# Optical Coupler Network Modeling and Parameter Estimation Based on a Generalized Tucker Train Decomposition

**DANILO S. ROCHA** [1], **FRANCISCO T. C. B. MAGALHÃES** [1,2], **GÉRARD FAVIER** [3],
**ANTONIO S. B. SOMBRA** [2], **AND GLENDO F. GUIMARÃES** [1]

[1] Photonics Laboratory, Federal Institute of Education, Science and Technology of Ceará, Fortaleza 60040-531, Brazil
[2] LOCEM, Federal University of Ceará, Fortaleza 60440-900, Brazil
[3] I3S Laboratory, Côte d'Azur University, 06903 Sophia Antipolis, France

Corresponding author: Danilo S. Rocha (danilo@fisica.ufc.br)

**ABSTRACT** Tensor models have been used extensively in signal processing applications to design different types of communication systems. This paper proposes, for the first time, the use of tensor models for optical communications. The signals of an optical dual-core coupler network are modeled as a multiway array (tensor), which satisfies a generalized Tucker train decomposition. This tensor model is then used to develop an estimation algorithm that enables the network parameters to be estimated from the input and output signals. The performance of this algorithm was evaluated by means of computer simulations, in terms of NMSE of the estimated parameters and convergence speed. For the tested configurations, good levels of NMSE with fast convergence were obtained, demonstrating the effectiveness of the proposed method as a promising tool for studying and designing optical devices, with a wide range of applications in the context of lightwave systems.

**INDEX TERMS** Alternating least squares, multidimensional signal processing, multilinear algebra, optical arrays, optical directional coupler, optical fiber devices, optical switches, parameter estimation, systems modeling, tensor analysis.

## I. INTRODUCTION

Over the last three decades, communications based on lightwave systems have developed rapidly. The worldwide demand for information technology has been increasing constantly and requires a continuous growth in connectivity of data center networks (DCNs). The scalability of a DCN is limited by the total capacity of the switching and multiplexing devices it uses. Since the fourth generation of lightwave systems, wavelength division multiplexing (WDM) techniques have been widely used to increase the bit rate [1].

The most common optical devices used in power switching are optical couplers. These are devices that function as optical beam splitters, and are used to design interferometric systems such as Mach-Zehnder (MZI) and Sagnac interferometers, as considered in several works [2]–[5]. The combination of several optical couplers in an arranged

The associate editor coordinating the review of this manuscript and approving it for publication was Gerard-Andre Capolino.

network provides the design of devices such as star couplers, routers, switches, in addition to the interferometers already mentioned.

Recently, new topologies using (star) couplers have been presented targeting gains in transmission rates compared to conventional architectures. The works [6], [7] designed flexible high port count stars for optical circuit switches (OCS). In [7], the authors presented the design of a flexible optical star by adding reconfigurable switching elements to the core of the topology, without significantly increasing the power consumption. The results showed an increase of 26–40% in the transmission rate per node. In [6], a WDM star coupler architecture was proposed for the construction of high-radius optical switches to improve DCN scalability.

To study such optical devices, it is necessary to solve the nonlinear Schroedinger's equations (NLS), since they govern the signal propagation in optical fibers and in fiber-based devices. However, NLS equations only have possible

D. S. Rocha *et al.*: Optical Coupler Network Modeling and Parameter Estimation Based on Generalized Tucker Train Decomposition

**IEEE** *Access*

analytical solutions in very particular situations; therefore, simulations based on numerical methods are an essential tool to study in order to gain improvements in optical systems. In the literature, there are many works involving couplers and other fiber-based devices, where it is quite common to use numerical methods such as the fourth-order Runge-Kutta (RK4) [2], [3], [8], [9], the split-step Fourier Method (SSFM) [10]–[12] and the finite-difference-time-domain (FDTD) [4], [13]–[15]. However, such methods require extensive simulations and a large amount of data to be analyzed, which may hinder the interpretation of the obtained results.

High order tensors (i.e., multiway arrays) and tensor decompositions [16] have become an important tool for designing different types of systems and applications that involve multidimensional data from areas such as chemometrics, psychometrics, numerical analysis, computer vision and telecommunications. During the last decade, new tensor decompositions have been proposed with the aim to alleviate the curse of dimensionality encountered with large-scale tensors, and also to design new wireless communication systems. Recently, tensor approaches have gained considerable space in signal processing [17], communication systems [18], machine learning [19], [20] and blind source separation [21].

Furthermore, tensor decompositions have some advantages over conventional matrix-based methods, like the uniqueness property of certain tensor models under milder conditions than those required by matrix approaches. However, to the best authors' knowledge, tensor models have not been applied in the context of optical communication systems until now.

This work proposes a novel tensor-based approach applied to optical communication systems. The main goal of this work is to consider tensor analysis as a new tool to study and model optical devices and to solve problems related to the design and signal processing of optical communication systems. The first scenario to be considered consists of an optical network composed of linear directional couplers that are interconnected in such a way that the input signals are switched along the network, providing output signals that depend on the system parameters. The signals at the end of the network form a third-order tensor with the dimensions *number of ports in a coupler × number of couplers × time (bit slot)* arranged as rows, columns, and tubes of the tensor, respectively. The proposed tensor model represents the output signals as a function of the input signals. The steps of the transmission throughout the network are represented by a fourth-order tensor that can be viewed as a generalization of a tensor train decomposition [22], whose the wagons (tensors that compose the train) satisfy a known tensor model, namely a generalized Tucker decomposition [23], [24]. However, the model that will be exploited in this paper differs from that introduced in [23] due to the presence of structural tensor factors, and our proposal exploits these structural constraints to estimate the system parameters.

From the proposed model, we derive cost functions to be minimized using an iterative algorithm to estimate the network parameters from input and output signals. The parameter estimation points to the possibility of designing optical devices/networks in such a way that they take as their starting point the expected response, like for an inverse design problem. Examples of devices that can be designed from the expected outputs include optical logic gates, routers, multiplexers, and power splitters. In this study, synthetic data are used to validate the proposed tensor model.

The main contributions of this work can be summarized as follows:

- this is a novel tensor modeling of optical devices. The main approach considers an array of optical couplers performing a 4 × 4 star coupler;
- the output signals of the network proposed here are described using a tensor train structure, which is composed of blocks that satisfy a generalized Tucker decomposition;
- the proposed tensor structure is extended to model the signals of a generic optical coupler network with various configurations;
- the proposed tensor model is used to develop an iterative estimation algorithm to estimate unknown network parameters from the input and output signals. This parameter estimation algorithm can be applied to design optical devices according to the expected outputs for a given application;
- the effectiveness and accuracy of the proposed algorithm to estimate the parameters of a generic coupler network are illustrated by means of computer simulations.

The rest of the paper is organized as follows. In Section II, we provide a theoretical background on tensor notations, basic operations involving matrices and tensors and some tensor decompositions. In Section III, the optical coupler and some applications are described. The mathematical model is presented for the low-power continuous wave signal. In Section IV, we propose the use of tensors for optical communications as a new tool to model and study optical devices. Section V introduces an estimation algorithm, based on the model proposed in Section IV; this algorithm provides estimates of the main parameters of the coupler networks under consideration. Numerical simulation results, perspectives and conclusions are presented respectively in Sections VI, VII and VIII.

## II. FUNDAMENTALS OF TENSORS

The theory of tensors is a branch of multilinear algebra. High order tensors can be viewed as a generalization of matrices, and they are represented as multi-dimensional arrays of order higher than two. Over the last two decades, tensor models have been extensively used for designing different types of communication systems. The main motivation for using tensor-based approaches is related to their ability to model multimodal data, with essential uniqueness properties under mild conditions.

**IEEE** Access

D. S. Rocha *et al.*: Optical Coupler Network Modeling and Parameter Estimation Based on Generalized Tucker Train Decomposition

**TABLE 1.** Tensor notations.

| Symbols | Definitions |
|---|---|
| $a \in \mathbb{C}$ | Scalar |
| $\mathbf{a} \in \mathbb{C}^I$ | Complex column vector of dimension $I$ |
| $\mathbf{A} \in \mathbb{C}^{I \times J}$ | Complex matrix of size $I \times J$ |
| $\mathbf{e}_n^{(N)}$ | $n$-th canonical basis vector of the Euclidean space $\mathbb{R}^N$ |
| $\mathbf{A}^T$ | Matrix transpose of $\mathbf{A}$ |
| $\mathbf{A}^\dagger$ | Moore-Penrose pseudo-inverse of $\mathbf{A}$ |
| $\mathbf{A}_{i\cdot}$ ($\mathbf{A}_{\cdot j}$) | $i$-th ($j$-th) row (column) of $\mathbf{A} \in \mathbb{C}^{I \times J}$ |
| $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ | $N$-order tensor of size $I_1 \times I_2 \times \cdots \times I_N$ |
| $a_{i_1, i_2, \cdots, i_N}$ | $(i_1, i_2, \cdots, i_N)$-th element of tensor $\mathcal{A}$ |
| $\mathcal{A}_{\cdots i_n \cdots}$ | Mode-$n$ tensor slice |
| $\mathbf{A}_{I_1 \cdots I_{N-1} \times I_N}$ | Tall mode-$n$ matrix unfolding of $\mathcal{A}$ |
| $vec(\cdot)$ | Vectorization operator |
| $diag(\cdot)$ | Diagonal matrix formed from the elements of the argument vector |
| $bdiag(\cdot)$ | Block-diagonal operator |
| $r_\mathbf{A}$ | Rank of $\mathbf{A}$ |
| $\det(\mathbf{A})$ | Determinant of $\mathbf{A}$ |
| $\otimes$ | Kronecker product |
| $\times_n$ | Mode-$n$ product |
| $*_m^n$ | Contraction operation over $m$-th mode of the first factor and $n$-th mode of the second factor |
| $\|\cdot\|_F$ | Frobenius norm |

In this section, we provide a brief overview of the mathematical background useful for this work, concerning basic tensor operations, and some tensor decompositions.

### A. NOTATION AND MAIN OPERATIONS
Table 1 summarizes the notation that will be used throughout this manuscript. Following, we discuss some details of the main concepts and operations.

#### 1) VECTORIZATION AND DIAGONALIZATION
The operator $vec(\cdot)$ transforms a matrix into a column vector by stacking the columns of its matrix argument. The operator $diag(\cdot)$ forms a diagonal matrix from its vector argument. Similarly, the operator $bdiag(\mathbf{A}_1, \ldots, \mathbf{A}_K) \triangleq bdiag(\mathbf{A}_k)$ forms a block-diagonal matrix composed of the $K$ matrices $\mathbf{A}_k$, with $k = 1, \ldots, K$, on the diagonal.

#### 2) KRONECKER PRODUCT
The Kronecker product of two matrices $\mathbf{A} \in \mathbb{C}^{I \times J}$ and $\mathbf{B} \in \mathbb{C}^{M \times N}$ is defined as $\mathbf{C} = \mathbf{A} \otimes \mathbf{B} \in \mathbb{C}^{IM \times JN}$, whose entries are $c_{(i-1)M+m,(j-1)N+n} = a_{i,j}b_{m,n}$. Given a set $\mathbb{S} = \{1, \ldots, N\}$ and the matrices $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times J_n}$, a multiple Kronecker product is denoted as $\bigotimes_{n \in \mathbb{S}} \mathbf{A}^{(n)} \triangleq \mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)} \otimes \cdots \otimes \mathbf{A}^{(N)} \in \mathbb{C}^{I_1 \cdots I_N \times J_1 \cdots J_N}$.

#### 3) FIBERS AND SLICES
Fibers are vectors obtained by fixing the indices of all modes of a tensor, except one. For example, a third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ has three kinds of fibers:
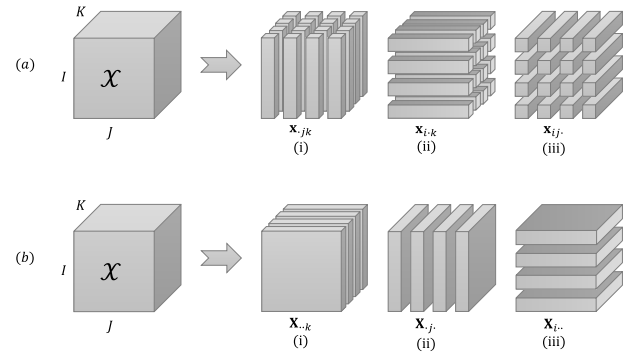


**FIGURE 1.** Tensor structure: (a) column, row and tube fibers of a third-order tensor; (b) frontal, lateral and horizontal matrix slices of a third-order tensor.

(i) columns ($\mathbf{x}_{\cdot jk} \in \mathbb{C}^I$) obtained by fixing the indices $j$ and $k$;
(ii) rows ($\mathbf{x}_{i \cdot k} \in \mathbb{C}^J$) obtained by fixing the indices $i$ and $k$;
(iii) tubes ($\mathbf{x}_{ij\cdot} \in \mathbb{C}^K$) obtained by fixing the indices $i$ and $j$.

Analogously, matrix slices are obtained by varying the indices of two modes and fixing all the others. For a third-order tensor $\mathcal{X}$, the three kinds of matrix slices are: (i) frontal slices ($\mathbf{X}_{\cdot \cdot k} \in \mathbb{C}^{I \times J}$); (ii) lateral slices ($\mathbf{X}_{\cdot j \cdot} \in \mathbb{C}^{I \times K}$); (iii) horizontal slices ($\mathbf{X}_{i \cdot \cdot} \in \mathbb{C}^{J \times K}$).

Fig. 1 illustrates the fibers and matrix slices for a third-order tensor. For a given $N$-th order tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$, tensor slices of order higher than 2 are obtained by varying $N_1$ indices, with $3 \leq N_1 < N$, while the other $N - N_1$ indices remain fixed, resulting in a $N_1$-th order tensor slice. Note that, if we assume $N_1 = 1$, we get fibers, and if we assume $N_1 = 2$, we get matrix slices.

#### 4) MATRIX UNFOLDING
Also called matricization, it is a matrix representation of a high-order tensor. The tall (or flat) mode-$n$ unfolding of $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ is obtained by mapping its elements into a matrix $\mathbf{A}_{I_{n+1} \cdots I_N I_1 \cdots I_{n-1} \times I_n}$ (or $\mathbf{A}_{I_n \times I_{n+1} \cdots I_N I_1 \cdots I_{n-1}}$), whose entries are $[\mathbf{A}_{I_{n+1} \cdots I_N I_1 \cdots I_{n-1} \times I_n}]_{\bar{i}, i_n}$, with

$$\bar{i} = (i_{n+1} - 1)I_{n+2} \cdots I_N I_1 \cdots I_{n-1} + \cdots + (i_{n-2} - 1)I_{n-1} + i_{n-1}. \quad (1)$$

From the above definition, we can see a mode-$n$ unfolding as a rearrangement of the elements of $\mathcal{A}$ obtained by varying index $i_n$ and keeping the other indices fixed, in such a way that the fibers of the $n$-th mode are placed along the rows (tall unfolding) or columns (flat unfolding). For example, let us consider a third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$. There are two flat mode-1 unfoldings, $\mathbf{X}_{I \times KJ}$ and $\mathbf{X}_{I \times JK}$, which consist of column fibers placed side by side, according to the order of combination of the last two modes. Here, we consider that the index that varies the fastest is the one associated with the last dimension, i.e., $[\mathbf{X}_{I \times KJ}]_{i,(k-1)J+j} = [\mathcal{X}]_{i,j,k}$ and $[\mathbf{X}_{I \times JK}]_{i,(j-1)K+k} = [\mathcal{X}]_{i,j,k}$. Equivalently, an unfolded matrix can be obtained by stacking the slices of a given mode, as shown in Fig. 2. Note that, in our notation, the subscript in
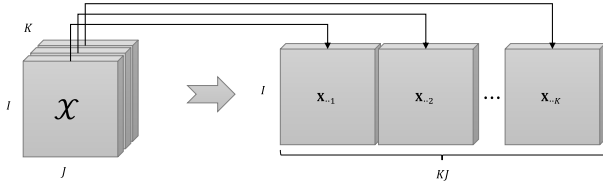
D. S. Rocha *et al.*: Optical Coupler Network Modeling and Parameter Estimation Based on Generalized Tucker Train Decomposition

**IEEE** *Access*



**FIGURE 2.** Matrix representation of the third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$. The second and third dimensions of $\mathcal{X}$ are combined to form the columns, yielding the matrix unfolding $\mathbf{X}_{I \times KJ}$.

the unfolded matrix indicates the order in which the modes are combined and, consequently, the size of the unfolded matrix.

