Semi-Blind Receivers for Joint Symbol and Channel Estimation in Space-Time-Frequency MIMO-OFDM Systems

Kefei Liu, *Student Member, IEEE*, João Paulo C. L. da Costa, *Member, IEEE*, H. C. So, *Senior Member, IEEE*, and André L. F. de Almeida, *Senior Member, IEEE*

***Abstract*—In wireless communications, increased spectral efficiency and low error rates can be achieved by means of space-time-frequency coded MIMO-OFDM systems. In this work, we consider a MIMO-OFDM transmit signal design combining space-frequency modulation with a time-varying linear precoding technique which allows spreading and multiplexing the transmitted symbols, in both space, time and frequency domains. For this system, we propose two closed-form semi-blind receivers that exploit differently the multilinear structure of the received signal, which is formulated as a nested PARAllel FACtor (PARAFAC) model. First, we devise a least squares Khatri-Rao factorization (LS-KRF) based receiver for joint channel and symbol estimation by making an efficient use of a short frame of pilot symbols. The LS-KRF receiver provides the same performance at a lower computational complexity compared to the alternating least squares (ALS) based receiver. For further reducing pilot overhead, we develop a simplified closed-form PARAFAC (S-CFP) receiver coupled with a pairing algorithm that yields an unambiguous estimation of the transmitted symbols without the need of a pilot frame. The uniqueness conditions, spectral efficiency and computational complexity of the LS-KRF and S-CFP with pairing receivers are analyzed and compared with the ALS receiver. It is shown that the S-CFP with pairing receiver has the same order of computational complexity as the ALS receiver. Meanwhile, simulation results show that our S-CFP with pairing receiver achieves the same or very similar performance of the competing receivers with extra pilot overhead at sufficiently high signal-to-noise ratio (SNR) conditions. On the other hand, it is slightly inferior to them in terms of channel estimation accuracy and bit error rate at lower SNRs.**

***Index Terms*—Closed-form PARAFAC, least squares Khatri-Rao factorization, MIMO-OFDM, nested PARAFAC, semi-blind receiver.**

I. Introduction

THE combination of multiple-input multiple-output (MIMO) systems and orthogonal frequency division multiplexing (OFDM) has been an important research topic [1]. In MIMO-OFDM, multiple transmit and receive antennas are employed to achieve high data rates via spatial multiplexing as well as improved link reliability in frequency-selective channels through space-time/space-frequency or space-time-frequency coding [2]–[7]. Matrix-based space-time/space-frequency coding methods for OFDM systems with blind or semi-blind detection have been proposed in the past few years in a number of works (See, e.g., [8], [9] and the references therein). The existing matrix-based solutions rely either on computationally demanding maximum likelihood detection strategies or on lower complexity detection strategies.

Multidimensional signaling schemes that utilize several signal dimensions such as space, time, frequency and constellation, are good candidates for increasing the data rate and/or improving the link reliability in future communication systems [10]–[20]. In a number of recent works, tensor decompositions are exploited in space-time coding/precoding based MIMO systems for blind channel identification. In [10], a linear space-time coding based on the Khatri-Rao matrix product has been suggested in an effort to achieve a flexible tradeoff between error performance in terms of bit or symbol error rate and transmission efficiency, for arbitrary transmit-receive antenna configuration or signal constellation. A joint blind channel estimation and symbol detection scheme is designed based on the parallel factor (PARAFAC) modeling [21], [22] of the received data. In [11], the idea of [10] for flat fading channels is extended to frequency selective channels. In [12], a three-dimensional (3-D) scheme that combines spatial multiplexing and space-time coding has been designed. The received signals obey a 3-D Tucker-2 tensor model [23]. Unlike [10], the number of data streams of [12] can be different from the number of transmit antennas. Moreover, by using a 3-D tensor coding instead of two matrix codings, the symbol detection at the receiver is appreciably simplified.

More general space-time codings include resource allocation [14]–[16], leading to 3-D constrained tensor models for the received signals. In the constrained tensor model, the constraint or allocation matrices define the allocation of user data streams and spreading codes to the transmit antennas, allowing variations in degree of spatial spreading and multiplexing. In [14], two allocation matrices are employed to control the spatial spreading of the data streams and the spatial reuse of the spreading codes. The received signals satisfy a 3-D constrained Tucker tensor model [23] whose core tensor is defined in terms of the two allocation matrices. The approach of [15] is a generalization of [14] in that the transceiver is defined as a joint stream-code-antenna

Manuscript received July 11, 2012; revised April 02, 2013 and April 02, 2013; accepted August 05, 2013. Date of publication August 15, 2013; date of current version September 27, 2013. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Rong-Rong Chen. The work described in this paper was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project no. CityU 120911), and by the CNPq/Brazil (Proc. 303238/2010-0).

K. Liu and H. C. So are with the Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong (e-mail: kefeiliu@ieee.org; hcso@ee.cityu.edu.hk).

J. P. C. L. da Costa is with the Department of Electrical Engineering, University of Brasília, 70910-900 Brasília, Brazil (e-mail: jpdacosta@unb.br).

A. L. F. de Almeida is with the GTEL-Wireless Telecom Research Group, Federal University of Ceará, Brazil (e-mail: andre@gtel.ufc.br).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2013.2278512

multiplexer that is decomposable into three allocation matrices. In [16], the received signals are formulated as a PARATUCK-2 tensor model [24]. The core tensor of this decomposition, which is composed of a precoding matrix and two allocation matrices, is exploited to control space-time multiplexing and diversity at the transmitter, while at the receiver the multilinear algebraic structure of the resultant PARATUCK-2 model is employed for joint blind symbol and channel estimation. The work of [16] has been generalized to space-time-frequency (STF) coding [19]. However, the common feature of these works is the use of iterative estimation algorithms at the receiver, which are computationally demanding, even for a moderate number of transmit antennas and/or data streams. Moreover, transmission efficiency and processing delay are key issues which are not addressed in the literature. Herein, we are particularly interested in non-iterative closed-form solutions for joint symbol and channel estimation in MIMO-OFDM systems that are more attractive from a practical viewpoint.

In this work, we consider a new STF coded MIMO-OFDM system where the transmit signal design combines frequency-domain modulation with a time-varying linear constellation precoding. This STF coding allows spreading and multiplexing the transmitted symbols, in both space, time and frequency domains, through the use of a fourth-order code tensor incorporating transmit antenna, time-slot, symbol period and subcarrier as modes. Assuming flat fading propagation channels, the received data are formulated as a three-way nested PARAFAC model. In [20], we have developed a closed-form receiver algorithm for joint channel and symbol estimation based on the least squares Khatri-Rao factorization (LS-KRF) [25]. It is shown that the LS-KRF receiver is computationally simpler than the traditional alternating least squares (ALS) solution [17] in solving the same problem.

This paper extends [20] in two new aspects. First, we present a parallel way to implement the LS-KRF receiver. The parallelized LS-KRF (P-LS-KRF) receiver remarkably saves the processing time required in LS-KRF via parallel computing [26], which renders it suitable for systems where processing delay is a key concern. Both the ALS and (P-) LS-KRF assume that the first row in the symbol matrix is known. This implies that a time-slot in each frame is dedicated to pilot symbol transmission, which results in a loss of transmission efficiency. As a second contribution, we relax this assumption and propose a simplified closed-form PARAFAC (S-CFP) receiver coupled with a pairing algorithm in an effort to improve the transmission efficiency. Compared with the CFP decomposition [27] that needs at least two simultaneous matrix diagonalizations (SMDs), S-CFP only requires one SMD to calculate the third factor matrix, while the other two factor matrices are estimated based on LS-KRF. Since the SMD is the most computationally expensive part in CFP, a significant reduction of complexity is achieved in our proposal. In order to correctly extract the symbol matrix, we devise a novel algorithm to remove the permutation and scaling ambiguities in the factor matrix estimates obtained by S-CFP. The S-CFP with pairing receiver effectively avoids the loss in transmission efficiency, and meanwhile it attains almost the same performance as the ALS and LS-KRF receiver with pilot overhead in terms of bit error rate (BER) and channel estimation accuracy for high signal-to-noise ratio (SNR) regimes.

The remainder of this paper is organized as follows. In Section II, the transmit and receive signal models of the system are described. In particular, the received data are formulated as a nested PARAFAC model. This model is exploited for the joint symbol and channel estimation in Section III, where we briefly review the ALS receiver. In Section IV, we describe our proposed LS-KRF and simplified CFP with pairing receivers. In Section V, the uniqueness conditions, spectral efficiency and computational complexity of the proposed receivers are analyzed and compared with those of the ALS receiver. Simulation results are provided in Section VI, and conclusion is drawn in Section VII.

Notation

The superscripts $^{\mathrm{T}}$, $^{\mathrm{H}}$, †, and * represent transpose, Hermitian transpose, pseudo-inverse and complex conjugate, respectively. The r-th column of $\boldsymbol{A} \in \mathbb{C}^{J\times R}$ is denoted by $\boldsymbol{A}(:,r) \in \mathbb{C}^{J\times 1}$. The operator $\mathrm{diag}(\boldsymbol{a})$ forms a diagonal matrix based on $\boldsymbol{a}$. The operator $\mathrm{vec}(\boldsymbol{A})$ converts $\boldsymbol{A}$ to a vector $\boldsymbol{a}$ by stacking its columns on top of each other, while $\mathrm{unvec}_{J\times R}(\boldsymbol{a})$ denotes the inverse vectorization operation that converts $\boldsymbol{a} \in \mathbb{C}^{JR\times 1}$ back to a $J\times R$ matrix. The operator $\mathrm{vecdiag}(\boldsymbol{D})$ forms a vector $\boldsymbol{d} \in \mathbb{C}^{R\times 1}$ from the diagonal elements of matrix $\boldsymbol{D} \in \mathbb{C}^{R\times R}$, while $D_j(\boldsymbol{A})$ is a diagonal matrix constructed from the j-th row of $\boldsymbol{A}$. The $\|\cdot\|$ denotes the Euclidean norm of a vector, while $\|\cdot\|_{\mathrm{F}}$ represents the Frobenius norm of a matrix or tensor, which is defined as the square root of the sum of the squared amplitudes of its elements.

Moreover, the Kronecker product and outer product operators are denoted by $\otimes$ and $\circ$, respectively. The Khatri-Rao product between two matrices $\boldsymbol{A} \in \mathbb{C}^{J\times R}$ and $\boldsymbol{B} = [\boldsymbol{b}_1 \ \ldots \ \boldsymbol{b}_R] \in \mathbb{C}^{K\times R}$, denoted by $\diamond$, is their column-wise Kronecker product

$$\boldsymbol{A} \diamond \boldsymbol{B} \doteq [\boldsymbol{A}(:,1) \otimes \boldsymbol{B}(:,1) \ldots \boldsymbol{A}(:,R) \otimes \boldsymbol{B}(:,R)]. \quad (1)$$

In this paper, the following property of the Khatri-Rao product is used

$$\boldsymbol{A}\mathrm{diag}(\boldsymbol{x}^{\mathrm{T}})\boldsymbol{B}^{\mathrm{T}} = \left(\boldsymbol{x}^{\mathrm{T}} \diamond \boldsymbol{A}\right)\boldsymbol{B}^{\mathrm{T}}, \quad (2)$$

where $\boldsymbol{x}$ is an R-D column vector.

The tensor operations are consistent with [28]: The r-mode unfolding of a tensor $\boldsymbol{\mathcal{T}} \in \mathbb{C}^{J_1\times J_2\times\cdots\times J_R}$, symbolized by $[\boldsymbol{\mathcal{T}}]_{(r)} \in \mathbb{C}^{J_r\times(J_1J_2\ldots J_{r-1}J_{r+1}\ldots J_R)}$, represents the matrix containing the r-mode vectors of $\boldsymbol{\mathcal{T}}$. The order of the columns is chosen in accordance with [28]. The r-mode product of $\boldsymbol{\mathcal{T}}$ and a matrix $\boldsymbol{U} \in \mathbb{C}^{K_r\times J_r}$ along the r-th mode is denoted as $\boldsymbol{\mathcal{T}} \times_r \boldsymbol{U} \in \mathbb{C}^{J_1\times J_2\ldots\times K_r\cdots\times J_R}$, which is obtained by multiplying the r-mode unfolding of $\boldsymbol{\mathcal{T}}$ from the left hand side (LHS) by $\boldsymbol{U}$.

II. System Model

Consider a MIMO-OFDM wireless communication system employing M transmit antennas and K receive antennas. The transmission time-frame is composed of a collection of N short time-slots of P symbol periods each. The block diagram of the transceiver system is shown in Fig. 1.

A. Transmit Signal Model

The transmit signal design combines the time-varying linear constellation precoding with the frequency-domain Vander-

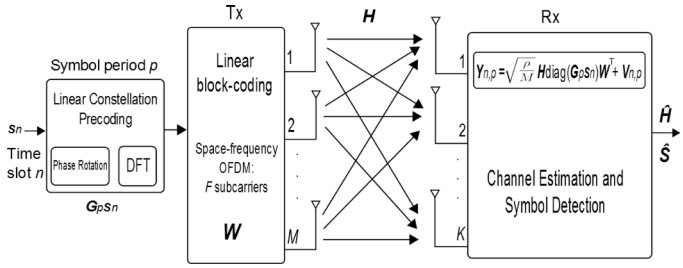


Fig. 1. System model block diagram.

monde spreading. The information symbol stream is first divided into blocks of symbols $\boldsymbol{s}_n \in \mathbb{C}^{M\times 1}, n = 1,\ldots,N$, each of which is linearly precoded across P symbol periods by means of a set of $M \times M$ unitary space-time modulation matrices $\{\boldsymbol{G}_p | p = 1,\ldots,P\}$. During the p-th symbol period of the n-th time-slot, $\boldsymbol{G}_p$ rotates the components of the symbol vector $\boldsymbol{s}_n$ and loads a linear combination of these components into the M transmit antennas.

