Semi-Blind Tensor-Based Channel Estimation For Double RIS MIMO Systems

Kabiru N. Aliyu*, Thanh Trung Le†‡, Karim Abed-Meraim†, Azzedine Zerguine*
*EE Dept. and the Center for Communication Systems and Sensing, KFUPM, Dhahran, 31261, Saudi Arabia,
emails: {g201705110, azzedine}@kfupm.edu.sa
†PRISME lab., Univ. Orléans, France, email: karim.abed-meraim@univ-orleans.fr
‡VNU University of Engineering and Technology, Hanoi, Vietnam, email: thanhletrung@vnu.edu.vn

***Abstract*—In this work, we propose a new joint (coupled) Tucker-2 decomposition method for semi-blind channel estimation in a double-reconfigurable intelligent surface (D-RIS)-assisted MIMO communications system. In the D-RIS-assisted system, two RISs are considered; where the first RIS is placed close to the transmitter, and the other is placed close to the receiver for optimal performance. We demonstrate that the received signals in flat fading D-RIS-assisted MIMO systems can be effectively constructed using a 3-way Tucker-2 tensor model. By utilizing the Tucker factorization on the Tucker-2 received signals model, all the three channels i.e., $\mathbf{H_T}$ channel between Tx-to-RIS 1, the $\mathbf{H_S}$ channel between RIS 1-to-RIS 2, and the $\mathbf{H_R}$ channel between RIS 2-to-Rx can be estimated effectively. Furthermore, compared with a single RIS (S-RIS) and the multiple RIS (M-RIS) case, we demonstrate how the system's performance is influenced by parameters related to the transceiver, training overhead, and the number of elements considered in the RISs. A numerical performance evaluation depicts a fast convergence and underscores the improvement in the spectral efficiency of the D-RIS system compared to S-RIS and M-RIS systems.**

***Index Terms*—D-RIS-assisted MIMO, joint Tucker-2 decomposition, semi-blind channel estimation.**

I. Introduction

The Fifth Generation (5G) and beyond wireless communication standards hold the potential to deliver improved wireless broadband and extensive connectivity with extremely low latency through advanced technology such as massive multiple-input multiple-output (MIMO) and millimeter-wave (mm-wave) communication systems. Nevertheless, these advancements come with significant power consumption and path loss challenges in ensuring users a consistent quality of service, particularly in challenging propagation environments [1]. In response to these limitations, potential and cost-effective remedies like developing reconfigurable intelligent surfaces (RISs) have been introduced [2]. The RIS is a 2D surface structure comprising a set of electrically tunable reflecting antenna arrays that make dynamic adjustments to their radiation patterns to improve the efficiency and capacity of the communications system.

The majority of previous research proposed the use of a single-RIS (S-RIS)-assisted system where the transmitter (Tx) and the receiver (Rx) communicate through an S-RIS-aided channel. Additionally, it has been demonstrated that optimal performance gain is achieved when the RIS is positioned in close proximity to the Tx/Rx [3]. On the other hand, some applications, like satellite-to-indoor communications, may require multiple deployments of RIS. Multiple deployments of RIS can be double RIS (D-RIS)-assisted where one RIS is positioned close to Tx and the other close to the Rx (see Fig. 1a) [4], [5] or a situation in which all the RISs are positioned at precise locations of the channel scatterers (see Fig. 1b) named Multi-RIS (M-RIS)-assisted [6]. The performance of the S-RIS, D-RIS, or M-RIS-aided system significantly relies on the number of the array elements on the RIS, the deployed position, and the quality of Channel State Information (CSI) estimation. This becomes challenging, particularly in a D-RIS-aided system where the channel contains three cascaded channels compared to two in a S-RIS-aided system.

A plethora of recent references, such as [5], [7]–[9], have focused on alternating least squares (ALS)-based channel estimation for both S-RIS and D-RIS systems. The authors in [7] proposed an ALS-based channel estimation for an S-RIS system, and their algorithm relied on the parallel tensor model of the received signal. Also, the authors in [8] consider ALS-based in S-RIS under a quasi-static block fading channel by controlling the training sequence overhead using a spatial correlation. Furthermore, the authors in [5], [9] proposed an ALS-based for D-RIS to exploit the 4-way tensor of the D-RIS system using a low-complexity Khatri-Rao factorization in [9] and using Tucker-2 tensor decomposition in [5]. Their method achieved a remarkable performance with a minimum overhead sequence. Also, the authors in [6] consider channel estimation and joint beamforming design for both the S-RIS and M-RIS systems, whereby all RISs were positioned in a precise location of channel scatterers see Fig 1b. However, blind and semi-blind channel estimation maintain the system's spectral efficiency [10] and have recently been proposed for RIS-aided systems. In [11], authors consider S-RIS and proposed a blind data recovery with the help of ALS using a constrained factor tensor decomposition. In [12], a generalized Parallel factor and Tucker tensor decomposition are proposed as a two-stage closed-form semi-blind joint channel and symbol estimation using Khatri-Rao and Kronecker factorization, respectively. Also, the authors in [13] proposed a semi-blind channel estimation approach for the S-RIS-aided channel, which helps

enhance the system's data rate.

In this work, we develop a new joint (coupled) Tucker-2 factorization for a semi-blind channel estimation method in the context of a D-RIS-assisted MIMO communication system. We show that the received signals in flat-fading D-RIS-assisted MIMO systems can be formulated as a joint Tucker-2 tensor model, allowing for a joint coupled Tucker-2 decomposition. By employing the coupled Tucker-2 factorization on the Tucker-2 received signals model, effective estimation of all three channels – Tx-to-RIS 1 ($\mathbf{H}_{\mathrm{T}}$), the channel between RIS 1-to-RIS 2 ($\mathbf{H}_{\mathrm{S}}$), and the channel between RIS 2-to-Rx ($\mathbf{H}_{\mathrm{R}}$) can be achieved. Additionally, in comparison to both S-RIS and M-RIS scenarios, we illustrate how various system parameters, such as the number of Tx/Rx antennas, training overhead, and the number of elements in the RISs, impact the system's performance. Furthermore, numerical performance evaluation shows fast convergence and highlights enhancements in spectral efficiency within the D-RIS system compared to both S-RIS and M-RIS configurations.