Unlike the mode-$n$ unfolding, which combines all modes of a tensor in rows (or columns) except for the $n$-th mode, it is possible to define a matrix unfolding where multiple modes are combined along the rows and columns of the resulting matrix. For this unfolding, let us define the set $\mathbb{S} = \{1, \cdots, N\}$ and the representations $I_{\mathbb{S}}$, $\bar{I}_{\mathbb{S}}$ and $i_{\mathbb{S}}$, which denote, respectively, short forms for the dimension, product of dimensions and the set of indices associated to the modes of the set $\mathbb{S}$. For instance, for $N = 3$, we have $I_{\mathbb{S}} = I_1 \times I_2 \times I_3$, $\bar{I}_{\mathbb{S}} = I_1 I_2 I_3$ and $i_{\mathbb{S}} = \{i_1, i_2, i_3\}$. Thus, defining $\mathbb{S}_1$ and $\mathbb{S}_2$ as ordered subsets of the set $\mathbb{S}$, such that $\mathbb{S}_1 \cup \mathbb{S}_2 = \mathbb{S}$, we have for $\mathcal{A} \in \mathbb{C}^{I_{\mathbb{S}}}$ [25]

$$\mathbf{A}_{\bar{I}_{\mathbb{S}_1} \times \bar{I}_{\mathbb{S}_2}} = \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} a_{i_1, \cdots, i_N} \left( \underset{n \in \mathbb{S}_1}{\otimes} \mathbf{e}_{i_n}^{(I_n)} \right) \left( \underset{n \in \mathbb{S}_2}{\otimes} \mathbf{e}_{i_n}^{(I_n)} \right)^T, \tag{2}$$

with $\bar{I}_{\mathbb{S}_1}$ and $\bar{I}_{\mathbb{S}_2}$ being any ordered combination of the dimensions of the tensor $\mathcal{A}$.

### 5) TENSOR-MATRIX MODE-$n$ PRODUCT

Given a $N$-th order tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$, let us consider a matrix $\mathbf{U} \in \mathbb{C}^{R_n \times I_n}$, whose number of columns is equal to the dimension of the $n$-th mode of $\mathcal{A}$. The mode-$n$ product of $\mathcal{A}$ with $\mathbf{U}$ yields a $N$-th order tensor $\mathcal{C} = \mathcal{A} \times_n \mathbf{U} \in \mathbb{C}^{I_1 \times \cdots \times I_{n-1} \times R_n \times I_{n+1} \times \cdots \times I_N}$ defined as

$$c_{i_1, \cdots, i_{n-1}, r_n, i_{n+1}, \cdots, i_N}$$
$$= \sum_{i_n=1}^{I_n} a_{i_1, \cdots, i_{n-1}, i_n, i_{n+1}, \cdots, i_N} u_{r_n, i_n}. \tag{3}$$

### 6) TENSOR-TENSOR MODE-$n$ PRODUCT

Given a set $\mathbb{S} = \{1, \cdots, N\}$, let $\mathbb{S}_t$ be an ordered subset of $\mathbb{S} - \{n\}$, with $1 \leq n \leq N$. Let us consider an $N_t$-th order tensor $\mathcal{T} \in \mathbb{C}^{R_n \times I_n \times I_{\mathbb{S}_t}}$, with $3 \leq N_t \leq N + 1$. The mode-$n$ product of the tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ and $\mathcal{T}$, denoted by $\mathcal{A} \times_n \mathcal{T}$, gives a tensor $\mathcal{C} \in \mathbb{C}^{I_1 \times \cdots \times I_{n-1} \times R_n \times I_{n+1} \times \cdots \times I_N}$ defined as

$$c_{i_1, \cdots, i_{n-1}, r_n, i_{n+1}, \cdots, i_N}$$
$$= \sum_{i_n=1}^{I_n} a_{i_1, \cdots, i_{n-1}, i_n, i_{n+1}, \cdots, i_N} t_{r_n, i_n, i_{\mathbb{S}_t}}. \tag{4}$$

Note that, by convention, the two above operations assume that the second mode of $\mathbf{U}$ (or $\mathcal{T}$) is equal to the $n$-th mode of $\mathcal{A}$. In both cases, the mode-$n$ product does not change the size of the resulting tensor, but provides a linear transformation on the mode-$n$ space of $\mathcal{A}$.

### 7) CONTRACTION OPERATION

Let us consider the tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ and $\mathcal{B} \in \mathbb{C}^{J_1 \times \cdots \times J_M}$ sharing a common dimension ($I_p = J_q = K$, with $1 \leq p \leq N$ and $1 \leq q \leq M$). The contraction of $\mathcal{A}$ with $\mathcal{B}$, denoted by $\mathcal{A} *_p^q \mathcal{B}$, is defined as the following sum over the common mode ($i_p = j_q = k$) [26]

$$c_{i_1, \cdots, i_{p-1}, j_1, \cdots, j_{q-1}, j_{q+1}, \cdots, j_M, i_{p+1}, \cdots, i_N}$$
$$= \sum_{k=1}^{K} a_{i_1, \cdots, i_{p-1}, k, i_{p+1}, \cdots, i_N} b_{j_1, \cdots, j_{q-1}, k, j_{q+1} \cdots j_M}, \tag{5}$$

which results in a $(N + M - 2)$-th order tensor $\mathcal{C} \in \mathbb{C}^{I_1 \times \cdots \times I_{p-1} \times J_1 \times \cdots \times J_{q-1} \times J_{q+1} \times \cdots \times J_M \times I_{p+1} \times \cdots \times I_N}$. Note that the contraction operation, unlike the mode-$n$ product, preserves all the modes of both involved tensors, except the common mode.

Contraction involving multiple modes of compatible dimensions is also possible [27]. For instance, a double contraction (along two modes) of $\mathcal{A}$ and $\mathcal{B}$ is denoted as $\mathcal{A} *_{p,k}^{q,l} \mathcal{B}$, where $1 \leq p, k \leq N$ and $1 \leq q, l \leq M$, with $I_p = J_q$ and $I_k = J_l$.

### B. TENSOR DECOMPOSITIONS

Factorization (or decomposition) of a tensor is an important tool for system modeling and solving problems such as parameter estimation, information recovering and missing data estimation for incomplete tensors, i.e. the completion problem. In some applications, tensor decompositions can be viewed as generalizations of matrix decompositions such as the singular value decomposition (SVD) for higher order tensors. In addition to enabling multidimensional data processing, some tensor decompositions are characterized by uniqueness properties that allow undetermined problems to be solved under conditions that are more relaxed than the conventional matrix approaches [16], [18], [28].

Among the most popular tensor decompositions, Tucker [29] and parallel factors analysis (PARAFAC) [30] are commonly used in signal processing applications. Several variants and generalizations such as PARATUCK [31], nested PARAFAC [32], coupled PARAFAC [33], generalized PARATUCK [23], nested Tucker [28] and coupled nested Tucker [24] have been derived in the context of wireless communication systems.

PARAFAC models have the important property of being essentially unique. By essential uniqueness we mean that the decompositions are unique up to arbitrary scaling and permutation of the columns of the factor matrices. On the other hand, Tucker models are not essentially unique, except under certain conditions like a priori knowledge of the core tensor. Despite this, Tucker model is one of the most flexible

**IEEE** *Access*

D. S. Rocha *et al.*: Optical Coupler Network Modeling and Parameter Estimation Based on Generalized Tucker Train Decomposition

tensor decompositions. Below, we recall some decompositions, focusing on Tucker-type ones.

### 1) TUCKER DECOMPOSITION

In 1966, L. Tucker [29] introduced the Tucker decomposition that consists of the factorization of a tensor into a core tensor of the same order that interacts with factor matrices. For a $N$-th order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$, the Tucker decomposition is defined as

$$\mathcal{X} = \mathcal{C} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \cdots \times_N \mathbf{A}^{(N)}, \qquad (6)$$

where $\mathcal{C} \in \mathbb{C}^{R_1 \times \cdots \times R_N}$ is the core tensor and $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R_n}$, with $n = 1, \ldots, N$, are the factor matrices. The scalar form is written as

$$x_{i_1, \cdots, i_N} = \sum_{r_1=1}^{R_1} \cdots \sum_{r_N=1}^{R_N} c_{r_1, \cdots, r_N} \prod_{n=1}^{N} a_{i_n, r_n}^{(n)}. \qquad (7)$$

Note that each factor $\mathbf{A}^{(n)}$ interacts with the core $\mathcal{C}$ by changing the dimension associated with the mode-$n$ product.

Special cases where a $N$-th order tensor has only $N_1$ factor matrices, with $N > N_1$, are called Tucker–$(N_1, N)$ models [25] and correspond to $\mathcal{X} = \mathcal{C} \times_{n=1}^{N_1} \mathbf{A}^{(n)}$. For the case $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$, with the core tensor $\mathcal{C} \in \mathbb{C}^{R_1 \times R_2 \times R_3}$ and the factors $\mathbf{A}^{(1)} \in \mathbb{C}^{I_1 \times R_1}$, $\mathbf{A}^{(2)} \in \mathbb{C}^{I_2 \times R_2}$ and $\mathbf{A}^{(3)} \in \mathbb{C}^{I_3 \times R_3}$, the third-order Tucker decomposition $\mathcal{X} = \mathcal{C} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \times_3 \mathbf{A}^{(3)}$ is illustrated in Fig. 3.

Some authors use the nomenclature Tucker-3 for the third-order Tucker decomposition. When one of the matrix factors of this decomposition is equal to the identity matrix, for example $\mathbf{A}^{(1)} = \mathbf{I}_{I_1}$, we say that we have a Tucker-2 model given by $\mathcal{X} = \mathcal{C} \times_2 \mathbf{A}^{(2)} \times_3 \mathbf{A}^{(3)}$, which is equivalent to a Tucker-$(2, 3)$ decomposition. Similarly, when two of the matrix factors are equal to identity matrices, for example $\mathbf{A}^{(1)} = \mathbf{I}_{I_1}$ and $\mathbf{A}^{(2)} = \mathbf{I}_{I_2}$, we have a Tucker-1, or Tucker-$(1, 3)$ decomposition, and $\mathcal{X}$ becomes $\mathcal{X} = \mathcal{C} \times_3 \mathbf{A}^{(3)}$.

The matrix representation of a decomposition can be deduced from (2). The Tucker decomposition defined in (6) has the following tall mode-$n$ matrix unfolding

$$\begin{aligned} \mathbf{X}_n &= \mathbf{X}_{I_{n+1} \cdots I_N I_1 \cdots I_{n-1} \times I_n} \\ &= (\mathbf{A}^{(n+1)} \otimes \cdots \otimes \mathbf{A}^{(N)} \otimes \mathbf{A}^{(1)} \otimes \cdots \otimes \mathbf{A}^{(n-1)}) \mathbf{C}_n \mathbf{A}^{(n)T}, \end{aligned} \qquad (8)$$

where $\mathbf{C}_n \in \mathbb{C}^{R_{n+1} \cdots R_N R_1 \cdots R_{n-1} \times R_n}$ is the corresponding tall mode-$n$ unfolding of $\mathcal{C}$. The flat mode-$n$ unfolding is obtained by transposing (8).

As previously mentioned, the Tucker model is not essentially unique. That means, the factors can be replaced by $\bar{\mathbf{A}}^{(n)} = \mathbf{A}^{(n)} \Delta_n$ with $\Delta_n \in \mathbb{C}^{R_n \times R_n}$ nonsingular, and the core tensor $\mathcal{C}$ replaced by $\bar{\mathcal{C}} = \mathcal{C} \times_{n=1}^{N} (\Delta_n)^{-1}$, without changing the tensor $\mathcal{X}$. However, uniqueness of this model is satisfied when the core tensor $\mathcal{C}$ is known. For more details, consult [34].
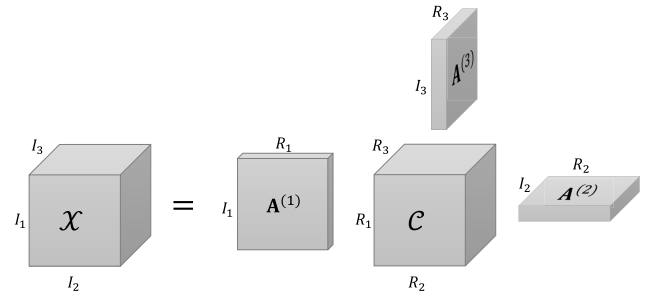


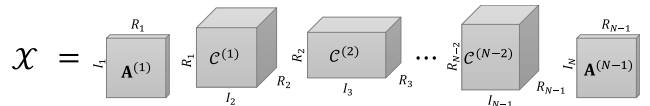**FIGURE 3.** Block-diagram of a Tucker decomposition for a third-order tensor.



**FIGURE 4.** Tensor-train decomposition of an $N$-th order tensor.

Some recent works have extended the Tucker model to represent high order tensors [23], [25], [28], [35]. The so-called nested Tucker decomposition can be viewed as a special case of a tensor-train (TT) decomposition defined in [22]. TT decomposition consists of a concatenation of third-order tensors that form a train of tensors as shown in Fig. 4. Two adjacent wagons (tensor factors) share an auxiliary mode.

The interaction between two neighbor factors relies on the contraction over their common mode. The $N$-th order TT of Fig. 4 is defined in scalar notation as

$$\begin{aligned} x_{i_1, i_2, \ldots, i_n} &= \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \cdots \sum_{r_{N-1}=1}^{R_{N-1}} a_{i_1, r_1}^{(1)} c_{r_1, i_2, r_2}^{(1)} c_{r_2, i_3, r_3}^{(2)} \cdots \\ &\quad \cdot c_{r_{N-2}, i_{N-1}, r_{N-1}}^{(N-2)} a_{i_N, r_{N-1}}^{(N-1)}. \end{aligned} \qquad (9)$$

### 2) GENERALIZED TUCKER DECOMPOSITION

Tucker models, as previously defined, are characterized by matrix factors. A generalized Tucker decomposition corresponds to a Tucker decomposition where some (or all) matrix factors are replaced by tensors, i.e., $\mathbf{A}^{(n)}$ in (6) is replaced by $\mathcal{A}^{(n)}$, resulting in tensor-tensor mode-$n$ products. Similarly to the Tucker model, a special case of the generalized Tucker model was introduced in [23] with some factors equal to identity matrices. For a given $N$-th order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$, the generalized Tucker–$(N_1, N)$ model, with $1 \leq N_1 < N$, is written as

$$\mathcal{X} = \mathcal{C} \times_1 \mathcal{A}^{(1)} \times_2 \mathcal{A}^{(2)} \cdots \times_{N_1} \mathcal{A}^{(N_1)}, \qquad (10)$$

with core tensor $\mathcal{C} \in \mathbb{C}^{R_1 \times \cdots \times R_{N_1} \times I_{N_1+1} \times \cdots \times I_N}$ and tensor factors $\mathcal{A}^{(n)} \in \mathbb{C}^{I_n \times R_n \times I_{\mathbb{S}_n}}$, where $\mathbb{S}_n$ is an ordered subset of the set $\mathbb{S} = \{N_1 + 1, \ldots, N\}$. The tensor factor $\mathcal{A}^{(n)}$ is a $N_n$-th order tensor, with $3 \leq N_n \leq N - N_1 + 2$.