Following [17], [20], we have

$$\boldsymbol{G}_p = \boldsymbol{\Theta}\mathrm{diag}(\boldsymbol{a}_p), \tag{3}$$

where $\boldsymbol{\Theta} \in \mathbb{C}^{M\times M}$ is the discrete Fourier transform (DFT) matrix, and $\boldsymbol{a}_p = [1\ e^{j\phi_p}\ \ldots\ e^{j(M-1)\phi_p}] \in \mathbb{C}^{1\times M}$ is a phase rotation vector, where ϕ_p denotes the random initial phase of the p-th symbol period at the transmitter. Consequently, all $\{\phi_p\}$ are unknown at the receiver except for ϕ_1, which is the reference phase whose value equals 0. As noticed in [17], [20], this model is a generalization of [29], [30] in the sense that the rotation vectors are time varying across different symbol periods. This feature allows channel reuse over P symbol periods and hence realizes time-domain modulation diversity.

The information symbol vector is then spread across a group of F neighboring frequency bins (subcarriers) over which the channel is considered to be invariant. Such a spreading operation is performed by a semi-unitary matrix with a Vandermonde structure, of which the (f, m) entry is given by [10], [17]

$$[\boldsymbol{W}]_{f,m} = e^{j(f-1)(m-1)\frac{2\pi}{M}}. \tag{4}$$

The above space-frequency coding operation is summarized as [10]

$$\boldsymbol{C}_{n,p} = \mathrm{diag}(\boldsymbol{G}_p\boldsymbol{s}_n)\boldsymbol{W}^{\mathrm{T}} \in \mathbb{C}^{M\times F}, \tag{5}$$

where $\boldsymbol{C}_{n,p}$ is the space-frequency code matrix transmitted during the p-th symbol period of the n-th time-slot. This space-frequency coding operation has the same structure as that proposed in [10] for space-time coding.

B. Receive Signal Model

Assuming that the channel is constant over a time-frame, the discrete-time baseband equivalent model for the received signal is given by

$$\begin{aligned}\boldsymbol{Y}_{n,p} &= \sqrt{\frac{\rho}{M}}\boldsymbol{H}\boldsymbol{C}_{n,p} + \boldsymbol{V}_{n,p} \\ &= \sqrt{\frac{\rho}{M}}\boldsymbol{H}\mathrm{diag}(\boldsymbol{G}_p\boldsymbol{s}_n)\boldsymbol{W}^{\mathrm{T}} + \boldsymbol{V}_{n,p},\end{aligned} \tag{6}$$

where $\boldsymbol{Y}_{n,p} \in \mathbb{C}^{K\times F}, n = 1,\ldots,N, p = 1,\ldots,P$, denotes the complex received signal matrix during the p-th symbol period of the n-th time-slot, and $\boldsymbol{V}_{n,p}$ contains the zero-mean circularly symmetric complex Gaussian (ZMCSCG) white noise. The channel matrix $\boldsymbol{H} \in \mathbb{C}^{K\times M}$ has uncorrelated ZMCSCG random entries with variance of 1, and ρ represents the SNR at each receive antenna.

Substituting (3) into (6), the receiver signal is expressed as

$$\boldsymbol{Y}_{n,p} = \sqrt{\frac{\rho}{M}}\boldsymbol{H}D_p(\boldsymbol{A}D_n(\boldsymbol{S})\boldsymbol{\Theta}^{\mathrm{T}})\boldsymbol{W}^{\mathrm{T}} + \boldsymbol{V}_{n,p}, \tag{7}$$

where $\boldsymbol{S} = [\boldsymbol{s}_1\ \ldots\ \boldsymbol{s}_N]^{\mathrm{T}} \in \mathbb{C}^{N\times M}$ denotes the symbol matrix which contains the N symbol vectors in a time-frame, and $\boldsymbol{A} \in \mathbb{C}^{P\times M}$ collects all phase rotation vectors throughout the P symbol periods, namely,

$$\boldsymbol{A} = \begin{bmatrix}\boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_P\end{bmatrix} = \begin{bmatrix}1 & 1 & \ldots & 1 \\ 1 & e^{j\phi_2} & \ldots & e^{j(M-1)\phi_2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{j\phi_P} & \ldots & e^{j(M-1)\phi_P}\end{bmatrix}. \tag{8}$$

Let $\boldsymbol{Y}_n = [\boldsymbol{Y}_{n,1}^{\mathrm{T}}\ \ldots\ \boldsymbol{Y}_{n,P}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{PK\times F}$. According to (2), it holds that

$$\boldsymbol{Y}_n = \sqrt{\frac{\rho}{M}}[(\boldsymbol{A}D_n(\boldsymbol{S})\boldsymbol{\Theta}^{\mathrm{T}}) \diamond \boldsymbol{H}]\boldsymbol{W}^{\mathrm{T}} + \boldsymbol{V}_n. \tag{9}$$

Likewise, letting $\boldsymbol{Y} = [\boldsymbol{Y}_1^{\mathrm{T}}\ \ldots\ \boldsymbol{Y}_N^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{NPK\times F}$ to collect the received data of all N time-slots yields

$$\boldsymbol{Y} = \sqrt{\frac{\rho}{M}}\,(\boldsymbol{T} \diamond \boldsymbol{H})\,\boldsymbol{W}^{\mathrm{T}} + \boldsymbol{V}. \tag{10a}$$

where

$$\boldsymbol{T} = (\boldsymbol{S} \diamond \boldsymbol{A})\boldsymbol{\Theta}^{\mathrm{T}} \in \mathbb{C}^{NP\times M}. \tag{10b}$$

The $\boldsymbol{Y}$ is an unfolded representation of a fourth-order received signal tensor of size $K \times F \times N \times P$, where the four dimensions correspond to space, frequency, time-slot and symbol period, respectively. Note that this tensor can be expressed in terms of r-mode product as

$$\boldsymbol{\mathcal{Y}} = \sqrt{\frac{\rho}{M}} \cdot \boldsymbol{\mathcal{I}}_M \times_1 \boldsymbol{H} \times_2 \boldsymbol{W} \times_3 \boldsymbol{T} + \boldsymbol{\mathcal{V}} \in \mathbb{C}^{K\times F\times NP}, \tag{11}$$

where $\boldsymbol{\mathcal{V}}$ is the noise tensor. Hereafter, the constant $\sqrt{\rho/M}$ is neglected for ease of presentation.

The tensor form of (10b) can also be written as

$$\boldsymbol{\mathcal{T}} = \boldsymbol{\mathcal{I}}_M \times_1 \boldsymbol{S} \times_2 \boldsymbol{A} \times_3 \boldsymbol{\Theta} \in \mathbb{C}^{N\times P\times M}, \tag{12}$$

where $\boldsymbol{\mathcal{I}}_M$ represents the identity tensor of size $M \times M \times M$.

It is now clear that the signal in (10) corresponds to the nested PARAFAC decomposition: (10a) is the outer PARAFAC part while (10b) is the inner PARAFAC part. According to the r-mode unfolding definition [28], we have $\boldsymbol{\mathcal{T}} = [\boldsymbol{\mathcal{T}}]_3^{\mathrm{T}}$ and $\boldsymbol{Y} = [\boldsymbol{\mathcal{Y}}]_2^{\mathrm{T}}$.

Given the noisy measurement $\boldsymbol{\mathcal{Y}}$, our objective is to estimate the symbol matrix $\boldsymbol{S}$ and channel matrix $\boldsymbol{H}$. Note that here we consider a semi-blind model since the frequency block-coding

TABLE I
ALS RECEIVER ALGORITHM

First stage
Initialization: Set $i = 0$; Randomly initialize $\widehat{\boldsymbol{T}}$;
(1.1) $i = i + 1$;
(1.2) Compute the LS estimate of $\boldsymbol{H}$:
$$\widehat{\boldsymbol{H}}(i) = [\boldsymbol{\mathcal{Y}}]_1 \cdot [(\boldsymbol{W} \diamond \widehat{\boldsymbol{T}}(i-1))^\dagger]^{\mathrm{T}};$$
(1.3) Compute the LS estimate of $\boldsymbol{T}$:
$$\widehat{\boldsymbol{T}}(i) = [\boldsymbol{\mathcal{Y}}]_3 \cdot [(\widehat{\boldsymbol{H}}(i) \diamond \boldsymbol{W})^\dagger]^{\mathrm{T}};$$
(1.4) Repeat steps (1.1)-(1.3) until convergence.

Second stage
Initialization: Set $j = 0$; Randomly initialize $\widehat{\boldsymbol{A}}$;
(2.1) $j = j + 1$;
(2.2) Compute the LS estimate of $\boldsymbol{S}$:
$$\widehat{\boldsymbol{S}}(j) = [\boldsymbol{\mathcal{T}}]_1 \cdot [(\widehat{\boldsymbol{A}}(j-1) \diamond \boldsymbol{\Theta})^\dagger]^{\mathrm{T}};$$
(2.3) Compute the LS estimate of $\boldsymbol{A}$:
$$\widehat{\boldsymbol{A}}(j) = [\boldsymbol{\mathcal{T}}]_2 \cdot [(\boldsymbol{\Theta} \diamond \widehat{\boldsymbol{S}}(j))^\dagger]^{\mathrm{T}};$$
(2.4) Repeat steps (2.1)-(2.3) until convergence.

matrix $\boldsymbol{W}$ and the DFT matrix $\boldsymbol{\Theta}$ are known. As in [17], we focus on two system configurations:

$$P < M \leq N, \tag{13a}$$
$$N < M \leq P. \tag{13b}$$

III. STATE-OF-THE-ART ALGORITHM: ALS RECEIVER

In the ALS receiver [17], it is assumed that prior to data transmission, a pilot symbol per transmit antenna is inserted at each time-slot, namely, the first row of $\boldsymbol{S}$ is known. The ALS algorithm is described in Table I. It consists in two iterative least squares (LS) based estimation stages: one is applied in the outer PARAFAC part of the nested PARAFAC model, and the other one is applied in the inner PARAFAC part.

Note that after the first stage the estimates of $\boldsymbol{H}$ and $\boldsymbol{T}$, namely, $\widehat{\boldsymbol{H}}$ and $\widehat{\boldsymbol{T}}$, are column scaled versions of their true values. This scaling ambiguity in $\widehat{\boldsymbol{T}}$ must be resolved before proceeding into the inner PARAFAC so that $\boldsymbol{S}$ and $\boldsymbol{A}$ can be correctly estimated. To this end, the ALS solution [17] assumes that one time-slot in each time-frame is dedicated to pilot symbol transmission, namely, the first row of $\boldsymbol{S}$ is known as a vector of all ones. Together with the fact that the first row of $\boldsymbol{A}$ is also available, we know the first row of $\boldsymbol{T}$ according to (10b). Then the scaling ambiguity in $\widehat{\boldsymbol{T}}$ can be removed.

IV. PROPOSED ALGORITHMS

In this section, we propose two closed-form receivers: one is the P-LS-KRF receiver and the other is the S-CFP receiver coupled with a pairing algorithm. Similar to ALS, the P-LS-KRF receiver needs the knowledge of the first row of $\boldsymbol{S}$, while in the S-CFP with pairing receiver, this requirement is relaxed to avoid the extra pilot overhead and to improve the transmission efficiency.

A. Parallelized LS-KRF Receiver

1) Parallelized LS-KRF Receiver: Assume $F \geq M$, the closed-form LS-KRF solution [20] can be used for semi-blind symbol and channel recovery. Note that this assumption is relevant in a wide range of scenarios of small to moderate number of transmit antennas M. Here, we introduce a parallel way to implement the LS-KRF solution, which is referred to as P-LS-KRF.

First, $\boldsymbol{T}$ and $\boldsymbol{H}$ are estimated via the outer PARAFAC model. The 2-mode unfolding of $\boldsymbol{\mathcal{Y}}$ gives

$$[\boldsymbol{\mathcal{Y}}]_2 = \boldsymbol{W} \cdot [(\boldsymbol{T} \diamond \boldsymbol{H})]^{\mathrm{T}}. \tag{14}$$

Multiplying both sides of (14) by the pseudo-inverse of $\boldsymbol{W}$ from the LHS and then taking the transpose, we obtain

$$\boldsymbol{T} \diamond \boldsymbol{H} = (\boldsymbol{W}^\dagger \cdot [\boldsymbol{\mathcal{Y}}]_2)^{\mathrm{T}} \in \mathbb{C}^{NPK \times M}. \tag{15}$$

The unvectorization of the m-th column of $\boldsymbol{T} \diamond \boldsymbol{H}$, $m = 1, \ldots, M$, yields a matrix $\boldsymbol{Q}_m \in \mathbb{C}^{NP \times K}$. According to (1), in the absence of noise, $\boldsymbol{Q}_m$ is a rank-1 matrix

$$\boldsymbol{Q}_m = \boldsymbol{H}(:, m) \circ \boldsymbol{T}(:, m). \tag{16}$$

Hence we can apply the truncated singular value decomposition (SVD) on $\boldsymbol{Q}_m$ to find $\boldsymbol{H}(:, m) \in \mathbb{C}^{NP \times 1}$ and $\boldsymbol{T}(:, m) \in \mathbb{C}^{K \times 1}$. Based on the rank-1 property, we have

$$\boldsymbol{Q}_m \approx \boldsymbol{u}_m \cdot \sigma_m \cdot \boldsymbol{v}_m^{\mathrm{H}}, \tag{17}$$

where σ_m is the largest singular value of the m-th unvectorized matrix, and $\boldsymbol{u}_m$ and $\boldsymbol{v}_m$ are the corresponding left and right singular vectors. The estimates of $\boldsymbol{H}(:, m)$ and $\boldsymbol{T}(:, m)$ are given by

$$\widehat{\boldsymbol{T}}(:, m) = \sqrt{\sigma_m}\boldsymbol{u}_m \in \mathbb{C}^{NP \times 1} \tag{18a}$$

and

$$\widehat{\boldsymbol{H}}(:, m) = \sqrt{\sigma_m}\boldsymbol{v}_m^* \in \mathbb{C}^{K \times 1}. \tag{18b}$$

Analogous to the ALS solution, the scaling ambiguity in $\widehat{\boldsymbol{T}}$ can be removed provided that the first row in $\boldsymbol{T}$ is known.

