

# Tensor-Based Channel Estimation Approach for One-Way Multi-Hop Relaying Communications

Shuangzhi Li<sup>1</sup>, Xiaomin Mu<sup>1</sup>, Xin Guo<sup>1</sup>, Jing Yang<sup>1,3</sup>, and Jiankang Zhang<sup>1,2</sup>

<sup>1</sup>School of Information Engineering, Zhengzhou University, Zhengzhou, China  
[e-mail :ieszli@gs.zzu.edu.cn, iexmmu@zzu.edu.cn, guoxin19880806@163.com, yangjing@haut.edu.cn, iejkzhang@zzu.edu.cn]

<sup>2</sup> National Mobile Communications Research Laboratory, Southeast University, Nanjing, China

<sup>3</sup>College of Information Science and Engineering, Henan University of Technology, Zhengzhou, China

\*Corresponding author: Jiankang Zhang

*Received August 3, 2015; revised September 16, 2015; accepted October 7, 2015;  
published December 31, 2015*

---

## Abstract

Multi-hop relaying communications have great potentials in improving transmission performance by deploying relay nodes. The benefit is critically dependent on the accuracy of the channel state information (CSI) of all the transmitting links. However, the CSI has to be estimated. In this paper, we investigate the channel estimation problem in one-way multi-hop MIMO amplify-and-forward (AF) relay system, where both the two-hop and three-hop communication link exist. Traditional point-to-point MIMO channel estimation methods will result in error propagation in estimating relay links, and separately tackling the channel estimation issue of each link will lose the gain as part of channel matrices involved in multiple communication links. In order to exploit all the available gains, we develop a novel channel estimation model by structuring different communication links using the PARAFAC and PARATUCK2 tensor analysis. Furthermore, a two-stage fitting algorithm is derived to estimate all the channel matrices involved in the communication process. In particular, essential uniqueness is further discussed. Simulation results demonstrate the advantage and effectiveness of the proposed channel estimator.

---

**Keywords:** multi-hop relaying communications, channel estimation, tensor, PARAFAC, PARATUCK2

---

This work is supported by the National Natural Science Foundation of China under grants 61271421, 61301150 and 61571401, the Specialized Research Fund for the Doctoral Program of Higher Education (SRFDP) under grant 20134101120001, the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University under grant 2016D02, the Excellent Young Teachers Development Fund of Zhengzhou University under grant 1421318085 as well as the Postdoctoral Science Foundation of Henan Province under grant 2013003.

## 1. Introduction

Multi-hop relaying communications have gained a lot of attention recently because of both its theoretical and practical importance in cooperative communication system [1]. As a fundamental technique to enable device-to-device communication, multi-hop relaying system is expected to realize high spectrum efficiency for next generation communication systems e.g., 5G wireless systems [2, 3]. Combined with the multiple-input and multiple-output (MIMO) techniques, the spectral efficiency of this promising technique can be further improved.

Referring to the strategies performed by the relay, relaying communications can be classified into various categories: amplify-and-forward (AF) relaying, decode-and-forward (DF) relaying, compressed-and-forward (CF) relaying and so on. In this work, we consider a multi-hop AF relaying system due to its simplicity and small system delay without additional decoding steps existed in DF relay system.

As its special cases, extensive methods have been proposed to analyze the performance of dual-hop MIMO AF relay system, including the optimal source and relay beamforming matrix [4, 5], optimal relay amplification matrix [6], energy efficient [7] on the assumption that the perfect channel state information (CSI) of all links have been acquired at the destination node. However, in practical communication system, the CSI has to be estimated by using different channel estimation methods. In [8, 9], an optimal transceiver design for multi-hop AF MIMO system under imperfect channel knowledge is studied, and it reveals that the achievable gain depends largely on the estimation accuracy of channel links in multi-hop system. And in [10], the impact of channel estimation error on outage loss of the multiple antennas multiple relaying network is investigated, which indicates that the effect of the channel estimation error has a great influence on the system, especially at low-to-medium SNR. Unlike traditional point-to-point MIMO systems, the destination node in cooperative system has to estimate the CSI of all hops due to the limited computation ability in AF relays.

In [11], least square (LS) based algorithm is proposed to estimate channel matrices of all links of one-way two-hop MIMO AF cooperative communication system. However, the channel estimation errors accumulate across the consecutive stages. For AF relay networks with single-antenna source, relay, and destination nodes, a modified transmitting frame is proposed in [12] for OFDM modulated relay system, which enables the relay node to have the ability of saving the pilot data as an ample to estimate the channel between the relay to the destination node, which in turn serves as known information to estimate the source-to-relay channel in a decision-directed (DD) way. However, this known channel information in this proposed DD algorithm is also the error-existing channel estimation value. A disintegrated channel estimation method has been presented in [13], which implies that the relay node has to be equipped with a channel estimator. However, it is not realistic for a low complexity non-regenerative relay node.

Exploiting the estimation of relay to the destination node and/or relay to relay node is available, the channel estimation issue in cooperative system can be transformed to the traditional MIMO channel estimation problems. However, the inherent estimation error in the estimation of channel matrices will propagate when estimating the following links, which will finally deteriorate the whole performance. By taking advantage of their unique characteristics in digging the hidden structure in mixed samples, a few tensor-based channel estimation

algorithms have shown to be efficient approaches for channel estimation and/or symbol detection in cooperative systems.

In [14, 15], a channel estimation algorithm with the aid of parallel factor (PARAFAC) model is investigated in one-way two-hop MIMO relay system, which allows the estimation of related channel matrices to be computed in a parallel way by resorting to tensor modeling. In addition, a unified PARAFAC received signal model for cooperative MIMO system is built based on which a blind receiver is proposed for jointly estimating the channel matrices and detecting the symbols [16]. Note that the works in [16] is on the condition that the same cluster has the same spatial signature, which is strict for the communication environment. The authors in [17] investigate the channel estimation algorithm for a two-way MIMO relay system. Since the algorithm in [17] exploits the channel reciprocity in a two-way relay system, its application in one-way MIMO relay systems is not straightforward. By introducing a space-time matrix at the source node, a PARAFAC-PARATUCK2 [18] tensor model is built for two-hop cooperative MIMO relay systems in [19].

The common feature of the aforementioned works focus on the two-hop cooperative system, it may not be adapted to scenario that more relays are employed to further increase the coverage of communication. In [20], the author proposes a trilinear coding structure of the MIMO AF system, which is capable of exploiting cooperative diversity by tensor-based signal processing. In [21], the author establishes a PARATUCK2-based framework for three-hop cooperative communication system. However, only considering the link of one-way three-hop link will lose diversity gain in real communication scenario. In this paper, a more general tensor-based channel estimator of this system is considered, which provides the destination node with full knowledge of all channel matrices involved in the system. The main contributions of this paper can be summarized as follows:

- A more general one-way multi-hop MIMO AF relay system is considered in this paper, where the path loss factor is taken into account in this paper to analyze the achievable diversity gain of additional one-way two-hop links existing in the one-way three-hop link system.
- By resorting to the multi-way analysis tool, a unified tensor model is established and a combined PARAFAC/PARATUCK2 alternative least square (ALS) fitting algorithm is proposed to obtain the CSI of all involved links.
- With the initial estimation of channel matrices, a linear minimum mean square error (LMMSE) approach is further derived to enhance the channel estimator performance.

The rest of this paper is organized as follows. In Section 2, we give a brief overview of the system model. The detailed proposed channel estimation algorithm is presented in Section 3. In Section 4, simulation results are given to verify the effectiveness of the proposed algorithm. Conclusions are given in Section 5.

*Notations and Properties:* Scalars, column vectors, matrices, and tensors are denoted by lower-case ( $a, b, \dots$ ), boldface lower-case ( $\mathbf{a}, \mathbf{b}, \dots$ ), boldface capital ( $\mathbf{A}, \mathbf{B}, \dots$ ), and calligraphic ( $\mathcal{A}, \mathcal{B}, \dots$ ) letters, respectively.  $\mathcal{A}_{::,k} \in \mathbb{C}^{I \times J}$  is the so-called front slice obtained by fixing the third dimension of a third-order tensor  $\mathcal{A} \in \mathbb{C}^{I \times J \times K}$ .  $\mathbf{A}^T$ ,  $\mathbf{A}^H$ ,  $\mathbf{A}^{-1}$ ,  $\mathbf{A}^\dagger$ ,  $\mathbf{A}_l$ , and  $\mathbf{A}_{-m}$  represent matrix transposition, Hermitian transposition, matrix inverse, the Moore-Penrose pseudo inverse, the  $l$ -th row and the  $m$ -th column of  $\mathbf{A} \in \mathbb{C}^{L \times M}$ , respectively. The operator  $\text{vec}(\cdot)$  forms a vector by stacking the columns of its matrix argument, while the  $\text{unvec}(\cdot)$  operator is the inverse process of  $\text{vec}(\cdot)$  to reshape a vector into a matrix. The operator  $\text{diag}(\cdot)$  forms a diagonal matrix from its vector argument, and  $\text{bd}[\mathbf{A}_1, \dots, \mathbf{A}_K]$  forms a

block-diagonal matrix with  $K$  matrix blocks.  $tr(\mathbf{A})$  and  $\|\mathbf{A}\|_F$  is the trace and the Frobenius norm of  $\mathbf{A}$ , and  $\mathbf{D}_i(\mathbf{A}) = \text{diag}(\text{vec}(\mathbf{A}_i))$  corresponds to the diagonal matrix with the  $i$ th row of  $\mathbf{A}$  forming its diagonal. The Kronecker and the Khatri-Rao (column-wise Kronecker) matrix products are denoted by  $\otimes$  and  $\diamond$ , respectively. For  $\mathbf{A} = [\mathbf{A}_1, \dots, \mathbf{A}_R] \in \mathbb{C}^{I \times R}$  and  $\mathbf{B} = [\mathbf{B}_1, \dots, \mathbf{B}_R] \in \mathbb{C}^{J \times R}$ , we have

$$\mathbf{A} \diamond \mathbf{B} = [\mathbf{A}_1 \otimes \mathbf{B}_1, \dots, \mathbf{A}_R \otimes \mathbf{B}_R] \quad (1)$$

We also use two following properties in this work:

$$\text{vec}\{\mathbf{ABC}\} = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}\{\mathbf{B}\} \quad (2)$$

$$\text{vec}\{\mathbf{A} \mathcal{D}_n(\mathbf{B}) \mathbf{C}^T\} = (\mathbf{C} \diamond \mathbf{A}) \mathbf{B}_n^T \quad (3)$$

## 2. System Model

We consider the one-way multi-hop MIMO AF relay system as illustrated in Fig. 1 where the direct link between the source node and the destination node is assumed to be negligible and hence ignored due to a large path loss, and the source node transmits information to the destination node with the aid of  $T$  MIMO AF relays, which should be chosen according to practical communication scenario, such as the path loss factor, the size of the cellular, etc.. Moreover, the relays operate in a half-duplex mode (i.e., each relay node does not receive and transmit signals simultaneously).