The generalized Tucker decomposition, like the Tucker one, is also not essentially unique when the core tensor is unknown, since its factors are unique up to nonsingular

D. S. Rocha *et al.*: Optical Coupler Network Modeling and Parameter Estimation Based on Generalized Tucker Train Decomposition

**IEEE**Access

transformations. Nevertheless, under the condition of a priori knowledge of the core tensor, the uniqueness can be satisfied. For more details on uniqueness of generalized Tucker decompositions, see [24], [34].

The matrix representation of a generalized Tucker model depends on the number of matrix and tensor factors, as well as on the set $\mathbb{S}_n$. Therefore, it is not possible to define a generic formulation for its matrix unfoldings, as defined in (8) for a Tucker model. However, these unfolding matrices can be derived for some particular cases of generalized Tucker models. Below, we describe a generalized Tucker-(2, 4) model, which will be useful in this paper. Let us take as an example a fourth-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3 \times I_4}$, the core tensor $\mathcal{C} \in \mathbb{C}^{R_1 \times I_2 \times I_3 \times R_4}$ and the tensor factors $\mathcal{A} \in \mathbb{C}^{I_1 \times R_1 \times I_2}$ and $\mathcal{B} \in \mathbb{C}^{I_4 \times R_4 \times I_3}$. A possible generalized Tucker-(2, 4) decomposition of $\mathcal{X}$ is given by

$$\mathcal{X} = \mathcal{C} \times_1 \mathcal{A} \times_4 \mathcal{B} \qquad (11)$$

or in scalar form

$$x_{i_1,i_2,i_3,i_4} = \sum_{r_1=1}^{R_1} \sum_{r_4=1}^{R_4} c_{r_1,i_2,i_3,r_4} a_{i_1,r_1,i_2} b_{i_4,r_4,i_3}. \qquad (12)$$

Let us define two useful unfoldings of this tensor model. The first one gives a matrix of dimension $I_2 I_3 I_4 \times I_1$ obtained by combining the second, third and fourth modes of $\mathcal{X}$ as the rows of the tall mode-1 unfolding given by

$$\mathbf{X}_{I_2 I_3 I_4 \times I_1} = \left[ \mathbf{I}_{I_2} \otimes bdiag\left(\mathbf{B}_{..i_3}\right) \right] \mathbf{C}_{I_2 I_3 R_4 \times I_2 R_1} \mathbf{A}_{I_2 R_1 \times I_1}, \qquad (13)$$

where $\mathbf{A}_{I_2 R_1 \times I_1}$ is a tall mode-1 unfolding of $\mathcal{A}$, $bdiag\left(\mathbf{B}_{..i_3}\right)$ results in a matrix of size $I_3 I_4 \times I_3 R_4$ and $\mathbf{C}_{I_2 I_3 R_4 \times I_2 R_1}$ is obtained from the transpose of the matrix $bdiag\left(\mathbf{C}_{.i_2 1.} \cdots \mathbf{C}_{.i_2 I_3.}\right) \in \mathbb{C}^{I_2 R_1 \times I_2 I_3 R_4}$, with $\mathbf{C}_{.i_2 i_3.} \in \mathbb{C}^{R_1 \times R_4}$. The second useful unfolding corresponds to a vectorization of $\mathcal{X}$ with dimension $I_2 I_1 I_3 I_4$, which is obtained by stacking its fibers as follows

$$\mathbf{x}_{I_2 I_1 I_3 I_4} = \left[ bdiag\left(\mathbf{A}_{..i_2}\right) \otimes bdiag\left(\mathbf{B}_{..i_3}\right) \right] \mathbf{c}_{I_2 R_1 I_3 R_4}, \qquad (14)$$

where $bdiag\left(\mathbf{A}_{..i_2}\right)$ results in a matrix of size $I_2 I_1 \times I_2 R_1$ and $\mathbf{c}_{I_2 R_1 I_3 R_4}$ denotes $vec(\mathcal{C}) \in \mathbb{C}^{I_2 R_1 I_3 R_4}$. For details on the demonstration of these unfoldings, see references [24], [34].

## III. OPTICAL COUPLERS CHARACTERISTICS

Optical couplers, also called directional couplers, are devices that have a wide range of applications in optical systems, due to their ability to provide switching, routing and modulation of optical signals, coherently splitting the pulse incident on one of the input ports, and directing the split signals towards the output ports [1], [9]. Fig. 5 shows a schematic of a dual-core fiber coupler. Such device consists of making the cores of two fibers close enough to allow a partial overlapping of the fundamental propagation modes in the region between the two cores. Thus, the coupling between the propagating fields can lead to the transfer of optical energy from one core
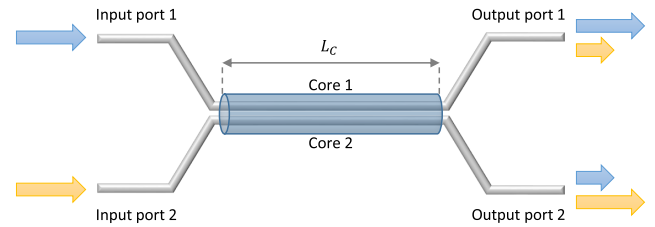


**FIGURE 5.** Schematic of a dual-core optical coupler, representing the division of the energy.

to another one. In this case, the spacing between the cores should be comparable to their diameters [1].

An optical coupler can be seen as a device that plays the role of a beam splitter, where the splitting ratio depends on the coupler length $L_c$ (coupling region between two cores). When the optical power is equally divided between two output ports, these devices are referred to as 50:50 or 3dB couplers.

The propagating pulses in optical fibers are affected by various linear and nonlinear effects, which depend on the system characteristics such as peak power, propagating distances or even the material with which the fiber is made of or doped with. In general, linear effects are negligible for short propagation distances, whereas nonlinear effects become quite significant when higher power levels are applied [1], [9]. Although nonlinearity has been widely exploited in the design of lightwave systems, in this first study on tensor-based optical systems, we have only considered configurations in which these effects are negligible, for the sake of simplicity.

Thus, for a case where low-power continuous wave (CW) beams are launched at the input ports of a dual-core coupler with amplitudes $x_1(t)$ and $x_2(t)$ (the subscript identifies a specific core and $t$ denotes the time), the outputs of the two ports at the coupler end are defined as [1], [36], [37]

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} \sqrt{\rho} & j\sqrt{1-\rho} \\ j\sqrt{1-\rho} & \sqrt{\rho} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}, \qquad (15)$$

where $\rho \in [0, 1]$ denotes the coupling constant of the coupler (splitting ratio). For instance, the 50:50 couplers have a coupling constant of $\rho = 0.5$. One can note that the $2 \times 2$ transfer matrix on the right-side of (15) satisfies the energy conservation and its determinant is equal to 1. Furthermore, a directional coupler introduces a phase shift between the signals arriving at the two output ports, as evidenced by the complex factor in the off-diagonal elements.

Studies and applications for optical systems composed of one or more directional couplers can be found in the literature. Interferometric systems, such as Mach-Zehnder and Sagnac interferometers [3], [5], [9], [11], [36], [37], are common examples of devices with cascade couplers used to implement Boolean functions and/or logic gates. In particular, the works in [38]–[40] provide some applications under a linear regime.

In addition to the interferometric devices, the use of a network of suitably arranged couplers is often employed to implement new optical components in communication systems, such as WDM links [1]. Among them, we can quote the

IEEE Access

D. S. Rocha *et al.*: Optical Coupler Network Modeling and Parameter Estimation Based on Generalized Tucker Train Decomposition

star couplers [6], [41], which have the function of combining the signals coming from the transmitter and forwarding the mixed signals towards multiple receivers. Originally, the role of a star coupler is to divide the input signals equally between its output ports. However, it is also possible to design a suitable combination of couplers with different splitting ratios to obtain specific signal intensities at each output port. This can eventually be used to implement Boolean functions. The simplest case is a $2 \times 2$ star coupler (two-input-two-output ports) and it has the same functionality as a 50:50 directional coupler. Higher-order star couplers can be obtained by combining several couplers. The complexity of such devices increases hugely with the number of ports. In [6], [7], there are some examples of star couplers applied to OCS-based systems.

In some examples above mentioned, the presence of multiple couplers can lead to signals with a multivariate nature. Therefore, multiple influences of the network parameters should be taken into account. In the next section, we introduce a tensor-based modeling of a network of couplers, by exploiting the multidimensional aspects of the involved signals. As an example of coupler network, a $4 \times 4$ star coupler yielding multidimensional signals is taken as a simple scenario to motivate and describe the proposed modeling.

## IV. TENSOR BASED OPTICAL COMMUNICATIONS: LINEAR COUPLER NETWORK MODELING

In this section, we present all the aspects related to the coupler network in question and propose a tensor model to represent the signals at the output ports. We introduce the details of the system parameters and the modeling adopted for the signal transmission step by step.

The main idea is to represent the signals of an optical couplers network, in a linear regime, as a multiway array by considering the multidimensional nature of the signals, i.e., using multiple modes, each one being associated with a dimension of the array.

The following key assumptions are made throughout this manuscript: (i) for all cases, linear dual-core couplers are used; (ii) propagation distances are assumed to be very short; (iii) propagation effects such as attenuation, dispersion and nonlinearity are negligible. These assumptions are made for the sake of simplicity and are reasonable since most couplers are only a few centimeters long. Furthermore, our approach is in accordance with the performance of star couplers in a reduced optical power regime. The modeling of networks under the presence of linear and non-linear effects based on a tensor approach will be addressed in a future work.

### A. PARAMETER DEFINITIONS

Fig. 6(a) illustrates the architecture of an optical network composed of four couplers arranged in a two-by-two array, which corresponds to a $4 \times 4$ star coupler. The signals transmitted through the input ports are combined by the couplers of the first layer of the network and then are directed to the couplers of the second layer. Thus, the transmission scheme is

composed of two "hops", corresponding to the two physical layers of the network. In addition, due to the cross connection between the couplers of the two hops, any input signal is divided among all output ports with intensities that depend on the combination of the coupling constants.

The presented scheme corresponds to the scenario used to define the parameters and describe the signals. For more complex cases ($N \times N$ star couplers with multiple layers of $N/2$ couplers in each layer), the parameters must be correctly designed. The parameters characterizing the network presented in Fig. 6(a) are defined below.

The number of hops is denoted by $I$ and the number of couplers at each hop is $N_i$, with $i \in \{1, \ldots, I\}$. Although we consider the same number of couplers at each hop ($N_1 = \cdots = N_I = N/2$, where $N$ is the total number of network inputs and outputs), the subscript $i$ is used to identify the couplers of a specific hop. Each dual-core coupler $n_i \in \{1, \ldots, N_i\}$ has two input and two output ports. All the couplers have the same number of ports, and we used the indices $p_i \in \{1, 2\}$ and $j_i \in \{1, 2\}$ to identify the input and output ports, respectively. In order to follow a consistent notation, $P_i = J_i = 2$ was defined as the total number of input and output ports of each coupler.

In the first hop ($i = 1$), the $p_1$-th input port of the $n_1$-th coupler is activated by the input signals $x_{p_1, n_1}^{(1)}(t)$, with $t = 1, \ldots, T$, where $T$ is the time dimension, denoting the size of the bit stream. The input signals form the third-order tensor $\mathcal{X}^{(1)} \in \mathbb{C}^{P_1 \times N_1 \times T}$.

Each coupler splits the incident power according to (15). Considering that the $n_i$-th coupler has a coupling constant $\rho_{n_i, i}$, we define the transfer tensor $\mathcal{K}^{(i)} \in \mathbb{C}^{J_i \times P_i \times N_i}$, corresponding to the $i$-th hop, whose mode-3 matrix slices are given by

$$\mathbf{K}_{\cdot \cdot n_i}^{(i)} = \begin{bmatrix} \sqrt{\rho_{n_i, i}} & j\sqrt{1 - \rho_{n_i, i}} \\ j\sqrt{1 - \rho_{n_i, i}} & \sqrt{\rho_{n_i, i}} \end{bmatrix}. \quad (16)$$

Fig. 6(b) illustrates the notation used for representing the input and output ports and the transfer matrix for each coupler.

The interconnection between the two layers of the network is represented by means of a fourth-order binary tensor $\mathcal{W} \in \mathbb{C}^{P_2 \times N_2 \times N_1 \times J_1}$, which assigns each output $(n_1, j_1)$ of the first hop to an input $(p_2, n_2)$ of the second hop. This tensor is composed of 0's and 1's in such a way that the elements equal to 1 determine the pairs $(p_2, n_2) - (n_1, j_1)$ that are connected. That means, for a given pair of couplers $(n_2, n_1)$, the element $w_{p_2, n_2, n_1, j_1}$ is equal to 1 if the ports $j_1$ of $n_1$ and $p_2$ of $n_2$ are connected. Otherwise, $w_{p_2, n_2, n_1, j_1}$ is equal to 0. Furthermore, the value 1 for the connection can be replaced by attenuation factors to represent propagation losses in fiber when large distances are considered. However, as previously mentioned, we assumed negligible propagation effects in this paper.

One can note that, for each value of $(n_1, j_1)$, the binary slice $\mathbf{W}_{\cdot \cdot n_1 j_1} \in \mathbb{C}^{P_2 \times N_2}$ is composed of 0's, except for a single element. Since each port has only one connection, the positions of the elements equal to 1 in each slice must be
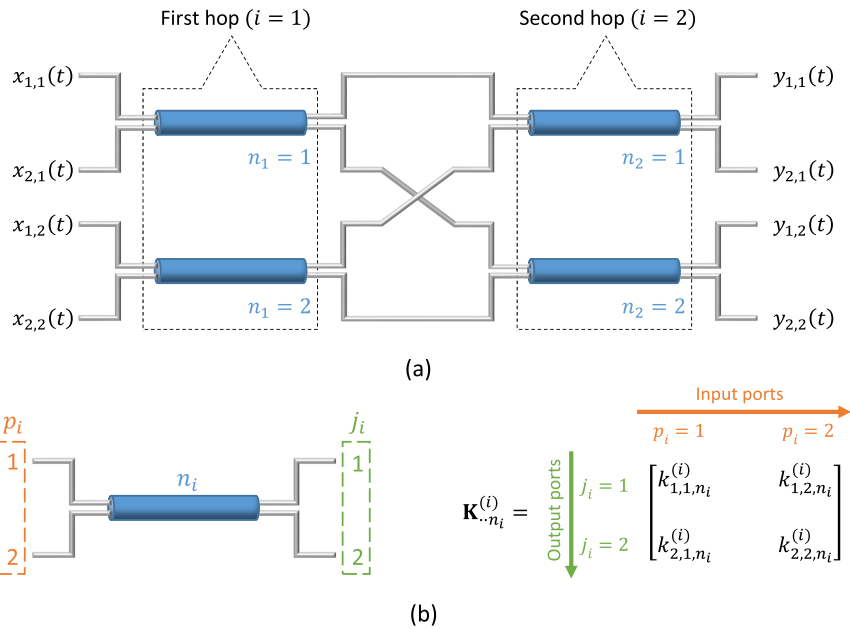
D. S. Rocha *et al.*: Optical Coupler Network Modeling and Parameter Estimation Based on Generalized Tucker Train Decomposition

**IEEE** *Access*



**FIGURE 6.** Optical coupler network: (a) 4 × 4 star coupler architecture, composed of four couplers arranged in a two-by-two array (configuration parameters: $I = N_i = J_i = P_i = 2$ and $i = 1, 2$); (b) transfer matrix dimensions for an individual dual-core coupler.
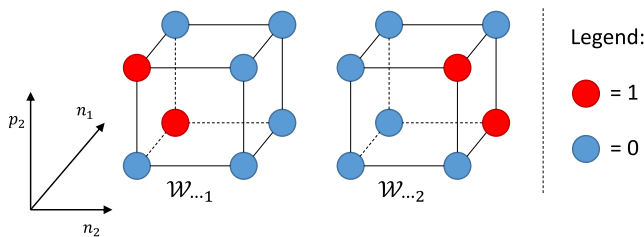


**FIGURE 7.** Didactic representation of the connection tensor $\mathcal{W}$ for the case $P_2 = N_2 = N_1 = J_1 = 2$.

**TABLE 2.** Parameters used in the model.

| Symbols | Description | Values |
|---------|-------------|--------|
| $I$ | Number of hops | – |
| $N$ | Number of input/output ports of the network | – |
| $N_i$ | Number of couplers in the $i$-th hop | $N/2$ |
| $P_i$ | Number of coupler inputs | 2 |
| $J_i$ | Number of coupler outputs | 2 |
| $T$ | Number of bit streams (time-slots) | – |

**TABLE 3.** Dimension of tensors used in the model.

| Description | Symbols | Dimension |
|-------------|---------|-----------|
| Input signals tensor | $\mathcal{X}^{(1)}$ | $P_1 \times N_1 \times T$ |
| Transfer tensor of the $i$-th hop | $\mathcal{K}^{(i)}$ | $J_i \times P_i \times N_i$ |
| Connection tensor | $\mathcal{W}^{(i)}$ | $P_{i+1} \times N_{i+1} \times N_i \times J_i$ |
| Output signals tensor of the $i$-th hop | $\mathcal{Y}^{(i)}$ | $J_i \times N_i \times T$ |

different. For the 4 × 4 star coupler in Fig. 6(a), the topology of the connections yields a binary tensor $\mathcal{W}$ with the elements in the positions $(p_2, n_2, n_1, j_1) = (1, 1, 1, 1), (1, 2, 1, 2),$ $(2, 1, 2, 1)$ and $(2, 2, 2, 2)$ equal to 1. Fig. 7 illustrates this tensor structure in a didactic way. For networks with multiple hops, new tensors should be defined in order to represent the link between two consecutive layers. Thus, analogous to the previously defined tensor $\mathcal{W}$, we can define, in a general way, a connection tensor $\mathcal{W}^{(i)} \in \mathbb{C}^{P_{i+1}, N_{i+1}, N_i, J_i}$ associated to the interconnection between the hops $i$ and $i + 1$.