The flow chart of the LS-KRF for the outer PARAFAC estimation is shown in Fig. 2. Note that in (16)–(18), the m-th, $m = 1, \ldots, M$, columns of $\widehat{\boldsymbol{H}}$ and $\widehat{\boldsymbol{T}}$ are computed from the m-th column of $\boldsymbol{T} \diamond \boldsymbol{H}$, and different columns operate independently and have no overlap in operands. This allows us to recover the M columns of $\widehat{\boldsymbol{H}}$ and $\widehat{\boldsymbol{T}}$ in a parallel way using M processors, which can reduce the processing time by a factor of M. Therefore, the parallelized LS-KRF is attractive in situations where low processing delay is needed and multiple processors are available.

With the newly obtained $\widehat{\boldsymbol{T}}$, we apply the LS-KRF solution to the inner PARAFAC model in a similar way. By multiplying both sides of (10b) by $\boldsymbol{\Theta}/M$ from the right hand side (RHS), we get

$$\widehat{\boldsymbol{S}} \diamond \widehat{\boldsymbol{A}} = \widehat{\boldsymbol{T}} \cdot \boldsymbol{\Theta}/M \in \mathbb{C}^{NP \times M}. \tag{19}$$

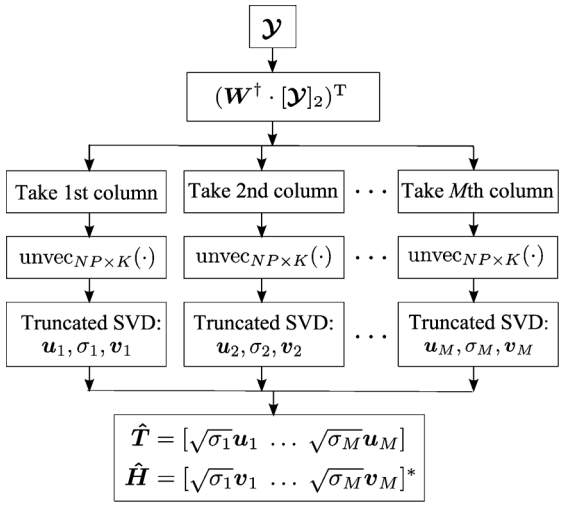


Fig. 2. Flow chart of LS-KRF for outer PARAFAC.

After unvectorizing the m-th column of $\left(\widehat{\boldsymbol{S}}\diamond\widehat{\boldsymbol{A}}\right)$ into a rank-1 matrix $\boldsymbol{Q}'_m\in\mathbb{C}^{N\times P}$, $m=1,\ldots,M$, the truncated SVD is exploited to find the dominant vectors $\boldsymbol{u}'_{\boldsymbol{m}}\in\mathbb{C}^{N\times 1}$, $\boldsymbol{v}'_{\boldsymbol{m}}\in\mathbb{C}^{P\times 1}$ and the largest singular value σ'_m. Consequently, we have

$$\widehat{\boldsymbol{S}}=[\sqrt{\sigma'_1}\boldsymbol{u}'_{\mathbf{1}}\ \ldots\ \sqrt{\sigma'_M}\boldsymbol{u}'_{\boldsymbol{M}}]\in\mathbb{C}^{N\times M} \tag{20}$$

$$\widehat{\boldsymbol{A}}=[\sqrt{\sigma'_1}\boldsymbol{v}'_{\mathbf{1}}\ \ldots\ \sqrt{\sigma'_M}\boldsymbol{v}'_{\boldsymbol{M}}]^*\in\mathbb{C}^{P\times M}. \tag{21}$$

The scaling ambiguity in $\widehat{\boldsymbol{S}}$ and $\widehat{\boldsymbol{A}}$ can be resolved by exploiting the knowledge of the first row of $\boldsymbol{S}$ or $\boldsymbol{A}$.

Similar to the outer PARAFAC part of the LS-KRF algorithm, the M columns in $\widehat{\boldsymbol{S}}$ and $\widehat{\boldsymbol{A}}$ can also be computed in parallel.

2) Accelerated ALS Receiver: In case of $F<M$, (14) is an underdetermined system of linear equations and $(\boldsymbol{T}\diamond\boldsymbol{H})$ cannot be uniquely recovered via (15). Therefore, the LS-KRF algorithm fails to work for the outer PARAFAC part. Nevertheless, the ALS algorithm may operate in this case. Here, we propose a fast ALS algorithm with lower complexity.

Since $NP\geq M$, the rank-M truncated SVD of $[\mathcal{Y}]_3\in\mathbb{C}^{NP\times KF}$ is given by

$$[\mathcal{Y}]_3=\boldsymbol{U}\cdot\boldsymbol{\Sigma}_M\cdot\boldsymbol{V}^{\mathrm{H}}, \tag{22}$$

where $\boldsymbol{U}\in\mathbb{C}^{NP\times M}$ and $\boldsymbol{V}\in\mathbb{C}^{M\times M}$ are (truncated) unitary matrices that respectively collect the left and right singular vectors corresponding to the M largest singular values in the $M\times M$ diagonal matrix $\boldsymbol{\Sigma}_M$.

In the absence of noise, we have

$$[\mathcal{Y}]_3=\boldsymbol{T}\cdot(\boldsymbol{H}\diamond\boldsymbol{W})^{\mathrm{T}}. \tag{23}$$

Comparing (23) with (22) and noticing that $\boldsymbol{T}$ is a tall matrix with full column rank, it follows that the columns of $\boldsymbol{T}$ and $\boldsymbol{U}$ span the same subspace in the noiseless case. Therefore, there exists an $M\times M$ non-singular transform matrix $\boldsymbol{\Phi}$ such that

$$\boldsymbol{T}=\boldsymbol{U}\cdot\boldsymbol{\Phi}. \tag{24}$$

TABLE II
ACCELERATED ALS RECEIVER ALGORITHM

First stage
Initialization: Set $i=0$; $\widehat{\boldsymbol{\Phi}}=\boldsymbol{I}_M$, $[\boldsymbol{\mathcal{X}}]_3=\boldsymbol{U}^{\mathrm{H}}[\mathcal{Y}]_3=\boldsymbol{\Sigma}_M\cdot\boldsymbol{V}^{\mathrm{H}}$;
(1.1) $i=i+1$;
(1.2) Compute the LS estimate of $\boldsymbol{H}$:
$$\widehat{\boldsymbol{H}}(i)=[\boldsymbol{\mathcal{X}}]_1\cdot[(\boldsymbol{W}\diamond\widehat{\boldsymbol{\Phi}}(i-1))^{\dagger}]^{\mathrm{T}};$$
(1.3) Compute the LS estimate of $\boldsymbol{\Phi}$:
$$\widehat{\boldsymbol{\Phi}}(i)=[\boldsymbol{\mathcal{X}}]_3\cdot[(\widehat{\boldsymbol{H}}(i)\diamond\boldsymbol{W})^{\dagger}]^{\mathrm{T}};$$
(1.4) Repeat steps (1.1)-(1.3) until convergence.

Substituting (24) into (23) and then multiplying both sides of the resultant equation from the LHS by $\boldsymbol{U}^{\mathrm{H}}$, we obtain

$$\boldsymbol{U}^{\mathrm{H}}[\mathcal{Y}]_3=\boldsymbol{\Phi}\cdot(\boldsymbol{H}\diamond\boldsymbol{W})^{\mathrm{T}}. \tag{25}$$

Defining a three-way tensor $\boldsymbol{\mathcal{X}}\in\mathbb{C}^{K\times F\times M}$ such that

$$[\boldsymbol{\mathcal{X}}]_3=\boldsymbol{U}^{\mathrm{H}}[\mathcal{Y}]_3=\boldsymbol{\Sigma}_M\cdot\boldsymbol{V}^{\mathrm{H}}, \tag{26}$$

it follows from (25) that in the absence of noise

$$\boldsymbol{\mathcal{X}}=\boldsymbol{\mathcal{I}}_M\times_1\boldsymbol{H}\times_2\boldsymbol{W}\times_3\boldsymbol{\Phi}. \tag{27}$$

We can apply the ALS algorithm on (27) for joint estimation of $\boldsymbol{H}$ and $\boldsymbol{\Phi}$. Since the number of unknowns in $\boldsymbol{\Phi}\in\mathbb{C}^{M\times M}$ is much smaller than that in $\boldsymbol{T}\in\mathbb{C}^{NP\times M}$, the new ALS receiver enjoys faster convergence and lower computational complexity than [17]. The accelerated ALS receiver algorithm is described in Table II.

Once $\widehat{\boldsymbol{T}}=\boldsymbol{U}\cdot\widehat{\boldsymbol{\Phi}}$ is obtained, $\boldsymbol{S}$ and $\boldsymbol{A}$ are estimated using the LS-KRF algorithm presented in Section IV-A.1.

B. Simplified Closed-Form PARAFAC With Pairing Receiver

In this subsection, a tilde $\tilde{}$ over a matrix is used to denote an estimate of the matrix which is obtained in the absence of noise and may be subject to permutation and/or scaling ambiguities. After applying the ALS or LS-KRF to the outer PARAFAC part, we get

$$\tilde{\boldsymbol{H}}=\boldsymbol{H}\boldsymbol{\Delta}_H \tag{28}$$

$$\tilde{\boldsymbol{T}}=\boldsymbol{T}\boldsymbol{\Delta}_T, \tag{29}$$

where $\boldsymbol{\Delta}_H,\boldsymbol{\Delta}_T\in\mathbb{C}^{M\times M}$ are diagonal matrices which contain the column scaling ambiguities in the corresponding factor matrix estimates and satisfy $\boldsymbol{\Delta}_H\boldsymbol{\Delta}_T=\boldsymbol{I}_M$.

In the ALS and LS-KRF receivers, before proceeding into the inner PARAFAC, $\boldsymbol{\Delta}_T$ must be estimated so that the column scaling ambiguities in $\tilde{\boldsymbol{T}}$ can be removed by multiplying the inverse of $\boldsymbol{\Delta}_T$ by its RHS. Otherwise, the RHS of (19), namely,

$$\tilde{\boldsymbol{T}}\cdot\boldsymbol{\Theta}=\boldsymbol{T}\boldsymbol{\Delta}_T\boldsymbol{\Theta}=(\boldsymbol{S}\diamond\boldsymbol{A})\boldsymbol{\Theta}^{\mathrm{T}}\boldsymbol{\Delta}_T\boldsymbol{\Theta} \tag{30}$$

loses the Khatri-Rao structure, because it is no longer equal to the Khatri-Rao product of two factor matrices, and consequently, $\boldsymbol{S}$ and $\boldsymbol{A}$ cannot be extracted unambiguously by employing the ALS or LS-KRF. To estimate $\boldsymbol{\Delta}_T$, both ALS and P-LS-KRF receivers need the knowledge of the first row in the symbol matrix $\boldsymbol{S}$. However, this implies that one of N time-slots

in each time-frame is used to transmit pilot symbols, which results in a loss in the transmission efficiency.

In the sequel, we relax the assumption on the knowledge of the first row of $\boldsymbol{S}$ so that the transmission efficiency can be improved. Substituting (10b) into (29) yields

$$\tilde{\boldsymbol{T}} = (\boldsymbol{S} \diamond \boldsymbol{A})\boldsymbol{\Theta}^{\mathrm{T}}\boldsymbol{\Delta}_T = (\boldsymbol{S} \diamond \boldsymbol{A})(\boldsymbol{\Delta}_T\boldsymbol{\Theta})^{\mathrm{T}} \in \mathbb{C}^{NP\times M}. \quad (31)$$

Note that the tensor form of (31) is given by

$$\tilde{\boldsymbol{\mathcal{T}}} = \boldsymbol{\mathcal{I}}_M \times_1 \boldsymbol{S} \times_2 \boldsymbol{A} \times_3 (\boldsymbol{\Delta}_T\boldsymbol{\Theta}) \in \mathbb{C}^{N\times P\times M}. \quad (32)$$

The estimation of $\boldsymbol{\Delta}_T$ is divided into two steps: estimation of $\boldsymbol{\Theta}_\Delta = \boldsymbol{\Delta}_T\boldsymbol{\Theta}$ as well as factor matrices $\boldsymbol{S}$ and $\boldsymbol{A}$ via an S-CFP decomposition, and estimation of $\boldsymbol{\Delta}_T$ from the estimate of $\boldsymbol{\Theta}_\Delta$.

1) Estimation of $\boldsymbol{\Theta}_\Delta$ via Simplified Closed-Form PARAFAC Decomposition: An efficient way to estimate $\boldsymbol{\Theta}_\Delta$, $\boldsymbol{S}$ and $\boldsymbol{A}$ is to apply the CFP decomposition [27] to the inner PARAFAC part. In CFP, estimation of the factors of the PARAFAC model is based on the higher-order SVD (HOSVD) low-rank approximation and SMD. Compared with the conventional methods for PARAFAC decomposition which is often conducted iteratively and computationally expensive, the CFP provides a relatively efficient implementation of the PARAFAC decomposition.

The HOSVD low-rank approximation of $\tilde{\boldsymbol{\mathcal{T}}}$ is given by [28]

$$\tilde{\boldsymbol{\mathcal{T}}} \approx \boldsymbol{\mathcal{S}}^{[\mathrm{s}]} \times_1 \boldsymbol{U}_1^{[\mathrm{s}]} \times_2 \boldsymbol{U}_2^{[\mathrm{s}]} \times_3 \boldsymbol{U}_3^{[\mathrm{s}]}, \quad (33)$$

where $\boldsymbol{\mathcal{S}}^{[\mathrm{s}]} \in \mathbb{C}^{p_1\times p_2\times M}$ is the core tensor with $p_1 = \min(N, M)$, $p_2 = \min(P, M)$, and $\boldsymbol{U}_1^{[\mathrm{s}]} \in \mathbb{C}^{N\times p_1}$, $\boldsymbol{U}_2^{[\mathrm{s}]} \in \mathbb{C}^{P\times p_2}$ and $\boldsymbol{U}_3^{[\mathrm{s}]} \in \mathbb{C}^{M\times M}$ represent the 1-mode, 2-mode and 3-mode singular vector matrices truncated to M columns, respectively.

