in Multiuser Uplink Cooperative Systems", *IEEE Signal Processing Letters*, vol. 19, no. 5, pp. 247-250, 2012.

[6] A. L. F. de Almeida, C. A. R. Fernandes and D. B. da Costa, "Multiuser Detection for Uplink DS-CDMA Amplify-and-Forward Relaying Systems", *IEEE Signal Processing Letters*, vol. 20, no. 7, pp. 697-700, Jul. 2013.

[7] L. R. Ximenes, G. Favier and A. L. F. de Almeida, "PARAFAC-PARATUCK Semi-Blind Receivers for Two-Hop Cooperative MIMO Relay Systems", *IEEE Transactions on Signal Processing*, vol. 62, no. 14, pp. 3604-3615, Jul. 2014.

[8] A. A. T. Peixoto and P. M. R. de Oliveira, "Tensor-Based Receiver With Adaptive Relay Activation for Two-Hop Cooperative Systems", *Proc. SBrT-2018, XXXVI Simpósio Brasileiro de Telecomunicações e Processamento de Sinais*, Campina Grande, Brazil, Sep. 16-19, 2018, pp. 804-808.

[9] L. R. Ximenes, G. Favier and A. L. F. de Almeida, "Semi-Blind Receivers for Non-Regenerative Cooperative MIMO Communications Based on Nested PARAFAC Modeling", *IEEE Transactions on Signal Processing*, vol. 63, no. 18, pp. 4985-4998, Sep. 2015.

[10] G. Favier, C. A. R. Fernandes and A. L.F. de Almeida, "Nested Tucker Tensor Decomposition with Application to MIMO Relay Systems Using Tensor Space-time Coding (TSTC)", *Signal Processing*, no. 128, pp. 318-331, Nov. 2016.

[11] P. M. R. de Oliveira and C. A. R. Fernandes, "PARATUCK-3 Semi-Blind Receivers for Three-Hop Cooperative MIMO Relay Systems", in *Proc. SBrT-2018, XXXV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais*, São Pedro, Brazil, Sep. 3-6, 2017, pp. 32-36.

[12] J. Du1, C. Yuan, P. Tian1 and H. Lin, "Channel Estimation for Multi-Input Multi-Output Relay Systems Using the PARATUCK2 Tensor Model", *IET Communications*, vol. 10, iss. 9, pp. 995-1002, 2016.

[13] S. Yang and J.-C. Belfiore, "Diversity of MIMO Multihop Relay Channels", submitted to *IEEE Transactions on Information Theory*, [Online]. Available: http://arxiv.org/abs/0708.0386v1.

[14] Y. Rong and Y. Hua, "Optimality of Diagonalization of Multi-Hop MIMO Relays", *IEEE Transactions on Wireless Communications*, vol. 8, pp. 6068-6077, Dec. 2009.

[15] M. R. A. Khandaker and Y. Rong, "Joint Source and Relay Optimization for Multiuser MIMO Relay Communication Systems", in *Proc. ICSPCS-2010, 4th International Conference on Signal Processing and Communication Systems*, Gold Coast, Australia, Dec. 13-15, 2010.

[16] N. D. Sidiropoulos and R. S. Budampati, "Khatri-Rao Space-Time Codes", *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2396-2407, Oct. 2002.

[17] A. L. F. de Almeida, G. Favier and L. R. Ximenes, "Space-Time-Frequency (STF) MIMO Communication Systems With Blind Receiver Based on a Generalized PARATUCK2 Model", *IEEE Transactions on Signal Processing*, vol. 61, no. 8, pp. 1895-1909, Apr. 2013.

[18] A. L. F. de Almeida, G. Favier and J. C. M. Mota, "Space-Time Spreading-Multiplexing for MIMO Wireless Communication Systems Using the PARATUCK-2 Tensor Model", *Elsevier Signal Processing*, vol. 89, no. 11, pp. 2103-2116, 2009.

[19] R. A. Harshman and M. E. Lundy, "Uniqueness Proof for a Family of Models Sharing Features of Tucker's Three-Mode Factor Analysis and PARAFAC/CANDECOMP", *Psychometrika*, vol. 61, no. 1, pp. 133-154, Mar. 1996.

[20] R. Boyer and P. Larzabal, "Sparsity-Based Estimation Bounds With Corrupted Measurements", *Elsevier Signal Processing*, vol. 143, pp. 86-93, 2018.

[21] H. L. Van Trees, K. L. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*, Wiley, 2007.

[22] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*, Pearson/Prentice Hall Upper Saddle River, NJ, 2005.

[23] Z. Ben-Haim, Y. Eldar, "A Lower Bound on the Bayesian MSE Based on the Optimal Bias Function", *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 5179-5196, 2009.

[24] C. Van Loan and N. Pitsianis, "Approximation with Kronecker Products", *Linear Algebra for Large Scale and Real-Time Applications*, pp. 293-314. Cornell University. 1993.

[25] B. Sokal, "Semi-Blind Receivers for Multi-Relaying MIMO Systems Using Rank-One Tensor Factorizations", *MSc Dissertation*, Universidade Federal do Ceará. 2017.