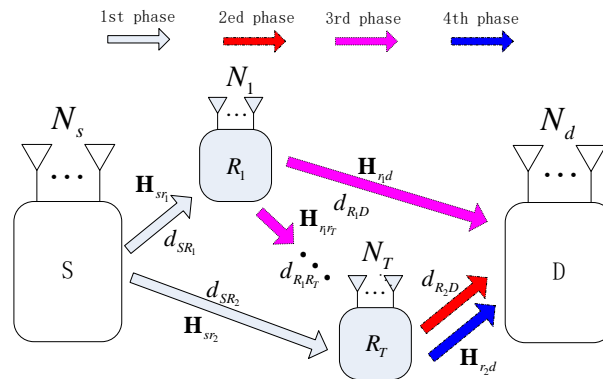


Fig. 1. One-way multi-hop MIMO AF relay system

We denote that the source node and the destination node are equipped with  $N_s \geq 2$  and  $N_d \geq 2$  antennas, respectively, while the  $i$ -th relay has  $N_i$  antennas,  $i = 1, 2, \dots, T$ .

In the following, for simplicity, we mainly consider the one-way three-hop MIMO AF relay system to present our algorithm, i.e.  $T = 2$ . But the generalization to the multi-hop MIMO relay system is just in a similar way. For example, we can firstly regard the cascade channel from the first relay node to the  $T$ -th relay node as a whole effective channel model  $\mathbf{H}_{r_1r_T} = \mathbf{H}_{r_1r_2} \mathbf{G}_2 \mathbf{H}_{r_2r_3} \cdots \mathbf{G}_{T-1} \mathbf{H}_{r_{T-1}r_T}$ , which can be modeled as various tensor model according to physical circumstance, like the methodology of using nested tensor model in [22, 23]. Through the proposed algorithm,  $\mathbf{H}_{r_1r_T}$  can be estimated in the first place; then, by means of corresponding fitting methods, the channel statement information contained in  $\mathbf{H}_{r_1r_T}$  can be further extracted.

Firstly, in order to derive the proposed channel estimators, we recast the formulation of the system model by resorting to tensor analysis. Let us define the overall training period as  $K$  time blocks, during which we assume the channels of all hops are quasi-static and each frame consists four phases. In  $k$ -th time block, the source node  $S$  transmits an orthogonal channel training sequences  $\mathbf{S}_0 \in \mathbb{C}^{N_s \times L_0}$  ( $L_0 \geq N_s$ ) with  $\mathbf{S}_0 \mathbf{S}_0^H = \mathbf{I}_{N_s}$  to the first relay node  $R_1$  and the second relay node  $R_2$  in the first phase. Then the received signal at the second relay node will be amplified with amplification matrix  $\mathbf{G}_2^{(k)}$  and forwarded to the destination node  $D$  in the second phase and  $R_1$  keeps silent. In the third phase,  $R_1$  will amplify its received signal with amplification matrix  $\mathbf{G}_1^{(k)}$  and transmit it to both  $R_2$  and  $D$ . During the fourth phase,  $R_2$  will transmit the data received in the previous phase to  $D$  with amplification matrix  $\mathbf{G}_2^{(k)}$  once again.

To sum up, there are five channel matrices in total to be estimated in the proposed system. Let us define the channel from source to  $R_1$  and  $R_2$  as  $\mathbf{H}_{sr_1} \in \mathbb{C}^{N_1 \times N_s}$  and  $\mathbf{H}_{sr_2} \in \mathbb{C}^{N_2 \times N_s}$ , respectively. From  $R_1$  to  $R_2$  and  $D$  are defined as  $\mathbf{H}_{r_1r_2} \in \mathbb{C}^{N_2 \times N_1}$  and  $\mathbf{H}_{r_1d} \in \mathbb{C}^{N_d \times N_1}$ , and from  $R_2$  to  $D$  is defined as  $\mathbf{H}_{r_2d} \in \mathbb{C}^{N_d \times N_2}$ . We assume that  $\mathbf{H}_{sr_1}$ ,  $\mathbf{H}_{r_1r_2}$ ,  $\mathbf{H}_{r_2d}$ ,  $\mathbf{H}_{sr_2}$  and  $\mathbf{H}_{r_1d}$  have complex Gaussian entries with zero means and variances of  $1/N_s$ ,  $1/N_1$ ,  $1/N_2$ ,  $1/(\eta_1^3 N_s)$  and  $1/(\eta_2^3 N_1)$ , respectively. The variances are set to normalize the effect of the number of transmit antennas to the received signal-to-noise (SNR) ratio [14]. And the  $S$ - $R_1$  link,  $R_1$ - $R_2$  link and  $R_2$ - $D$  link are assumed to have equal distance, i.e.  $d_{SR_1} = d_{R_1R_2} = d_{R_2D} = d_0$ . Assuming  $R_1$  and  $R_2$  lie between the source node and the destination node, and  $R_1$  is closer to  $S$  than  $R_2$ , while  $R_2$  is closer to  $D$ , and the distance of  $S$ - $R_2$  link and the  $R_1$ - $D$  link are denoted by  $d_{SR_2}$  and  $d_{R_1D}$ , respectively. Then the normalized parameters  $\eta_1 = d_{SR_2}/d_0$  and  $\eta_2 = d_{R_1D}/d_0$  satisfy  $1 \leq \eta_1 \leq 2$  and  $1 \leq \eta_2 \leq 2$ . With a path loss factor of 3, the above channel matrices can be constructed.

Thus the received signal at the destination node in phase 2 (i.e.  $S$ - $R_2$ - $D$  link), 3 (i.e.  $S$ - $R_1$ - $D$  link) and 4 (i.e.  $S$ - $R_1$ - $R_2$ - $D$  link) of  $k$ -th time block can be given by:

$$\mathbf{Y}_{d,k}(t+1) = \mathbf{H}_{r_2d} \mathbf{G}_2^{(k)} \mathbf{H}_{sr_2} \mathbf{S}_0 + \mathbf{H}_{r_2d} \mathbf{G}_2^{(k)} \mathbf{V}_{r_2}^{(k)}(t) + \mathbf{V}_d^{(k)}(t+1) \in \mathbb{C}^{N_d \times L_0}, \quad (4)$$

$$k = 1, \dots, K$$

$$\mathbf{Y}_{d,k}(t+2) = \mathbf{H}_{r_1d} \mathbf{G}_1^{(k)} \mathbf{H}_{sr_1} \mathbf{S}_0 + \mathbf{H}_{r_1d} \mathbf{G}_1^{(k)} \mathbf{V}_{r_1}^{(k)}(t) + \mathbf{V}_d^{(k)}(t+2) \in \mathbb{C}^{N_d \times L_0}, \quad (5)$$

$$k = 1, \dots, K$$

$$\mathbf{Y}_{d,k}(t+3) = \mathbf{H}_{r_2d} \mathbf{G}_2^{(k)} \mathbf{H}_{r_1r_2} \mathbf{G}_1^{(k)} \mathbf{H}_{sr_1} \mathbf{S}_0 + \mathbf{H}_{r_2d} \mathbf{G}_2^{(k)} \mathbf{H}_{r_1r_2} \mathbf{G}_1^{(k)} \mathbf{V}_{r_1}^{(k)}(t) \quad (6)$$

$$+ \mathbf{H}_{r_2d} \mathbf{G}_2^{(k)} \mathbf{V}_{r_2}^{(k)}(t+2) + \mathbf{V}_d^{(k)}(t+3) \in \mathbb{C}^{N_d \times L_0}, k = 1, \dots, K$$

where  $\mathbf{V}_{r_1}^{(k)}(t)$ ,  $\mathbf{V}_{r_2}^{(k)}(t)$  and  $\mathbf{V}_d^{(k)}(t)$  are noise matrices at  $R_1$ ,  $R_2$  and  $D$  corresponding to the  $k$ -th time block  $t$ -th transmission phase.

### 3. Proposed Tensor-based Channel Estimation Algorithm

In order to extract the channel matrices involved in the received signal model (4), (5) and (6), we resort to the tensor unfolding algorithm, which is useful when we want to isolate a matrix

in a tensor model. In the following, we will present the core equations leading to the development of the proposed tensor-based channel estimator.

### 3.1 Model of Two-Hop Link

In this section, we first establish the tensor model of one-way two-hop (i.e.  $S$ - $R_1$ - $D$  and  $S$ - $R_2$ - $D$ ) communication link based on PARAFAC model [24]. For simplicity, the amplification matrix design pattern will be similar to the pattern in [21], in which the rows of matrices  $\mathbf{E} \in \mathbb{C}^{K \times N_1}$  and  $\mathbf{F} \in \mathbb{C}^{K \times N_2}$  contain AF coefficients of  $R_1$  and  $R_2$  in different time blocks. More specifically, the amplification matrix at  $k$ -th time block at  $R_1$  and  $R_2$  are  $\mathbf{G}_1^{(k)} = \mathcal{D}_k(\mathbf{E})$  and  $\mathbf{G}_2^{(k)} = \mathcal{D}_k(\mathbf{F})$ , respectively. Note that, since our work is focused on channel estimation issues in multi-hop AF relaying system, the AF matrices  $\mathbf{G}_1$  and  $\mathbf{G}_2$  cannot be optimized at the channel estimation phases. Thus, in order to be fair for the comparison, we followed the setting as the references [14, 15, 19, 21, 22] selected as benchmark. In addition, in the scenario where multiple single antenna relay nodes operate in a distributed manner [23], it is reasonable to assume the relay amplification matrix  $\mathbf{G}$  is diagonal. During the normal communication period, however, the relay amplification matrix does not need to be diagonal. For this work, the use of non-diagonal AF matrices in the proposed approach is left for future work. Note also that, once the channels are estimated, the design of full AF matrices can be done.