From the r-mode unfolding of (32) and (33), it follows that the columns of $\boldsymbol{U}_r^{[\mathrm{s}]}$, $r = 1, 2, 3$, and those of $\boldsymbol{S}$, $\boldsymbol{A}$ or $\boldsymbol{\Theta}_\Delta$, span the same space in the noiseless case. Therefore, there exist $M \times M$ non-singular transform matrices $\boldsymbol{T}_r$[27] such that $\boldsymbol{S} = \boldsymbol{U}_1^{[\mathrm{s}]} \cdot \boldsymbol{T}_1$ when $N \geq M$ or $\boldsymbol{S}$ has full column rank, $\boldsymbol{A} = \boldsymbol{U}_2^{[\mathrm{s}]} \cdot \boldsymbol{T}_2$ for $P \geq M$, and $\boldsymbol{\Theta}_\Delta = \boldsymbol{U}_3^{[\mathrm{s}]} \cdot \boldsymbol{T}_3$.

In addition to $\boldsymbol{\Theta}$ that has full column rank, for the system configurations of our interest, either the factor matrix $\boldsymbol{S}$ or $\boldsymbol{A}$ has full column rank according to (13). Hereafter, we consider the first system configuration in (13a), namely, $P < M \leq N$, where $\boldsymbol{S}$ has full column rank while $\boldsymbol{A}$ does not. For the second system configuration of $N < M \leq P$, it is only required to swap the roles of $\boldsymbol{S}$ and $\boldsymbol{A}$. Therefore, we have $\boldsymbol{S} = \boldsymbol{U}_1^{[\mathrm{s}]} \cdot \boldsymbol{T}_1$ and $\boldsymbol{\Theta}_\Delta = \boldsymbol{U}_3^{[\mathrm{s}]} \cdot \boldsymbol{T}_3$.

In order to estimate $\boldsymbol{S}$ and $\boldsymbol{\Theta}_\Delta$, it only remains to estimate $\boldsymbol{T}_1$ and $\boldsymbol{T}_3$. Defining the tensors $\boldsymbol{\mathcal{S}}_2$ and $\boldsymbol{\mathcal{A}}_2$ as

$$\boldsymbol{\mathcal{S}}_2 = \boldsymbol{\mathcal{S}}^{[\mathrm{s}]} \times_2 \boldsymbol{U}_2^{[\mathrm{s}]} \in \mathbb{C}^{M\times P\times M}, \quad (34)$$

$$\boldsymbol{\mathcal{A}}_2 = \boldsymbol{\mathcal{I}}_M \times_2 \boldsymbol{A} \in \mathbb{C}^{M\times P\times M}, \quad (35)$$

it follows from (32) and (33) that

$$\boldsymbol{\mathcal{S}}_2 \times_1 \boldsymbol{T}_1^{-1} \times_3 \boldsymbol{T}_3^{-1} = \boldsymbol{\mathcal{A}}_2. \quad (36)$$

Letting $\boldsymbol{S}_{2,(p)}, \boldsymbol{A}_{2,(p)} \in \mathbb{C}^{M\times M}$, $p = 1, \ldots, P$, to represent the p-th slices of $\boldsymbol{\mathcal{S}}_2$ and $\boldsymbol{\mathcal{A}}_2$, respectively, we have

$$\begin{aligned} &\boldsymbol{S}_{2,(p)} \times_1 \boldsymbol{T}_1^{-1} \times_2 \boldsymbol{T}_3^{-1} = \boldsymbol{A}_{2,(p)} \\ &\Rightarrow \boldsymbol{T}_1^{-1} \cdot \boldsymbol{S}_{2,(p)} \cdot \boldsymbol{T}_3^{-\mathrm{T}} = \boldsymbol{A}_{2,(p)}. \end{aligned} \quad (37)$$

Note that the slices of $\boldsymbol{\mathcal{A}}_2$ along the 2-mode are all diagonal matrices. Hence (37) means that the transform matrices $\boldsymbol{T}_1$ and $\boldsymbol{T}_3$ jointly diagonalize all the P slices of $\boldsymbol{\mathcal{S}}_2$.

In order to express (37) in a symmetric form, we multiply $\boldsymbol{S}_{2,(p)}$ for all $p = 1, \ldots, P$, by the inverse of $\boldsymbol{S}_{2,(q)}$ from the LHS:

$$\begin{aligned} \boldsymbol{S}_{2,(p)}^{\mathrm{LHS}} &= \left(\boldsymbol{S}_{2,(q)}^{-1} \cdot \boldsymbol{S}_{2,(p)}\right)^{\mathrm{T}} \\ &= \boldsymbol{T}_3 \cdot \boldsymbol{A}_{2,(p)}^{\mathrm{T}} \cdot \boldsymbol{T}_1^{\mathrm{T}} \cdot \boldsymbol{T}_1^{-\mathrm{T}} \cdot \boldsymbol{A}_{2,(q)}^{-\mathrm{T}} \cdot \boldsymbol{T}_l^{-1} \\ &= \boldsymbol{T}_3 \cdot \underbrace{\boldsymbol{A}_{2,(p)} \cdot \boldsymbol{A}_{2,(q)}^{-1}}_{\boldsymbol{A}_{2,(p|q)}} \cdot \boldsymbol{T}_3^{-1}, \end{aligned} \quad (38)$$

where the constant index q is chosen such that $\boldsymbol{S}_{2,(q)}$ is best-conditioned and easily invertible. Note that $\boldsymbol{A}_{2,(p|q)}$ is also a diagonal matrix for each n, which implies that $\boldsymbol{S}_{2,(p)}^{\mathrm{LHS}}$ for $p = 1, \ldots, P$, are jointly diagonalized by $\boldsymbol{T}_3$. Consequently, the transform matrix $\boldsymbol{T}_3$ can be estimated by joint diagonalization of the set of matrices $\mathbb{S} = \left\{\boldsymbol{S}_{2,(p)}^{\mathrm{LHS}} | p = 1, \ldots, P, p \neq q\right\}$. As remarked in [31], determination of $\boldsymbol{T}_3$ from only one matrix in $\mathbb{S}$ is not satisfactory because it amounts to discarding the information contained in the other matrices of $\mathbb{S}$. Moreover, it may happen that individual $\boldsymbol{S}_{2,(p)}^{\mathrm{LHS}}$ is defective but that the whole set $\mathbb{S}$ has well determined common eigenvectors. Hence, joint diagonalization of $\mathbb{S}$ instead of the eigenvalue decomposition of its individual member yields a more accurate and robust estimate of $\boldsymbol{T}_3$.

Once $\boldsymbol{T}_3$ and $\boldsymbol{\Theta}_\Delta$ are estimated, $\boldsymbol{S}$ and $\boldsymbol{A}$ are recovered based on the LS-KRF in Section IV-A.1. Note that in CFP [27], two SMDs are needed, one for estimating $\boldsymbol{T}_1$, and the other for $\boldsymbol{T}_3$. Here we simplify this process by solving only one SMD to estimate $\boldsymbol{T}_3$. In order to further reduce the computational complexity, by noting that the matrices in $\mathbb{S}$ are non-defective, we use the computationally more efficient method in [32] for joint matrix diagonalization instead of that used in [27], namely, the one proposed in [33]. As shown in [32], by utilizing both the shear and unitary transformations for SMD instead of using only the unitary transformation for simultaneous Schur decomposition, the modified SMD method enjoys a quadratic convergence rate and requires much fewer iterations than the one in [33]. Since the SMD is the primary computationally intensive part in CFP, the simplified CFP significantly reduces the computational complexity while empirically shows no noticeable performance degradation.

Note that the obtained factor matrices suffer from column scaling and permutation ambiguities, which is an inherent characteristic of the PARAFAC decomposition. In other words, the factor estimates are column scaled and permutated versions of their corresponding true factor matrices, namely

$$\tilde{\boldsymbol{\Theta}}_\Delta\boldsymbol{\Pi} = \boldsymbol{\Delta}_T\boldsymbol{\Theta}\boldsymbol{\Delta}_\Theta \quad (39)$$

$$\tilde{\boldsymbol{A}}\boldsymbol{\Pi} = \boldsymbol{A}\boldsymbol{\Delta}_A \quad (40)$$

$$\tilde{\boldsymbol{S}}\boldsymbol{\Pi} = \boldsymbol{S}\boldsymbol{\Delta}_S, \quad (41)$$

where $\mathbf{\Pi}$ is an $M \times M$ permutation matrix that contains the column permutation ambiguities in the inner PARAFAC, while $\mathbf{\Delta}_\Theta$, $\mathbf{\Delta}_A$ and $\mathbf{\Delta}_S$ are diagonal matrices which represent the column scaling ambiguities in the corresponding estimates and satisfy

$$\mathbf{\Delta}_S \mathbf{\Delta}_A \mathbf{\Delta}_\Theta = \boldsymbol{I}_M. \tag{42}$$

2) Estimation of $\mathbf{\Delta}_T$ and $\boldsymbol{S}$ via Pairing Algorithm: To unambiguously extract $\boldsymbol{S}$, we need to estimate $\mathbf{\Pi}$ and $\mathbf{\Delta}_S$. Since the first row in $\boldsymbol{A}$ is known, $\mathbf{\Delta}_A$ can be readily obtained. According to (42), $\mathbf{\Delta}_S$ can be found once we have the estimate of $\mathbf{\Delta}_\Theta$.

First, we determine $\mathbf{\Pi}$ by exploiting the Vandermonde structure of $\boldsymbol{A}$ in (8). To this end, we observe that there is a one-to-one correspondence (bijective mapping) between the permutation matrix and the following permutation function:

$$\pi : \{1, 2, \ldots, M\} \mapsto \{1, 2, \ldots, M\}, \tag{43}$$

which can be expressed in two-line form by

$$\begin{bmatrix} 1 & 2 & \ldots & M \\ \pi(1) & \pi(2) & \ldots & \pi(M) \end{bmatrix}. \tag{44}$$

For an arbitrary permutation function π, a permutation matrix $\mathbf{\Pi}$ is constructed as

$$\mathbf{\Pi} = [\, \boldsymbol{e}_{\pi(1)} \quad \boldsymbol{e}_{\pi(2)} \quad \ldots \quad \boldsymbol{e}_{\pi(M)} \,], \tag{45}$$

where $\boldsymbol{e}_{\pi(m)}$ denotes a column vector of length M with 1 in the $\pi(m)$-th position and 0 elsewhere. Thereby, estimation of $\mathbf{\Pi}$ is equivalent to determination of π.

Since the first row in $\boldsymbol{A}$ is known, the column scaling ambiguity of $\tilde{\boldsymbol{A}}$ can be readily removed. Therefore, hereafter we assume that $\tilde{\boldsymbol{A}}$ only suffers from permutation ambiguity, namely, $\tilde{\boldsymbol{A}}\mathbf{\Pi} = \boldsymbol{A}$. First of all, $\pi(1)$ is determined as

$$\pi(1) = \underset{m_1 \in \{1,\ldots,M\}}{\arg\min} \|\tilde{\boldsymbol{A}}(:, m_1) - \mathbf{1}_P\|^2, \tag{46}$$

where $\mathbf{1}_P$ is the first column of $\boldsymbol{A}$ which is a length-P column vector of all ones.

Next, we determine $\pi(m)$, $m = 2, \ldots, M$. For each candidate value $m_2 \in \{1, \ldots, M\} \backslash \pi(1)$ of $\pi(2)$, we first reconstruct $\boldsymbol{A}$ from $\tilde{\boldsymbol{A}}(:, m_2) = [1 \ \tilde{a}_{2,m_2} \ \ldots \ \tilde{a}_{P,m_2}]^{\mathrm{T}}$ as

$$\widehat{\boldsymbol{A}}_{m_2} = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ 1 & \tilde{a}_{2,m_2} & \ldots & \tilde{a}_{2,m_2}^{M-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \tilde{a}_{P,m_2} & \ldots & \tilde{a}_{P,m_2}^{M-1} \end{bmatrix}, \tag{47}$$

and then find the associated estimate of $\pi(m)$, $m = 3, \ldots, M$, by pairing up the columns of $\tilde{\boldsymbol{A}}$ and $\widehat{\boldsymbol{A}}_{m_2}$:

$$\begin{aligned} \tilde{\boldsymbol{A}}(:, m_3) &\Leftrightarrow \widehat{\boldsymbol{A}}_{m_2}(:, 3), \\ &\vdots \\ \tilde{\boldsymbol{A}}(:, m_M) &\Leftrightarrow \widehat{\boldsymbol{A}}_{m_2}(:, M), \end{aligned} \tag{48}$$

where $\{m_2, m_3, \ldots, m_M\}$ is a permutation of $\{1, \ldots, M\} \setminus \pi(1)$. The pairing up can be accomplished with the greedy algorithm [34] or by successive comparison.

Denoting $\mathbf{\Pi}_{m_2} = \left[\boldsymbol{e}_{\pi(1)} \ \boldsymbol{e}_{m_2} \ \ldots \ \boldsymbol{e}_{m_M}\right]$ and

$$\tilde{\boldsymbol{A}}\mathbf{\Pi}_{m_2} = \left[\tilde{\boldsymbol{A}}(:, \pi(1)) \ \tilde{\boldsymbol{A}}(:, m_2) \ \tilde{\boldsymbol{A}}(:, m_3) \ \ldots \ \tilde{\boldsymbol{A}}(:, m_M)\right],$$

$\pi(2)$ is determined as

$$\pi(2) = \underset{m_2 \in \{1,\ldots,M\} \backslash \pi(1)}{\arg\min} \|\tilde{\boldsymbol{A}}\mathbf{\Pi}_{m_2} - \widehat{\boldsymbol{A}}_{m_2}\|_{\mathrm{F}}^2. \tag{49}$$

Once $\pi(2)$ is determined, $\pi(m)$, $m = 3, \ldots, M$, will also be known.

We then proceed to estimate $\mathbf{\Delta}_\Theta$. Note in (39) that the column scaling ambiguity in $\tilde{\boldsymbol{T}}$ has been converted into the row scaling ambiguity in $\tilde{\mathbf{\Theta}}_\Delta$. Consequently, the $\tilde{\mathbf{\Theta}}_\Delta$ obtained after applying CFP to inner PARAFAC is subject to both column and row scaling ambiguities. Hereafter, we propose an algorithm to jointly estimate $\mathbf{\Delta}_\Theta$ and $\mathbf{\Delta}_T$.