Finally, the signals coming out of a given hop have the same structure as the input signals $\mathcal{X}^{(1)}$. Thus, the signals at the $j_i$-th port of the $n_i$-th coupler are given by $y_{j_i,n_i}^{(i)}(t)$ that compose the third-order tensor $\mathcal{Y}^{(i)} \in \mathbb{C}^{J_i \times N_i \times T}$. All the parameters associated to the tensors defined in this section and their dimensions are summarized in Tables 2 and 3.

## B. TENSOR MODELING OF THE COUPLER NETWORK
In this subsection, we describe the power transmission throughout the coupler network under consideration,

detailing each step in order to obtain a model of the output signals tensor. Although the tensors in the previous subsection were defined based on the 4 × 4 star coupler as shown in Fig. 6(a), the signal model proposed here is valid for any $N \times N$ two-hop coupler network. Furthermore, later on in this subsection, we provide an extension of the proposed tensor model for an arbitrary value of $I$, which could be applied to networks composed of multiple layers of couplers. Initially we use a scalar notation to define the signals and, in the following subsection, we introduce the tensor notation.

The global transmission can be viewed as a $(2I - 1)$-step process. In the first step, corresponding to the first hop

($i = 1$), $T$ symbols are transmitted by using the $P_1$ ports ($p_1 = 1, \ldots, P_1$) of the $N_1$ couplers ($n_1 = 1, \ldots, N_1$), composing the input signals tensor $\mathcal{X}^{(1)}$. For a given coupler $n_1$, the entries of $\mathcal{X}^{(1)}$ are defined as the time-dependent vector $\mathbf{x}^{(1)}_{\cdot,n_1}(t) = \left[ x^{(1)}_{1,n_1}(t) \cdots x^{(1)}_{P_1,n_1}(t) \right]^T$. These signals are combined by the couplers in the first hop and then, by applying the transfer tensor $\mathcal{K}^{(1)}$, we write the signals at the output of first hop as

$$\mathbf{y}^{(1)}_{\cdot n_1}(t) = \mathbf{K}^{(1)}_{\cdot\cdot n_1} \mathbf{x}^{(1)}_{\cdot n_1}(t) \in \mathbb{C}^{J_1}, \tag{17}$$

which can be rewritten in scalar notation as a summation over $p_1$

$$y^{(1)}_{j_1,n_1}(t) = \sum_{p_1=1}^{P_1} k^{(1)}_{j_1,p_1,n_1} x^{(1)}_{p_1,n_1}(t). \tag{18}$$

In the second step of the transmission, the signals coming out of the first layer are directed to the input ports of the second layer according to the connection tensor $\mathcal{W}^{(1)}$. As the role of this step is to allocate the outputs $j_1$ to the appropriate inputs $p_2$, we get a simple change of indices by applying the connection tensor $\mathcal{W}^{(1)}$. Thus, the signals arriving at the $n_2$-th coupler are given by

$$x^{(2)}_{p_2,n_2}(t) = \sum_{n_1=1}^{N_1} \sum_{j_1=1}^{J_1} w^{(1)}_{p_2,n_2,n_1,j_1} y^{(1)}_{j_1,n_1}(t)$$
$$= \sum_{n_1=1}^{N_1} \sum_{j_1=1}^{J_1} \sum_{p_1=1}^{P_1} w^{(1)}_{p_2,n_2,n_1,j_1} k^{(1)}_{j_1,p_1,n_1} x^{(1)}_{p_1,n_1}(t). \tag{19}$$

Analogous to the first step, we consider the transfer tensor $\mathcal{K}^{(2)}$, which leads to a summation over $p_2$. Then, the tensor $\mathcal{Y}^{(2)}$ of the output signals in the second hop, is defined as follows

$$y^{(2)}_{j_2,n_2}(t) = \sum_{p_2=1}^{P_2} k^{(2)}_{j_2,p_2,n_2} x^{(2)}_{p_2,n_2}(t)$$
$$= \sum_{p_2=1}^{P_2} \sum_{n_1=1}^{N_1} \sum_{j_1=1}^{J_1} \sum_{p_1=1}^{P_1} k^{(2)}_{j_2,p_2,n_2} w^{(1)}_{p_2,n_2,n_1,j_1}$$
$$\cdot k^{(1)}_{j_1,p_1,n_1} x^{(1)}_{p_1,n_1}(t). \tag{20}$$

The above equation shows that the output signals tensor $\mathcal{Y}^{(2)}$ results from the successive transformations of the input signals $\mathcal{X}^{(1)}$ performed by the tensor train $\{\mathcal{K}^{(1)}, \mathcal{W}^{(1)}, \mathcal{K}^{(2)}\}$, which represents the three-step transmission of a two-hop ($I = 2$) network.

Based on (20), let us define a fourth-order tensor $\mathcal{G}^{(2)} \in \mathbb{C}^{J_2 \times N_2 \times N_1 \times P_1}$ given by

$$g^{(2)}_{j_2,n_2,n_1,p_1} = \sum_{p_2=1}^{P_2} \sum_{j_1=1}^{J_1} k^{(2)}_{j_2,p_2,n_2} w^{(1)}_{p_2,n_2,n_1,j_1} k^{(1)}_{j_1,p_1,n_1}, \tag{21}$$

which represents the whole transformation performed by the system over $\mathcal{X}^{(1)}$. Thus, the signals in (20) can be rewritten in the following compact way

$$y^{(2)}_{j_2,n_2}(t) = \sum_{n_1=1}^{N_1} \sum_{p_1=1}^{P_1} g^{(2)}_{j_2,n_2,n_1,p_1} x^{(1)}_{p_1,n_1}(t). \tag{22}$$

Equations (21)-(22) define in a scalar notation the model proposed for the signals of a $N \times N$ two-hop coupler network.

In order to extend this modeling to the multi-hop case ($I > 2$), we can exploit a standard structure for all transmission steps. For instance, to describe a third hop (and so on), a new tensor $\mathcal{W}^{(2)} \in \mathbb{C}^{P_3 \times N_3 \times N_2 \times J_2}$, associated to the connection topology between the couplers of the second and third layers, and a new transfer tensor $\mathcal{K}^{(3)} \in \mathbb{C}^{J_3 \times P_3 \times N_3}$ are used. A straightforward reasoning from (20) leads to the following output signals tensor $\mathcal{Y}^{(3)}$ at the third layer of couplers

$$y^{(3)}_{j_3,n_3}(t) = \sum_{p_3=1}^{P_3} \sum_{n_2=1}^{N_2} \sum_{j_2=1}^{J_2} \sum_{p_2=1}^{P_2} \sum_{n_1=1}^{N_1} \sum_{j_1=1}^{J_1} \sum_{p_1=1}^{P_1} k^{(3)}_{j_3,p_3,n_3}$$
$$\cdot w^{(2)}_{p_3,n_3,n_2,j_2} k^{(2)}_{j_2,p_2,n_2} w^{(1)}_{p_2,n_2,n_1,j_1}$$
$$\cdot k^{(1)}_{j_1,p_1,n_1} x^{(1)}_{p_1,n_1}(t). \tag{23}$$

An extension to the general case with $I$ hops is straightforward since any hop to be added after the first one follows the same structure. In this way, for $i = 1, \ldots, I - 1$, let us define the following tensor $\mathcal{H}^{(i)} \in \mathbb{C}^{J_{i+1} \times N_{i+1} \times N_i \times J_i}$

$$h^{(i)}_{j_{i+1},n_{i+1},n_i,j_i} = \sum_{p_{i+1}=1}^{P_{i+1}} k^{(i+1)}_{j_{i+1},p_{i+1},n_{i+1}} w^{(i)}_{p_{i+1},n_{i+1},n_i,j_i}. \tag{24}$$

Then, the whole transformation performed by the system over $\mathcal{X}^{(1)}$ can be represented as the tensor $\mathcal{G}^{(I)} \in \mathbb{C}^{J_I \times N_I \times N_1 \times P_1}$ defined as

$$g^{(I)}_{j_I,n_I,n_1,p_1} = \sum_{n_{I-1}=1}^{N_{I-1}} \sum_{j_{I-1}=1}^{J_{I-1}} \cdots$$
$$\cdot \sum_{n_2=1}^{N_2} \sum_{j_2=1}^{J_2} \sum_{j_1=1}^{J_1} h^{(I-1)}_{j_I,n_I,n_{I-1},j_{I-1}} \cdots$$
$$\cdot h^{(2)}_{j_3,n_3,n_2,j_2} h^{(1)}_{j_2,n_2,n_1,j_1} k^{(1)}_{j_1,p_1,n_1}. \tag{25}$$

Thus, the signals at the output of the $I$-th hop are given by

$$y^{(I)}_{j_I,n_I}(t) = \sum_{n_1=1}^{N_1} \sum_{p_1=1}^{P_1} g^{(I)}_{j_I,n_I,n_1,p_1} x^{(1)}_{p_1,n_1}(t). \tag{26}$$

In the case $I = 3$, using the definitions (24)–(26), the output signals (23) can be rewritten in the following compact form

$$y^{(3)}_{j_3,n_3}(t) = \sum_{n_1=1}^{N_1} \sum_{p_1=1}^{P_1} g^{(3)}_{j_3,n_3,n_1,p_1} x^{(1)}_{p_1,n_1}(t). \tag{27}$$

Below, we present this model in a tensor notation, which will be used later for developing the parameter estimation algorithm.

D. S. Rocha *et al.*: Optical Coupler Network Modeling and Parameter Estimation Based on Generalized Tucker Train Decomposition

IEEE *Access*

## C. TENSOR NOTATION

The tensors in the previous subsection are defined by operations already discussed in Subsection II-A. In (22), as well as in (26) and (27), the operation between $\mathcal{G}^{(i)}$ and $\mathcal{X}^{(1)}$ is a double-contraction over the common modes $n_1$ and $p_1$. The tensor notation can also be used to define the output signals tensor $\mathcal{Y}^{(I)}$ of a multi-hop network as follows

$$\mathcal{Y}^{(I)} = \mathcal{G}^{(I)} *_{4,3}^{1,2} \mathcal{X}^{(1)} \in \mathbb{C}^{J_I \times N_I \times T}, \qquad (28)$$

where $*_{4,3}^{1,2}$ denotes a double-contraction carried out over the fourth mode of $\mathcal{G}^{(I)}$ and first mode of $\mathcal{X}^{(1)}$, and the third mode of $\mathcal{G}^{(I)}$ and second mode of $\mathcal{X}^{(1)}$.

Thus, the tensor $\mathcal{G}^{(I)}$ accumulates all the steps of the transmission scheme. The interaction between $\mathcal{W}^{(i)}$ and $\mathcal{K}^{(i+1)}$ can be interpreted as a mode-1 product using (24). Then, $\mathcal{H}^{(i)}$ satisfies a generalized Tucker-(1, 4) decomposition given by

$$\mathcal{H}^{(i)} = \mathcal{W}^{(i)} \times_1 \mathcal{K}^{(i+1)} \in \mathbb{C}^{J_{i+1} \times N_{i+1} \times N_i \times J_i}. \qquad (29)$$

Similarly, (25) shows that the successive hops are denoted by the interactions between $\mathcal{H}^{(i+1)}$ and $\mathcal{H}^{(i)}$ which are described by double contractions over their common modes. Thus, $\mathcal{G}^{(I)}$ can be defined in a tensor notation as

$$\mathcal{G}^{(I)} = \mathcal{H}^{(I-1)} *_{4,3}^{1,2} \mathcal{H}^{(I-2)} *_{4,3}^{1,2} \cdots$$
$$*_{4,3}^{1,2} \mathcal{H}^{(2)} *_{4,3}^{1,2} \mathcal{H}^{(1)} \times_4 \mathcal{K}^{(1)'}. \qquad (30)$$

The tensor $\mathcal{K}^{(1)'} \in \mathbb{C}^{P_1 \times J_1 \times N_1}$ is defined from $\mathcal{K}^{(1)} \in \mathbb{C}^{J_1 \times P_1 \times N_1}$ by transposing its mode-3 slices, i.e. $\mathbf{K}^{(1)'}_{..n_1} = \mathbf{K}^{(1)}_{..n_1}{}^T$, with $n_1 = 1, \ldots, N_1$. The above definition is made to satisfy the convention of a mode-$n$ product, in which the second mode of a factor is equal to the $n$-th mode of the core.

The generic tensor $\mathcal{G}^{(I)}$, defined in (30), can be viewed as a generalization of a TT decomposition, as defined in Subsection II-B, whose wagons $\mathcal{H}^{(i)}$ satisfy generalized Tucker decompositions, in addition to sharing two common modes between two consecutive wagons. Fig. 8 shows a block-diagram of the tensor model of the output signals $\mathcal{Y}^{(I)}$ for a multi-hop coupler network, which can be viewed as a TT structure similar to the one shown in Fig. 4. Since there is no representation for arrays of order higher than three, the fourth-order tensor $\mathcal{W}^{(i)}$ is roughly represented as a bigger block. The structure of $\mathcal{W}^{(i)}$ has already been presented in Fig. 7.

The present study aims to propose a method to estimate the parameters of the model such as the tensors $\mathcal{K}^{(i)}$ and $\mathcal{W}^{(i)}$, which are strongly linked to the network design. Estimation of such parameters can be a very useful tool for some applications. Depending on the kind of device (switcher, router, star coupler, logic gate, etc.) to be implemented, the unknown parameters must be correctly designed using the input and output signals. In the next section, the proposed tensor model is exploited to develop an algorithm to estimate the design parameters.