The structure of this paper is as follows: Section II presents two subsections: the system model and the tensor-based model. The tensor-based model illustrates the Tucker model and briefly revises the Tucker-2 decomposition of the received signal. Section III details our proposed joint (coupled) Tucker-2 model and semi-blind channel estimation algorithm based on coupled Tucker-2 tensor decomposition. The numerical evaluation is presented in Section VI.

Notation: $\mathbf{A}$, $(a, b, ...)$ and $(\mathbf{a}, \mathbf{b}, ...)$ are the matrix, the scalars, and column vectors, respectively. In addition, $\mathbf{A}^{\mathsf{T}}, \mathbf{A}^{*}, \mathbf{A}^{\mathsf{H}}, \mathbf{A}^{+}, \otimes$, and $\odot$ are the transpose, the complex conjugate, the conjugate transpose (Hermitian), the Moore–Penrose inverse, the Kronecker product, and the Khatri-Rao product, respectively. diag$\{\mathbf{a}\}$ creates a diagonal matrix $\mathbf{A}$ with elements of $\mathbf{a}$ positioned in its main diagonal. vec$\{\mathbf{A}\}$ means create a vector by stacking columns of $\mathbf{A}$ on top of each other and unvec$\{\mathbf{a}\}$ is the reversal of the vec operator. $\boldsymbol{\mathcal{A}} \times_n \mathbf{B}$ denotes the product of n-mode tensor $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{I_1 \times I_2 \times ... \times I_N}$ with $\mathbf{B} \in \mathbb{C}^{J \times I_n}$ matrix. The mode-n unfolding matrix of $\boldsymbol{\mathcal{A}}$ is denoted as $[\boldsymbol{\mathcal{A}}]_{(n)}$ and $\cup_n$ represents the concatenation along the nth dimension. Also, the following identities are utilized: vec$\{\mathbf{ABC}\} = (\mathbf{C}^{\mathsf{T}} \otimes \mathbf{A})vec\{\mathbf{B}\}$; vec$\{\mathbf{A}$diag$(\boldsymbol{\lambda})\mathbf{B}\} = (\mathbf{B}^{\mathsf{T}} \odot \mathbf{A})\boldsymbol{\lambda}$; $(\mathbf{AC}) \odot (\mathbf{BD}) = (\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \odot \mathbf{D})$, where all matrices possess compatible dimensions.

II. Problem Formulation

A. System Model

In this work, we consider a MIMO communication systems equipped with a D-RIS between N_t transmit antennas and N_r receive antennas as shown in Fig. 1a. The RIS 1 has M_1 reflecting elements and is placed near the Tx, hereas the RIS 2 has M_2 reflecting elements and is placed near the Rx. It is assumed that L symbol frames were transmitted, such that $L = I \cdot K$, and there are no direct communications between Tx-to-Rx, Tx-to-RIS 2, and RIS 1-to-Rx owing to the high propagation pathloss phenomena.

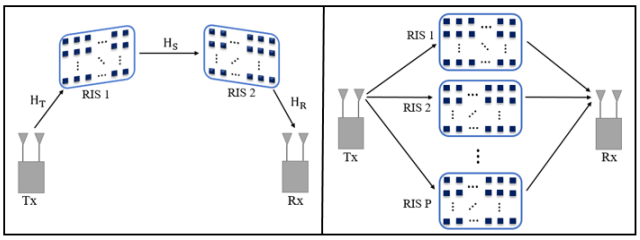

Fig. 1. (a) D-RIS assisted, (b) multi-RIS MIMO communication system.

The received signal $\mathbf{y}_{i,k}$ at the (i,k)-th subframe with $i = 1, 2, ..., I$ and $k = 1, 2, ..., K$ can be expressed as:

$$\mathbf{y}_{i,k} = \mathbf{H}_{\mathrm{R}} \boldsymbol{\Phi}_i \mathbf{H}_{\mathrm{S}} \boldsymbol{\Psi}_i \mathbf{H}_{\mathrm{T}} \mathbf{x}_k + \mathbf{n}_{i,k} \in \mathbb{C}^{N_r \times 1}, \tag{1}$$

where $\mathbf{x}_k \in \mathbb{C}^{N_t \times 1}$ is the transmitted signal; $\mathbf{n}_{i,k} \in \mathbb{C}^{N_r \times 1}$ is the additive white Gaussian circularly symmetric complex-valued noise with zero mean and variance σ_n^2; $\mathbf{H}_{\mathrm{T}} \in \mathbb{C}^{M_1 \times N_t}$ is the channel between Tx to RIS 1; $\mathbf{H}_{\mathrm{R}} \in \mathbb{C}^{N_r \times M_2}$ is the channel between RIS 2 to Rx; and $\mathbf{H}_{\mathrm{S}} \in \mathbb{C}^{M_2 \times M_1}$ is the channel between RIS 1 to RIS 2, respectively. Let $\boldsymbol{\Psi}_i = \mathrm{diag}(\boldsymbol{\psi}_i) \in \mathbb{C}^{M_1 \times M_1}$ with $|[\psi_i]_{m_1}| = 1/\sqrt{M_1}$ and $\boldsymbol{\Phi}_i = \mathrm{diag}(\boldsymbol{\phi}_i) \in \mathbb{C}^{M_2 \times M_2}$ with $|[\phi_i]_{m_2}| = 1/\sqrt{M_2}$ be the i-th diagonal reflection matrix of RIS 1 and RIS 2, respectively; where $\boldsymbol{\psi}_i \in \mathbb{C}^{M_1 \times 1}$ and $\boldsymbol{\phi}_i \in \mathbb{C}^{M_2 \times 1}$, respectively. From (1), we stack the vectors $\{\mathbf{y}_{i,k}\}_{k=1}^{K}$ into a matrix $\mathbf{Y}_i = [\mathbf{y}_{i,1}, ..., \mathbf{y}_{i,K}] \in \mathbb{C}^{N_r \times K}$ and expressed as:

$$\mathbf{Y}_i = \mathbf{H}_{\mathrm{R}} \boldsymbol{\Phi}_i \mathbf{H}_{\mathrm{S}} \boldsymbol{\Psi}_i \mathbf{H}_{\mathrm{T}} \mathbf{X} + \mathbf{N}_i, \tag{2}$$