Then, at the destination node  $D$ , by multiplying both sides of (4), (5) with  $\mathbf{S}_0^H$ , respectively, yields the  $k$ -th frontal slice of the third-order tensor:

$$\tilde{\mathbf{M}}_k = \mathbf{M}_k + \mathbf{V}_k(1) \quad (7)$$

$$\tilde{\mathbf{T}}_k = \mathbf{T}_k + \mathbf{V}_k(2) \quad (8)$$

where  $\mathbf{M}_k = \mathbf{H}_{r_2d} \mathcal{D}_k(\mathbf{F}) \mathbf{H}_{sr_2}$  and  $\mathbf{T}_k = \mathbf{H}_{r_1d} \mathcal{D}_k(\mathbf{E}) \mathbf{H}_{sr_1}$  are the matrix-of-interest.  $\tilde{\mathbf{M}}_k$  and  $\tilde{\mathbf{T}}_k$  are the noise version of  $\mathbf{M}_k$  and  $\mathbf{T}_k$ , respectively, while  $\mathbf{V}_k(1)$ ,  $\mathbf{V}_k(2)$  are the effective noise matrices at the destination node corresponding to the second phase and third phase of the  $k$ -th training block, which can be denoted as follows:

$$\mathbf{V}_k(1) = \mathbf{H}_{r_2d} \mathcal{D}_k(\mathbf{F}) \mathbf{V}_{r_2}^{(k)}(t) \mathbf{S}_0^H + \mathbf{V}_d^{(k)}(t+1) \mathbf{S}_0^H \quad (9)$$

$$\mathbf{V}_k(2) = \mathbf{H}_{r_1d} \mathcal{D}_k(\mathbf{E}) \mathbf{V}_{r_1}^{(k)}(t) \mathbf{S}_0^H + \mathbf{V}_d^{(k)}(t+2) \mathbf{S}_0^H \quad (10)$$

We take the case of  $S$ - $R_2$ - $D$  link for illustrating the tensor-based channel estimation procedure; nevertheless, the extension to the  $S$ - $R_1$ - $D$  link is rather straightforward. Considering the  $k$ -th time-block, the received signals at destination via the  $S$ - $R_2$ - $D$  link are stored in the  $k$ -th frontal slice  $\mathbf{M}_k$  of the third-order received signal tensor  $\mathcal{M} = [\mathbf{M}_1 \cup_3 \mathbf{M}_2 \cup_3 \dots \mathbf{M}_K] \in \mathbb{C}^{N_d \times N_s \times K}$ , the operator  $\cup_r$  represents the concatenation of two matrices along the  $r$ -th mode. Each entry of  $\mathcal{M}$  is given by

$$\mathcal{M}(i, j, k) = \sum_{m=1}^{N_2} \mathbf{H}_{r_2d}(i, m) \mathbf{F}(k, m) \mathbf{H}_{sr_2}(m, j) \quad (11)$$

Note that (11) represents the typical PARAFAC tensor model [24], with  $\mathbf{H}_{r_2d}$ ,  $\mathbf{F}$  and  $\mathbf{H}_{sr_2}$  as matrix factors. In order to isolate the channel matrices involved in the  $S$ - $R_2$ - $D$  link,  $\mathbf{H}_{sr_2}$  and  $\mathbf{H}_{r_2d}$ , there are two unfolding of tensor  $\mathcal{M}$  by stacking frontal slice  $\mathbf{M}_k$  ( $k = 1, \dots, K$ ) and its transpose  $\mathbf{M}_k^T$  on top of each other as follows,

$$[\mathcal{M}]_{(1)} = \begin{bmatrix} \mathbf{M}_1 \\ \vdots \\ \mathbf{M}_K \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{r_2d} \mathcal{D}_1(\mathbf{F}) \\ \vdots \\ \mathbf{H}_{r_2d} \mathcal{D}_K(\mathbf{F}) \end{bmatrix} \mathbf{H}_{sr_2} = (\mathbf{F} \diamond \mathbf{H}_{r_2d}) \mathbf{H}_{sr_2} \quad (12)$$

$$[\mathcal{M}]_{(2)} = \begin{bmatrix} \mathbf{M}_1^T \\ \vdots \\ \mathbf{M}_K^T \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{sr_2}^T \mathcal{D}_1(\mathbf{F}) \\ \vdots \\ \mathbf{H}_{sr_2}^T \mathcal{D}_K(\mathbf{F}) \end{bmatrix} \mathbf{H}_{r_2d}^T = (\mathbf{F} \diamond \mathbf{H}_{sr_2}^T) \mathbf{H}_{r_2d}^T \quad (13)$$

And, we can have the noise version of  $[\mathcal{M}]_{(1)}$  and  $[\mathcal{M}]_{(2)}$ ,

$$[\tilde{\mathcal{M}}]_{(1)} = \begin{bmatrix} \mathbf{M}_1 \\ \vdots \\ \mathbf{M}_K \end{bmatrix} + \begin{bmatrix} \mathbf{V}_1(1) \\ \vdots \\ \mathbf{V}_K(1) \end{bmatrix} = [\mathcal{M}]_{(1)} + \mathbf{V}^{(1)}(1) \quad (14)$$

$$[\tilde{\mathcal{M}}]_{(2)} = \begin{bmatrix} \mathbf{M}_1^T \\ \vdots \\ \mathbf{M}_K^T \end{bmatrix} + \begin{bmatrix} \mathbf{V}_1^T(1) \\ \vdots \\ \mathbf{V}_K^T(1) \end{bmatrix} = [\mathcal{M}]_{(2)} + \mathbf{V}^{(2)}(1) \quad (15)$$

To extract the channel matrices  $\mathbf{H}_{sr_1}$  and  $\mathbf{H}_{r_1d}$ , similar to the steps (11)-(15), we can obtain another PARAFAC tensor model  $\mathcal{T} = [\mathbf{T}_1 \cup_3 \mathbf{T}_2 \cdots \cup_3 \mathbf{T}_K]$ , and its two tensor unfolding, which consist of matrix-of-interest, can be denoted as follows,

$$[\tilde{\mathcal{T}}]_{(1)} = \begin{bmatrix} \mathbf{T}_1 \\ \vdots \\ \mathbf{T}_K \end{bmatrix} + \begin{bmatrix} \mathbf{V}_1(2) \\ \vdots \\ \mathbf{V}_K(2) \end{bmatrix} = (\mathbf{E} \diamond \mathbf{H}_{r_1d}) \mathbf{H}_{sr_1} + \mathbf{V}^{(1)}(2) \quad (16)$$

$$[\tilde{\mathcal{T}}]_{(2)} = \begin{bmatrix} \mathbf{T}_1^T \\ \vdots \\ \mathbf{T}_K^T \end{bmatrix} + \begin{bmatrix} \mathbf{V}_1^T(2) \\ \vdots \\ \mathbf{V}_K^T(2) \end{bmatrix} = (\mathbf{E} \diamond \mathbf{H}_{sr_1}^T) \mathbf{H}_{r_1d}^T + \mathbf{V}^{(2)}(2) \quad (17)$$

### 3.2 Model of Three-hop Link

During the fourth phase, the training signals are transmitted to destination node through the channel  $\mathbf{H}_{sr_1}$ ,  $\mathbf{H}_{r_1r_2}$  and  $\mathbf{H}_{r_2d}$ . Compared to the PARAFAC tensor model, a more flexible tensor model should be adopted to capture this characteristic. By multiplying both sides of received signal (6) with  $\mathbf{S}_0^H$ , we have

$$\tilde{\mathbf{Q}}_k = \mathbf{Q}_k + \mathbf{V}_k(3) \quad (18)$$

where

$$\mathbf{Q}_k = \mathbf{H}_{r_2d} \mathcal{D}_k(\mathbf{F}) \mathbf{H}_{r_1r_2} \mathcal{D}_k(\mathbf{E}) \mathbf{H}_{sr_1} \quad (19)$$

$$\mathbf{V}_k(3) = \mathbf{H}_{r_2d} \mathcal{D}_k(\mathbf{F}) \mathbf{H}_{r_1r_2} \mathcal{D}_k(\mathbf{E}) \mathbf{V}_{r_1}^{(k)}(t) \mathbf{S}_0^H + \mathbf{H}_{r_2d} \mathcal{D}_k(\mathbf{F}) \mathbf{V}_{r_2}^{(k)}(t+2) \mathbf{S}_0^H + \mathbf{V}_d^{(k)}(t+3) \quad (20)$$

By storing the set of  $K$  matrices in (19) together along the direction of the index  $k$  (the third dimension), we obtain a third-order tensor  $\mathcal{Q} \in \mathbb{C}^{N_d \times N_s \times K}$ , whose  $(i, j, k)$ -th element is given by

$$Q(i, j, k) = \sum_{m=1}^{N_2} \sum_{r=1}^{N_1} \mathbf{H}_{r_2d}(i, m) \mathbf{F}(k, m) \mathbf{H}_{r_1r_2}(m, r) \mathbf{E}(k, r) \mathbf{H}_{sr_1}(r, j) \quad (21)$$

$$\tilde{\mathcal{Q}}_{\cdot, \cdot, k} = \mathcal{Q}_{\cdot, \cdot, k} + \mathcal{V}'_{\cdot, \cdot, k} \quad (22)$$

for all  $i=1, \dots, N_d$ ,  $j=1, \dots, N_s$  and  $k=1, \dots, K$ .  $\mathcal{V}' \in \mathbb{C}^{N_d \times N_s \times K}$  is the noise tensor stacked with

$\mathbf{V}_k(3)$  as  $\mathbf{Q}_k$ . Note that (21) follows the typical PARATUCK2 tensor model [25], in which the matrices  $\mathbf{H}_{r_2d}$  and  $\mathbf{H}_{sr_1}$  are associated with the first and second dimension of  $\mathbf{Q}$ .  $\mathbf{E}$  and  $\mathbf{F}$  are interaction matrices defining the linear combination profile between the  $N_2$  columns of  $\mathbf{H}_{r_2d}$  and the  $N_1$  columns of  $\mathbf{H}_{sr_1}^T$ , while  $\mathbf{H}_{r_1r_2}$  is the core matrix of PARATUCK2 model.