Let $\boldsymbol{\delta}_T = [\delta_{T,1} \ \ldots \ \delta_{T,M}]^{\mathrm{T}}$ and $\boldsymbol{\delta}_\Theta = [\delta_{\Theta,1} \ \ldots \ \delta_{\Theta,M}]^{\mathrm{T}}$ be vectors that collect the diagonal elements in $\mathbf{\Delta}_T$ and $\mathbf{\Delta}_\Theta$, respectively. From (39), it follows that

$$\tilde{\mathbf{\Theta}}_\Delta \mathbf{\Pi} ./ \mathbf{\Theta} = \boldsymbol{\delta}_T \boldsymbol{\delta}_\Theta^{\mathrm{T}}, \tag{50}$$

where ./ denotes entrywise division of two matrices with the same dimensions. Note that the RHS of (50) is a rank-1 matrix. Therefore, we can apply the rank-1 truncated SVD on (50) to find $\boldsymbol{\delta}_T$ and $\boldsymbol{\delta}_\Theta$. Suppose after SVD, the largest singular value is λ and the corresponding left and right singular vectors are denoted as $\boldsymbol{u}$ and $\boldsymbol{v}$, respectively, then the estimates of $\boldsymbol{\delta}_T$ and $\boldsymbol{\delta}_\Theta$ are given by

$$\widehat{\boldsymbol{\delta}}_T = \sqrt{\lambda}\boldsymbol{u}, \quad \widehat{\boldsymbol{\delta}}_\Theta = \sqrt{\lambda}\boldsymbol{v}^*. \tag{51}$$

V. Identifiability, Spectral Efficiency and Computational Complexity

A. Identifiability Conditions

We investigate the necessary and sufficient conditions under which the symbol matrix $\boldsymbol{S}$ and channel matrix $\boldsymbol{H}$ can be uniquely and correctly recovered in the absence of noise.

1) ALS Receiver: Consider

$$\boldsymbol{\mathcal{Y}} = \sqrt{\frac{\rho}{M}} \cdot \boldsymbol{\mathcal{I}}_M \times_1 \boldsymbol{H} \times_2 \boldsymbol{W} \times_3 \boldsymbol{T}, \tag{52}$$

or its equivalent matrix form, e.g., (14) or (23), as a system of complex-valued nonlinear equations where the entries in $\boldsymbol{T}$ and $\boldsymbol{H}$ are unknown variables. The following proposition connects the conditions under which the estimates of $\boldsymbol{T}$ and $\boldsymbol{H}$ generated by the ALS algorithm converge to the true values and the conditions for the uniqueness of the solutions to the system of nonlinear equations of (52).

Proposition 1: Assume that in the absence of noise both $\boldsymbol{W} \diamond \boldsymbol{T}$ and $\boldsymbol{H} \diamond \boldsymbol{W}$ have full rank, and the sequence of $\widehat{\boldsymbol{T}}$ and $\widehat{\boldsymbol{H}}$ generated by the ALS algorithm has limit point. The $\widehat{\boldsymbol{T}}$ and $\widehat{\boldsymbol{H}}$ converge to the true values if and only if the solution to the system of

$F \cdot NPK$ complex-valued nonlinear equations in (52) is unique under the constraint of $\boldsymbol{T}(1,:) = \boldsymbol{T}_0(1,:)$, where $\boldsymbol{T}_0(1,:)$ is the first row of $\boldsymbol{T}$ defined in (10b), which is known as explained in Section III.

The proof is given as follows.

Necessity: Assume there exist two different solutions $\boldsymbol{x}_1 = \text{vec}([\boldsymbol{H}_1, \boldsymbol{T}_1])$ and $\boldsymbol{x}_2 = \text{vec}([\boldsymbol{H}_2, \boldsymbol{T}_2])$. It can be verified that both $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are the critical points of the residual error function $f = \|\boldsymbol{\mathcal{Y}}_0 - \boldsymbol{\mathcal{I}}_M \times_1 \boldsymbol{H} \times_2 \boldsymbol{W} \times_3 \boldsymbol{T}\|_{\text{F}}$ such that $\nabla f(\boldsymbol{x}_1) = \nabla f(\boldsymbol{x}_2) = \boldsymbol{0}$. According to [35], both $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are possible points of convergence of the ALS estimate sequence.

Sufficiency: Since both $\boldsymbol{W} \diamond \boldsymbol{T}$ and $\boldsymbol{H} \diamond \boldsymbol{W}$ have full rank, according to Theorem 3.8 of [35], $\widehat{\boldsymbol{T}}$ and $\widehat{\boldsymbol{H}}$ converge to the critical point. On the other hand, f is convex and hence the solution to the system of nonlinear equations is the only critical point.

This completes the proof.

For unique determination of $\boldsymbol{T}$ and $\boldsymbol{H}$, (52) must be (over-)determined, which means that the number of unknowns is less than or equal to that of equations. By taking into account the constraint $\bar{\boldsymbol{T}}(1,:) = \boldsymbol{T}(1,:)$, the number of unknown complex variables is equal to $(NP + K - 1)M$. On the other hand, there are $F \cdot NPK$ equations in the system of complex-valued nonlinear equations. The necessary uniqueness condition of the ALS algorithm is now ready.

Corollary 1: For the ALS estimates of $\boldsymbol{T}$ and $\boldsymbol{H}$ to converge to their true values, it must satisfy

$$FNPK \geq (NP + K - 1)M. \tag{53}$$

Note that for $F \geq M$, the second matrix unfolding of (52), namely, (14), can be considered as a determined or an overdetermined system of linear equations when the Khatri-Rao structure in $\boldsymbol{T} \diamond \boldsymbol{H}$ is ignored. Therefore, the sufficient condition in Proposition 1 is satisfied.

Corollary 2: The estimates of $\boldsymbol{T}$ and $\boldsymbol{H}$ in the ALS algorithm converge to the true value, if

$$F \geq M. \tag{54}$$

For the inner PARAFAC part, since $\boldsymbol{\Theta}$ is an $M \times M$ square matrix of full rank, the system of complex-valued linear equations of (12) is always (over-)determined, in which case the convergence of the ALS algorithm is not an issue.

To sum up, the identifiability condition for the ALS receiver is stipulated by (53).

2) Accelerated ALS Receiver: By applying similar analysis to Section V-A.1, a necessary condition for the accelerated ALS algorithm is given by

$$FK \geq M + K - 1. \tag{55}$$

Regarding the sufficient condition, the condition established in (54) applies in the accelerated ALS receiver as well.

3) (Parallelized) LS-KRF Receiver: The necessary and sufficient uniqueness condition of the (parallelized) LS-KRF receiver algorithm is given by

$$F \geq M. \tag{56}$$

Remark 1: Both conditions (55) and (56) are special cases of (53), which indicate that in exchange for the reduction in computational cost, the accelerated ALS and LS-KRF receiver algorithms require more strict uniqueness conditions.

4) S-CFP with Pairing Receiver: In the outer PARAFAC part, the (accelerated) ALS or LS-KRF algorithm is exploited. The necessary and sufficient uniqueness conditions stipulated by (53)–(56) apply.

In the inner PARAFAC part, the S-CFP algorithm is exploited, and the conditions under which the three-way PARAFAC decomposition is essentially unique, namely, unique up to permutation and scaling of the columns of the factor matrices, apply. According to [36], a sufficient condition for the essential uniqueness of the inner PARAFAC decomposition of (12) is given by

$$k_{\boldsymbol{S}} + k_{\boldsymbol{A}} + k_{\boldsymbol{\Delta}} \geq 2M + 2, \tag{57}$$

where $k_{\boldsymbol{A}}$ denotes the Kruskal-rank of $\boldsymbol{A}$, which is defined as the largest integer k such that any set of k columns of $\boldsymbol{A}$ are linearly independent.

Note that in our system design, the DFT matrix $\boldsymbol{\Theta}$ has full k-rank that is equal to its size M. Moreover, since the rows of $\boldsymbol{A}$ and columns of $\boldsymbol{S}$ are drawn independently from a continuous distribution, according to [37] they also have full k-rank with probability one. Applying the uniqueness condition of (57) to the inner PARAFAC part of (12) yields

$$\min(N, M) + \min(P, M) \geq M + 2. \tag{58}$$

Under the system configurations of our interest in (13), (58) simplifies to $\min(N, P) \geq 2$. We see that compared with the ALS and LS-KRF receivers that rely on the knowledge of the first row of $\boldsymbol{S}$, the S-CFP with pairing receiver algorithm without such knowledge requires almost no extra constraints on the signal dimensions for unique recovery of the symbol and channel parameters.

B. Spectral Efficiency

In the MIMO-OFDM system with the ALS/LS-KRF and S-CFP with pairing receiver algorithms, $M(N-1)$ and MN symbols in each time frame are respectively sent over F subcarriers and N time-slots each with P symbol periods. The resultant spectral efficiency is $M(N-1)/(NPF)$ and $MN/(NPF)$ symbol/symbol-period/subcarrier, respectively.

Compared with the ALS and LS-KRF receivers, the S-CFP with pairing algorithm increases the spectral efficiency by $1/(N-1)$. This improvement is nontrivial particularly for a small N. For example, when $N = 4$, $P = 3$, $M = 4$ and $F = 3$, the spectral efficiency of the former receiver is 1/3 symbol/symbol-period/subcarrier, whereas the spectral efficiency of the latter is 4/9, which has an improvement by 33.3%.

C. Computational Complexity

In Table III, the numbers of flops required in each computational steps of the (accelerated) ALS, LS-KRF and S-CFP with pairing receivers are tabulated. We see that that the complexity of both the ALS and S-CFP with pairing receivers is polynomial in terms of the data size. The ratio of the number of flops

TABLE III
NUMBER OF FLOPS REQUIRED IN THE (ACCELERATED) ALS, LS-KRF AND S-CFP WITH PAIRING RECEIVERS

	Outer PARAFAC		Inner PARAFAC	
ALS	Compute $\boldsymbol{W} \diamond \widehat{\boldsymbol{T}}$ and $\widehat{\boldsymbol{H}} \diamond \boldsymbol{W}$	$(NP+K)FM$	Compute $\widehat{\boldsymbol{A}} \diamond \boldsymbol{\Theta}$ and $\boldsymbol{\Theta} \diamond \boldsymbol{S}$	$(P+N)M^2$
	and their pseudoinverses	$\mathcal{O}\left([NP+K]FM^2\right)$	and their pseudoinverses	$\mathcal{O}\left([P+N]M^3\right)$
	Multiply with $[\boldsymbol{\mathcal{Y}}]_1$ and $[\boldsymbol{\mathcal{Y}}]_3$	$2NPMKF$	Multiply with $[\boldsymbol{\mathcal{T}}]_1$ and $[\boldsymbol{\mathcal{T}}]_2$	$2PNM^2$
Subtotal	$\mathcal{O}\left(NPMF[M+K]\right)$		$\mathcal{O}\left([P+N]M^3\right)$	
Accelerated ALS	Compute $\boldsymbol{W} \diamond \widehat{\boldsymbol{T}}$ and $\widehat{\boldsymbol{H}} \diamond \boldsymbol{W}$	$(M+K)FM$	Same as ALS	
	and their pseudoinverses	$\mathcal{O}\left([M+K]FM^2\right)$		
	Multiply with $[\boldsymbol{\mathcal{Y}}]_1$ and $[\boldsymbol{\mathcal{Y}}]_3$	$2M^2KF$		
Subtotal	$\mathcal{O}\left(M^2F[M+K]\right)$		$\mathcal{O}\left([P+N]M^3\right)$	
LS-KRF	Estimate $\boldsymbol{T} \diamond \boldsymbol{H}$ in (15)	$NPMKF$	Estimate $\boldsymbol{S} \diamond \boldsymbol{A}$ in (19)	$\mathcal{O}\left(NPM^2\right)$
	Rank-1 SVD of $\{\boldsymbol{Q}_m\}_{m=1}^M \in \mathbb{C}^{NP\times K}$	$\mathcal{O}\left(NPK \cdot M\right)$ [38]	Rank-1 SVD of $\{\boldsymbol{Q}'_m\}_{m=1}^M \in \mathbb{C}^{N\times P}$	$\mathcal{O}\left(NP \cdot M\right)$
Subtotal	$\mathcal{O}\left(NPMKF\right)$		$\mathcal{O}\left(NPK \cdot M\right)$	
Total	$\mathcal{O}\left(NPMKF\right)$			
S-CFP with pairing	Same as LS-KRF or accelerated ALS		Low-rank approximation of $\tilde{\boldsymbol{\mathcal{T}}}$ in (33)	$\mathcal{O}\left(NPM[N+P+M]\right)$
			Find condition numbers $\{\boldsymbol{S}_{2,(p)}\}_{p=1}^P$	$\mathcal{O}\left(M^2 \cdot P\right)$ [39]
			Estimate $\boldsymbol{T}_3$ via SMD	$\mathcal{O}\left(\min(N,P)M^3\right)$
			Estimate $\boldsymbol{S}$ and $\boldsymbol{A}$ via LS-KRF	$\mathcal{O}\left(NPM^2\right)$
			Determine π	$\mathcal{O}\left(PM^3\right)$
			Estimate $\boldsymbol{\Delta}_\Theta$	$\mathcal{O}\left(M^2\right)$
Subtotal	$\mathcal{O}\left(NPMKF\right)$ or $\mathcal{O}\left(M^2F[M+K]\right)$		$\mathcal{O}\left(NPM\left[N+P+2M+M^2/N\right]\right)$	

required in the S-CFP with pairing receiver to that required in the first stage of the ALS receiver is given by

$$\text{ratio} = \mathcal{O}\left(\frac{NPM\left(N+P+2M+M^2/N+KF\right)}{NPMKF}\right) = \mathcal{O}\left(\frac{N+P+2M+M^2/N}{KF}+1\right). \quad (59)$$

We see that the S-CFP with pairing receiver has the same order of computational complexity as the ALS receiver.