## V. PARAMETER ESTIMATION ALGORITHM

In the previous section, a TT decomposition was highlighted to model the output signals of an $N \times N$ coupler network composed of multiple layers of optical couplers. By exploiting the proposed model, an iterative algorithm can be developed to estimate the tensor factors that contain the design parameters of the network using the tensors $\mathcal{X}^{(1)}$ and $\mathcal{Y}^{(I)}$ of the input and output signals. Although we have developed a tensor model for a generic $I$-hop coupler network, in this section, we will consider a two-hop coupler network, as illustrated in Fig. 6. Thus, for $I = 2$, (29) and (30) become

$$\mathcal{G}^{(2)} = \mathcal{H}^{(1)} \times_4 \mathcal{K}^{(1)'}$$
$$= \mathcal{W}^{(1)} \times_1 \mathcal{K}^{(2)} \times_4 \mathcal{K}^{(1)'} \in \mathbb{C}^{J_2 \times N_2 \times N_1 \times P_1}, \qquad (31)$$

where the connection tensor $\mathcal{W}^{(1)} \in \mathbb{C}^{P_2 \times N_2 \times N_1 \times J_1}$ and the transfer tensors $\mathcal{K}^{(2)} \in \mathbb{C}^{J_2 \times P_2 \times N_2}$ and $\mathcal{K}^{(1)} \in \mathbb{C}^{J_1 \times P_1 \times N_1}$ are to be estimated. The tensor (31) satisfies a structured generalized Tucker-(2, 4) decomposition, as introduced in (11), with the following correspondences

$$(\mathcal{X}, \mathcal{C}, \mathcal{A}, \mathcal{B}) \longleftrightarrow \left( \mathcal{G}^{(2)}, \mathcal{W}^{(1)}, \mathcal{K}^{(2)}, \mathcal{K}^{(1)'} \right)$$
$$(I_1, I_2, I_3, I_4, R_1, R_4) \longleftrightarrow (J_2, N_2, N_1, P_1, P_2, J_1). \qquad (32)$$

The decomposition (31) can be viewed as a special case of (11), due to the constraints on the structure of the core tensor $\mathcal{W}^{(1)}$ and the factors $\mathcal{K}^{(2)}$ and $\mathcal{K}^{(1)}$, as discussed in Subsection IV-A. These constraints play an important role for the parameter estimation and they will be taken into account in the estimation algorithm. For the sake of simplicity, we will denote $(\mathcal{W}^{(1)}, \mathcal{G}^{(2)}, \mathcal{X}^{(1)}, \mathcal{Y}^{(2)}) = (\mathcal{W}, \mathcal{G}, \mathcal{X}, \mathcal{Y})$.

### A. ALS ESTIMATION

The proposed estimation algorithm has two stages. The first one consists of estimating the tensor $\mathcal{G}$ by using an LS estimator derived from the input-output relationship (28). The second stage estimates the parameters $\mathcal{W}$, $\mathcal{K}^{(2)}$ and $\mathcal{K}^{(1)}$ from the estimated tensor $\mathcal{G}$ by using a three-step alternating leas-squares (tri-ALS) method, derived from the generalized Tucker model (31). The ALS estimation method [16], [30] is based on an alternating minimization of LS cost functions, in an iterative way. That means, starting from the initial values, the estimates of unknown parameters are refined at each iteration, until a convergence criterion is reached. The steps of the proposed ALS algorithm are defined below.

A contracted form of the double contraction in (28) can be obtained by combining the first two modes of $\mathcal{X}$ and the last two modes of $\mathcal{G}$, yielding the following mode-3 product

$$\mathcal{Y} = \mathcal{G}_{J_2 \times N_2 \times P_1 N_1} \times_3 \mathbf{X}_{T \times P_1 N_1}. \qquad (33)$$

Note that this rewriting of $\mathcal{Y}$ highlights a Tucker-(1, 3) decomposition. The following matrix unfolding can be deduced from (8)

$$\mathbf{Y}_{J_2 N_2 \times T} = \mathbf{G}_{J_2 N_2 \times P_1 N_1} \mathbf{X}_{P_1 N_1 \times T}, \qquad (34)$$

which leads to the LS estimator

$$\widehat{\mathbf{G}}_{J_2 N_2 \times P_1 N_1} = \mathbf{Y}_{J_2 N_2 \times T} \left( \mathbf{X}_{P_1 N_1 \times T} \right)^\dagger. \qquad (35)$$
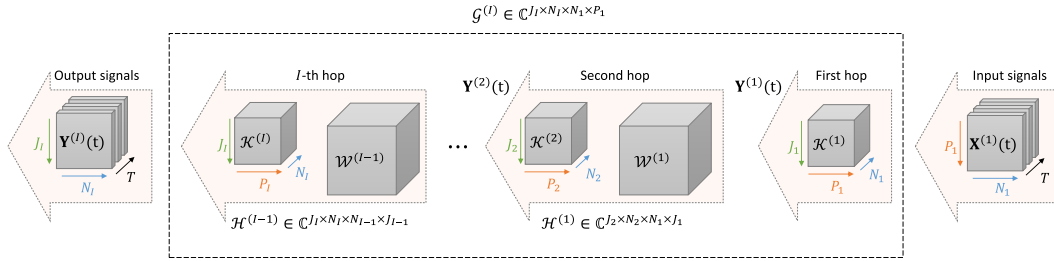
**FIGURE 8.** Block-diagram of the tensor train structure of the signal model. Each wagon $\mathcal{H}^{(i)}$, composed of the tensors $\mathcal{W}^{(i)}$ and $\mathcal{K}^{(i+1)}$, is associated to a hop in the transmission scheme.

After obtaining the estimate $\widehat{\mathcal{G}}$ under the matrix form $\widehat{\mathbf{G}}_{J_2N_2 \times P_1N_1}$, other unfolded versions of $\mathcal{G}$ can be obtained by means of a reshaping operation (which corresponds to other combinations of modes). As $\mathcal{G}$ satisfies a generalized Tucker-(2, 4) decomposition similar to (11), one can exploit the unfoldings (13)–(14) to derive estimators for $\mathcal{W}$, $\mathcal{K}^{(2)}$ and $\mathcal{K}^{(1)}$.

Thus, from the unfolding (14) and the correspondences (32), a vectorized form of $\mathcal{G}$ is given by

$$
\begin{aligned}
\mathbf{g}_{N_2J_2N_1P_1} \\
= \left[ bdiag\left( \mathbf{K}_{\cdot\cdot n_2}^{(2)} \right) \otimes bdiag\left( \mathbf{K}_{\cdot\cdot n_1}^{(1)\ T} \right) \right] \mathbf{w}_{N_2P_2N_1J_1}.
\end{aligned}
\tag{36}
$$

A tall mode-1 unfolding of $\mathcal{G}$ can be derived from (13) as

$$
\begin{aligned}
\mathbf{G}_{N_2N_1P_1 \times J_2} \\
= \left[ \mathbf{I}_{N_2} \otimes bdiag\left( \mathbf{K}_{\cdot\cdot n_1}^{(1)\ T} \right) \right] \mathbf{W}_{N_2N_1J_1 \times N_2P_2} \mathbf{K}_{N_2P_2 \times J_2}^{(2)}.
\end{aligned}
\tag{37}
$$

In a similar way, from a simple permutation of modes, a tall mode-4 unfolding of $\mathcal{G}$ is given by

$$
\begin{aligned}
\mathbf{G}_{N_1N_2J_2 \times P_1} \\
= \left[ \mathbf{I}_{N_1} \otimes bdiag\left( \mathbf{K}_{\cdot\cdot n_2}^{(2)} \right) \right] \mathbf{W}_{N_1N_2P_2 \times N_1J_1} \mathbf{K}_{N_1J_1 \times P_1}^{(1)}.
\end{aligned}
\tag{38}
$$

The unfoldings (36)–(38) are used to define the following LS cost functions to be alternately minimized with respect to $\mathcal{W}$, $\mathcal{K}^{(2)}$ and $\mathcal{K}^{(1)}$, respectively, while the other factors are fixed with their previous estimated values

$$
\begin{aligned}
\underset{\mathcal{W}}{\arg\min} \; \| \widehat{\mathbf{g}}_{N_2J_2N_1P_1} - \Big[ bdiag\big( \widehat{\mathbf{K}}_{\cdot\cdot n_2}^{(2)} \big)_{(it-1)} \\
\otimes bdiag\big( \widehat{\mathbf{K}}_{\cdot\cdot n_1}^{(1)\ T} \big)_{(it-1)} \Big] \mathbf{w}_{N_2P_2N_1J_1} \|_2^2,
\end{aligned}
\tag{39}
$$

$$
\begin{aligned}
\underset{\mathcal{K}^{(2)}}{\arg\min} \; \| \widehat{\mathbf{G}}_{N_2N_1P_1 \times J_2} - \Big[ \mathbf{I}_{N_2} \otimes bdiag\big( \widehat{\mathbf{K}}_{\cdot\cdot n_1}^{(1)\ T} \big)_{(it-1)} \Big] \\
\cdot \widehat{\mathbf{W}}_{N_2N_1J_1 \times N_2P_2 \; (it)} \mathbf{K}_{N_2P_2 \times J_2}^{(2)} \|_F^2,
\end{aligned}
\tag{40}
$$

$$
\begin{aligned}
\underset{\mathcal{K}^{(1)}}{\arg\min} \; \| \widehat{\mathbf{G}}_{N_1N_2J_2 \times P_1} - \Big[ \mathbf{I}_{N_1} \otimes bdiag\big( \widehat{\mathbf{K}}_{\cdot\cdot n_2}^{(2)} \big)_{(it)} \Big] \\
\cdot \widehat{\mathbf{W}}_{N_1N_2P_2 \times N_1J_1 \; (it)} \mathbf{K}_{N_1J_1 \times P_1}^{(1)} \|_F^2,
\end{aligned}
\tag{41}
$$

where *it* denotes the iteration number. The estimates of $\mathcal{W}$, $\mathcal{K}^{(2)}$ and $\mathcal{K}^{(1)}$ are obtained by alternately minimizing the cost

functions (39)-(41). The three-step ALS algorithm is then given by

$$
\begin{aligned}
\widehat{\mathbf{w}}_{N_2P_2N_1J_1 \; (it)} = \Bigg( bdiag\big( \widehat{\mathbf{K}}_{\cdot\cdot n_2}^{(2)} \big)_{(it-1)} \\
\otimes bdiag\big( \widehat{\mathbf{K}}_{\cdot\cdot n_1}^{(1)\ T} \big)_{(it-1)} \Bigg)^{\dagger} \widehat{\mathbf{g}}_{N_2J_2N_1P_1},
\end{aligned}
\tag{42}
$$

$$
\begin{aligned}
\widehat{\mathbf{K}}_{N_2P_2 \times J_2 \; (it)}^{(2)} = \Bigg( \Big[ \mathbf{I}_{N_2} \otimes bdiag\big( \widehat{\mathbf{K}}_{\cdot\cdot n_1}^{(1)\ T} \big)_{(it-1)} \Big] \\
\cdot \widehat{\mathbf{W}}_{N_2N_1J_1 \times N_2P_2 \; (it)} \Bigg)^{\dagger} \widehat{\mathbf{G}}_{N_2N_1P_1 \times J_2},
\end{aligned}
\tag{43}
$$

$$
\begin{aligned}
\widehat{\mathbf{K}}_{N_1J_1 \times P_1 \; (it)}^{(1)} = \Bigg( \Big[ \mathbf{I}_{N_1} \otimes bdiag\big( \widehat{\mathbf{K}}_{\cdot\cdot n_2}^{(2)} \big)_{(it)} \Big] \\
\cdot \widehat{\mathbf{W}}_{N_1N_2P_2 \times N_1J_1 \; (it)} \Bigg)^{\dagger} \widehat{\mathbf{G}}_{N_1N_2J_2 \times P_1}.
\end{aligned}
\tag{44}
$$

In this way, the unknown factors $\mathcal{W}$, $\mathcal{K}^{(2)}$ and $\mathcal{K}^{(1)}$ are estimated under their unfolded forms $\mathbf{w}_{N_2P_2N_1J_1}$, $\mathbf{K}_{N_2P_2 \times J_2}^{(2)}$ and $\mathbf{K}_{N_1J_1 \times P_1}^{(1)}$, respectively. At the end of each iteration, a stop criterion is tested, allowing a refinement of the estimates until reaching convergence. The stop criterion to decide the convergence of the iterative estimation algorithm is based on the reconstruction of the tensor $\mathcal{G}$ using the estimates $\widehat{\mathcal{W}}$, $\widehat{\mathcal{K}}^{(1)}$ and $\widehat{\mathcal{K}}^{(2)}$, and the difference of the Frobenius norm of the estimation error between two successive iterations, i.e.

$$
\left( \| \widehat{\mathcal{G}} - \bar{\mathcal{G}} \|_F^2 \right)_{it} - \left( \| \widehat{\mathcal{G}} - \bar{\mathcal{G}} \|_F^2 \right)_{it-1} \leq \text{threshold},
\tag{45}
$$

where $\widehat{\mathcal{G}}$ and $\bar{\mathcal{G}}$ denote respectively the tensor estimated from $\mathcal{X}$ and $\mathcal{Y}$ using (35) and the reconstructed tensor using (21) with the estimates $\widehat{\mathcal{W}}$, $\widehat{\mathcal{K}}^{(1)}$ and $\widehat{\mathcal{K}}^{(2)}$. The threshold value in (45) is chosen according to the desired refinement.

The proposed algorithm, derived from a generalized Tucker decomposition, allows to estimate the unknown parameters, which can be affected by scaling ambiguities. Indeed, no elements of $\mathcal{K}^{(1)}$ and $\mathcal{K}^{(2)}$, nor the core tensor $\mathcal{W}$, are a priori known. On the other hand, these tensor factors have a strong constraint on their structure as discussed in Subsection IV-A. These constraints are exploited in order to eliminate the ambiguities.

Thus, we propose that, at each iteration, the estimate $\widehat{\mathcal{W}}$ has a structure similar to that of Fig. 7 and therefore its elements are projected onto the binary set $\{0, 1\}$. To eliminate

D. S. Rocha *et al.*: Optical Coupler Network Modeling and Parameter Estimation Based on Generalized Tucker Train Decomposition

**IEEE** *Access*

the ambiguities on the estimates $\widehat{\mathcal{K}}^{(1)}$ and $\widehat{\mathcal{K}}^{(2)}$, the structural conditions of their slices are ensured by keeping the energy conservation in the transfer matrices, i.e. the determinant of each slice $\widehat{\mathbf{K}}^{(i)}_{..n_i}$ must be equal to 1. The operations used to ensure the tensor structures as well as all steps of the proposed ALS-based estimation algorithm are summarized in Table 4.

### B. IDENTIFIABILITY CONDITIONS AND COMPLEXITY ANALYSIS

The system parameter identifiability with the proposed algorithm depends on the uniqueness of the LS solutions. Indeed, for computing the pseudo-inverses in (42)-(44), as well as for the computing $\widehat{\mathcal{G}}$ in (35), some conditions must be satisfied to ensure the uniqueness of the left- or right-inverses. Necessary conditions are directly linked to the rank of the matrices. In (35), the argument of the pseudo-inverse operator has dimensions $P_1N_1 \times T$ and must be right-invertible, i.e. it must be full row rank. In (42)-(44), the arguments have dimensions $N_2J_2N_1P_1 \times N_2P_2N_1J_1$, $N_2N_1P_1 \times N_2P_2$ and $N_1N_2J_2 \times N_1J_1$, respectively, and must be left-invertible, i.e. they must be full column rank. This implies the following necessary (but not sufficient) conditions

$$T \geq P_1N_1, \quad \frac{P_1}{J_1} \geq \frac{P_2}{J_2}, \ N_1 \geq \frac{P_2}{P_1}, \ N_2 \geq \frac{J_1}{J_2}. \quad (46)$$

These conditions lead to some constraints on the values of the network parameters. The last three conditions are directly satisfied for the values of $P_i$ and $J_i$ defined in Table 2 and by considering at least one coupler at each hop. The first condition gives a constraint on the number of time-slots. For example, a $4 \times 4$ star coupler leads to $N_1 = 2$ and therefore a minimum sequence of 4 bits must be considered.