From the set of slices  $\tilde{\mathbf{Q}}_k$ , in order to estimate  $\mathbf{H}_{sr_1}$  and  $\mathbf{H}_{r_2d}$ , we can define two matrix unfolding  $[\tilde{\mathbf{Q}}]_{(2)} \in \mathbf{C}^{KN_d \times N_s}$  and  $[\tilde{\mathbf{Q}}]_{(3)} \in \mathbf{C}^{KN_s \times N_d}$  in the following manner,

$$[\tilde{\mathbf{Q}}]_{(2)} = \begin{bmatrix} \tilde{\mathbf{Q}}_1 \\ \vdots \\ \tilde{\mathbf{Q}}_k \\ \vdots \\ \tilde{\mathbf{Q}}_K \end{bmatrix}, \quad [\tilde{\mathbf{Q}}]_{(3)} = \begin{bmatrix} \tilde{\mathbf{Q}}_1^T \\ \vdots \\ \tilde{\mathbf{Q}}_k^T \\ \vdots \\ \tilde{\mathbf{Q}}_K^T \end{bmatrix}. \quad (23)$$

From (23), we get the following factorization for  $[\tilde{\mathbf{Q}}]_{(2)}$  and  $[\tilde{\mathbf{Q}}]_{(3)}$ ,

$$[\tilde{\mathbf{Q}}]_{(2)} = (\mathbf{I}_K \otimes \mathbf{H}_{r_2d}) \mathbf{\Omega}_2 \mathbf{H}_{sr_1} + \tilde{\mathbf{V}} \quad (24)$$

$$[\tilde{\mathbf{Q}}]_{(3)} = (\mathbf{I}_K \otimes \mathbf{H}_{sr_1}^T) \mathbf{\Omega}_3 \mathbf{H}_{r_2d}^T + \tilde{\mathbf{V}} \quad (25)$$

where

$$\mathbf{\Omega}_2 = \begin{bmatrix} \mathcal{D}_1(\mathbf{F}) \mathbf{H}_{r_1r_2} \mathcal{D}_1(\mathbf{E}) \\ \vdots \\ \mathcal{D}_K(\mathbf{F}) \mathbf{H}_{r_1r_2} \mathcal{D}_K(\mathbf{E}) \end{bmatrix} \in \mathbf{C}^{KN_2 \times N_1} \quad (26)$$

$$\tilde{\mathbf{V}} = [\mathbf{V}_1^T(3), \dots, \mathbf{V}_K^T(3)]^T \quad (27)$$

and

$$\mathbf{\Omega}_3 = \begin{bmatrix} \mathcal{D}_1(\mathbf{E}) \mathbf{H}_{r_1r_2}^T \mathcal{D}_1(\mathbf{F}) \\ \vdots \\ \mathcal{D}_K(\mathbf{E}) \mathbf{H}_{r_1r_2}^T \mathcal{D}_K(\mathbf{F}) \end{bmatrix} \in \mathbf{C}^{KN_1 \times N_2} \quad (28)$$

$$\hat{\mathbf{V}} = [\mathbf{V}_1(3), \dots, \mathbf{V}_K(3)]^T \quad (29)$$

To estimate  $\mathbf{H}_{r_1r_2}$ , the third matrix unfolding of  $\tilde{\mathbf{Q}} \in \mathbf{C}^{N_d \times N_s \times K}$  can be defined as

$$\begin{aligned} [\tilde{\mathbf{Q}}]_1 &= [\text{vec}(\tilde{\mathbf{Q}}_1), \dots, \text{vec}(\tilde{\mathbf{Q}}_K)] \\ &= (\mathbf{H}_{sr_1}^T \otimes \mathbf{H}_{r_2d}) \text{diag}(\text{vec}(\mathbf{H}_{r_1r_2})) (\mathbf{E}^T \diamond \mathbf{F}^T) + \tilde{\mathbf{V}} \end{aligned} \quad (30)$$

where  $\tilde{\mathbf{V}} = [\text{vec}(\mathbf{V}_1(3)), \dots, \text{vec}(\mathbf{V}_K(3))]$ . Note that from equation (30), to estimate the channel matrix  $\mathbf{h}_{r_1r_2} = \text{vec}(\mathbf{H}_{r_1r_2})$ , using the property (3), we can have

$$\tilde{\mathbf{q}}_1 = \text{vec}([\tilde{\mathbf{Q}}]_1) = \mathbf{\Omega}_1 \mathbf{h}_{r_1r_2} + \tilde{\mathbf{v}} \quad (31)$$

and

$$\mathbf{\Omega}_1 = [(\mathbf{E}^T \diamond \mathbf{F}^T)^T \diamond (\mathbf{H}_{sr_1}^T \otimes \mathbf{H}_{r_2d})] \in \mathbf{C}^{KN_s N_d \times N_1 N_2} \quad (32)$$

$$\tilde{\mathbf{v}} = [\text{vec}(\mathbf{V}_1(3))^T, \dots, \text{vec}(\mathbf{V}_K(3))^T]^T \quad (33)$$



### 3.3 Combined ALS (Comb-ALS) Channel Estimation Algorithm

From the previous discussion, both the channel matrices of  $\mathbf{H}_{sr_1}$  and  $\mathbf{H}_{r_2d}$  are involved in the two-hop link and three-hop link, corresponding to different tensor models. How to extract the channel matrices effectively is a challenge. In this section, we propose a combined channel estimator to estimate the CSI by jointly exploit the benefit of PARAFAC and PARATUCK2.

Equation (31) leads to the following LS estimation problem,

$$\begin{aligned}\hat{\mathbf{h}}_{r_2} &= \underbrace{\arg \min}_{\mathbf{h}_{r_2}} \left\| \tilde{\mathbf{q}}_1 - \mathbf{\Omega}_1 \mathbf{h}_{r_2} \right\|_F^2 \\ &= (\mathbf{\Omega}_1^H \mathbf{\Omega}_1)^{(-1)} \mathbf{\Omega}_1^H \tilde{\mathbf{q}}_1\end{aligned}\quad (34)$$

Combining (16) and (24) the LS estimation of  $\mathbf{H}_{sr_1}$  can be obtained by solving the following problem,

$$\hat{\mathbf{H}}_{sr_1} = \underbrace{\arg \min}_{\mathbf{H}_{sr_1}} \left\| \begin{bmatrix} [\tilde{\mathcal{T}}]_{(1)} \\ [\tilde{\mathcal{Q}}]_{(2)} \end{bmatrix} - \begin{bmatrix} \mathbf{E} \diamond \mathbf{H}_{r_2d} \\ (\mathbf{I}_K \otimes \mathbf{H}_{r_2d}) \mathbf{\Omega}_2 \end{bmatrix} \mathbf{H}_{sr_1} \right\|_F^2 \quad (35)$$

The solution is given by

$$\hat{\mathbf{H}}_{sr_1} = \underbrace{\begin{bmatrix} \mathbf{E} \diamond \mathbf{H}_{r_2d} \\ (\mathbf{I}_K \otimes \mathbf{H}_{r_2d}) \mathbf{\Omega}_2 \end{bmatrix}}_{\mathbf{H}_{eff1}}^\dagger \begin{bmatrix} [\tilde{\mathcal{T}}]_{(1)} \\ [\tilde{\mathcal{Q}}]_{(2)} \end{bmatrix} \quad (36)$$

With a similar procedure, we combine (15) and (25) to establish the following LS optimization problem for  $\mathbf{H}_{r_2d}$ ,

$$\hat{\mathbf{H}}_{r_2d}^T = \underbrace{\arg \min}_{\mathbf{H}_{r_2d}^T} \left\| \begin{bmatrix} [\tilde{\mathcal{M}}]_{(2)} \\ [\tilde{\mathcal{Q}}]_{(3)} \end{bmatrix} - \begin{bmatrix} \mathbf{F} \diamond \mathbf{H}_{sr_1}^T \\ (\mathbf{I}_K \otimes \mathbf{H}_{sr_1}^T) \mathbf{\Omega}_3 \end{bmatrix} \mathbf{H}_{r_2d}^T \right\|_F^2 \quad (37)$$

The solution of (37) is given by

$$\hat{\mathbf{H}}_{r_2d}^T = \underbrace{\begin{bmatrix} \mathbf{F} \diamond \mathbf{H}_{sr_1}^T \\ (\mathbf{I}_K \otimes \mathbf{H}_{sr_1}^T) \mathbf{\Omega}_3 \end{bmatrix}}_{\mathbf{H}_{eff2}}^\dagger \begin{bmatrix} [\tilde{\mathcal{M}}]_{(2)} \\ [\tilde{\mathcal{Q}}]_{(3)} \end{bmatrix} \quad (38)$$

From (14) and (17), we can obtain the LS approximation of  $\mathbf{H}_{sr_2}$  and  $\mathbf{H}_{r_1d}$  as follows,

$$\begin{aligned}\hat{\mathbf{H}}_{sr_2} &= \underbrace{\arg \min}_{\mathbf{H}_{sr_2}} \left\| [\tilde{\mathcal{M}}]_{(1)} - (\mathbf{F} \diamond \mathbf{H}_{r_2d}) \mathbf{H}_{sr_2} \right\|_F^2 \\ &= (\mathbf{F} \diamond \mathbf{H}_{r_2d})^\dagger [\tilde{\mathcal{M}}]_{(1)}\end{aligned}\quad (39)$$

$$\begin{aligned}\hat{\mathbf{H}}_{r_1d}^T &= \underbrace{\arg \min}_{\mathbf{H}_{r_1d}^T} \left\| [\tilde{\mathcal{T}}]_{(2)} - (\mathbf{E} \diamond \mathbf{H}_{sr_1}^T) \mathbf{H}_{r_1d}^T \right\|_F^2 \\ &= (\mathbf{E} \diamond \mathbf{H}_{sr_1}^T)^\dagger [\tilde{\mathcal{T}}]_{(2)}\end{aligned}\quad (40)$$

**Identifiability:** Both PARAFAC and PARATUCK2 decompositions are essentially unique under some mild conditions, which allows a joint estimation of channel matrices of all links. Note that  $\mathbf{H}_{sr_1}$  and  $\mathbf{H}_{r_2d}$  are identifiable from (36) and (38) if  $\mathbf{H}_{eff1}$  and  $\mathbf{H}_{eff2}$  are full column-rank, which requires  $2KN_d \geq N_1$  and  $2KN_s \geq N_2$ . From equations (34), (39), (40), we

can have  $KN_s N_d \geq N_1 N_2$ ,  $KN_d \geq N_2$  and  $KN_s \geq N_1$ . Combining these conditions yields the following lower bound on the number of time blocks,

$$K \geq \max \left( \left\lceil \frac{N_1 N_2}{N_d N_s} \right\rceil, \left\lceil \frac{N_1}{2N_d} \right\rceil, \left\lceil \frac{N_2}{2N_s} \right\rceil, \left\lceil \frac{N_2}{N_d} \right\rceil, \left\lceil \frac{N_1}{N_s} \right\rceil, 2 \right) \quad (41)$$

where  $\lceil x \rceil$  is equal to the smallest integer which is greater than or equal to  $x$ .

Let  $\{\hat{\mathbf{H}}_{sr_1}, \hat{\mathbf{H}}_{r_1r_2}, \hat{\mathbf{H}}_{r_2d}, \hat{\mathbf{H}}_{sr_2}, \hat{\mathbf{H}}_{r_1d}\}$  be an alternative set of matrices yielding the same tensor  $\mathcal{M}$ ,  $\mathcal{T}$  and  $\mathcal{Q}$ . If the channel matrices of all the involved links are full rank and the identifiability conditions (41) are satisfied, then  $\{\hat{\mathbf{H}}_{sr_1}, \hat{\mathbf{H}}_{r_1r_2}, \hat{\mathbf{H}}_{r_2d}, \hat{\mathbf{H}}_{sr_2}, \hat{\mathbf{H}}_{r_1d}\}$  are essentially unique. In this case, we have  $\hat{\mathbf{H}}_{sr_1} = \Delta_{sr_1} \mathbf{H}_{sr_1}$ ,  $\hat{\mathbf{H}}_{r_2d} = \mathbf{H}_{r_2d} \Delta_{r_2d}$ ,  $\hat{\mathbf{H}}_{r_1r_2} = \Delta_{r_1r_2}^{(2)} \mathbf{H}_{r_1r_2} \Delta_{r_1r_2}^{(1)}$ ,  $\hat{\mathbf{H}}_{sr_2} = \Delta_{sr_2} \mathbf{H}_{sr_2}$  and  $\hat{\mathbf{H}}_{r_1d} = \mathbf{H}_{r_1d} \Delta_{r_1d}$ , where the following relation holds,

$$(\Delta_{sr_1} \Delta_{r_1r_2}^{(1)}) \otimes (\Delta_{r_2d} \Delta_{r_1r_2}^{(2)}) = \mathbf{I}_{N_1 N_2} \quad (42)$$

$$\Delta_{sr_1} \Delta_{r_1d} = \mathbf{I}_{N_1} \quad (43)$$

$$\Delta_{sr_2} \Delta_{r_2d} = \mathbf{I}_{N_2} \quad (44)$$

Since these ambiguities compensate each other, the compound channel is strictly unique  $\mathbf{H}_c = [\mathbf{H}_{r_1d} \mathbf{H}_{sr_1}; \mathbf{H}_{r_2d} \mathbf{H}_{sr_2}; \mathbf{H}_{r_2d} \mathbf{H}_{r_1r_2} \mathbf{H}_{sr_1}] = [\hat{\mathbf{H}}_{r_1d} \hat{\mathbf{H}}_{sr_1}; \hat{\mathbf{H}}_{r_2d} \hat{\mathbf{H}}_{sr_2}; \hat{\mathbf{H}}_{r_2d} \hat{\mathbf{H}}_{r_1r_2} \hat{\mathbf{H}}_{sr_1}]$ . From (41), we can see that there is a tradeoff can be achieved between the number of antenna and the period of required training symbols, which is meaningful in the scenario that the destination antenna number is less than the relay node antenna numbers as the traditional LS-based channel training method cannot be used. In addition, if we know the first row of  $\mathbf{H}_{r_1d}$  and  $\mathbf{H}_{r_2d}$ , these scaling ambiguities can be eliminated by means of the following equations:

$$\Delta_{r_1d} = \mathcal{D}_1^{(-1)}(\mathbf{H}_{r_1d}) \mathcal{D}_1(\hat{\mathbf{H}}_{r_1d}^\infty) \quad (45)$$

$$\Delta_{r_2d} = \mathcal{D}_1^{(-1)}(\mathbf{H}_{r_2d}) \mathcal{D}_1(\hat{\mathbf{H}}_{r_2d}^\infty) \quad (46)$$

where  $\hat{\mathbf{H}}_{\bullet}^\infty$  denotes the estimated values of  $\mathbf{H}_{\bullet}$  at the convergence of Comb-ALS algorithms.

In practice, the knowledge of the first row of  $\mathbf{H}_{r_1d}$  and  $\mathbf{H}_{r_2d}$  can be obtained by a simple supervised procedure [19, 22]. In this paper, for purpose of performance evaluation, the scaling ambiguities affecting the estimation of the channel matrices are removed by assuming the first row of  $\mathbf{H}_{r_1d}$  and  $\mathbf{H}_{r_2d}$  contain all one elements, similar to [14, 15, 19, 21, 23]. First,

we find  $\Delta_{r_1d} = \mathcal{D}_1(\hat{\mathbf{H}}_{r_1d}^\infty)$  and  $\Delta_{r_2d} = \mathcal{D}_1(\hat{\mathbf{H}}_{r_2d}^\infty)$ . Then, applying property  $(\mathbf{A}\mathbf{B}) \otimes (\mathbf{C}\mathbf{D}) = (\mathbf{A} \otimes \mathbf{C})(\mathbf{B} \otimes \mathbf{D})$  yields  $(\Delta_{sr_1} \otimes \Delta_{r_2d})(\Delta_{r_1r_2}^{(1)} \otimes \Delta_{r_1r_2}^{(2)}) = \mathbf{I}_{N_1 N_2}$  and from (42)-(44) we obtain  $\Delta_{r_1r_2}^{(1)} = \Delta_{sr_1}^{-1} = \Delta_{r_1d}$  and  $\Delta_{sr_2} = \Delta_{r_1r_2}^{(2)} = \Delta_{r_2d}^{-1}$ .

Note that these uniqueness results are naturally applicable if the  $S$ - $R_1$ - $D$  link and  $S$ - $R_2$ - $D$  link are too weak or not available (e.g. due to shadowing). Thus, we do not have to estimate  $\mathbf{H}_{sr_2}$ ,  $\mathbf{H}_{r_2d}$  and  $\mathbf{H}_{r_1d}$  any more, and we can model  $S$ - $R_1$ - $R_2$ - $D$  link by using the PARATUCK2 tensor model. Then the uniqueness issue can be simplified to the works in [21],

$$K \geq \max \left( \left\lceil \frac{N_1 N_2}{N_d N_s} \right\rceil, \left\lceil \frac{N_1}{N_d} \right\rceil, \left\lceil \frac{N_2}{N_s} \right\rceil, 2 \right) \quad (47)$$

If the  $R_1$ - $R_2$  link is too weak or not available, our model will be simplified to the work in typical two-hop AF relay system, the uniqueness condition simplifies to the scenario in works [14, 15].

### 3.4 Linear Minimal Mean Square Error (LMMSE) Estimation

**Table 1.** Proposed two-stage fitting algorithm

<p>First Stage (Comb-ALS algorithm)</p> <p>Step 1.1. Initialize the algorithm with given <math>\mathbf{E}</math> and <math>\mathbf{F}</math> and random <math>\hat{\mathbf{H}}_{sr_1}^{(i=0)}</math>, <math>\hat{\mathbf{H}}_{r_1d}^{(i=0)}</math>, <math>\hat{\mathbf{H}}_{sr_2}^{(i=0)}</math>, <math>\hat{\mathbf{H}}_{r_2d}^{(i=0)}</math>; set <math>\delta(0) = \infty</math> and <math>i=1</math>;</p> <p>Step 1.2. Update <math>\hat{\mathbf{H}}_{r_1r_2}^{(i)} = \text{unvec}(\hat{\mathbf{h}}_{r_1r_2}^{(i)}) \in \mathbb{C}^{N_2 \times N_1}</math> as (34) using <math>\hat{\mathbf{H}}_{sr_1}^{(i-1)}</math> and <math>\hat{\mathbf{H}}_{r_2d}^{(i-1)}</math>;</p> <p>Step 1.3. Update <math>\hat{\mathbf{H}}_{sr_1}^{(i)}</math> as (36) using <math>\hat{\mathbf{H}}_{r_1d}^{(i-1)}</math>, <math>\hat{\mathbf{H}}_{r_2d}^{(i-1)}</math> and <math>\hat{\mathbf{H}}_{r_1r_2}^{(i)}</math>; update <math>\hat{\mathbf{H}}_{r_2d}^{(i)}</math> as (38) using <math>\hat{\mathbf{H}}_{sr_2}^{(i-1)}</math>, <math>\hat{\mathbf{H}}_{sr_1}^{(i)}</math> and <math>\hat{\mathbf{H}}_{r_1r_2}^{(i)}</math>;</p> <p>Step 1.4. Update <math>\hat{\mathbf{H}}_{sr_2}^{(i)}</math> and <math>\hat{\mathbf{H}}_{r_1d}^{(i)}</math> from (39) and (40), using <math>\hat{\mathbf{H}}_{r_2d}^{(i)}</math> and <math>\hat{\mathbf{H}}_{sr_1}^{(i)}</math>, respectively.</p> <p>Step 1.5. Calculate <math>\delta(i) = \ \hat{\mathbf{H}}_c - \mathbf{H}_c\ _F^2</math>. If <math>[\delta(i-1) - \delta(i)] / \delta(i) \leq \varepsilon</math>, then end. Otherwise, let <math>i = i + 1</math> and go to step 1.2.</p> <p>Step 1.6. Eliminate scaling ambiguity: <math>\hat{\mathbf{H}}_{sr_1} \leftarrow \Delta_{sr_1} \hat{\mathbf{H}}_{sr_1}^\infty</math>, <math>\hat{\mathbf{H}}_{r_1d} \leftarrow \hat{\mathbf{H}}_{r_1d}^\infty \Delta_{r_1d}^{(-1)}</math>, <math>\hat{\mathbf{H}}_{r_2d} \leftarrow \hat{\mathbf{H}}_{r_2d}^\infty \Delta_{r_2d}^{(-1)}</math>, <math>\hat{\mathbf{H}}_{sr_2} \leftarrow \Delta_{sr_2} \hat{\mathbf{H}}_{sr_2}^\infty</math> and <math>\hat{\mathbf{H}}_{r_1r_2} = \Delta_{r_2d} \hat{\mathbf{H}}_{r_1r_2}^\infty \Delta_{r_1d}^{(-1)}</math>, where <math>\Delta_{r_1d}</math> and <math>\Delta_{r_2d}</math> are defined in (45) and (46).</p> <p>Second Stage (LMMSE-based processing)</p> <p>Step 2.1. Calculate <math>\mathbf{T}_{sr_1}</math> as (50) using <math>\hat{\mathbf{H}}_{r_2d}</math> and <math>\hat{\mathbf{H}}_{r_1r_2}</math>, update <math>\check{\mathbf{H}}_{sr_1} = \mathbf{T}_{sr_1}^H [\tilde{\mathbf{Q}}]_2</math>;</p> <p>Step 2.2. Calculate <math>\mathbf{T}_{r_2d}</math> as (51) using <math>\check{\mathbf{H}}_{sr_1}</math> and <math>\hat{\mathbf{H}}_{r_1r_2}</math>, update <math>\check{\mathbf{H}}_{r_2d} = (\mathbf{T}_{r_2d}^H [\tilde{\mathbf{Q}}]_3)^T</math>;</p> <p>Step 2.3. Calculate <math>\mathbf{T}_{r_1r_2}</math> as (52) using <math>\check{\mathbf{H}}_{sr_1}</math> and <math>\check{\mathbf{H}}_{r_2d}</math>, update <math>\check{\mathbf{h}}_{r_1r_2} = \mathbf{T}_{r_1r_2}^H \tilde{\mathbf{q}}_1</math>, and then <math>\check{\mathbf{H}}_{r_1r_2} = \text{unvec}(\check{\mathbf{h}}_{r_1r_2}) \in \mathbb{C}^{N_2 \times N_1}</math>.</p>
--

In this section, an improved estimation of  $\mathbf{H}_{sr_1}$ ,  $\mathbf{H}_{r_1r_2}$  and  $\mathbf{H}_{r_2d}$  can be derived by the LMMSE approach, which will further inhibit the error propagation by project the sampled data to the desired channel subspace by assuming a linear system model [14, 15, 27]. Obviously, the covariance matrix of the effective noise in (27), (29) and (33) is non-white any more. Note that the subsequent LMMSE signal processing method has been used in paper [14, 15] in two-hop PARAFAC tensor model channel estimation algorithm. In this paper, we will firstly derive the corresponding procedures for PARATUCK2-based LMMSE algorithm of  $S$ - $R_1$ - $R_2$ - $D$  three-hop link. Table 1 briefly summarizes the proposed algorithm, referred to two-stage fitting algorithm. The basic ideas are as follows: after an initial estimation of  $\hat{\mathbf{H}}_{r_1r_2}$  and  $\hat{\mathbf{H}}_{r_2d}$  by the Comb-ALS algorithm, an improved estimation of  $\hat{\mathbf{H}}_{sr_1}$  can be obtained by the following LMMSE approach as  $\check{\mathbf{H}}_{sr_1} = \mathbf{T}_{sr_1}^H [\tilde{\mathbf{Q}}]_2$ , where  $\mathbf{T}_{sr_1} \in \mathbb{C}^{KN_d \times N_1}$  is the weight matrix. The mean square error of channel estimation can be written as,

$$\begin{aligned} \mathbf{J}_2 &= \mathbf{E} \left[ \text{tr} \left( (\check{\mathbf{H}}_{sr_1} - \mathbf{H}_{sr_1}) (\check{\mathbf{H}}_{sr_1} - \mathbf{H}_{sr_1})^H \right) \right] \\ &= \left[ \mathbf{T}_{sr_1}^H \Phi - \mathbf{I}_{N_1} \right] \left[ \mathbf{T}_{sr_1}^H \Phi - \mathbf{I}_{N_1} \right]^H + \mathbf{T}_{sr_1}^H \hat{\Theta} \mathbf{T}_{sr_1} \end{aligned} \quad (48)$$

where  $\Phi = (\mathbf{I}_K \otimes \mathbf{H}_{r_2d}) \Omega_2$ ,  $\hat{\Theta} = bd \left[ \hat{\Theta}_1, \hat{\Theta}_2, \dots, \hat{\Theta}_K \right]$  and

$$\begin{aligned}\hat{\mathbf{\Theta}}_k &= \mathbb{E}[\tilde{\mathbf{V}}_k \tilde{\mathbf{V}}_k^H] \\ &= N_s \mathbf{H}_{r_2d} \mathcal{D}_k(\mathbf{F}) \mathbf{H}_{r_1r_2} \mathcal{D}_k(\mathbf{E}) \mathcal{D}_k^H(\mathbf{E}) \mathbf{H}_{r_1r_2}^H \mathcal{D}_k^H(\mathbf{F}) \mathbf{H}_{r_2d}^H \\ &\quad + N_s \mathbf{H}_{r_2d} \mathcal{D}_k(\mathbf{F}) \mathcal{D}_k^H(\mathbf{F}) \mathbf{H}_{r_2d}^H + N_s \mathbf{I}_{N_d}\end{aligned}\quad (49)$$

The weight matrix minimizing (48) is given by,

$$\mathbf{T}_{s_{r_1}} = (\mathbf{\Phi} \mathbf{\Phi}^H + \hat{\mathbf{\Theta}})^{(-1)} \mathbf{\Phi} \quad (50)$$

Similarly, we can also obtain an improved estimation of  $\mathbf{H}_{r_2d}$  and  $\mathbf{H}_{r_1r_2}$  by the LMMSE approach as  $\tilde{\mathbf{H}}_{r_2d} = (\mathbf{T}_{r_2d}^H [\tilde{\mathbf{Q}}]_3)^T$  and  $\tilde{\mathbf{h}}_{r_1r_2} = \mathbf{T}_{r_1r_2}^H \tilde{\mathbf{q}}_1$ , where  $\mathbf{T}_{r_2d} \in \mathbb{C}^{KN_s \times N_2}$  and  $\mathbf{T}_{r_1r_2} \in \mathbb{C}^{KN_d N_s \times N_1 N_2}$  is the weight matrix with,

$$\mathbf{T}_{r_2d} = N_d (N_d \mathbf{\Psi} \mathbf{\Psi}^H + N_2 \hat{\mathbf{C}})^{(-1)} \mathbf{\Psi} \quad (51)$$

$$\mathbf{T}_{r_1r_2} = (\mathbf{\Omega}_1 \mathbf{\Omega}_1^H + N_1 \hat{\mathbf{\Gamma}})^{(-1)} \mathbf{\Omega}_1 \quad (52)$$

where  $\mathbf{\Psi} = (\mathbf{I}_K \otimes \mathbf{H}_{sr_1}^T) \mathbf{\Omega}_3$ ,  $\hat{\mathbf{C}} = bd[\hat{\mathbf{C}}_1, \dots, \hat{\mathbf{C}}_K]$ ,  $\hat{\mathbf{\Gamma}} = bd[\hat{\mathbf{\Gamma}}_1, \dots, \hat{\mathbf{\Gamma}}_K]$  and

$$\begin{aligned}\hat{\mathbf{C}}_k &= \frac{N_d}{N_2} \text{tr}\{\mathcal{D}_k^H(\mathbf{E}) \mathbf{H}_{r_1r_2}^H \mathcal{D}_k^H(\mathbf{F}) \mathcal{D}_k(\mathbf{F}) \mathbf{H}_{r_1r_2} \mathcal{D}_k(\mathbf{E})\} \mathbf{I}_{N_s} \\ &\quad + \frac{N_d}{N_2} \text{tr}\{\mathcal{D}_k^H(\mathbf{F}) \mathcal{D}_k(\mathbf{F})\} + N_d \mathbf{I}_{N_s}\end{aligned}\quad (53)$$

$$\begin{aligned}\hat{\mathbf{\Gamma}}_k &= \frac{\text{tr}(\mathcal{D}_k(\mathbf{E}) \mathcal{D}_k^H(\mathbf{E}))}{N_1} [\mathbf{I}_{N_s} \otimes (\mathbf{H}_{r_2d} \mathcal{D}_k(\mathbf{F}) \mathcal{D}_k^H(\mathbf{F}) \mathbf{H}_{r_2d}^H)] \\ &\quad + \mathbf{I}_{N_s} \otimes (\mathbf{H}_{r_2d} \mathcal{D}_k(\mathbf{F}) \mathcal{D}_k^H(\mathbf{F}) \mathbf{H}_{r_2d}^H) + \mathbf{I}_{N_s N_d}\end{aligned}\quad (54)$$

### 3.5 Complexity Analysis

The computational cost at each iteration of the proposed algorithm is given in **Table 2**. For the sake of simplicity, the expressions in **Table 2** approximate the number of floating-point operations per iteration in considering the dominant cost associated with SVD computation, which is used to calculate the matrix pseudo-inverses. According to [22, 26], for a full column-rank matrix of dimensions  $M \times N$ , the SVD computation cost is  $O(MN^2)$ . The overall complexity of the Comb-ALS based algorithm depends on the number of iterations. Generally, the required iterations for convergence of the proposed Comb-ALS is less than 12, which will be further analyzed in Section 4. In addition, the subsequent LMMSE approach imposes  $O(K^3(N_d^3 + N_s^3 N_d^3 + N_s^3))$  extra operations due to matrix inversions.

In comparison with the PARATUCK-based algorithm [21], we can conclude from **Table 2** that the complexity per iteration scales similarly to the proposed algorithm, but the proposed algorithm is a little higher than [21]. One comes from the estimation of the additional channel links  $\mathbf{H}_{sr_2}$  and  $\mathbf{H}_{r_1d}$ , which is  $O(KN_s N_1^2 + KN_d N_2^2)$ . The other difference results from the extra sampled data provided in two-hop links when estimating  $\mathbf{H}_{sr_1}$  and  $\mathbf{H}_{r_2d}$ , which is  $O(KN_d N_1^2 + KN_s N_2^2)$ .

**Table 2.** Algorithmic complexity of the Comb-ALS and PARATUCK2-based algorithms

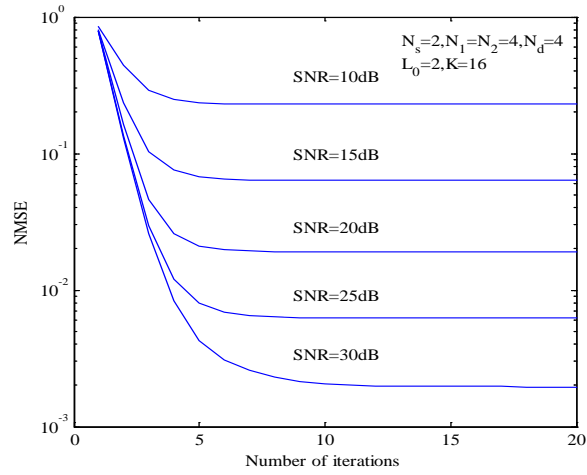
Algorithm	Number of floating-point operations $O(\bullet)$ per iteration
Comb-ALS	$K[(2N_d + N_s)N_1^2 + (2N_s + N_d)N_2^2 + N_s N_d N_1^2 N_2^2]$
PARATUCK2-based [21]	$K[N_d N_1^2 + N_s N_2^2 + N_s N_d N_1^2 N_2^2]$

### 4. Simulation Results

In this section, some simulation results are provided to demonstrate the performance of the proposed scheme. Each obtained result is an average over  $R = 5000$  independent Monte Carlo simulations. For each run, the channels are generated from an i.i.d Rayleigh generator while the users symbols are generated from an i.i.d. distribution and are modulated using QPSK. And the proposed algorithm is performed following the procedure in **Table 1** with  $\varepsilon = 1 \times 10^{(-5)}$ . In particular, we compare the proposed algorithm with algorithm in [21] and the typical LS estimator. As for the LS estimator, we adopt the sequential estimation method in [21] to estimate  $\mathbf{H}_{sr_1}$ ,  $\mathbf{H}_{r_1r_2}$  and  $\mathbf{H}_{r_2d}$ , the channel matrix  $\mathbf{H}_{r_1d}$  is estimated from the received signal model of  $R_2$ - $D$  link while  $\mathbf{H}_{sr_2}$  is draw from the received signal of link  $S$ - $R_2$ - $D$  based on the estimated  $\mathbf{H}_{r_2d}$ . Note that in order to ensure the uniqueness of the LS estimation, the antenna number should satisfy  $N_d \geq N_2 \geq N_1$ . A factor of  $\sqrt{K}$  is used to scale the training sequences  $\mathbf{S}_0$  in the LS training process to ensure a fair comparison [14]. In addition, in order to make a fair comparison with the algorithm in [21], we adopt the method in [21] to remove the ambiguities in the Step 1.6 of the proposed algorithm. From **Fig. 2** to **Fig. 6**, we consider a special communication scenario, where the source node, the two relay nodes and the destination node are in a line with equal distance, that is to say,  $\eta_1 = \eta_2 = 2$ . The effect of the path loss parameter on the achievable diversity gain will be further discussed in **Fig. 7** and **Fig. 8**.

The estimator’s performance is evaluated in terms of the normalized mean square error (NMSE) of the estimated channel matrices and the bit error rate (BER). For each channel realization, the NMSE of channel estimation for different algorithms is calculated as  $\|\hat{\mathbf{H}} - \mathbf{H}\|_F^2 / \|\mathbf{H}\|_F^2$  for the channel  $\mathbf{H}$ , where  $\hat{\mathbf{H}}$  is the estimated value. The SNR level at the first relay node and the second relay node is assumed to be 20dB above the SNR level at the destination, and the BER performance is obtained with a zero forcing (ZF) receiver. During the data transmission period, unlike the training period, the amplify matrices are generated based on DFT matrix. From (4)-(6), the ZF solution can be written as,

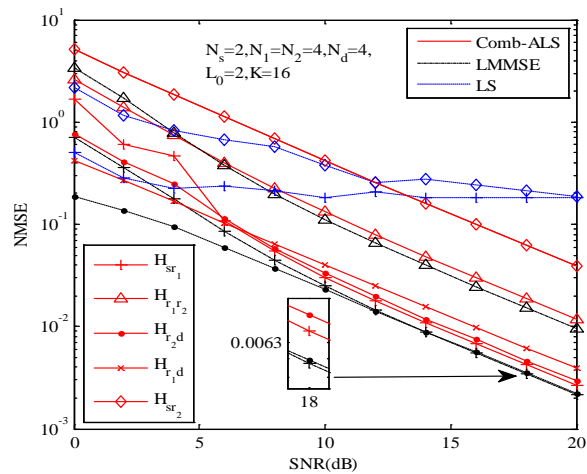
$$\hat{\mathbf{S}}_{ZF}(t) = \begin{bmatrix} \hat{\mathbf{H}}_{r_2d} \mathbf{G}_2 \hat{\mathbf{H}}_{sr_2} \\ \hat{\mathbf{H}}_{r_1d} \mathbf{G}_1 \hat{\mathbf{H}}_{sr_1} \\ \hat{\mathbf{H}}_{r_2d} \mathbf{G}_2 \hat{\mathbf{H}}_{r_1r_2} \mathbf{G}_1 \hat{\mathbf{H}}_{sr_1} \end{bmatrix}^* \begin{bmatrix} \mathbf{Y}_d(t+1) \\ \mathbf{Y}_d(t+2) \\ \mathbf{Y}_d(t+3) \end{bmatrix} \tag{55}$$



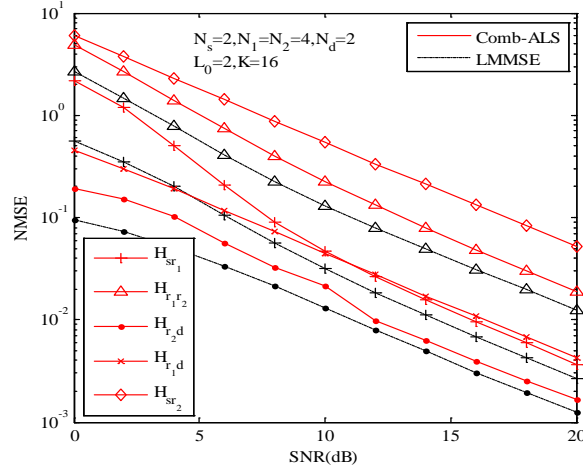
**Fig. 2.** Convergence of the proposed Comb-ALS algorithm

We begin by evaluating the convergence of the proposed Comb-ALS algorithm. Considering the first propagation scenario, where the parameters are set to  $N_s = 2$ ,  $N_1 = N_2 = 4$ ,  $N_d = 6$ ,  $L_0 = 2$  and  $K = 16$ . The NMSE value of estimated compound channel matrix  $\mathbf{H}_c$  is plotted as a function of the iterations for various SNR values. It can be seen from **Fig. 2** that the algorithm rapidly converges, usually within 12 iterations.

The NMSE performance of all the channel matrices of the proposed scheme is shown in **Fig. 3** and **Fig. 4**, where the solid line, the dash-dot line and the dot line represent the proposed Comb-ALS scheme, LMMSE scheme and LS scheme, respectively. In **Fig. 3**, the system parameters are  $N_s = 2$ ,  $N_1 = N_2 = 4$ ,  $N_d = 4$ ,  $L_0 = 2$  and  $K = 16$ . As can be seen from **Fig. 3**, the NMSE decreases as SNR increases. As expected, the estimation of  $\mathbf{H}_{sr_1}$ ,  $\mathbf{H}_{r_1r_2}$  and  $\mathbf{H}_{r_2d}$  is improved by carrying out the additional LMMSE estimation. At high SNR level, the estimation of  $\mathbf{H}_{sr_1}$  and  $\mathbf{H}_{r_2d}$  yields better performance than other links, the reason is that both



**Fig. 3.** NMSE versus SNR for the antenna number of destination node equals to that of the relay nodes



**Fig. 4.** NMSE versus SNR for the antenna number of destination node less than that of relay nodes

the channel matrices are involved in all the commutation phases and additional samples can be provided to improve the performance. In addition, compared to the typical LS channel estimation scheme, the proposed algorithm yields a more accurate performance. In fact, the estimation of  $\mathbf{H}_{sr_1}$  and  $\mathbf{H}_{sr_2}$  in LS scheme has an error floor. The reason is that the estimation of  $\mathbf{H}_{r_2d}$  and  $\mathbf{H}_{r_1r_2}$  in the LS scheme presents an error propagation for the estimation of  $\mathbf{H}_{sr_1}$  and  $\mathbf{H}_{sr_2}$ , thus affecting the total NMSE performance. In **Fig. 4**, we investigate the scenario in which the antenna number of the destination node is less than that of destination node, i.e.  $N_d = 2$ . In this scenario, the sequential LS-based channel estimation algorithm exploited in [14, 15, 21] cannot be used as the channel estimation problem will be a rank-deficient problem when the antenna number at the destination node is less than that of relay nodes. From **Fig. 4**, we can also see that the NMSE performance of all links decreases when SNR increases.

The impact of antenna number at the destination node on the system BER performance is shown in **Fig. 5**. The other system parameters are fixed to  $N_s = 2$ ,  $N_1 = N_2 = 4$ ,  $L_0 = 2$ ,  $K = 16$ . The dash-dot line represents  $N_d = 2$  while the solid line represents the scheme of  $N_d = 4$ , respectively. It can be seen that the BER performance of proposed Comb-ALS algorithm is considerably improved as the antenna number of destination node increases, especially at high SNR levels. As the proposed communication process consists two additional communication links, i.e. the  $S-R_1-D$  link and the  $S-R_2-D$  link, a diversity gain can be further obtained compared to the PARATUCK2-based one-way three-hop channel estimation scheme. For instance, when  $N_d = 4$ , at target BER of  $10^{-2}$ , the SNR gap between the LMMSE case and Comb-ALS case is more than 1dB.

In **Fig. 6**, we investigate the number of relay antennas on the BER performance of the ZF detector using the proposed channel estimator. The other system parameters are set to  $N_s = 2$ ,  $N_d = 6$ ,  $L_0 = 2$  and  $K = 20$ . The dash-dot line represents the number of antenna at relay nodes is  $N_1 = N_2 = 2$  while the solid line represents  $N_1 = N_2 = 4$ , respectively. It can be seen that the

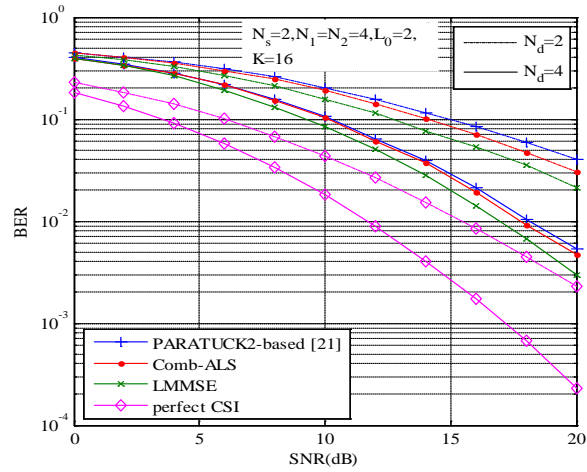


Fig. 5. BER performance versus the antenna number at the destination node

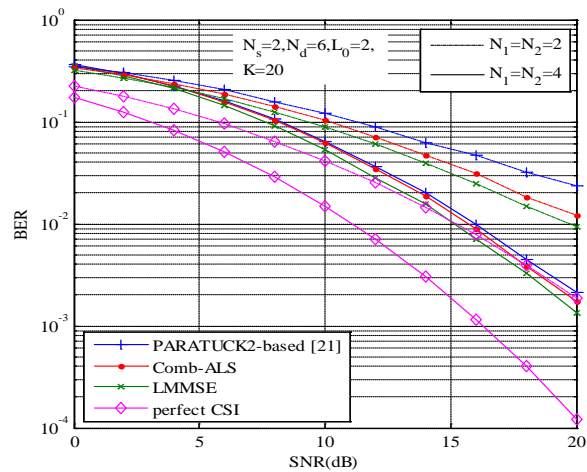
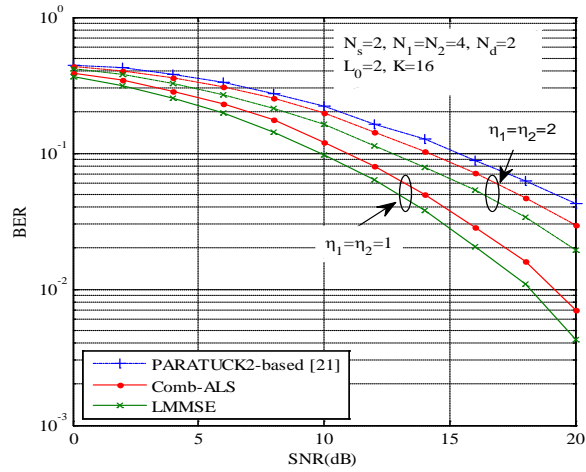


Fig. 6. BER performance versus the antenna number at the relay nodes

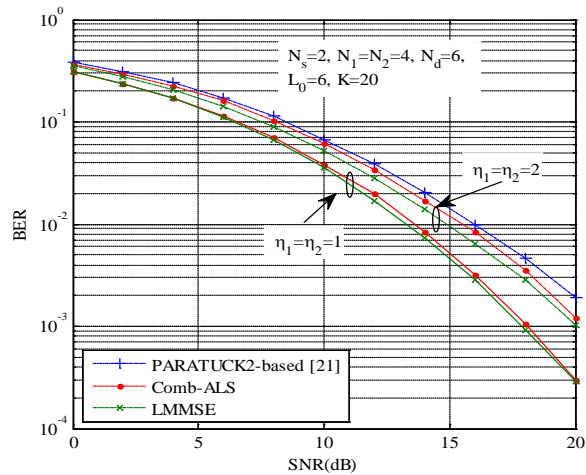
BER performance is considerably improved as the number of antenna at relay nodes is increased, owing to the higher spatial diversity. As expected, we can also see that additional LMMSE process can further improve the BER performance. But with the number of antenna at relay nodes increasing, the gap between the proposed scheme and the scheme in [21] is narrowing.