Notes in Table III:

a) Since the SVD of an $J \times L$ ($J \leq L$) matrix takes $O\left(J^2L\right)$ flops, calculation of $\boldsymbol{U}_r$, $r = 1,2,3$, from the SVD of the r-mode matrix unfoldings of $\tilde{\boldsymbol{\mathcal{T}}}$ costs $\mathcal{O}\left(NPM[N+P+M]\right)$ flops. Calculation of the core tensor $\boldsymbol{\mathcal{S}}$ via the r-mode products of $\tilde{\boldsymbol{\mathcal{T}}}$ and the Hermitian transpose (inverse) of $\boldsymbol{U}_r$, $r = 1,2,3$, costs $NPM(N+P+M)$ flops. Therefore, the overall HOSVD costs $\mathcal{O}\left(NPM[N+P+M]\right)$ flops.

b) Each iteration of the SMD [32] requires a shear transformation, which involves $\mathcal{O}\left(PM\right)$ flops, and a unitary transformation, corresponding to $\mathcal{O}\left(P\right)$ flops. Since the algorithm converges at a quadratic rate and typically within a couple of sweeps [32], the total number of flops required is $\mathcal{O}\left(PM \cdot M(M-1)/2\right) = \mathcal{O}\left(PM^3\right)$, where $M(M-1)/2$ is the number of iterations in a *sweep*.

In case of the second system configuration in (13a), the total number of flops required is $\mathcal{O}\left(NM^3\right)$.

c) i) Determination of $\pi(1)$ via (46) costs PM flops; ii) Determination of $\pi(m)$, $m = 2,\ldots,M$: For each candidate value of $\pi(2)$, (47) takes $P(M-2)$ flops, the pairing step in (48) using the successive comparison method takes $P \cdot M(M-3)/2$ flops, and calculation of the residual error in (49) takes PM flops. Therefore, to determine $\pi(m)$, $m = 2,\ldots,M$, it costs $(M-1) \cdot [P(M-2)+P \cdot M(M-3)/2+PM] = \mathcal{O}\left(PM^3\right)$ flops.

In total, it costs $\mathcal{O}\left(PM^3\right)$ to determine π.

VI. SIMULATION RESULTS

The BER, the normalized mean square error (NMSE) of the estimated channel defined as

$$\text{NMSE} = \frac{\|\widehat{\boldsymbol{H}} - \boldsymbol{H}\|_{\text{F}}^2}{\|\boldsymbol{H}\|_{\text{F}}^2}, \quad (60)$$

and the mean processing time are used as the performance measures of the proposed P-LS-KRF and S-CFP semi-blind receivers. The SNR is defined as $\text{SNR} = \|\boldsymbol{Y}_0\|_{\text{F}}^2/\|\boldsymbol{V}\|_{\text{F}}^2$. The additive noise power is scaled to produce different SNRs. For each SNR, the result represents an average of 20000 independent Monte Carlo runs. At each run, the transmitted symbols are drawn from a quadrature phase-shift keying (QPSK) sequence, and the channel coefficients are white ZMCSCG distributed with unit power.

In the ALS algorithm, the outer PARAFAC iteration stops when the relative square reconstruction change (RSRC) between two successive iterations given by

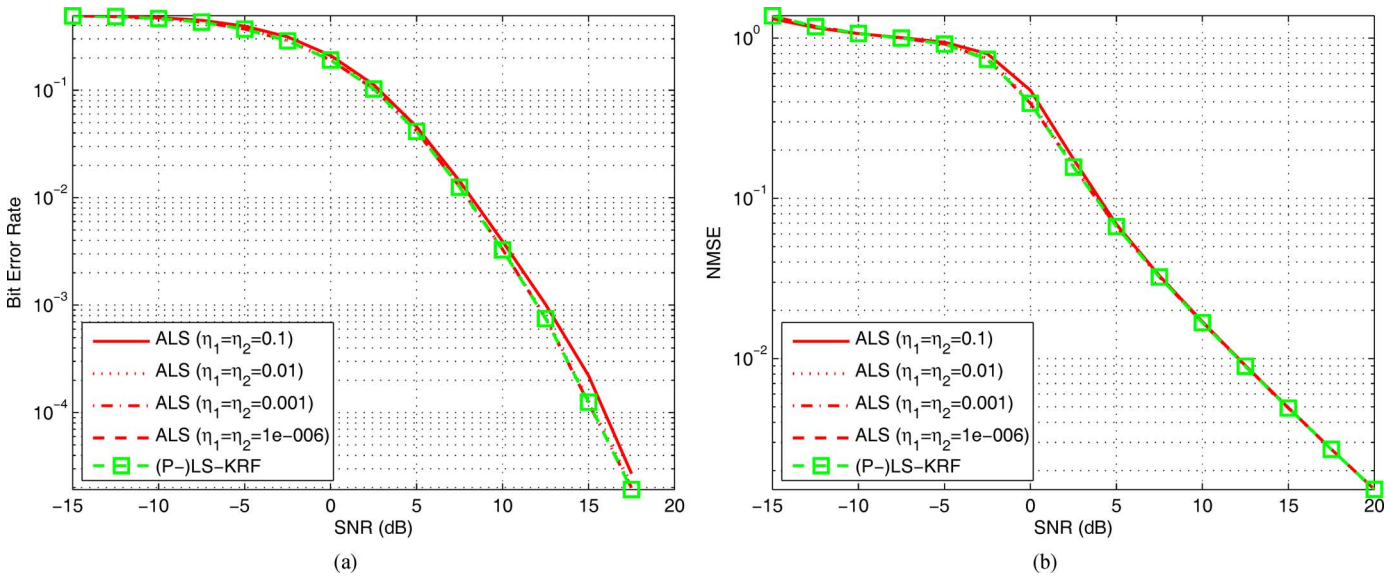
$$\text{RSRC}_{\text{out}}(k) = \frac{\|\widehat{\boldsymbol{\mathcal{Y}}}_{k,0} - \widehat{\boldsymbol{\mathcal{Y}}}_{k-1,0}\|_{\text{F}}^2}{\|\widehat{\boldsymbol{\mathcal{Y}}}_{k,0}\|_{\text{F}}^2} \quad (61)$$

is smaller than a predefined threshold η_1. In a similar manner, the inner PARAFAC iteration stops when the RSRC between two successive iterations given by

$$\text{RSRC}_{\text{in}}(k) = \frac{\|\widehat{\boldsymbol{\mathcal{T}}}_{k,0} - \widehat{\boldsymbol{\mathcal{T}}}_{k-1,0}\|_{\text{F}}^2}{\|\widehat{\boldsymbol{\mathcal{T}}}_{k,0}\|_{\text{F}}^2} \quad (62)$$

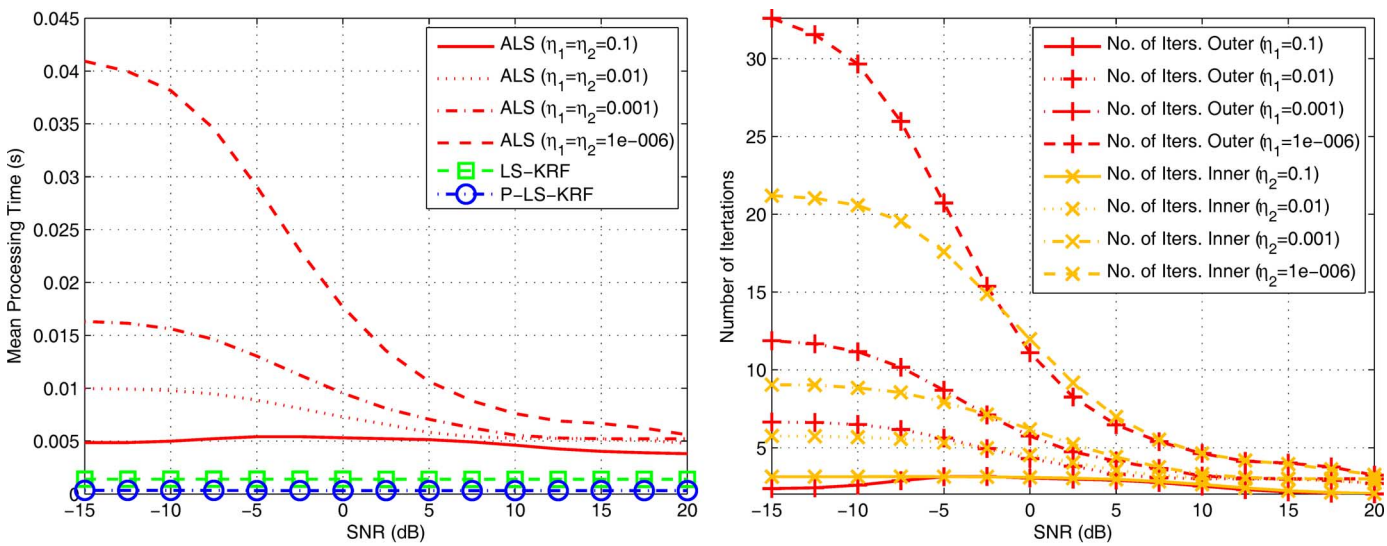
is smaller than another threshold η_2.

Unless stated otherwise we consider only one case of the first system configuration in (13a), namely, $K = 3, M = 5, F = 5, N = 8$, and $P = 4$. For such a system, the spectral efficiency is $M(N-1)/(NPF) = 7/32$ for the LS-KRF/ALS receivers and $M/(PF) = 1/4$ symbol/symbol period/subcarrier for the S-CFP with pairing receiver. Note that the transmission system is flexible in the sense that M, P and/or F can be adjusted (provided that the identifiability conditions established in Section V-A are satisfied) to realize a tradeoff between performance and spectral efficiency. In particular, with the increase of



(a) (b)

Fig. 3. Comparison of ALS and (P-)LS-KRF receivers in terms of BER and NMSE versus SNR. (a) BER versus SNR. (b) Channel NMSE versus SNR.



(a) (b)

Fig. 4. Processing time of ALS, LS-KRF and P-LS-KRF receivers as well as number of iterations required in ALS for various SNRs. (a) Mean processing time versus SNR. (b) Number of iterations required in ALS versus SNR.

M and decrease of P and/or F, the spectral efficiency increases, but meanwhile a performance degradation is expected.

A. ALS Versus (P-)LS-KRF and Accelerated ALS Receivers: With Transmission Overhead

First, we evaluate the proposed (P-)LS-KRF receiver with the knowledge of the first row in the symbol matrix $\boldsymbol{S}$. In Fig. 3 (a) and 3 (b), the ALS and (P-)LS-KRF receivers are compared in terms of BER and NMSE for different SNRs. For the ALS receiver, the convergence thresholds $\eta_1 = \eta_2$ vary from 10^{-6} to 0.1. We see that the ALS with $\eta_1 = \eta_2 = 0.1$ is inferior to the (P-)LS-KRF receiver. However, its performance improves as η_i, $i = 1, 2$, decreases and converges to that of the LS-KRF receiver when $\eta_i \leq 10^{-3}$, $i = 1, 2$.

In Fig. 4 (a), the mean processing times in seconds of the ALS, LS-KRF and P-LS-KRF receivers are shown. We see that the mean processing times required in both LS-KRF versions do not change with SNR since they are non-iterative, whereas that of the ALS receiver decreases with SNR since it is an iterative algorithm and the number of iterations required decreases with SNR, as shown in Fig. 4 (b). As η_i, $i = 1, 2$, increases, the mean processing time required by the ALS algorithm decreases, but is still greater than that of the LS-KRF even for $\eta_1 = \eta_2 = 0.1$. When $\eta_1 = \eta_2 = 0.001$, depending on the SNR, 70%–90% of the processing time is saved by the LS-KRF. Note also that the P-LS-KRF receiver that computes the M columns of the factor matrices in parallel further reduces the processing time of the LS-KRF by a factor of approximately 4.3. Here $4.3 < M = 5$ is due to the non-parallelized part of (15) and (19) in LS-KRF.

We then proceed to evaluate the performance of the accelerated ALS receiver. The parameter settings are the same as those in Fig. 3 except that $F = 3$. Simulation results show that, the accelerated ALS receiver has the same performance as the ALS in

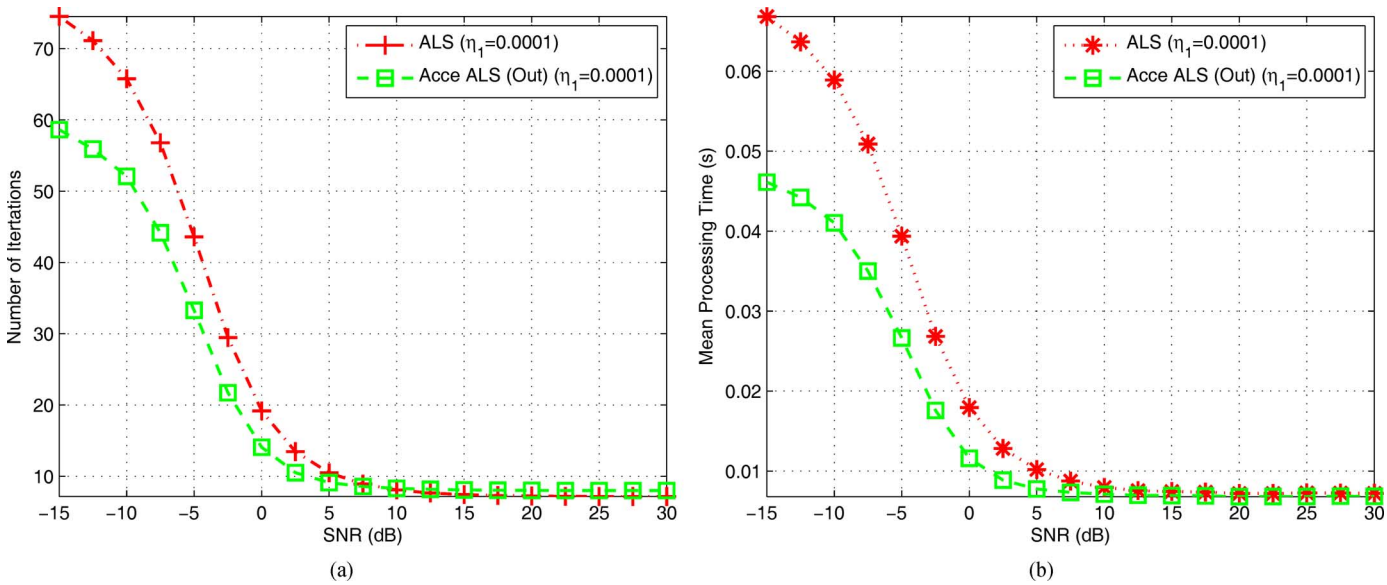


Fig. 5. Comparison of ALS and accelerated ALS receivers in terms of required number of iterations and processing time for various SNRs. (a) Number of iterations required in outer PARAFAC part. (b) Mean processing time in outer PARAFAC part.