The computational complexity of the proposed ALS algorithm is essentially linked to the computation of pseudo-inverses in (42)–(44). Since the terms within parentheses are block-diagonals, the individual complexities are greatly reduced, leading to non-dominant costs. Thus, we establish the computational complexity of the proposed algorithm from the cost of matrix multiplications, which yields $\mathcal{O}(N_1^2N_2^2P_1P_2J_1J_2) + \mathcal{O}(N_2^2N_1P_1P_2J_2) + \mathcal{O}(N_1^2N_2P_1J_1J_2)$. Assuming $N_1 = N_2 = N$ and $P_1 = P_2 = J_1 = J_2 = P$, it becomes $\mathcal{O}(N^4P^4)$.

## VI. NUMERICAL SIMULATION RESULTS

In this section, we present some numerical experimental results to validate the proposed tensor model for coupler networks, and to evaluate the effectiveness of the proposed estimation algorithm. Signals from nonreturn-to-zero (NRZ) CW beams with an on-off keying (OOK) modulation were generated. The power $P_0 = 10$ mW was used to transmit each bit 1 in the data stream. In the next subsections, several configurations, defined in terms of bit stream size, coupling constant and topology connection, are discussed. Details of each experiment are presented. As previously mentioned, linear and nonlinear propagation effects are ignored.

### A. TRANSMISSION EXPERIMENT FOR A $4 \times 4$ STAR COUPLER

Firstly, we present a transmission experiment for the purpose of model validation. We plotted the output signals of a $4 \times 4$ star coupler (see Fig. 6(a)) using two different methods. In the first one, the signals are modeled with the proposed tensor model in the scalar form (20). In the second one, for comparison, we used an integrated optical circuit simulator (OptiSystem from Optiwave Systems Inc.) to simulate the same star coupler. The simulation setup used in the OptiSystem software is shown in Fig. 9.

Fig. 10(a) shows the bit sequences transmitted from each input port of the star coupler. A size $T = 4$ was used for the streams, sending a single bit 1 from each port during each time slot. The goal is to illustrate the power division at the output ports. As mentioned earlier, a star coupler should split any input power between all the output ports equally.

The signals intensity at each output depends on the combination of the splitting ratio of the couplers. Several configurations were tested. Just for illustration, here are two different configurations. Configuration 1: all the four couplers have the same coupling constant and are 3dB couplers, i.e. $\rho_{1,1} = \rho_{2,1} = \rho_{1,2} = \rho_{2,2} = 0.5$. This configuration results in a balanced distribution (same intensity) of the incoming power to all outputs. Configuration 2: each coupler has a different coupling constant (unbalanced distribution). The values of the coupling constants applied were $\{\rho_{1,1}, \rho_{2,1}, \rho_{1,2}, \rho_{2,2}\} = \{0.2, 0.4, 0.7, 0.9\}$.

Fig. 10(b) and 10(c) show the normalized transmission power $|y_{j_2,n_2}(t)/x_{p_1,n_1}(t)|^2$ for the two configurations. The balanced star coupler (configuration 1) (Fig. 10(b)) shows that the output powers correspond to 25% of the input power, regardless of the port to which the signal was transmitted. With an unbalanced star coupler (configuration 2, Fig. 10(c)), a different power fraction is directed to each output port and depends on where the signal is sent from.

In Fig. 10(b) and 10(c), the output signals simulated using the proposed tensor model were similar to the signals provided by the OptiSystem software, thus validating the proposed model.

### B. PERFORMANCE EVALUATION OF THE ESTIMATION ALGORITHM

The proposed ALS-based estimation algorithm was also evaluated. The tensors $\mathcal{W}$, $\mathcal{K}^{(1)}$ and $\mathcal{K}^{(2)}$ were estimated using the input and output tensors $\mathcal{X}$ and $\mathcal{Y}$. To show the effectiveness and applicability of the proposed algorithm to design optical networks, several configuration schemes were tested by means of Monte Carlo simulations. The performance of the estimations was evaluated using the normalized mean squared error (NMSE) of the estimated parameters, defined as

$$\text{NMSE} = 10 \log_{10} \left( \frac{1}{MC} \sum_{mc=1}^{MC} \frac{\|\mathcal{T}_{mc} - \widehat{\mathcal{T}}_{mc}\|_F^2}{\|\mathcal{T}_{mc}\|_F^2} \right), \quad (47)$$

IEEE *Access*

D. S. Rocha *et al.*: Optical Coupler Network Modeling and Parameter Estimation Based on Generalized Tucker Train Decomposition

**TABLE 4.** ALS-based estimation algorithm.

---

**Stage 1: from the measurements of the input and output tensors $\mathcal{X}$ and $\mathcal{Y}$**

1. Calculate the estimate $\widehat{\mathcal{G}}$ from (35):
$$\widehat{\mathbf{G}}_{J_2 N_2 \times P_1 N_1} = \mathbf{Y}_{J_2 N_2 \times T} \left( \mathbf{X}_{P_1 N_1 \times T} \right)^{\dagger}$$
and then determine the other unfoldings of the tensor $\mathcal{G}$ by means of reshaping operations
$$\widehat{\mathbf{g}}_{N_2 J_2 N_1 P_1}, \widehat{\mathbf{G}}_{N_2 N_1 P_1 \times J_2}, \widehat{\mathbf{G}}_{N_1 N_2 J_2 \times P_1} \leftarrow reshape \left( \widehat{\mathbf{G}}_{J_2 N_2 \times P_1 N_1} \right).$$

**Stage 2: from the estimate $\widehat{\mathcal{G}}$ obtained in the first stage**

2. Randomly initialize $\rho_{n_i, i} \in [0, 1]$ to construct the tensors $\widehat{\mathcal{K}}^{(1)}_{(it=0)}$ and $\widehat{\mathcal{K}}^{(2)}_{(it=0)}$ from (16).

3. Update the iteration number: $it = it + 1$.

4. Calculate the estimate $\widehat{\mathcal{W}}$ from (42):
$$\widehat{\mathbf{w}}_{N_2 P_2 N_1 J_1 \ (it)} = \left( bdiag \left( \widehat{\mathbf{K}}^{(2)}_{\cdot \cdot n_2} \right)_{(it-1)} \otimes bdiag \left( \widehat{\mathbf{K}}^{(1)}_{\cdot \cdot n_1}{}^{T} \right)_{(it-1)} \right)^{\dagger} \widehat{\mathbf{g}}_{N_2 J_2 N_1 P_1}.$$

5. Project the elements of $\widehat{\mathcal{W}}$ onto the set $\{0, 1\}$.

6. Calculate the estimate $\widehat{\mathcal{K}}^{(2)}$ from (43):
$$\widehat{\mathbf{K}}^{(2)}_{N_2 P_2 \times J_2 \ (it)} = \left( \left[ \mathbf{I}_{N_2} \otimes bdiag \left( \widehat{\mathbf{K}}^{(1)}_{\cdot \cdot n_1}{}^{T} \right)_{(it-1)} \right] \widehat{\mathbf{W}}_{N_2 N_1 J_1 \times N_2 P_2 \ (it)} \right)^{\dagger} \widehat{\mathbf{G}}_{N_2 N_1 P_1 \times J_2}.$$

7. Calculate the estimate $\widehat{\mathcal{K}}^{(1)}$ from (44):
$$\widehat{\mathbf{K}}^{(1)}_{N_1 J_1 \times P_1 \ (it)} = \left( \left[ \mathbf{I}_{N_1} \otimes bdiag \left( \widehat{\mathbf{K}}^{(2)}_{\cdot \cdot n_2} \right)_{(it)} \right] \widehat{\mathbf{W}}_{N_1 N_2 P_2 \times N_1 J_1 \ (it)} \right)^{\dagger} \widehat{\mathbf{G}}_{N_1 N_2 J_2 \times P_1}.$$

8. Reconstruct the tensor $\bar{\mathcal{G}}$ using $\widehat{\mathbf{w}}_{N_2 P_2 N_1 J_1 \ (it)}$, $\widehat{\mathbf{K}}^{(2)}_{N_2 P_2 \times J_2 \ (it)}$ and $\widehat{\mathbf{K}}^{(1)}_{N_1 J_1 \times P_1 \ (it)}$ in (21).

9. Return to the Step 3 until the convergence in (45) is reached.

10. Recover the tensor structure – forcing equal diagonal elements and energy conservation:
$$k^{(i)}_{1,1,n_i}, k^{(i)}_{2,2,n_i} \leftarrow \left( k^{(i)}_{1,1,n_i} + k^{(i)}_{2,2,n_i} \right)/2,$$
$$k^{(i)}_{1,2,n_i}, k^{(i)}_{2,1,n_i} \leftarrow \left( k^{(i)}_{1,2,n_i} + k^{(i)}_{2,1,n_i} \right)/2,$$
$$\widehat{\mathbf{K}}^{(i)}_{\cdot \cdot n_i} \leftarrow \left( \det \widehat{\mathbf{K}}^{(i)}_{\cdot \cdot n_i} \right)^{-1/2} \widehat{\mathbf{K}}^{(i)}_{\cdot \cdot n_i}.$$
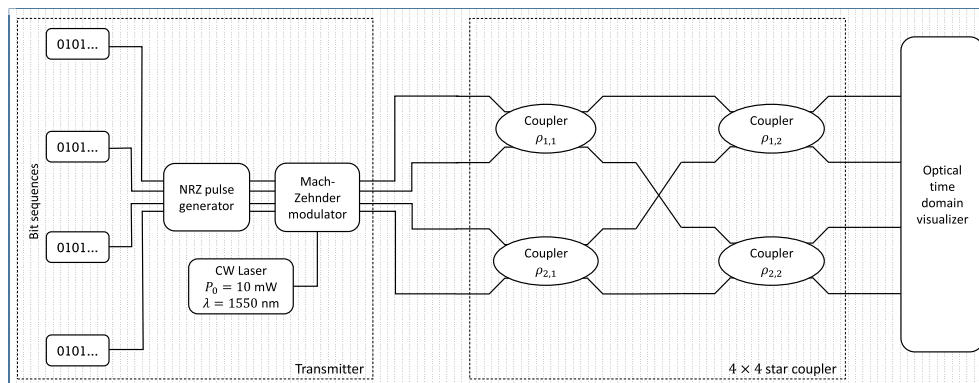
---



**FIGURE 9.** Simulation setup for a 4 × 4 star coupler.

with $MC = 5 \times 10^3$ denoting the number of Monte Carlo runs. $\mathcal{T}_{mc}$ and $\widehat{\mathcal{T}}_{mc}$ represent, respectively, the simulated and estimated tensors at the $mc$-th Monte Carlo run. The speed of convergence was also evaluated. The convergence is confirmed when the criterion defined in (45) is less than or equal to $1.0 \times 10^{-5}$.

First, we applied the ALS-based algorithm to estimate the transfer tensors $\mathcal{K}^{(1)}$ and $\mathcal{K}^{(2)}$ in the two configurations described in Subsection VI-A. The cross connection between the layers of the 4 × 4 star coupler in both cases are the same. Hence, in this first test, the tensor $\mathcal{W}$ was considered known by the algorithm and the Steps 4 and 5 in Table 4 were suppressed. The NMSE results are plotted as a function of $T$ in order to analyze the impact of the size of the transmitted bit sequences on the estimation. Moreover, $T$ must take values greater than or equal to 4 for the 4 × 4 star coupler in order to fulfill the identifiability condition (46).

D. S. Rocha *et al.*: Optical Coupler Network Modeling and Parameter Estimation Based on Generalized Tucker Train Decomposition
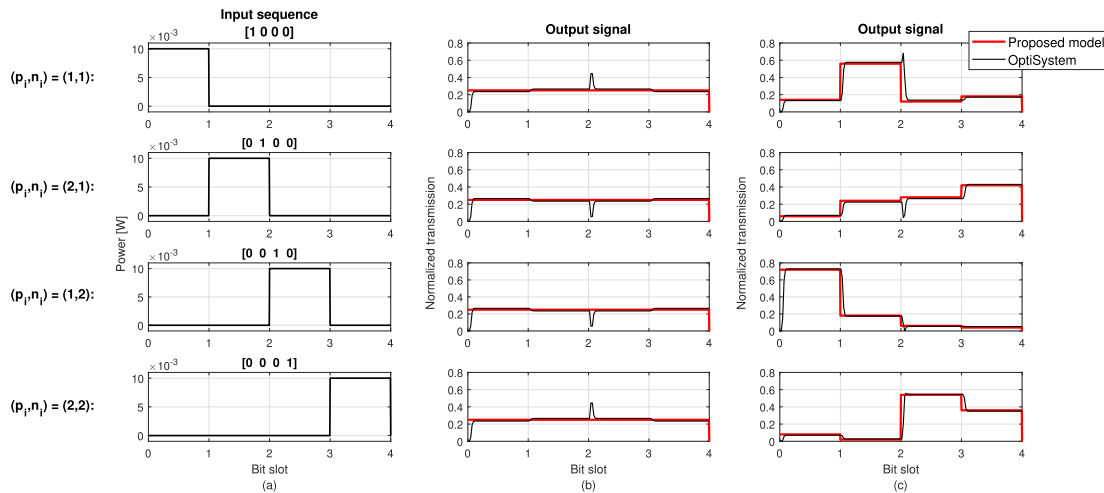
**IEEE** *Access*



**FIGURE 10.** Transmission experiment result. (a) Input bit sequence; (b) output signals for the balanced star coupler—configuration 1; and (c) output signals for an unbalanced star coupler—configuration 2.
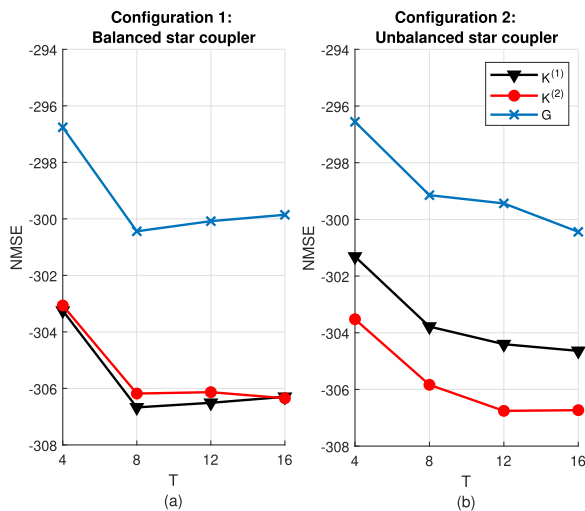


**FIGURE 11.** NMSE of the transfer tensors versus *T* for (a) the balanced and (b) unbalanced 4 × 4 star coupler.



**FIGURE 12.** Number of iterations versus *T* for 4 × 4 star coupler.

Fig. 11 shows the NMSE of $\mathcal{K}^{(1)}$, $\mathcal{K}^{(2)}$ and $\mathcal{G}$ versus $T$, for $T \in \{4, 8, 12, 16\}$. At each Monte Carlo run, new bit sequences are randomly generated. Generally speaking, the NMSE levels obtained were quite low for the configurations used here, indicating a good approximation of the estimated parameters. Furthermore, there is a decrease in NMSE when larger bit stream sizes were used, pointing to a better approximation of the unknown parameters for greater values of $T$. Although the estimation of the tensors $\mathcal{K}^{(1)}$ and $\mathcal{K}^{(2)}$ with (44) and (43) do not depend directly on $T$, the estimation of $\mathcal{G}$ depends on $T$ and, as demonstrated by the blue curve in Fig. 11, this improves when more input-output signals measurements are used in (35).

Fig. 12 shows the number of iterations needed to achieve the convergence versus $T$ averaged along all the Monte Carlo samples. As expected, since $\mathcal{K}^{(1)}$ and $\mathcal{K}^{(2)}$ are independent
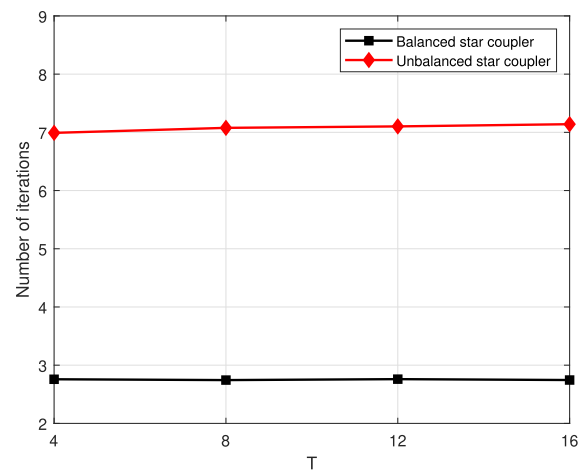
of $T$, the results show that the size of the bit sequence does not impact the convergence speed of the algorithm, i.e. the required number of iterations remains constant with an increase of $T$. This can be seen as an advantage, since better performances are obtained with a higher value of $T$, without degrading the convergence rate.

## C. PERFORMANCE COMPARISON FOR N × N COUPLER NETWORKS

In the following results, we evaluate the performance of the algorithm for more complex cases, by estimating the parameters of larger coupler networks. Fig. 13 illustrates a generic structure for $N \times N$ coupler networks composed of two layers, with $N = 2N_i$, where $N_i \geq 2$ denotes the number of couplers in the $i$-th layer. Configurations with $N_i = 2, 3, 4$ and 5 couplers in each layer are considered, leading to the analysis of $4 \times 4$, $6 \times 6$, $8 \times 8$ and $10 \times 10$ networks. The proposed tensor model, initially described for a $4 \times 4$ star coupler, was
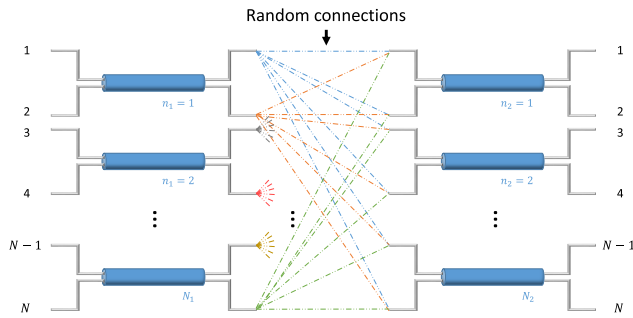
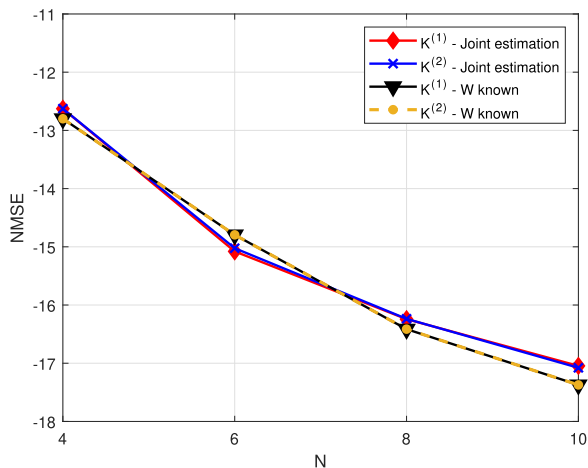**FIGURE 13. Schematic of an *N* × *N* coupler network.**



**FIGURE 14. NMSE of the transfer tensors versus *N* for an *N* × *N* coupler network.**



**FIGURE 15. Error rate of topology connection versus *N* for an *N* × *N* coupler network.**

extended to the coupler networks under consideration. With a greater number of couplers in each layer, the possibilities of connections between the couplers increase. Each output port of a given coupler $n_1$ in the first layer can connect to any input of any coupler $n_2$ in the second layer. Remembering that the output signals depend on the connection between the couplers and it is impossible to obtain star coupler configurations for cases where $N_i > 2$ when only two layers are used.

The connection topology was randomly drawn at each Monte Carlo run of the simulations and the coupling constants $\rho_{n_i,i}$ were also randomly chosen within the range [0, 1], thus allowing several configurations to be analyzed. Since $\mathcal{K}^{(1)}$, $\mathcal{K}^{(2)}$ and $\mathcal{W}$ have random entries, we estimate these three tensors without knowledge of the connection topology ($\mathcal{W}$), unlike the experiments carried out in Subsection VI-B. We used random bit sequences with $T = 16$, in order to satisfy the condition (46) for all the cases.

Fig. 14 shows the NMSE of $\mathcal{K}^{(1)}$ and $\mathcal{K}^{(2)}$ versus $N$ (number of input and output ports). First of all, the NMSE, averaged throughout the random configurations, is noticeably degraded comparatively to the one in Fig. 11, where fixed configurations were used in all Monte Carlo runs. This increase of NMSE can be explained by the fact that local minima can be reached, since some combinations of
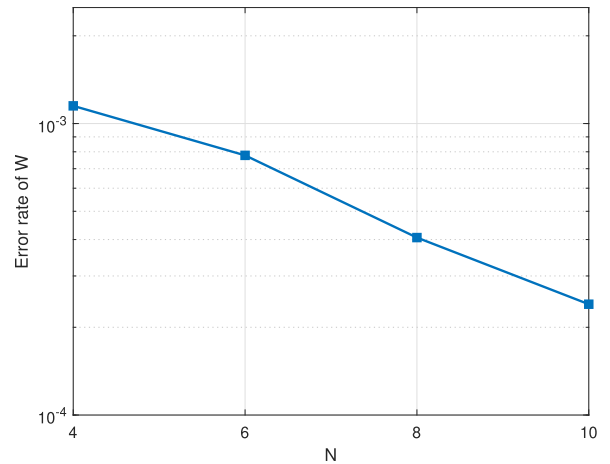
network parameters can lead to similar output signals. However, an improvement in the estimation performance is seen when the number of ports (and couplers) is increased. A larger and more complex network reduces the possibilities of configurations with similar outputs, leading to estimates with greater accuracy.

Still in Fig. 14, the estimation performance for the case where the random tensor $\mathcal{W}$ is known is plotted in comparison with the case of joint estimation of the three tensors. Notably, the lack of knowledge of the topology $\mathcal{W}$ does not impact the performance of the algorithm.

At each MC run, the estimate $\widehat{\mathcal{W}}$ is compared to the original tensor $\mathcal{W}$. Furthermore, each frontal matrix slice of $\mathcal{W}$ has only one entry equal to 1. Thus, connections incorrectly estimated are counted and then the error rate is defined as

$$\text{error rate} = \frac{1}{MC}\sum_{mc=1}^{MC}\frac{\omega_{mc}}{P_2N_2N_1J_1}, \quad (48)$$

where $\omega_{mc}$ denotes the number of incorrect estimations at each MC run, i.e. $\widehat{w}_{p_2,n_2,n_1,j_1} \neq w_{p_2,n_2,n_1,j_1}$, $\forall(p_2, n_2, n_1, j_1)$. Fig. 15 shows the error rate of $\mathcal{W}$ versus $N$. As expected, the results follow the same trend of improvement as observed in Fig. 14 when $N$ is increased, illustrating the effectiveness of the estimation algorithm.

The proposed iterative algorithm was also evaluated in terms of convergence speed. Fig. 16 shows the NMSE of $\mathcal{K}^{(1)}$ and $\mathcal{K}^{(2)}$ versus the number of iterations. The results show that, for all values of $N$, the convergence is achieved in a few iterations. One can also note that with an increase of $N$, the NMSE decreases, corroborating the results shown in Fig. 14. On the other hand, it is possible to notice that the slope of the curve increases more slowly for greater values of $N$, indicating the need for more iterations to achieve convergence in more complex networks. On average, the number of iterations increases when $N$ is increased, with a faster convergence for less complex networks. However, the variation between convergence speeds is small (few iterations) when $N$ increases,
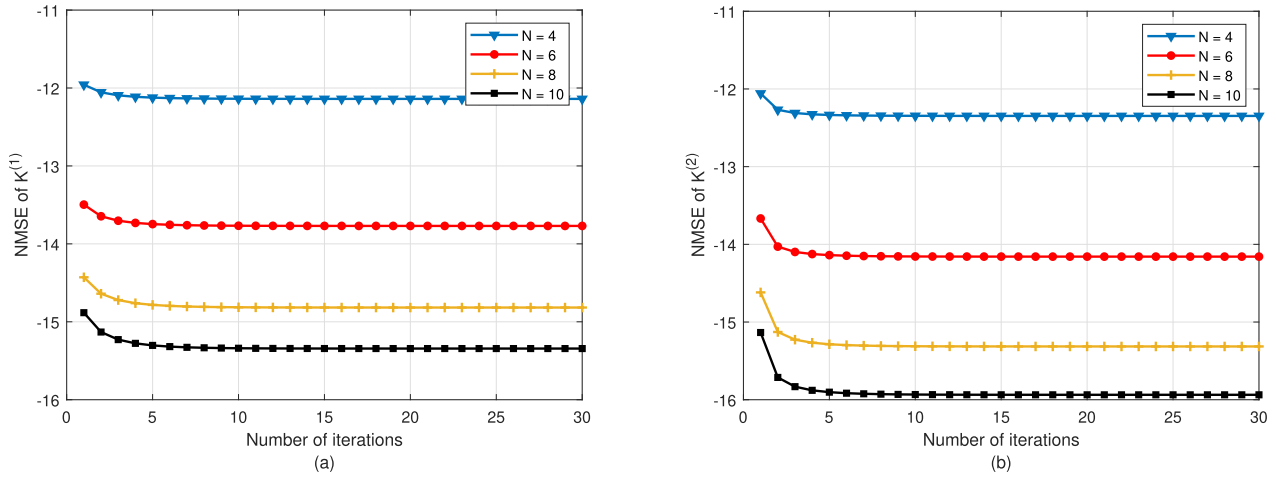
D. S. Rocha *et al.*: Optical Coupler Network Modeling and Parameter Estimation Based on Generalized Tucker Train Decomposition

**IEEE** *Access*

**FIGURE 16. Speed convergence for different configurations with respect to the NMSE of the transfer tensors.**
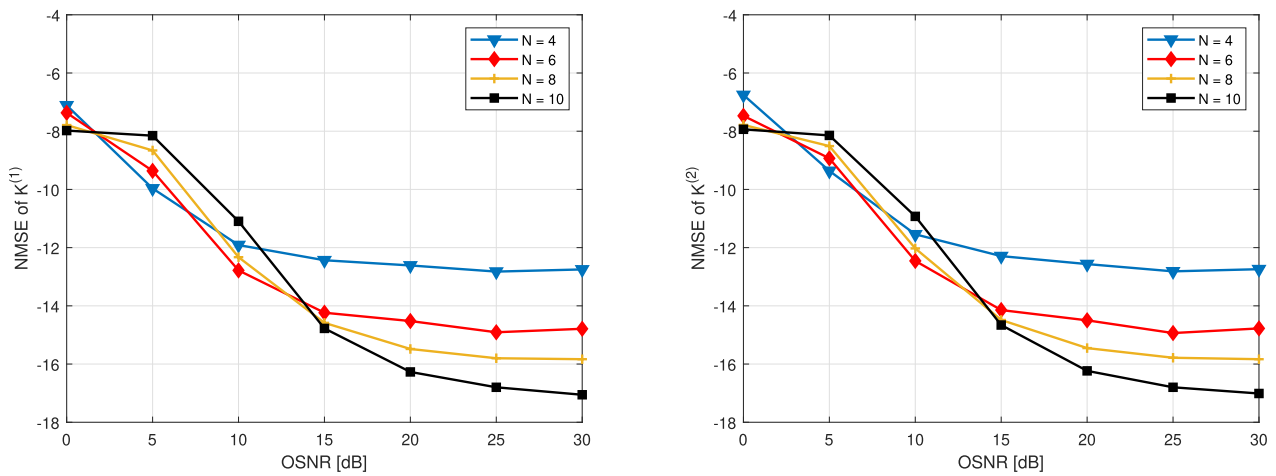


**FIGURE 17. NMSE of the transfer tensors versus OSNR for an $N \times N$ coupler network.**

demonstrating the robustness of the proposed algorithm when considering the complexity of the network.

## D. PERFORMANCE EVALUATION IN THE PRESENCE OF ADDITIVE NOISE

In this subsection, we evaluate the accuracy and convergence of the ALS algorithm in the presence of additive noise. The purpose of this experiment is to analyze the behavior of the estimation algorithm in a noisy environment corresponding to a modification of (28) as

$$\tilde{\mathcal{Y}} = \mathcal{G} *_{4,3}^{1,2} \mathcal{X} + \mathcal{N}, \qquad (49)$$

where $\tilde{\mathcal{Y}}$ denotes the noisy version of $\mathcal{Y}$ and $\mathcal{N} \in \mathbb{C}^{J_2 \times N_2 \times T}$ is a Gaussian noise tensor added at the end node of the network with a noise power of $N_0$. At each run, $N_0$ was fixed according to the desired value of optical signal-to-noise ratio (OSNR), i.e. $N_0 = P_0/\text{OSNR}$, where OSNR levels from 0 dB to 30 dB were used. The results in Fig. 17, 18 and 19 show,

respectively, the NMSE of $\mathcal{K}^{(1)}$ and $\mathcal{K}^{(2)}$, the error rate of $\mathcal{W}$ and the number of iterations versus OSNR for the different networks discussed in the previous subsection.

As expected, Fig. 17 and 18 show a degradation in the accuracy of the estimates of $\mathcal{K}^{(1)}$, $\mathcal{K}^{(2)}$ and $\mathcal{W}$ for low values of OSNR, which corresponds to noisier situations. Fig. 19 indicates that there is a fast convergence for these levels of OSNR. This leads us to conclude that in adverse situations, the convergence criterion adopted is quickly reached due to the existence of local minima as mentioned in Subsection VI-C. In addition, the estimation of the unknown parameters improves significantly when the OSNR level increases, at the cost of an increase in the number of iterations. Higher OSNR values show a trend of improvement in estimates and speed of convergence for all tested cases. Regarding the performance comparison between the different tested networks, the results in Fig. 17, 18 and 19 show a better performance when $N$ is increased, as discussed in the previous subsection.
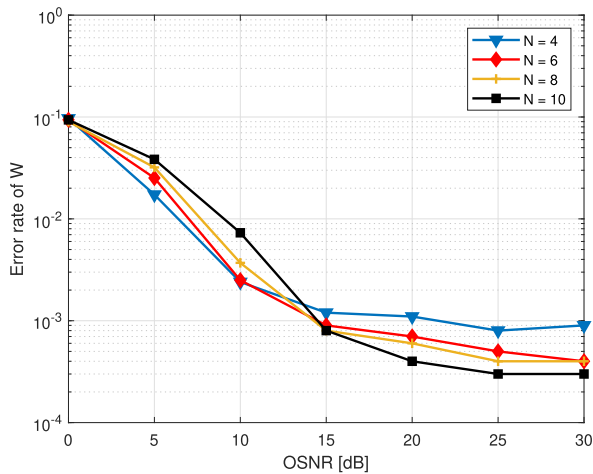
**FIGURE 18.** Error rate of $\mathcal{W}$ versus OSNR for $N \times N$ coupler network.
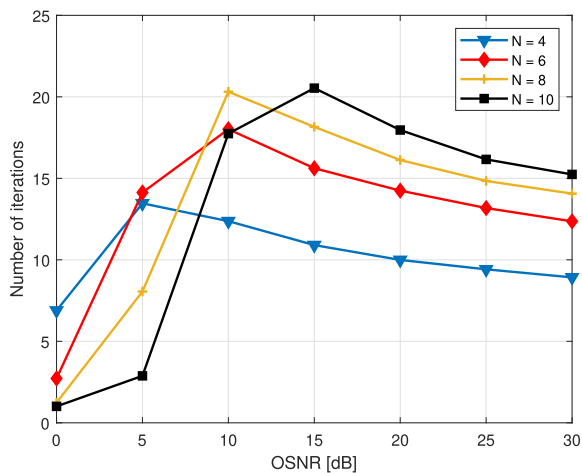


**FIGURE 19.** Number of iterations versus OSNR for $N \times N$ coupler network.

## VII. PERSPECTIVES

The present work is as an original approach for the study and modeling of optical devices. In this section, we briefly point out some perspectives for future works exploiting tensor modeling in the field of optical communications.

Although we proposed a tensor model for a generic coupler network in Section IV, numerical experiments were performed only for two-hop coupler networks. We aim to extend the applications of tensor models to more complex optical networks in order to address higher-order star couplers and/or interferometric devices by associating several couplers in cascade.

The application of tensor methods for modeling multi-core couplers is also a possible future work. With a multi-core coupler, the interactions between adjacent cores can occur in different ways depending on the number of cores and their arrangement.

In signal processing for optical communications such as WDM and DWDM systems, tensor decompositions promise to be an interesting tool to develop receiver algorithms that

yield good performances for channel estimation and transmitted symbols recovery, surpassing the results obtained by conventional methods. As can be seen in the works [17], [18], [24], among others, tensor-based approaches are already widely applied to solve signal processing tasks in the context of wireless communications.

The study of the mentioned devices and networks with nonlinear effects is also considered as a perspective. For this, tensor models must be appropriate to represent nonlinear systems. Tensor models have already been applied to model non-linear systems, such as the references [42]–[44], which exploit Volterra models to model nonlinear channels in communication systems. These works can be considered as a way for future studies.

## VIII. CONCLUSION

In this paper, we have proposed the use of tensor tools as a new method to study and design optical networks. The scenario considered for this first study is based on a $4 \times 4$ star coupler, which can be extended to more complex networks composed of cascaded optical couplers. A tensor model was developed to describe the signals transmitted throughout the network under a linear regime. After modeling the signals for a $4 \times 4$ star coupler, the proposed model was extended to a generic coupler network. The signals were modeled as a third-order tensor that follows a tensor train structure composed of wagons that satisfy a structured generalized Tucker decomposition. With the knowledge of the input and output data, the proposed tensor model of a two-hop coupler network was applied to develop an iterative estimation algorithm to estimate the unknown parameters of the network. A study of the estimation performance in the presence of an additive Gaussian noise was also carried out. Tensor models appear to be a promising approach to study and design optical network structures. As a perspective, real data will be processed for validating our approach.

## REFERENCES

[1] P. A. Govind, *Applications of Nonlinear Fiber Optics*. Amsterdam, The Netherlands: Elsevier, 2006.

[2] J. C. Sales, A. F. G. F. Filho, A. C. Ferreira, J. R. R. Sousa, C. S. Sobrinho, J. W. M. Menezes, G. F. Guimarães, and A. S. B. Sombra, "All-optical XOR and OR by Mach–Zehnder interferometer engineered photonic crystal fibers," *Opt. Laser Technol.*, vol. 94, pp. 128–137, Sep. 2017.

[3] G. S. B. Filho, D. G. Correia, W. B. de Fraga, and G. F. Guimarães, "Obtaining optical logic gates–OR, XOR, AND and logic functions using asymmetric Mach–Zehnder interferometer based on photonic crystal fiber," *Opt. Laser Technol.*, vol. 97, pp. 370–378, Dec. 2017.

[4] S. Kumar, L. Singh, and N.-K. Chen, "Design of all-optical universal gates using plasmonics Mach–Zehnder interferometer for WDM applications," *Plasmonics*, vol. 13, no. 4, pp. 1277–1286, Aug. 2018.

[5] F. T. C. B. Magalhães, H. A. Pinho, D. N. S. Cavalcante, A. C. Ferreira, J. W. M. Menezes, G. F. Guimarães, and W. B. Fraga, "Unbalance of the Sagnac interferometer through nonlinear asymmetry," *J. Electromagn. Waves Appl.*, vol. 30, no. 9, pp. 1227–1239, Jun. 2016.

[6] J. L. Benjamin and G. Zervash, "Parallel star-coupler OCS architectures using distributed hardware schedulers," in *Proc. Photon. Switching Comput. (PSC)*, Sep. 2018, pp. 1–3.

[7] A. C. Funnell, D. Butler, and G. Zervas, "Reconfigurable optical star network architecture for multicast media production data centres," *Opt. Switching Netw.*, vol. 36, Feb. 2020, Art. no. 100556.

D. S. Rocha *et al.*: Optical Coupler Network Modeling and Parameter Estimation Based on Generalized Tucker Train Decomposition

IEEE *Access*

[8] J. R. R. Sousa, A. F. G. F. Filho, A. C. Ferreira, G. S. Batista, C. S. Sobrinho, A. M. Bastos, M. L. Lyra, and A. S. B. Sombra, "Generation of logic gates based on a photonic crystal fiber Michelson interferometer," *Opt. Commun.*, vol. 322, pp. 143–149, Jul. 2014.

[9] A. Araújo, A. Oliveira, F. Martins, A. Coelho, Jr., W. Fraga, and J. Nascimento, "Two all-optical logic gates in a single photonic interferometer," *Opt. Commun.*, vol. 355, pp. 485–491, Nov. 2015.

[10] T. Uthayakumar, R. Vasantha Jayakantha Raja, and K. Porsezian, "Realization of all-optical logic gates through three core photonic crystal fiber," *Opt. Commun.*, vol. 296, pp. 124–131, Jun. 2013.

[11] A. C. Ferreira, A. G. Coêlho, J. R. R. Sousa, C. S. Sobrinho, F. T. C. B. Magalhães, G. F. Guimarães, J. C. Sales, J. W. M. Menezes, and A. S. B. Sombra, "PAM–ASK optical logic gates in an optical fiber Sagnac interferometer," *Opt. Laser Technol.*, vol. 77, pp. 116–125, Mar. 2016.

[12] A. Ghadi and S. Sohrabfar, "All-optical multiple logic gates based on spatial optical soliton interactions," *IEEE Photon. Technol. Lett.*, vol. 30, no. 6, pp. 569–572, Mar. 15, 2018.

[13] W. Liu, D. Yang, G. Shen, H. Tian, and Y. Ji, "Design of ultra compact all-optical XOR, XNOR, NAND and OR gates using photonic crystal multimode interference waveguides," *Opt. Laser Technol.*, vol. 50, pp. 55–64, Sep. 2013.

[14] A. Mohebzadeh-Bahabady and S. Olyaee, "All-optical NOT and XOR logic gates using photonic crystal nano-resonator and based on an interference effect," *IET Optoelectron.*, vol. 12, no. 4, pp. 191–195, 2018.

[15] H. M. E. Hussein, T. A. Ali, and N. H. Rafat, "New designs of a complete set of photonic crystals logic gates," *Opt. Commun.*, vol. 411, pp. 175–181, Mar. 2018.

[16] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.

[17] A. Cichocki, D. Mandic, L. de Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H. A. Phan, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 145–163, Mar. 2014.

[18] A. L. F. de Almeida, G. Favier, J. C. M. Mota, and J. P. C. L. da Costa, "Overview of tensor decompositions with applications to communications," in *Signals and Images: Advances and Results in Speech, Estimation, Compression, Recognition, Filtering, and Processing*. Boca Raton, FL, USA: CRC Press, 2016, ch. 12, pp. 325–356.

[19] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3551–3582, Jan. 2017.

[20] Y. Ji, Q. Wang, X. Li, and J. Liu, "A survey on tensor techniques and applications in machine learning," *IEEE Access*, vol. 7, pp. 162950–162990, 2019.

[21] G. Zhou and A. Cichocki, "Canonical polyadic decomposition based on a single mode blind source separation," *IEEE Signal Process. Lett.*, vol. 19, no. 8, pp. 523–526, Aug. 2012.

[22] I. V. Oseledets, "Tensor-train decomposition," *SIAM J. Sci. Comput.*, vol. 33, no. 5, pp. 2295–2317, Sep. 2011.

[23] G. Favier and A. L. F. de Almeida, "Tensor space-time-frequency coding with semi-blind receivers for MIMO wireless communication systems," *IEEE Trans. Signal Process.*, vol. 62, no. 22, pp. 5987–6002, Nov. 2014.

[24] D. S. Rocha, C. A. R. Fernandes, and G. Favier, "MIMO multi-relay systems with tensor space-time coding based on coupled nested tucker decomposition," *Digit. Signal Process.*, vol. 89, pp. 170–185, Jun. 2019.

[25] G. Favier and A. L. de Almeida, "Overview of constrained PARAFAC models," *EURASIP J. Adv. Signal Process.*, vol. 2014, no. 1, pp. 1–25, Dec. 2014.

[26] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky, "Hyperdeterminants," *Adv. Math.*, vol. 96, no. 2, pp. 226–263, Dec. 1992.

[27] A. Cichocki, "Era of big data processing: A new approach via tensor networks and tensor decompositions," in *Proc. Int. Workshop Smart Info-Media Syst. Asia (SISA)*, Nagoya, Japan, 2013.

[28] G. Favier, C. A. R. Fernandes, and A. L. F. de Almeida, "Nested Tucker tensor decomposition with application to MIMO relay systems using tensor space–time coding (TSTC)," *Signal Process.*, vol. 128, pp. 318–331, Nov. 2016.

[29] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, Sep. 1996.

[30] R. A. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an 'explanatory' multimodal factor analysis," in *Proc. UCLA Work. Papers Phonetics*, vol. 16, 1970, pp. 1–84.

[31] G. Favier, M. N. Costa, 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, 2012.

[32] A. L. F. de Almeida and G. Favier, "Double Khatri–Rao space-time-frequency coding using semi-blind PARAFAC based receiver," *IEEE Signal Process. Lett.*, vol. 20, no. 5, pp. 471–474, May 2013.

[33] M. Sorensen and L. De Lathauwer, "Coupled tensor decompositions for applications in array signal processing," in *Proc. 5th IEEE Int. Workshop Comput. Adv. Multi-Sensor Adapt. Process. (CAMSAP)*, Dec. 2013, pp. 228–231.

[34] D. S. Rocha, "Nested tensor decomposition applied to cooperative MIMO communication systems," Ph.D. dissertation, Dept. Teleinformatics Eng., Federal Univ. Ceará, Fortaleza, Brazil, 2019.

[35] D. S. Rocha, G. Favier, and C. A. R. Fernandes, "Tensor coding for three-hop MIMO relay systems," in *Proc. IEEE Symp. Comput. Commun. (ISCC)*, Jun. 2018, pp. 00384–00389.

[36] A. Kumar, S. Kumar, and S. K. Raghuwanshi, "Implementation of XOR/XNOR and AND logic gates by using Mach–Zehnder interferometers," *Optik*, vol. 125, no. 19, pp. 5764–5767, 2014.

[37] S. K. Raghuwanshi, A. Kumar, and S. Kumar, "1×4 signal router using three Mach–Zehnder interferometers," *Opt. Eng.*, vol. 52, no. 3, pp. 1–10, 2013.

[38] C. Peng, J. Li, H. Liao, Z. Li, C. Sun, J. Chen, and Q. Gong, "Universal linear-optical logic gate with maximal intensity contrast ratios," *ACS Photon.*, vol. 5, no. 3, pp. 1137–1143, Mar. 2018.

[39] R. M. Younis, N. F. F. Areed, and S. S. A. Obayya, "Fully integrated AND and OR optical logic gates," *IEEE Photon. Technol. Lett.*, vol. 26, no. 19, pp. 1900–1903, Oct. 1, 2014.

[40] K. Fasihi, "Design and simulation of linear logic gates in the two-dimensional square-lattice photonic crystals," *Optik*, vol. 127, no. 11, pp. 4669–4674, Jun. 2016.

[41] R. Chipalkatti, Z. Zhang, and A. S. Acampora, "Protocols for optical star-coupler network using WDM: Performance and complexity study," *IEEE J. Sel. Areas Commun.*, vol. 11, no. 4, pp. 579–589, May 1993.

[42] A. Y. Kibangou and G. Favier, "Matrix and tensor decompositions for identification of block-structured nonlinear channels in digital transmission systems," in *Proc. IEEE 9th Workshop Signal Process. Adv. Wireless Commun.*, Jul. 2008, pp. 281–285.

[43] A. Y. Kibangou and G. Favier, "Blind equalization of nonlinear channels using a tensor decomposition with code/space/time diversities," *Signal Process.*, vol. 89, no. 2, pp. 133–143, Feb. 2009.

[44] C. A. Fernandes, G. Favier, and J. C. M. Mota, "Blind identification of multiuser nonlinear channels using tensor decomposition and precoding," *Signal Process.*, vol. 89, no. 12, pp. 2644–2656, Dec. 2009.

**DANILO S. ROCHA** was born in Fortaleza, Brazil, in 1986. He received the degree in physics and the Ph.D. degree in teleinformatics engineering from the Federal University of Ceará (UFC), Brazil, in 2008 and 2019, respectively. From May 2017 to April 2018, he was at the I3S Laboratory, Sophia Antipolis, France, as an exchange (sandwich) Ph.D. student. He is currently a Professor at the Federal Institute of Education, Science and Technology of Ceará (IFCE), where he works as a Researcher at the Photonics Laboratory. His research interests include signal processing for communication systems, including tensor models/decompositions for MIMO communications, cooperative communications, and optical devices and networks.

IEEE *Access*

D. S. Rocha *et al.*: Optical Coupler Network Modeling and Parameter Estimation Based on Generalized Tucker Train Decomposition

**FRANCISCO T. C. B. MAGALHÃES** was born in Fortaleza, Ceará, Brazil, in 1979. He received the bachelor's degree in physics from the Federal University of Ceará, in 2003, and the master's degree in telecommunications engineering from the Federal Institute of Education, Science and Technology of Ceará (IFCE), Fortaleza, in 2016. He is a Researcher at the Photonics Laboratory, IFCE, and a Professor at the IFCE (Sobral Campus). His research interests include optical interferometers, nonlinear optics, fiber optics, and numerical simulations.

**ANTONIO S. B. SOMBRA** received the degree in physics, the degree in physics and chemistry, and the master's degree in physics from the Federal University of Ceará, in 1981, 1983, and 1984, respectively, and the Ph.D. degree in physics from the Federal University of Pernambuco, in 1990. He is currently a Full Professor at the Federal University of Ceará. He has experience in electrical engineering, focusing on dielectric, piezoelectric, and ferroelectric materials, acting on the subjects: optical communications, optical fiber devices, optical fibers, biomaterials, and collagen.

**GÉRARD FAVIER** received the Engineering Diploma degrees from the École Nationale Supérieure de Chronométrie et Micromécanique (ENSCM), Besançon, and the École Nationale Supérieure de l'Aéronautique et de l'Espace (ENSAE), Toulouse, in 1973 and 1974, respectively, and the Dr.-Ing. (Ph.D.) and State Doctorate degrees from the University of Nice Sophia Antipolis, France, in 1977 and 1981, respectively. He is a coauthor or author of more than 300 papers published in international scientific conferences and journals, and 20 books or book chapters. In 1976, he joined the French National Center for Scientific Research (CNRS). He is currently an Emeritus Research Director of CNRS at the I3S Laboratory, Sophia Antipolis, France. From 1995 to 1999, he was the Director of the I3S Laboratory. His research interests include nonlinear system modeling and identification, tensor models/decompositions, and tensor approaches for MIMO communication systems.

**GLENDO F. GUIMARÃES** received the degree in physics from the Federal University of Ceará, in 2004, the M.Sc. degree in physics from the Federal University of Pernambuco, in 2006, and the Ph.D. degree in teleinformatics engineering from the Federal University of Ceará, in 2011. He is currently a Professor at the Federal Institute of Education, Science and Technology of Ceará, where he is also a Professor of the Graduate Program in Telecommunications Engineering. He is the Leader of the Photonics Laboratory of the Federal Institute of Education, Science and Technology of Ceará. He has experience in nonlinear optics, optical sensing, optical communications, and photonic devices. He also has research interests in various analytical and numerical problems related to nonlinear optics.

● ● ●