In Fig. 7 and Fig. 8, the impact of the distribution of relay nodes on the achievable gain of the proposed communication scheme is investigated in two special communication scenarios. In the best scenario  $\eta_1 = \eta_2 = 1$ , which means both the first relay node and the second relay node lie in the middle of the SD link, thus has the smallest path loss effect. Obviously, with the quality of the two-hop link increasing, the achievable gain of the system is superior to the scenario when  $\eta_1 = \eta_2 = 2$ .





**Fig. 7.** BER performance versus different  $(\eta_1, \eta_2)$  in the case that the antenna number at the destination node less than the relays



**Fig. 8.** BER performance versus different  $(\eta_1, \eta_2)$  in the case that the antenna number at the destination node greater than the relays

### 5. Conclusion

We have proposed a novel channel estimation algorithm for multi-hop relaying communication systems, where both the one-way two-hop communication link and one-way three-hop link exist. The proposed algorithm provides the destination node with full knowledge of all channel matrices involved in the communication, suffering from smaller channel estimation error. Compared to existing one-way three-hop approaches, the proposed scheme is able to make a more efficient use of cooperative diversity by jointly estimating the channel matrices of all links via PARAFAC and PARATUCK2 tensor model in an iterative way. Our simulation results verify the effectiveness of the proposed algorithms in terms of the convergence speed, NMSE and BER performance.

## References

- [1] M. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: Challenges, solutions, and future directions," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 86-92, May, 2014. [Article \(CrossRef Link\)](#)
- [2] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective," *IEEE Wireless Communications*, vol. 21, no. 3, pp. 118-127, June, 2014. [Article \(CrossRef Link\)](#)
- [3] C. Xing, F. Gao, and Y. Zhou, "A framework for transceiver designs for multi-hop communications with covariance shaping constraints," *IEEE Transactions on Signal Processing*, vol. 63, no. 15, pp. 3930-3945, August, 2015. [Article \(CrossRef Link\)](#)
- [4] Y. Rong, "Optimal joint source and relay beamforming for MIMO relays with direct link," *IEEE Communications Letters*, vol. 14, no. 5, pp. 390-392, May, 2010. [Article \(CrossRef Link\)](#)
- [5] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear non-regenerative multicarrier MIMO relay communication systems," *IEEE Transactions on Signal Processing*, vol. 57, no. 12, pp. 4837-4851, December, 2009. [Article \(CrossRef Link\)](#)
- [6] A. S. Behbahani, R. Merched, and A. M. Eltaswil, "Optimizations of MIMO relay network," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 5062-5073, October, 2007. [Article \(CrossRef Link\)](#)
- [7] F. Héliot, "Low-complexity energy-efficient joint resource allocation for two-hop MIMO-AF systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 6, pp. 3088-3099, June, 2014. [Article \(CrossRef Link\)](#)
- [8] C. Xing, S. Ma, Z. Fei, Y. Wu, and H. V. Poor, "A general robust linear transceiver design for multi-hop amplify-and-forward MIMO relaying systems," *IEEE Transactions on Signal Processing*, vol. 61, no. 5, pp. 1196-1209, March, 2013. [Article \(CrossRef Link\)](#)
- [9] C. Xing, S. Ma, and Y. Zhou, "Matrix-monotonic optimization for MIMO systems," *IEEE Transactions on Signal Processing*, vol. 63, no. 2, pp. 334-348, January, 2015. [Article \(CrossRef Link\)](#)
- [10] L. Wang, Y. Cai, W. Yang, W. Yan, and J. Song, "Effect of outdated channel estimates on multiple antennas multiple relaying networks," *KSI Transactions on Internet and Information Systems*, vol. 9, no. 5, pp. 1682-1701, May, 2015. [Article \(CrossRef Link\)](#)
- [11] T. Kong and Y. Hua, "Optimal design of source and relay pilots for MIMO relay channel estimation," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4438-4446, September, 2011. [Article \(CrossRef Link\)](#)
- [12] J. S. Sheu, J. K. Lain, and W. H. Wang, "On Channel estimation of orthogonal frequency-division multiplexing amplify-and-forward cooperative relaying systems," *IET Communications*, vol. 7, no. 4, pp. 325-334, March, 2013. [Article \(CrossRef Link\)](#)
- [13] O. Amin, B. Gedik, and M. Uysal, "Channel estimation for amplify-and-forward relaying: cascaded against disintegrated estimators," *IET Communications*, vol. 4, no. 10, pp. 1207-1216, July, 2010. [Article \(CrossRef Link\)](#)
- [14] Y. Rong, M. R. A. Khandaker, and Y. Xiang, "Channel estimation of dual-hop MIMO relay system via parallel factor analysis," *IEEE Transactions on Wireless Communications*, vol. 11, no. 6, pp. 2224-2233, June, 2012. [Article \(CrossRef Link\)](#)
- [15] J. Du, C. Yuan, and J. Zhang, "Low complexity PARAFAC-based channel estimation for non-regenerative MIMO relay systems," *IET Communications*, vol. 8, no. 12, pp. 2193-2199, August, 2014. [Article \(CrossRef Link\)](#)
- [16] C. A. Rolim Fernandes, A. L. F. De Almeida, and D. Benevides Da Costa, "Unified tensor modeling for blind receivers in multiuser uplink cooperative systems," *IEEE Signal Processing Letters*, vol. 19, no. 5, pp. 247-250, May, 2012. [Article \(CrossRef Link\)](#)
- [17] F. Roemer and M. Haardt, "Tensor-based channel estimation and iterative refinements for two-way relaying with multiple antennas and spatial reuse," *IEEE Transactions on Signal Processing*, vol. 58, no. 11, pp. 5720-5735, November, 2010. [Article \(CrossRef Link\)](#)

- [18] R. A. Harshman, "Foundations of the PARAFAC procedure: model and conditions for an "explanatory" multi-mode factor analysis," *UCLA Working Papers in Phonetics*, vol. 16, no. 1, pp. 1-84, 1970.
- [19] L. R. Ximenes, G. Favier, A. L. F. De Almeida, and Y. C. B. Silva, "PARAFAC-PARATUCK semi-blind receivers for two-hop cooperative MIMO relay systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 14, pp. 3604-3615, July, 2014. [Article \(CrossRef Link\)](#)
- [20] I. V. Cavalcante, A. L. F. De Almeida, and M. Haardt, "Tensor-based approach to channel estimation in amplify-and-forward MIMO relaying systems," in *Proc. of 2014 IEEE 8th Sensor Array and Multichannel Signal Processing Workshop*, pp. 445-448, June 22-25, 2014. [Article \(CrossRef Link\)](#)
- [21] X. Han, A. L. F. De Almeida, and Z. Yang, "Channel estimation for MIMO multi-relay systems using a tensor approach," *Eurasip Journal on Advances in Signal Processing*, vol. 2014, no. 1, pp. 1-14, December, 2014. [Article \(CrossRef Link\)](#)
- [22] 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, September, 2015. [Article \(CrossRef Link\)](#)
- [23] J. Du, C. Yuan, and J. Zhang, "Semi-blind parallel factor based receiver for joint symbol and channel estimation in amplify-and-forward multiple-input multiple-output relay systems," *IET Communications*, vol. 9, no. 6, pp. 737-744, April, 2015. [Article \(CrossRef Link\)](#)
- [24] R. Bro, "Multi-way analysis in the food industry: Models, algorithms and applications," *Ph.D. dissertation*, Univ. Amsterdam, Amsterdam, The Netherlands, 1998.
- [25] Tamara G. Kolda and Brett W. Bader, *Tensor Decompositions and Applications*. Siam Review, pp. 48-50, 2009. [Article \(CrossRef Link\)](#)
- [26] G. H. Golub and V. C. F. Loan, *Matrix Computations*. Baltimore, MD, USA: Johns Hopkins University Press, 1996.
- [27] J. Zhang, X. Mu, E. Chen, and S. Yang, "Decision-directed channel estimation based on iterative linear minimum mean square error for orthogonal frequency division multiplexing systems," *IET Communications*, vol. 3, no. 7, pp. 1136-1143, July, 2009. [Article \(CrossRef Link\)](#)



**Shuangzhi Li** received his B.E. Degree in Electronic Information Engineering from Zhengzhou University, Zhengzhou, China in 2012. He is currently working toward the Ph. D. Degree with the School of Information Engineering, Zhengzhou University, China. His research interests include signal processing in cooperation communications and MIMO systems.



**Xiaomin Mu** received her B.E. Degree from the Beijing Institute of Technology, Beijing, China in 1982. She is currently a full professor with the School of Information Engineering, Zhengzhou University. She has published many papers in the field of signal processing and co-authored two books. Her research interests include signal processing in communication systems, wireless communications and cognitive radio.



**Xin Guo** received her B.E. Degree in Communication Engineering from The PLA Information Engineering University, Zhengzhou, China in 2010. She is currently working toward the Ph. D. Degree with the School of Information Engineering, Zhengzhou, China. Her research interests include image signal processing, pattern recognition and optimization algorithms.



**Jing Yang** received her Ph.D. Degree from Beijing Institute of Technology, Beijing, China in 2011. Since then, she joined the College of Information Science and Engineering, Henan University of Technology. Now, she is a lecturer. Her research interests include cooperative communication systems and rateless codes.



**Jiankang Zhang** received the B.Sc. Degree in Mathematics and Applied Mathematics from Beijing University of Posts and Telecommunications in 2006, and the Ph. D. Degree in Communication and Information Systems from Zhengzhou University in 2012. Since then, he has been a lecturer in School of Information Engineering, Zhengzhou University. From September 2009 to December 2011 and from January 2013 to May 2013, Dr. Zhang was a visiting researcher in Electronics and Computer Science, the University of Southampton, UK. His research interests are in the areas of wireless communications and signal processing, including channel estimation, multi-user detection, beamforming/precoding and optimization algorithms.