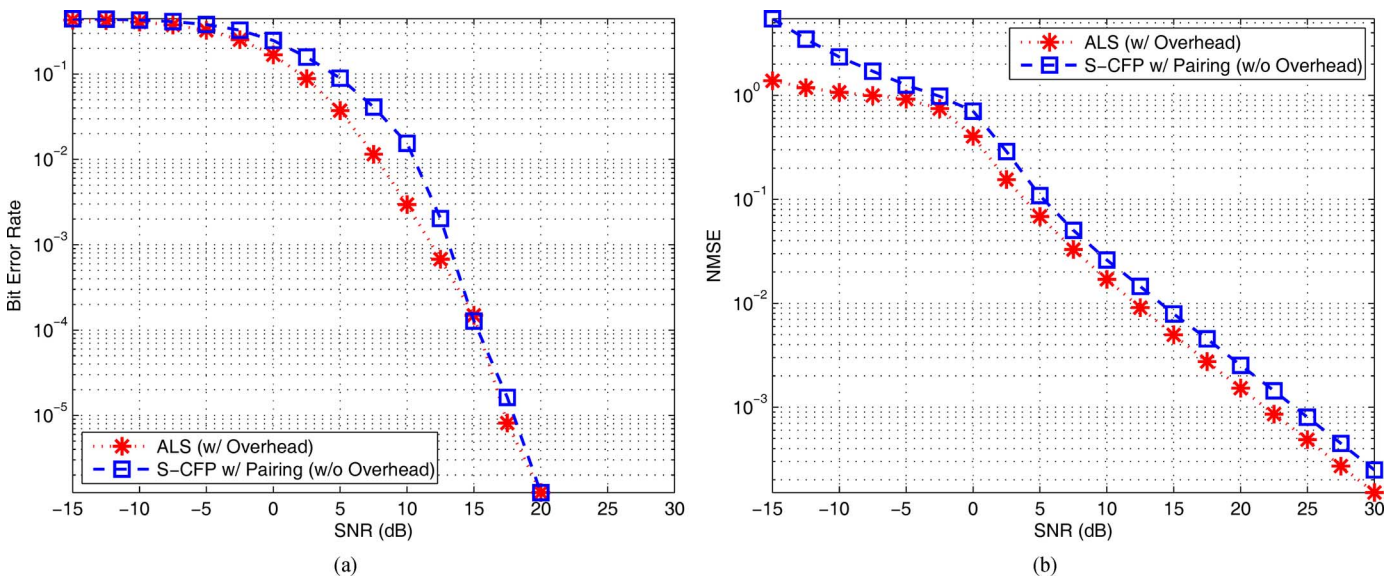


Fig. 6. Comparison of ALS and S-CFP with pairing receivers in terms of BER and NMSE versus SNR. (a) BER versus SNR. (b) Channel NMSE versus SNR.

terms of BER and NMSE for the same η_i. In Fig. 5, we see that the accelerated ALS algorithm applied in the outer PARAFAC part requires less iterations and processing time than its non-accelerated counterpart, with the mean processing time being reduced by 19.2%–41.8% depending on the SNRs. This together with the use of the LS-KRF algorithm in the inner PARAFAC part jointly contribute to the reduction of the whole mean processing time by 39.4%–59.5%. For a larger data size, the time saving is more significant but the results are not included here due to space limitation.

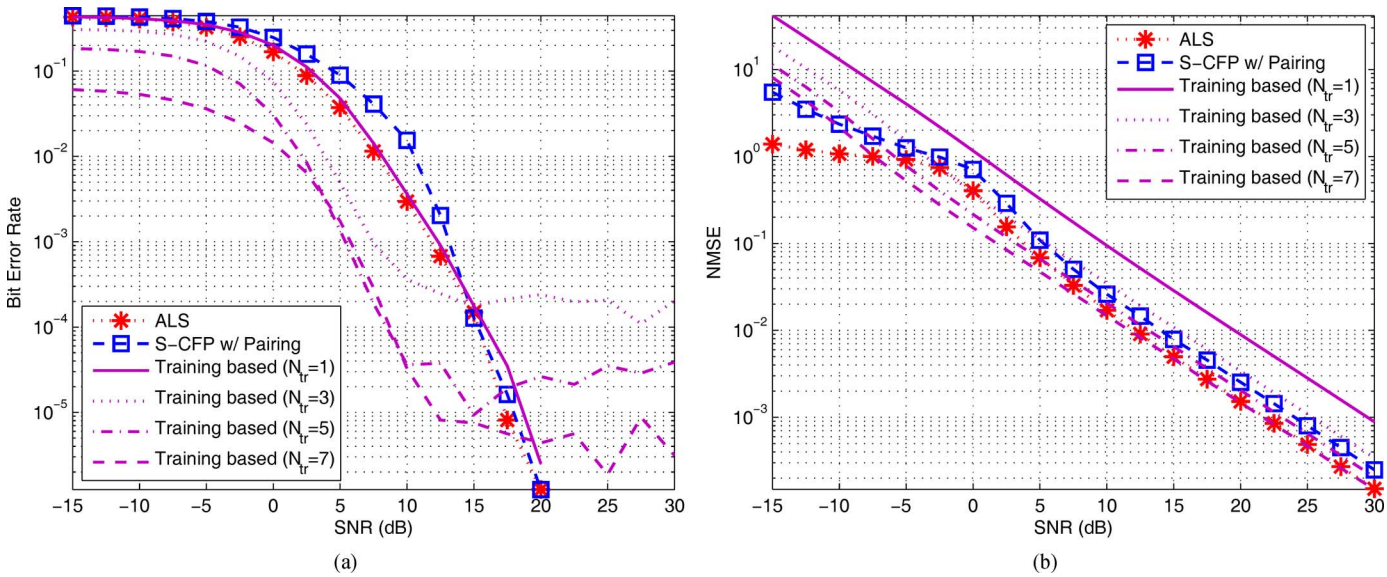
B. ALS Receiver Versus S-CFP with Pairing Receiver: Without Transmission Overhead

For $F \geq M$, in the S-CFP with pairing receiver, the LS-KRF and S-CFP with pairing algorithms are respectively employed in the outer and inner PARAFAC parts. For comparison, the ALS receiver which needs knowledge of the first row in $\boldsymbol{S}$ is used as the performance benchmark.

In Fig. 6, we see that for low-to-medium SNRs less than 10 dB, the S-CFP with pairing receiver is inferior to the ALS receiver. However, the performance gap decreases with SNR, and for high SNRs, their performance is very close to each other. In terms of computational complexity, the comparison in Fig. 8 shows that the mean processing time required in the S-CFP with pairing receiver is no more than 3 to 5 times of that required in the ALS receiver. This corroborates the analysis in Section V-C, which shows that the S-CFP with pairing and ALS receivers require the same order of computational complexity. Considering that the former achieves a gain of 12.5% in transmission efficiency, it is preferred for high SNRs.

C. Comparison with Training-Based Receiver

In the training based receiver, the symbol matrix $\boldsymbol{S} = [\boldsymbol{S}_{\rm tr}^{\rm T}, \boldsymbol{S}_{\rm d}^{\rm T}]^{\rm T} \in \mathbb{C}^{N \times M}$ consists of two parts: $\boldsymbol{S}_{\rm tr} \in \mathbb{C}^{N_{\rm tr} \times M}$ collects the training symbols that are known to the receiver, while $\boldsymbol{S}_{\rm d} \in \mathbb{C}^{N_{\rm d} \times M}$ contains unknown data symbols to be



(a) (b)

Fig. 7. Comparison of ALS, S-CFP with pairing and training based receivers in terms of BER and NMSE versus SNR. (a) BER versus SNR. (b) Channel NMSE versus SNR.

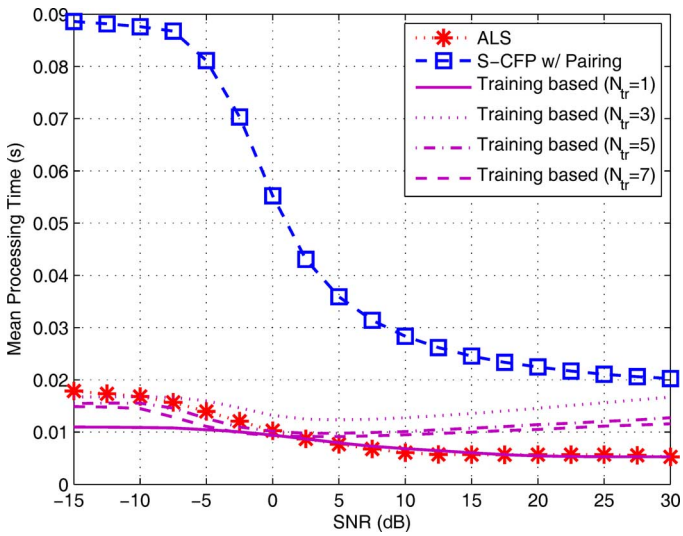


Fig. 8. Comparison of ALS, S-CFP with pairing and training based receivers in terms of processing time versus SNR.

decoded, with $N = N_{\rm tr} + N_{\rm d}$. A popular approach consists in first estimating the channel matrix $\boldsymbol{H}$ using $\boldsymbol{\mathcal{S}}_{\rm tr}$ and then decoding $\boldsymbol{\mathcal{S}}_{\rm d}$ using the estimated channel matrix $\widehat{\boldsymbol{H}}$. For our nested PARAFAC receiver model, the training based receiver algorithm for estimating $\boldsymbol{H}$ and $\boldsymbol{\mathcal{S}}_{\rm d}$ is summarized in Table IV.

The spectral efficiency is $M(N - N_{\rm tr})/(NPF)$ symbol/symbol-period/subcarrier. To establish the uniqueness condition for the training-based receiver, we consider

$$[\boldsymbol{\mathcal{Y}}_{\rm tr}]_1 = \boldsymbol{H} \cdot (\boldsymbol{W} \diamond \boldsymbol{T}_{\rm tr})^{\rm T} \in \mathbb{C}^{K \times N_{\rm tr} PF} \tag{63}$$

as a system of $N_{\rm tr}PKF$ complex-valued equations. By applying similar analysis to Section V-A.1, the necessary uniqueness condition is stipulated by

$$N_{\rm tr}PKF \geq (N_{\rm tr}P + K)M, \tag{64}$$

or equivalently, by

$$N_{\rm tr} \geq \frac{KM}{P(KF - M)}. \tag{65}$$

TABLE IV
TRAINING BASED RECEIVER ALGORITHM

Estimation of $\boldsymbol{H}$ using $\boldsymbol{\mathcal{S}}_{\rm tr}$ and $\boldsymbol{\mathcal{Y}}_{\rm tr}$:

Initialization: Set $i = 0$; Randomly initialize $\widehat{\boldsymbol{A}}$, $\widehat{\boldsymbol{T}}_{\rm tr} = (\boldsymbol{\mathcal{S}}_{\rm tr} \diamond \widehat{\boldsymbol{A}})\boldsymbol{\Theta}^{\rm T}$;

(1.1) $i = i + 1$;

(1.2) Compute the LS estimate of $\boldsymbol{H}$:

$$\widehat{\boldsymbol{H}}(i) = [\boldsymbol{\mathcal{Y}}_{\rm tr}]_1 \cdot [(\boldsymbol{W} \diamond \widehat{\boldsymbol{T}}_{\rm tr}(i-1))^{\dagger}]^{\rm T};$$

(1.3) Compute the LS estimate of $\boldsymbol{T}_{\rm tr}$:

$$\widehat{\boldsymbol{T}}_{\rm tr}(i) = \left([\boldsymbol{\mathcal{Y}}_{\rm tr}]_3 \cdot [(\widehat{\boldsymbol{H}}(i) \diamond \boldsymbol{W})^{\dagger}]^{\rm T}\right);$$

(1.4) Refine the estimate of $\boldsymbol{T}_{\rm tr}$:

$$\widehat{\boldsymbol{A}}(i) = [\widehat{\boldsymbol{T}}_{\rm tr}(i)]_2 \cdot [(\boldsymbol{\Theta} \diamond \boldsymbol{\mathcal{S}}_{\rm tr})^{\dagger}]^{\rm T}; \quad \widehat{\boldsymbol{T}}_{\rm tr}(i) = (\boldsymbol{\mathcal{S}}_{\rm tr} \diamond \widehat{\boldsymbol{A}}(i))\boldsymbol{\Theta}^{\rm T};$$

(1.5) Repeat steps (1.1)-(1.4) until convergence.

Estimation of $\boldsymbol{\mathcal{S}}_{\rm d}$ from $\widehat{\boldsymbol{H}}$:

(2.1) Estimation of $\boldsymbol{T}_{\rm d}$:

$$\widehat{\boldsymbol{T}}_{\rm d} = [\boldsymbol{\mathcal{Y}}_{\rm d}]_3 \cdot [(\widehat{\boldsymbol{H}} \diamond \boldsymbol{W})^{\dagger}]^{\rm T};$$

(2.2) With $\widehat{\boldsymbol{T}}_{\rm d}$, $\boldsymbol{\mathcal{S}}_{\rm d}$ is estimated as

$$\widehat{\boldsymbol{\mathcal{S}}}_{\rm d} = [\widehat{\boldsymbol{T}}_{\rm d}]_1 \cdot [(\widehat{\boldsymbol{A}} \diamond \boldsymbol{\Theta})^{\dagger}]^{\rm T},$$

or via the LS-KRF described in Section IV-A1.