$$\mathbf{Z}_i \triangleq \mathbf{Y}_i \mathbf{X}^{\mathsf{H}} = \mathbf{H}_{\mathrm{R}} \boldsymbol{\Phi}_i \mathbf{H}_{\mathrm{S}} \boldsymbol{\Psi}_i \mathbf{H}_{\mathrm{T}} + \mathbf{N}_i \mathbf{X}^{\mathsf{H}} \in \mathbb{C}^{N_r \times N_t}, \tag{3}$$

where $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2 ..., \mathbf{x}_K] \in \mathbb{C}^{N_t \times K}$ and it is assumed that $\mathbf{X}\mathbf{X}^{\mathsf{H}} = \mathbf{I}_{N_t}$ for $K \geq N_t$ (under the assumption that the pilot matrix $\mathbf{X}$ is orthogonal). Our primary target in the following section is to estimate the channels $\mathbf{H}_{\mathrm{T}}, \mathbf{H}_{\mathrm{S}}$, and $\mathbf{H}_{\mathrm{R}}$.

B. Tensor-based Model

If we stack I matrices $\{\mathbf{Y}_i\}_{i=1}^{I}$ and $\{\mathbf{Z}_i\}$ into two third-order tensors $\boldsymbol{\mathcal{Y}} \in \mathbb{C}^{N_r \times K \times I}$ (i.e., $\boldsymbol{\mathcal{Y}}(:,:,i) = \mathbf{Y}_i \in \mathbb{C}^{N_r \times K}$) and $\boldsymbol{\mathcal{Z}} \in \mathbb{C}^{N_r \times N_t \times I}$ (i.e., $\boldsymbol{\mathcal{Z}}(:,:,i) = \mathbf{Z}_i \in \mathbb{C}^{N_r \times N_t}$), then $\boldsymbol{\mathcal{Y}}$ and $\boldsymbol{\mathcal{Z}}$ admit the following Tucker representations:

$$\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{G}} \times_1 \mathbf{H}_{\mathrm{R}} \times_2 \mathbf{H}_{\mathrm{T}}^{\mathsf{T}}, \tag{4}$$

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{G}} \times_1 \mathbf{H}_{\mathrm{R}} \times_2 (\mathbf{X}^{\mathsf{T}} \mathbf{H}_{\mathrm{T}}^{\mathsf{T}}), \tag{5}$$

where

$$\boldsymbol{\mathcal{G}} = [\boldsymbol{\Phi}_1 \mathbf{H}_{\mathrm{S}} \boldsymbol{\Psi}_1, \cup_3, ..., \cup_3, \boldsymbol{\Phi}_I \mathbf{H}_{\mathrm{S}} \boldsymbol{\Psi}_I] \in \mathbb{C}^{M_2 \times M_1 \times I}. \tag{6}$$

Accordingly, taking the Tucker-2 decomposition of $\boldsymbol{\mathcal{Z}}$ can result in $\mathbf{H}_{\mathrm{S}}$, $\mathbf{H}_{\mathrm{R}}$ and $\mathbf{H}_{\mathrm{T}}$ as presented in [5] which is expressed as follows;

$$\{\hat{\mathbf{H}}_{\mathrm{R}}, \hat{\mathbf{H}}_{\mathrm{T}}, \hat{\mathbf{H}}_{\mathrm{S}}\} = \underset{\mathbf{H}_{\mathrm{R}}, \mathbf{H}_{\mathrm{T}}, \mathbf{H}_{\mathrm{S}}}{\arg\min} \left\| \boldsymbol{\mathcal{Z}} - \boldsymbol{\mathcal{G}} \times_1 \mathbf{H}_{\mathrm{R}} \times_2 \mathbf{H}_{\mathrm{T}}^{\mathsf{T}} \right\|_{\mathrm{F}}^2, \tag{7}$$

The above problem involves joint optimization, which is nonconvex in nature and can be solved using an alternating

minimization method where one variable is solved at a time while the others remain fixed. The individual channel estimate can be represented as:

$$\hat{\mathbf{H}}_{\mathrm{R}} = \arg\min_{\mathbf{H}_{\mathrm{R}}} \left\| [\boldsymbol{\mathcal{Z}}]_{(1)} - \mathbf{H}_{\mathrm{R}} \mathbf{F}_{\mathrm{R}}(\mathbf{H}_{\mathrm{T}}, \mathbf{H}_{\mathrm{S}}) \right\|_{\mathrm{F}}^2, \tag{8}$$

$$\hat{\mathbf{H}}_{\mathrm{T}} = \arg\min_{\mathbf{H}_{\mathrm{T}}} \left\| [\boldsymbol{\mathcal{Z}}]_{(2)} - \mathbf{H}_{\mathrm{T}}^{\mathsf{T}} \mathbf{F}_{\mathrm{T}}(\mathbf{H}_{\mathrm{R}}, \mathbf{H}_{\mathrm{S}}) \right\|_{\mathrm{F}}^2, \tag{9}$$

$$\hat{\mathbf{h}}_{\mathrm{S}} = \arg\min_{\mathbf{h}_{\mathrm{S}}} \left\| \boldsymbol{z}_{(3)} - \mathbf{F}_{\mathrm{S}}(\mathbf{H}_{\mathrm{T}}, \mathbf{H}_{\mathrm{R}}) \mathbf{h}_{\mathrm{S}} \right\|_2^2, \tag{10}$$

where $\mathbf{F}_{\mathrm{R}}(\mathbf{H}_{\mathrm{T}}, \mathbf{H}_{\mathrm{S}}) = [\boldsymbol{\mathcal{G}}]_{(1)} (\mathbf{I}_I \otimes \mathbf{H}_{\mathrm{T}}^{\mathsf{T}})^{\mathsf{T}} \in \mathbb{C}^{M_2 \times IN_t}$; $\mathbf{F}_{\mathrm{T}}(\mathbf{H}_{\mathrm{R}}, \mathbf{H}_{\mathrm{S}}) = [\boldsymbol{\mathcal{G}}]_{(2)} (\mathbf{I}_I \otimes \mathbf{H}_{\mathrm{R}}^{\mathsf{T}})^{\mathsf{T}} \in \mathbb{C}^{M_1 \times IN_r}$; $\mathbf{F}_{\mathrm{S}}(\mathbf{H}_{\mathrm{T}}, \mathbf{H}_{\mathrm{R}}) = \left((\mathbf{H}_{\mathrm{T}}^{\mathsf{T}} \otimes \mathbf{H}_{\mathrm{R}}) \odot (\boldsymbol{\Psi} \odot \boldsymbol{\Phi})^{\mathsf{T}} \right) \in \mathbb{C}^{IN_r N_t \times M_1 M_2}$; $\hat{\mathbf{h}}_{\mathrm{S}} = \mathrm{vec}\{\hat{\mathbf{H}}_{\mathrm{S}}\} \in \mathbb{C}^{M_1 M_2}$; $\boldsymbol{z}_{(3)} = \mathrm{vec}\{[\boldsymbol{\mathcal{Z}}]_{(3)}\} \in \mathbb{C}^{IN_t N_r}$ and $[\boldsymbol{\mathcal{Z}}]_{(n)}, n \in \{1, 2, 3\}$ is the n-mode unfolding of $\boldsymbol{\mathcal{Z}}$ which is written as:

$$[\boldsymbol{\mathcal{Z}}]_{(1)} = \mathbf{H}_{\mathrm{R}} [\boldsymbol{\mathcal{G}}]_{(1)} (\mathbf{I}_I \otimes \mathbf{H}_{\mathrm{T}}^{\mathsf{T}})^{\mathsf{T}} + [\boldsymbol{\mathcal{N}}]_{(1)} \in \mathbb{C}^{N_r \times IN_t}, \tag{11}$$

$$[\boldsymbol{\mathcal{Z}}]_{(2)} = \mathbf{H}_{\mathrm{T}}^{\mathsf{T}} [\boldsymbol{\mathcal{G}}]_{(2)} (\mathbf{I}_I \otimes \mathbf{H}_{\mathrm{R}})^{\mathsf{T}} + [\boldsymbol{\mathcal{N}}]_{(2)} \in \mathbb{C}^{N_t \times IN_r}, \tag{12}$$

$$[\boldsymbol{\mathcal{Z}}]_{(3)} = [\boldsymbol{\mathcal{G}}]_{(3)} (\mathbf{H}_{\mathrm{T}}^{\mathsf{T}} \otimes \mathbf{H}_{\mathrm{R}})^{\mathsf{T}} + [\boldsymbol{\mathcal{N}}]_{(3)} \in \mathbb{C}^{I \times N_t N_r}. \tag{13}$$

However, for a unique solution of (8), (9) and (10) this requires that $\mathbf{F}_{\mathrm{R}}$ and $\mathbf{F}_{\mathrm{T}}$ have a full column-rank such that $IN_t \geq M_2$ and $IN_r \geq M_1$; whereas $\mathbf{F}_{\mathrm{S}}$ has full row-rank implying $IM_1M_2 \geq N_rN_t$. Therefore, authors in [5] concluded that the number of channel coefficients for S-RIS and D-RIS denoted as $\mathcal{H}_{\text{S-RIS}}$ and $\mathcal{H}_{\text{D-RIS}}$ can be written as:

$$\mathcal{H}_{\text{S-RIS}} = N_t M + N_r M \tag{14}$$

$$\mathcal{H}_{\text{D-RIS}} = N_t M_1 + M_1 M_2 + N_r M_2. \tag{15}$$

The work in [5] recasts (4) to the PARAFAC model in [9].

III. Proposed Semi-Blind Channel Estimation

This section presents our new proposed coupled Tucker-2 received signal model and semi-blind channel estimation algorithm based on a new coupled Tucker-2 tensor decomposition. Assume that we have $\mathbf{X} = [\mathbf{X}_p \ \mathbf{X}_d] \in \mathbb{C}^{N_t \times (K_p + K_d)}$ where $\mathbf{X}_p \in \mathbb{C}^{N_t \times K_p}$ is the pilot matrix and $\mathbf{X}_d \in \mathbb{C}^{N_t \times K_d}$ is the unknown matrix. From (5), we obtain

$$\boldsymbol{\mathcal{Y}}_p = \boldsymbol{\mathcal{G}} \times_1 \mathbf{H}_{\mathrm{R}} \times_2 (\mathbf{X}_p^{\mathsf{T}} \mathbf{H}_{\mathrm{T}}^{\mathsf{T}}) + \boldsymbol{\mathcal{N}}_p \in \mathbb{C}^{N_r \times K_p \times I}, \tag{16}$$

$$\boldsymbol{\mathcal{Y}}_d = \boldsymbol{\mathcal{G}} \times_1 \mathbf{H}_{\mathrm{R}} \times_2 (\mathbf{X}_d^{\mathsf{T}} \mathbf{H}_{\mathrm{T}}^{\mathsf{T}}) + \boldsymbol{\mathcal{N}}_d \in \mathbb{C}^{N_r \times K_d \times I}. \tag{17}$$

For short, let denote $\mathbf{U}_{T_p} = \mathbf{X}_p^{\mathsf{T}} \mathbf{H}_{\mathrm{T}}^{\mathsf{T}} \in \mathbb{C}^{K_p \times M_1}$, $\mathbf{P}_d = \mathbf{X}_d^{\mathsf{T}} \mathbf{X}_p$, and $\mathbf{U}_{T_d} = \mathbf{P}_d \mathbf{U}_{T_p} \in \mathbb{C}^{K_d \times M_1}$. Expressions (16) and (17) can be recast into the following joint (coupled) Tucker-2 decomposition of $\boldsymbol{\mathcal{Y}}_p$ and $\boldsymbol{\mathcal{Y}}_d$

$$\begin{cases} \boldsymbol{\mathcal{Y}}_p = \boldsymbol{\mathcal{G}} \times_1 \mathbf{H}_{\mathrm{R}} \times_2 \mathbf{U}_{T_p} + \boldsymbol{\mathcal{N}}_p, \\ \boldsymbol{\mathcal{Y}}_d = \boldsymbol{\mathcal{G}} \times_1 \mathbf{H}_{\mathrm{R}} \times_2 \mathbf{U}_{T_d} + \boldsymbol{\mathcal{N}}_d. \end{cases} \tag{18}$$

Here, $\boldsymbol{\mathcal{G}}$ and $\mathbf{H}_R$ represent the common core tensor and the common loading factor, respectively. Two matrices $\mathbf{U}_{T_p}$ and $\mathbf{U}_{T_d}$ are the individual factors of $\boldsymbol{\mathcal{Y}}_p$ and $\boldsymbol{\mathcal{Y}}_d$, respectively. By performing the joint Tucker-2 decomposition (18), we can obtain the matrices $\mathbf{H}_{\mathrm{R}}$, $\mathbf{U}_{T_p}$, $\mathbf{U}_{T_d}$ and common core $\boldsymbol{\mathcal{G}}$ directly. This can be reformulated as an optimization problem:

$$\begin{aligned} \arg\min_{\boldsymbol{\mathcal{G}}, \mathbf{H}_{\mathrm{R}}, \mathbf{U}_{T_p}, \mathbf{U}_{T_d}} & \left\| \boldsymbol{\mathcal{Y}}_p - \boldsymbol{\mathcal{G}} \times_1 \mathbf{H}_{\mathrm{R}} \times_2 \mathbf{U}_{T_p} \right\|_F^2 \\ & + \left\| \boldsymbol{\mathcal{Y}}_d - \boldsymbol{\mathcal{G}} \times_1 \mathbf{H}_{\mathrm{R}} \times_2 \mathbf{U}_{T_d} \right\|_F^2. \end{aligned} \tag{19}$$

To solve (19), we propose the following iterative procedure

$$\begin{aligned} \mathbf{H}_{\mathrm{R}}^{(i)} = \arg\min_{\mathbf{H}_{\mathrm{R}}} & \left\| [\boldsymbol{\mathcal{Y}}_p]_{(1)} - \mathbf{H}_R [\boldsymbol{\mathcal{G}}]_{(1)}^{(i-1)} (\mathbf{I} \otimes \mathbf{U}_{T_p}^{(i-1)})^{\top} \right\|_F^2 \\ & + \left\| [\boldsymbol{\mathcal{Y}}_d]_{(1)} - \mathbf{H}_R [\boldsymbol{\mathcal{G}}]_{(1)}^{(i-1)} (\mathbf{I}_{M_2} \otimes \mathbf{U}_{T_d}^{(i-1)})^{\top} \right\|_F^2, \end{aligned} \tag{20}$$

$$\mathbf{U}_{T_p}^{(i)} = \arg\min_{\mathbf{U}_{T_p}} \left\| [\boldsymbol{\mathcal{Y}}_p]_{(2)} - \mathbf{U}_{T_p} [\boldsymbol{\mathcal{G}}]_{(2)}^{(i-1)} (\mathbf{I}_K \otimes \mathbf{H}_R^{(i)})^{\top} \right\|_{\mathrm{F}}^2, \tag{21}$$

$$\mathbf{U}_{Td}^{(i)} = \arg\min_{\mathbf{U}_{T_d}} \left\| [\boldsymbol{\mathcal{Y}}_d]_{(2)} - \mathbf{U}_{T_d} [\boldsymbol{\mathcal{G}}]_{(2)}^{(i-1)} (\mathbf{I}_K \otimes \mathbf{H}_R^{(i)})^{\top} \right\|_{\mathrm{F}}^2, \tag{22}$$

$$\begin{aligned} [\boldsymbol{\mathcal{G}}^{(i)}]_{(1)} = \arg\min_{\boldsymbol{\mathcal{G}}} & \left\| [\boldsymbol{\mathcal{Y}}_p]_{(1)} - \mathbf{H}_R^{(i)} [\boldsymbol{\mathcal{G}}]_{(1)} (\mathbf{I} \otimes \mathbf{U}_{T_p}^{(i)})^{\top} \right\|_F^2 \\ & + \left\| [\boldsymbol{\mathcal{Y}}_d]_{(1)} - \mathbf{H}_R^{(i)} [\boldsymbol{\mathcal{G}}]_{(1)} (\mathbf{I}_{M_2} \otimes \mathbf{U}_{T_d}^{(i)})^{\top} \right\|_F^2. \end{aligned} \tag{23}$$

Note that, we obtain the common core $\boldsymbol{\mathcal{G}}^{(i)}$ by reshaping the matrix $[\boldsymbol{\mathcal{G}}^{(i)}]_{(1)}$ at each iteration. Their closed-form solutions are provided in Algorithm 1. Subsequently, upon obtaining $\hat{\mathbf{H}}_R, \hat{\mathbf{U}}_{T_p}, \hat{\mathbf{U}}_{T_d}$ and $\hat{\boldsymbol{\mathcal{G}}}$, we proceed to estimate $\hat{\mathbf{H}}_{\mathrm{T}} = \hat{\mathbf{U}}_{T_p}^{\mathsf{T}} \mathbf{X}_p^{\mathsf{T}}$ and $\hat{\mathbf{H}}_{\mathrm{S}} \approx \boldsymbol{\Phi}^{-1} \hat{\boldsymbol{\mathcal{G}}} \boldsymbol{\Psi}^{-1}$. Ultimately, the cascaded end-to-end channel estimate can be expressed as:

$$\hat{\mathbf{H}}_e = \hat{\mathbf{H}}_{\mathrm{R}} \boldsymbol{\Phi} \hat{\mathbf{H}}_{\mathrm{S}} \boldsymbol{\Psi} \hat{\mathbf{H}}_{\mathrm{T}} \in \mathbb{C}^{N_r \times N_t}. \tag{24}$$

Remarks: The individual channel estimates can be expressed in relation to the actual channels as: $\hat{\mathbf{H}}_{\mathrm{R}} \approx \mathbf{H}_{\mathrm{R}} \boldsymbol{\Delta}_{\mathrm{R}}$, $\hat{\mathbf{H}}_{\mathrm{T}} \approx \boldsymbol{\Delta}_{\mathrm{T}} \mathbf{H}_{\mathrm{T}}$, $\mathbf{H}_{\mathrm{S}} \approx \boldsymbol{\Delta}_{\mathrm{R}}^{-1} \hat{\mathbf{H}}_{\mathrm{S}} \boldsymbol{\Delta}_{\mathrm{T}}^{-1}$, where $\boldsymbol{\Delta}$ is the diagonal matrix containing the scaling ambiguities which normally disappear in the end-to-end cascaded channel $\hat{\mathbf{H}}_e = \hat{\mathbf{H}}_{\mathrm{R}} \hat{\mathbf{H}}_{\mathrm{S}} \hat{\mathbf{H}}_{\mathrm{T}} \in \mathbb{C}^{N_r \times N_t}$ because the reflection matrices $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ are known at the receiver side and the permutation ambiguities disappeared [9]. Additionally, the proposed method presents an approach to estimate the unknown data matrix $\mathbf{X}_d$ from $\mathbf{U}_{T_d}$ and $\mathbf{U}_{T_p}$. Specifically, we first determine the matrix $\mathbf{P}_d = \mathbf{U}_{T_d} \mathbf{U}_{T_d}^{+}$ and then simply estimate $\mathbf{X}_d = (\mathbf{P}_d \mathbf{X}_p^{+})^{\mathsf{T}}$.

IV. Numerical Evaluation

This section presents a performance assessment of the proposed semi-blind coupled Turker-2 approach for a D-RIS-aided compared to the model proposed in [6] based on spectral efficiency which is defined as:

$$\text{SE (bps/Hz)} = \log 2 \left[\det \left(\mathbf{I}_P + \text{SNR} \cdot \hat{\mathbf{H}}_e^{\mathsf{H}} \hat{\mathbf{H}}_e \right) \right], \tag{25}$$

Algorithm 1: Joint (Coupled) Tucker-2 Decomposition

Input: $\mathcal{Y}_p$, $\mathcal{Y}_d$, and a small regularized parameter ρ
Output: $\mathcal{G}, \mathbf{H}_R, \mathbf{U}_{T_p}$, and $\mathbf{U}_{T_d}$
Initialization:
$\mathcal{G}^{(0)}, \mathbf{H}_R^{(0)}, \mathbf{U}_{T_p}^{(0)}$, and $\mathbf{U}_{T_d}^{(0)}$ are generated at random
$\mathbf{Z}_p^{(0)} = (\mathbf{I} \otimes \mathbf{U}_{T_p}^{(0)})^\top$; $\mathbf{Z}_d^{(0)} = (\mathbf{I} \otimes \mathbf{U}_{T_d}^{(0)})^\top$
Main Program:
For $i = 1, 2, \ldots$ **do**
// Estimate the common factor $\mathbf{H}_R$
$\mathbf{W}_p^{(i)} = \left[\mathcal{G}^{(i-1)}\right]_{(1)} \mathbf{Z}_p^{(i-1)}$
$\mathbf{W}_d^{(i)} = \left[\mathcal{G}^{(i-1)}\right]_{(1)} \mathbf{Z}_d^{(i-1)}$
$\mathbf{Y}^{(i)} = \left[\mathcal{Y}_p\right]_{(1)} (\mathbf{W}_p^{(i)})^\top + \left[\mathcal{Y}_d\right]_{(1)} (\mathbf{W}_d^{(i)})^\top$
$\mathbf{H}_R^{(i)} = \mathbf{Y}^{(i)} \left(\mathbf{W}_p^{(i)} (\mathbf{W}_p^{(i)})^\top + \mathbf{W}_d^{(i)} (\mathbf{W}_d^{(i)})^\top + \rho \mathbf{I}_{M_2}\right)^{-1}$
// Estimate two factors $\mathbf{U}_{T_p}$ and $\mathbf{U}_{T_d}$
$\mathbf{Q}^{(i)} = \left[\mathcal{G}^{(i-1)}\right]_{(2)} (\mathbf{I}_K \otimes \mathbf{H}_R^{(i)})^\top$
$\mathbf{U}_{T_p}^{(i)} = \left[\mathcal{Y}_p\right]_{(2)} (\mathbf{Q}^{(i)})^+$
$\mathbf{U}_{T_d}^{(i)} = \left[\mathcal{Y}_d\right]_{(2)} (\mathbf{Q}^{(i)})^+$
// Estimate the common core $\mathcal{G}$
$\mathbf{Z}_p^{(i)} = (\mathbf{I} \otimes \mathbf{U}_{T_p}^{(i)})^\top$
$\mathbf{Z}_d^{(i)} = (\mathbf{I} \otimes \mathbf{U}_{T_d}^{(i)})^\top$
$\mathbf{Z}^{(i)} = (\mathbf{H}_R^{(i)})^+ \left(\left[\mathbf{Y}_p\right]_{(1)} (\mathbf{Z}_p^{(i)})^\top + \left[\mathbf{Y}_d\right]_{(1)} (\mathbf{Z}_d^{(i)})^\top\right)$
$\left[\mathcal{G}^{(i)}\right]_{(1)} = \mathbf{Z}^{(i)} \left(\mathbf{Z}_p^{(i)} (\mathbf{Z}_p^{(i)})^\top + \mathbf{Z}_d^{(i)} (\mathbf{Z}_d^{(i)})^\top + \rho \mathbf{I}\right)^{-1}$
End For

where P is the number of RIS considered in the system; $\hat{\mathbf{H}}_e$ is the cascaded end-to-end channel estimate. The channel matrices $\mathbf{H}_\mathrm{R}, \mathbf{H}_\mathrm{S}$, and $\mathbf{H}_\mathrm{T}$ are generated randomly and assumed to be Gaussian independent and identically distributed (i.i.d.) with zero mean and unit variance. A finite alphabet sequences are considered as transmitted information. The Signal-to-Noise Ratio (SNR) is defined as: $\mathrm{SNR} = \mathbb{E}\{\|\boldsymbol{\mathcal{Z}} - \boldsymbol{\mathcal{N}}\|_\mathrm{F}^2\} / \mathbb{E}\{\|\boldsymbol{\mathcal{N}}\|_F^2\}$. The reflection matrices $\boldsymbol{\Psi} \in \mathbb{C}^{I_1 \times M_1}$ and $\boldsymbol{\Phi} \in \mathbb{C}^{I_2 \times M_2}$ are updated by random selection of the I_1 and I_2 rows of the normalized M_1 and M_2-DFT matrices, such that $I_1 \leq M_1$ and $I_2 \leq M_2$ as in [5].

Figures 2 and 3 illustrate the performance evaluation of the proposed joint Tucker-2 decomposition of $\mathcal{Y}_p$ and $\mathcal{Y}_d$ using the following metric:

$$\mathrm{Error}(\boldsymbol{\mathcal{Z}}_{gt}, \boldsymbol{\mathcal{Z}}_{es}) = \|\boldsymbol{\mathcal{Z}}_{gt} - \boldsymbol{\mathcal{Z}}_{es}\|_\mathrm{F} / \|\boldsymbol{\mathcal{Z}}_{gt}\|_F, \qquad (26)$$

where $\boldsymbol{\mathcal{Z}}_{gt}$ and $\boldsymbol{\mathcal{Z}}_{es}$ are ground truth and estimate, respectively. The convergence of the algorithm is decided based on the threshold δ (i.e., $\mathrm{Error}(i) - \mathrm{Error}(i-1) \leq \delta$). The setup considered in this experiment consists of $I = 40$ training overhead; $N_r = 4, N_t = 2, M_1 = 30, M_2 = 10$ and $\delta = 10^{-2}$. The legends of Fig. 2 and Fig. 3 depict the common core $\mathcal{G}$; individual factor of $\mathcal{Y}_p$ and individual factor of $\mathcal{Y}_d$. As can be observed from Fig. 2, the proposed algorithm performs very well and effectively in the presence of noise. Moreover, in terms of convergence, as depicted in Fig. 3, the proposed algorithm attains a steady state in almost two iterations.

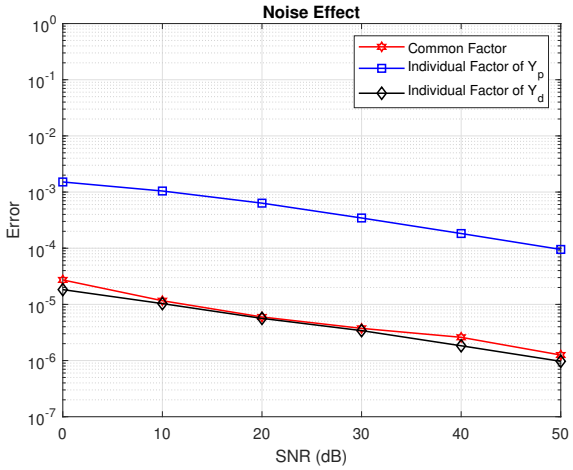

Fig. 2. Error versus SNR.

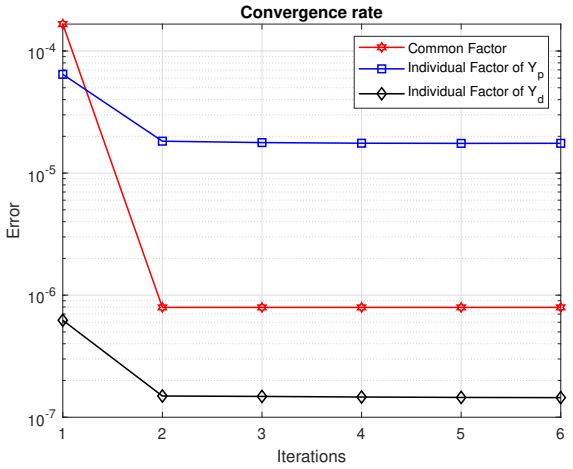

Fig. 3. Error versus number of iterations.

Figure 4 shows the results of the proposed joint Tucker-2 approach for D-RIS in terms of the normalized mean square error (NMSE) of the effective channels defined as $\mathrm{NMSE} = \mathbb{E}\{\|\mathbf{H}_e - \hat{\mathbf{H}}_e\|_\mathrm{F}^2\} / \mathbb{E}\{\|\mathbf{H}_e\|_F^2\}$. In this experiment, the setup consists of $I = 20$ training overhead, $N_r = 4, N_t = 2$, and different elements distributions for M_1 and M_2 in the RIS 1 and RIS 2 in D-RIS as indicated in the legend of the Fig. 4. We can observe a significant NMSE performance improvement when $M_1 = M_2 = 20$ compared to when $M_1 = 10, M_2 = 30$. This indicates that excellent performance can be achieved with careful distribution of elements between RIS 1 and RIS 2 of the D-RIS system.

As far as the spectral efficiency is concerned, Fig. 5 presents the performance of the proposed joint Tucker-2 semi-blind approach for D-RIS compared to the work presented in [6] for S-RIS and M-RIS systems, whereby all RISs were positioned

in same scenario as reported in [6]. In this experiment, $I = 20$ training overhead is used for all cases. Furthermore, the

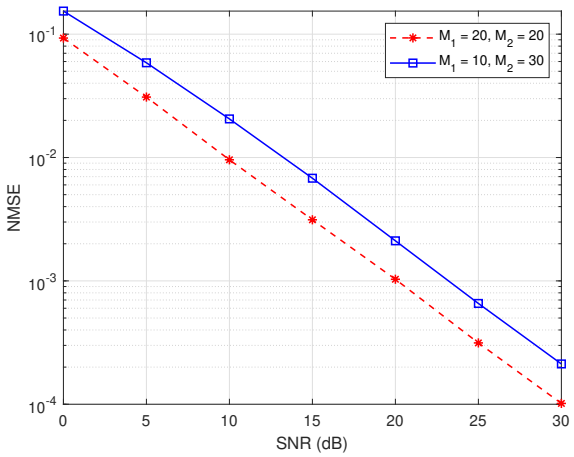

Fig. 4. NMSE versus SNR

number of S-RIS elements, M, is considered to be distributed between RIS 1 and RIS 2 in both the D-RIS and M-RIS scenarios, i.e., $M = M_1 + M_2$. As an example, assume that $M = 40$ elements, then from (14) and (15), S-RIS has channel coefficient $\mathcal{H}_{\text{S-RIS}} = 240$, whereas D-RIS and M-RIS have channel coefficients $\mathcal{H}_{\text{D-RIS}} = 520$ with $M_1 = M_2 = M/2 = 20$ elements. This shows that the D-RIS and M-RIS systems have more channel coefficients than the S-RIS system. Moreover, as highlighted in [5], when $N_r \approx N_t$ and $\mathcal{H}_{\text{D-RIS}} > \mathcal{H}_{\text{S-RIS}}$, the S-RIS requires less training overhead. However, with less training overhead (in this case, $I = 20$), we can observe that the proposed semi-blind algorithm displays a better SE of about 35.9 bps/Hz than S-RIS and M-RIS scenarios with a SE of 16 bps/Hz and 25 bps/Hz, respectively, for an SNR= 30 dB. Also, the performance of our algorithm improves to SE 37 bps/Hz with an increase in N_r (see Fig. 5). Nevertheless, considering $M_1 = 10$, and $M_2 = 30$, the D-RIS channel number of coefficients decreases ($\mathcal{H}_{\text{D-RIS}} = 440$), which affects our algorithm's performance but still achieves SE of 30 bps/Hz better than its counterpart (see Fig. 5). In Fig. 6, we can see that an increase in training overhead ($I = 40$) has no significant impact on S-RIS performance compared to M-RIS and D-RIS, which shows a remarkable performance improvement.

V. Conclusion

This paper considers D-RIS-aided MIMO systems and proposes a new semi-blind channel estimation based on a joint (coupled) Tucker-2 tensor decomposition approach. We demonstrated that our proposed algorithm performs better in the presence of noise and converges in a few iterations. Moreover, we demonstrated that a perfect end-to-end channel can be estimated with a small training overhead, which results in improved spectral efficiency of a D-RIS system in comparison with S-RIS and M-RIS scenarios.

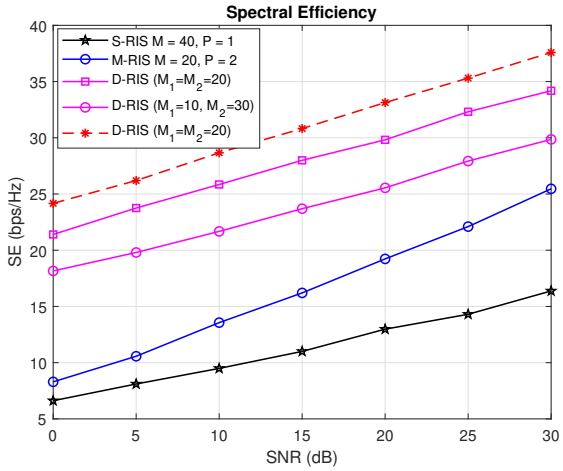

Fig. 5. SE versus SNR for $I = 20$ training overhead for all cases. $N_r = 4, N_t = 2$ for solid line, $N_r = 6, N_t = 2$ dashed line.

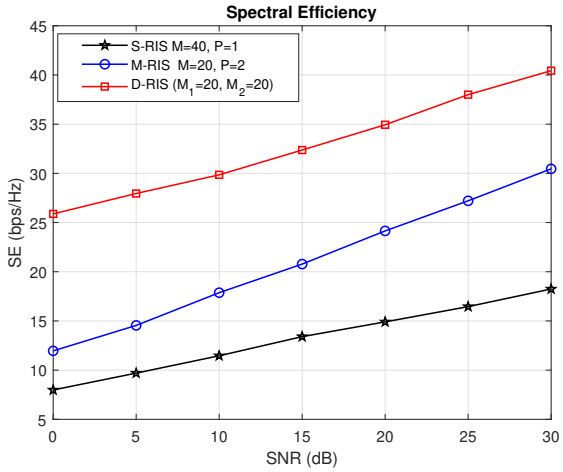

Fig. 6. SE versus SNR.$N_r = 4, N_t = 2$.

Acknowledgment

The authors would like to acknowledge the support provided by the Deanship of Research Oversight and Coordination at King Fahd University of Petroleum & Minerals for funding under the Interdisciplinary Research Center for Communication Systems and Sensing through project No. **INCS2404**.

References

[1] Q.-U.-A. Nadeem, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini, "Intelligent reflecting surface assisted wireless communication: Modeling and channel estimation," *arXiv preprint arXiv:1906.02360*, 2019.

[2] J. Du, X. Luo, X. Li, M. Zhu, K. M. Rabie, and F. Kara, "Semi-blind joint channel estimation and symbol detection for RIS-empowered multiuser mmWave systems," *IEEE Communications Letters*, vol. 27, no. 1, pp. 362–366, 2022.

[3] E. Björnson and L. Sanguinetti, "Power scaling laws and near-field behaviors of massive mimo and intelligent reflecting surfaces," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 1306–1324, 2020.

[4] B. Zheng, C. You, and R. Zhang, "Efficient channel estimation for double-IRS aided multi-user MIMO system," *IEEE Transactions on Communications*, vol. 69, no. 6, pp. 3818–3832, 2021.

[5] K. Ardah, S. Gherekhloo, A. L. de Almeida, and M. Haardt, "Double-RIS versus single-RIS aided systems: Tensor-based MIMO channel estimation and design perspectives," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5183–5187, 2022.

[6] K. d. A. Benicio, B. Sokal, and A. L. de Almeida, "Channel estimation and joint beamforming design for multi-IRS MIMO systems," in *2021 Brazilian Symposium on Telecommunications and Signal Processing (SBrT)*, pp. 1–5, 2021.

[7] G. T. de Araújo, A. L. de Almeida, and R. Boyer, "Channel estimation for intelligent reflecting surface assisted MIMO systems: A tensor modeling approach," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 3, pp. 789–802, 2021.

[8] Z. Wang, L. Liu, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis," *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, pp. 6607–6620, 2020.

[9] S. Gherekhloo, K. Ardah, A. L. De Almeida, M. Maleki, and M. Haardt, "Nested PARAFAC tensor-based channel estimation method for double RIS-aided MIMO communication systems," in *31st European Signal Processing Conference (EUSIPCO)*, pp. 1674–1678, 2023.

[10] K. N. Aliyu, A. Lawal, K. Abed-Meraim, and A. Zerguine, "Fast subspace-based semi-blind channel estimation for MIMO-OFDM communications," in *31st European Signal Processing Conference (EUSIPCO)*. IEEE, pp. 1460–1463, 2023.

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

[12] G. T. de Araújo, P. R. Gomes, A. L. de Almeida, G. Fodor, and B. Makki, "Semi-blind joint channel and symbol estimation in IRS-assisted multiuser MIMO networks," *IEEE Wireless Communications Letters*, vol. 11, no. 7, pp. 1553–1557, 2022.

[13] A. S. Alwakeel and A. Elzanaty, "Semi-blind channel estimation for intelligent reflecting surfaces in massive MIMO systems," *IEEE Access*, vol. 10, pp. 127 783–127 797, 2022.