For unique recovery of $\boldsymbol{H}$ and $\boldsymbol{\mathcal{S}}$, the required minimum number of time-slots in each time-frame is $N_{\rm tr} = 1$ according to (65). In Fig. 7, the proposed S-CFP with pairing, ALS and training based receivers with $N_{\rm tr} = 1, 3, 5, 7$, are compared in terms of BER and channel NMSE. We see that as $N_{\rm tr}$ increases, the training based receiver consistently improves in both BER and channel estimation accuracy. For low-to-medium SNRs less than 10 dB, the performance improvement is significant even $N_{\rm tr} = 1$, namely, only one time-slot in each time-frame is used as training data. Therefore, the training based receiver is preferred for low-to-medium SNRs. However, with the increase of the SNR the performance gap becomes smaller and when the SNR rises above 20 dB, their performance is very close to each other. Moreover, Fig. 8 shows that the mean processing

time required in the training based receiver is comparable to that in the ALS receiver, although the former slightly varies with N_{tr} and reaches its maximum when N_{tr} is equal to $N/2$. Therefore, the S-CFP with pairing receiver needs the same order of processing time as the training based receivers as well. In particular, for high SNRs, the mean processing time of the former is no more than 1.2–2 times that in the latter. Considering that the latter suffers from a loss in transmission efficiency to various degrees, our proposed S-CFP with pairing receiver is a preferred choice for high SNRs.

VII. Conclusion

We have proposed two semi-blind receivers, namely, the least squares Khatri-Rao factorization (LS-KRF) and simplified closed-form PARAFAC decomposition (S-CFP) coupled with a pairing algorithm, for joint symbol and channel estimation. The LS-KRF receiver is a closed-form solution that provides performance stability as well as lower computational complexity compared to the alternating least squares (ALS) based algorithm, while the S-CFP with pairing receiver allows for higher transmission efficiency since it does not require the use of pilot symbols for symbol detection. The uniqueness conditions, spectral efficiency and computational complexity of the LS-KRF and S-CFP with pairing receivers are analyzed and compared with the ALS receiver algorithm. It is shown that the S-CFP with pairing receiver has the same order of computational complexity as the ALS receiver. Moreover, simulation results show that our S-CFP with pairing receiver attains nearly the same performance as the ALS and training based receivers with extra pilot overhead at sufficiently high signal-to-noise ratio conditions.

Acknowledgment

The authors would like to thank the anonymous reviewers for their constructive comments that greatly improved the quality of the manuscript.

References

[1] G. L. Stuber, J. R. Barry, S. W. McLaughlin, Y. Li, M. A. Ingram, and T. G. Pratt, "Broadband MIMO-OFDM wireless communications," *Proc. IEEE*, vol. 92, no. 2, pp. 271–294, Feb. 2004.

[2] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Aug. 1998.

[3] D. Agrawal, V. Tarokh, A. Naguib, and N. Seshadri, "Space-time coded OFDM for high data-rate wireless communication over wideband channels," in *Proc. 48th IEEE Veh. Technol. Conf. (VTC '1998)*, Ottawa, Canada, May 1998, vol. 3, pp. 2232–2236.

[4] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, May 1999.

[5] H. Bolcskei and A. J. Paulraj, "Space-frequency coded broadband OFDM systems," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC '2000)*, Chicago, IL, USA, Sep. 2000, pp. 1–6.

[6] W. Su, Z. Safar, and K. J. R. Liu, "Full-rate full-diversity space-frequency codes with optimum coding advantage," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 229–249, Jan. 2005.

[7] L. Shao and S. Roy, "Rate-one space-frequency block codes with maximum diversity for MIMO-OFDM," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1674–1687, Apr. 2005.

[8] N. Sarmadi, S. Shahbazpanahi, and A. B. Gershman, "Blind channel estimation in orthogonally coded MIMO-OFDM systems: A semidefinite relaxation approach," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2354–2364, Jun. 2009.

[9] J. Vía, I. Santamaría, J. Pérez, and L. Vielva, "A new subspace method for blind estimation of selective MIMO-STBC channels," *Wireless Commun. Mobile Comput.*, vol. 10, no. 11, pp. 1478–1492, 2010.

[10] N. D. Sidiropoulos and R. S. Budampati, "Khatri-Rao space-time codes," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2396–2407, 2002.

[11] R. S. Budampati and N. D. Sidiropoulos, "Khatri-Rao space-time codes with maximum diversity gains over frequency-selective channels," in *Proc. Sens. Array and Multichannel Signal Process. Workshop*, Rosslyn, VA, USA, Aug. 2002, pp. 432–436.

[12] A. L. F. de Almeida, G. Favier, and J. C. M. Mota, "Space-time multiplexing codes: A tensor modeling approach," presented at the IEEE 7th Workshop Signal Process. Adv. Wireless Commun. (SPAWC), Cannes, France, Jul. 2006.

[13] A. L. F. de Almeida, G. Favier, C. A. R. Fernandes, and J. C. M. Mota, "A trilinear decomposition approach for space-time-frequency multiple-access wireless systems," presented at the IEEE 8th Workshop on Signal Process. Adv. Wireless Commun. (SPAWC), Helsinki, Finland, Jun. 2007.

[14] A. L. F. de Almeida, G. Favier, and J. C. M. Mota, "Constrained tensor modeling approach to blind multiple-antenna CDMA schemes," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2417–2428, Jun. 2008.

[15] A. L. F. de Almeida, G. Favier, and J. C. M. Mota, "A constrained factor decomposition with application to MIMO antenna systems," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2429–2442, Jun. 2008.

[16] 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," *Signal Process.*, vol. 89, no. 11, pp. 2103–2116, Nov. 2009.

[17] A. L. F. de Almeida, "Blind joint detection and channel estimation in space-frequency diversity systems using time-varying linear constellation precoding," presented at the Brazilian Telecommun. Symp. (SBrT'11), Curitiba, Paraná, Brazil, Oct. 2011.

[18] M. N. da Costa, G. Favier, A. L. F. de Almeida, and J. M. T. Romano, "Tensor space-time (TST) coding for MIMO wireless communication systems," *Signal Process.*, vol. 92, no. 4, pp. 1079–1092, Apr. 2012.

[19] A. L. F. de Almeida, G. Favier, and L. Ximenes, "Space-time-frequency (STF) MIMO communication systems with blind receiver based on a generalized PARATUCK2 model," *IEEE Trans. Signal Process.*, vol. 61, no. 8, pp. 1895–1909, Apr. 2013.

[20] K. Liu, J. P. C. L. da Costa, A. L. F. de Almeida, and H. C. So, "A closed-form solution to semi-blind joint symbol and channel estimation in MIMO-OFDM systems," in *Proc. IEEE Int. Conf. Signal Process., Commun. Comput. (ICSPCC'2012)*, Hong Kong, Aug. 2012, pp. 191–196.

[21] R. A. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an explanatory multimodal factor analysis," *UCLA Working Papers in Phonet.*, vol. 16, no. 1, pp. 1–84, 1970.

[22] R. Bro, "PARAFAC. Tutorial and applications," *Chemom. Intell. Lab. Syst.*, vol. 38, no. 2, pp. 149–171, Oct. 1997.

[23] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966.

[24] R. Harshman and M. Lundy, "Uniqueness proof for a family of models sharing features of Tucker's three-mode factor analysis and PARAFAC/CANDECOMP," *Psychometrika*, vol. 61, no. 1, pp. 133–154, 1996.

[25] F. Roemer and M. Haardt, "Tensor-based channel estimation and iterative refinements for two-way relaying with multiple antennas and spatial reuse," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5720–5735, Nov. 2010.

[26] K. Hwang and Z. Xu, *Scalable Parallel Computing: Technology, Architecture, Programming*. New York, NY, USA: McGraw-Hill, 1998.

[27] F. Roemer and M. Haardt, "A closed-form solution for multilinear PARAFAC decompositions," in *Proc. 5th IEEE Sens. Array and Multichannel Signal Process. Workshop (SAM'2008)*, Darmstadt, Germany, Jul. 2008, pp. 487–491.

[28] L. de Lathauwer, B. de Moor, and J. Vanderwalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253–1278, 2000.

[29] Y. Xin, Z. Wang, and G. B. Giannakis, "Space-time diversity systems based on linear constellation precoding," *IEEE Trans. Wireless Commun.*, vol. 2, no. 2, pp. 294–309, Feb. 2003.

[30] R. Vishwanath and M. R. Bhatnagar, "Optimum linear constellation precoding for space time wireless systems," *Wireless Pers. Commun.*, vol. 40, no. 4, pp. 511–521, Mar. 2007.

[31] J.-F. Cardoso and A. Souloumiac, "Jacobi angles for simultaneous diagonalization," *SIAM J. Matrix Anal. Appl.*, vol. 17, no. 1, pp. 161–164, 1996.

[32] T. Fu and X. Gao, "Simultaneous diagonalization with similarity transformation for non-defective matrices," in *Proc. Int. Conf. Acoust, Speech, Signal Process. (ICASSP '2006)*, Toulouse, France, May 2006, vol. 4, pp. 1137–1140.

[33] M. Haardt and J. A. Nossek, "Simultaneous Schur decomposition of several non-symmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems," *IEEE Trans. Signal Process.*, vol. 46, no. 1, pp. 161–169, Jan. 1998.
[34] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numer. Math.*, vol. 1, no. 1, pp. 269–271, 1959.
[35] N. Li, S. Kindermann, and C. Navasca, "Some convergence results on the regularized alternating least-squares method for tensor decomposition," *Linear Algebra Appl.*, vol. 438, no. 2, pp. 796–812, Jan. 2011.
[36] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra Appl.*, vol. 18, no. 2, pp. 95–138, 1977.
[37] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, "Blind PARAFAC receivers for DS-CDMA systems," *IEEE Trans. Signal Process.*, vol. 48, no. 3, pp. 81–823, Mar. 2000.
[38] V. Rokhlin, A. Szlam, and M. Tygert, "A randomized algorithm for principal component analysis," *SIAM J. Matrix Anal. Appl.*, vol. 31, no. 3, pp. 1100–1124, 2009.
[39] W. W. Hager, "Condition estimates," *SIAM J. Sci. Statist. Comput.*, vol. 5, no. 2, pp. 311–316, 1984.

Kefei Liu (S'12) recently received his Ph.D. degree in electronic engineering from City University of Hong Kong. He received the B.Sc. degree from Wuhan University and M.Sc. degree from Beihang University, both in mathematics, in 2006 and 2009, respectively.

From February to August 2012, he studied at the Department of Electrical Engineering, University of Brasilia and the Institute for Circuit Theory and Signal Processing, Technical University of Munich, as a visiting research student. His research interests are statistical signal and array processing, model order selection, prewhitening, parameter estimation and multilinear algebra, centering about wireless communication and radar applications.

João Paulo C. L. da Costa (M'11) was born in Fortaleza—CE, Brazil, on May 22, 1981. He received the Diploma degree in electronic engineering in 2003 from the Military Institute of Engineering (IME), Rio de Janeiro, Brazil, the M.Sc. degree in telecommunications in 2006 from the University of Brasília (UnB), and the Doktor-Ingenieur (Ph.D.) degree in electrical engineering and information technology *(magna cum laude)* in 2010 from Ilmenau University of Technology (TU Ilmenau), Germany. During his Ph.D. studies, he was a scholarship holder of the National Counsel of Technological and Scientific Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq) of the Brazilian Government and also a captain of the Brazilian Army.

Currently, he is a professor with the Department of Electrical Engineering, UnB. He has published more than 50 papers in refereed journals and conference proceedings. He is a member of the Graduate Program in Electrical Engineering (PPGEE), of the Laboratory of Automation and Robotics (LARA) and of the Microwave and Wireless Research Group (MWSL) at UnB. He is responsible for the Laboratory of Array Signal Processing (LASP) at UnB. He has been visiting professor at the University of Erlangen-Nuremberg, Munich Technical University, University of Seville, and Harvard University. His research interests are in the areas of multidimensional array signal processing, MIMO communication systems, microphone array signal processing, multilinear algebra, principal component analysis (PCA), and model order selection.

Dr. da Costa received three conference Best Paper awards at the International Congress on Ultra Modern Telecommunications and Control Systems (ICUMT'12) and the international conference on forensic computer science (ICoFCS'12 and ICoFCS'13).

H. C. So (S'90–M'95–SM'07) was born in Hong Kong. He received the B.Eng. degree from the City University of Hong Kong and the Ph.D. degree from The Chinese University of Hong Kong, both in electronic engineering, in 1990 and 1995, respectively.

From 1990 to 1991, he was an electronic engineer with the Research and Development Division, Everex Systems Engineering Ltd., Hong Kong. During 1995–1996, he worked as a Postdoctoral Fellow at The Chinese University of Hong Kong. From 1996 to 1999, he was a Research Assistant Professor in the Department of Electronic Engineering, City University of Hong Kong, where he is currently an Associate Professor. His research interests include statistical signal processing, fast and adaptive algorithms, signal detection, parameter estimation, and source localization.

Dr. So has been on the editorial boards of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, *Signal Processing*, *Digital Signal Processing*, and ISRN *Applied Mathematics*, as well as a member of the Signal Processing Theory and Methods Technical Committee of the IEEE Signal Processing Society.

André L. F. de Almeida (M'08–SM'13) received the B.Sc. and M.Sc. degrees in electrical engineering from the Federal University of Ceará, Brazil, in 2001 and 2003, respectively, and the double Ph.D. degree in sciences and teleinformatics engineering from the University of Nice, Sophia Antipolis, France, and the Federal University of Ceará, Fortaleza, Brazil, in 2007.

He is currently an Assistant Professor with the Department of Teleinformatics Engineering of the Federal University of Ceará. During fall 2002, he was a visiting researcher at Ericsson Research Labs, Stockholm, Sweden. From 2007 to 2008, he held a one-year research position at the I3S Laboratory, CNRS, France. In 2008, he was awarded a CAPES/COFECUB research fellowship with the I3S Laboratory, CNRS, France. In 2010, he was appointed a productivity research fellow from the Brazilian National Council for Scientific and Technological Development (CNPq). In spring 2012, he was a visiting professor at the University of Nice Sophia-Antipolis. His research interests lie in the area of signal processing for communications, and include blind identification, source separation, MIMO systems, tensor decompositions, and multilinear algebra applied to communications and data analysis.

Dr. de Almeida is affiliated with the IEEE Signal Processing for Communications and Networking (SPCOM) Technical Committee of the IEEE Signal Processing Society. He serves as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, *Circuits, Systems & Signal Processing*, and the KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEMS.