DOUBLE-RIS VERSUS SINGLE-RIS AIDED SYSTEMS: TENSOR-BASED MIMO CHANNEL ESTIMATION AND DESIGN PERSPECTIVES

Khaled Ardah, Sepideh Gherekhloo, André L. F. de Almeida, Martin Haardt

ABSTRACT

Reconfigurable intelligent surfaces (RISs) have been proposed recently as new technology to tune the wireless propagation channels in real-time. However, most of the current works assume single-RIS (S-RIS)-aided systems, which can be limited in some application scenarios where a transmitter might need a multi-RIS-aided channel to communicate with a receiver. In this paper, we consider a double-RIS (D-RIS)-aided MIMO system and propose an alternating least-squares-based channel estimation method by exploiting the Tucker2 tensor structure of the received signals. Using the proposed method, the cascaded MIMO channel parts can be estimated separately, up to trivial scaling factors. Compared with the S-RIS systems, we show that if the RIS elements of an S-RIS system are distributed carefully between the two RISs in a D-RIS system, the training overhead can be reduced and the estimation accuracy can also be increased. Therefore, D-RIS systems can be seen as an appealing approach to further increase the coverage, capacity, and efficiency of future wireless networks compared to S-RIS systems.

***Index Terms*—** Double RIS, TUCKER2 decomposition, channel estimation, RIS reflection design.

1. INTRODUCTION

Reconfigurable intelligent surfaces (RISs) have been proposed recently as a cost-effective technology for reconfiguring the propagation channels in wireless communication systems [1]. An RIS is a 2D surface equipped with a large number of tunable units that can be realized using, e.g., inexpensive antennas or metamaterials and controlled in real-time to influence the communication channels without generating their own signals. Recently, RIS-aided communications have attracted great attention [2], due to their potential of improving the efficiency, the communication range, and the capacity of wireless communication systems.

The authors gratefully acknowledge the support of the German Research Foundation (DFG) under contracts no. HA 2239/6-2, HA 2239/14-1, ZH 640/2-1, and the support of CAPES/PRINT (Grant no. 88887.311965/2018-00). The research of André L. F. de Almeida is partially supported by the CNPq (Grant no. 306616/2016-5).

K. Ardah, S. Gherekhloo,, and M. Haardt are with the Communications Research Laboratory (CRL), TU Ilmenau, Ilmenau, Germany (e-mail: {khaled.ardah, sepideh.gherekhloo, martin.haardt}@tu-ilmenau.de). A. L. F. de Almeida is with the Wireless Telecom Research Group (GTEL), Federal University of Ceará, Fortaleza, Brazil (e-mail: andre@gtel.ufc.br).

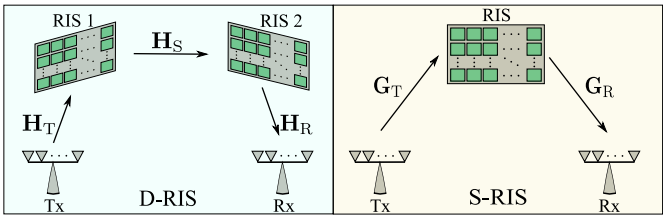

Fig. 1. D-RIS versus S-RIS aided MIMO communications.

Most of the current works, e.g., in [3–7], assume single-RIS (S-RIS)-aided systems, where a transmitter (Tx) communicates with one receiver (Rx), or more, via a single RIS-aided channel. However, in many application scenarios, e.g., in urban areas or satellite-to-indoor communications, the Tx might need a multi-RIS-aided channel to establish successful communication with the Rx. Moreover, in S-RIS systems, it was shown that the RIS should be either deployed closer to the Tx or closer to the Rx to achieve the best performance gain [8]. This fundamental result gives rise to double-RIS (D-RIS) systems, where one RIS is deployed closer to the Tx and another is deployed closer to the Rx. In such systems, channel estimation (CE) becomes more problematic since the cascaded (effective) channel contains three parts, not only two as in S-RIS systems (see Fig. 1).

In this paper, we consider a D-RIS aided MIMO system and propose an efficient CE method by exploiting the tensor structure of the received signals [9–11]. Specifically, we first show that the received signals in flat-fading D-RIS aided MIMO systems can be arranged in a 3-way tensor that admits a Tucker2 decomposition [9]. Accordingly, an alternating least-squares (ALS)-based method is proposed, where the Tx-to-RIS 1 channel (denoted by $\mathbf{H}_T$), the-RIS 1-to-RIS 2 channel (denoted by $\mathbf{H}_S$), and the RIS 2-to-Rx channel (denoted by $\mathbf{H}_R$) can be estimated separately, up to trivial scaling factors. We compare the proposed ALS method for D-RIS systems to the ALS method for S-RIS systems proposed in [7] in terms of the minimum training overhead and the estimation accuracy. It is shown that if the RIS elements in the S-RIS system are distributed carefully between the two RISs in the D-RIS system, the training overhead can be reduced and the estimation accuracy can also be increased. It is worth mentioning that the considered D-RIS system can also resemble communication scenarios where the RISs at both communicating ends are co-located with the transceivers, as it has been

proposed in [12]. In this case, the $\mathbf{H}_{\mathrm{T}}$ and $\mathbf{H}_{\mathrm{R}}$ channels can assumed to be known by a careful transceivers design.

2. D-RIS SYSTEM MODEL

In this section[1], we consider a D-RIS-aided MIMO communication system as depicted on the left-side of Fig. 1, where a Tx with M_{T} antennas is communicating with an Rx with M_{R} antennas via a D-RIS-aided channel. Here, RIS 1 is assumed to be close to the Tx and has N_1 reflecting elements, while RIS 2 is assumed to be close to the Rx and has N_2 reflecting elements. We assume that the Tx-to-Rx, the Tx-to-RIS 2, and the RIS 1-to-Rx channels are unavailable due to blockage or too weak due to high pathloss.

Let $\mathbf{H}_{\mathrm{T}} \in \mathbb{C}^{N_1 \times M_{\mathrm{T}}}$ be the Tx-to-RIS 1 channel, $\mathbf{H}_{\mathrm{S}} \in \mathbb{C}^{N_2 \times N_1}$ be the RIS 1-to-RIS 2 channel, and $\mathbf{H}_{\mathrm{R}} \in \mathbb{C}^{M_{\mathrm{R}} \times N_2}$ be the RIS 2-to-Rx channel. To estimate these channels, we conduct a channel-training procedure, which occupies $L = I \cdot K$ time symbols. The received signal at the (i,k)th time-symbol, $i \in \{1,\dots,I\}$ and $k \in \{1,\dots,K\}$, is given as

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

where $\boldsymbol{\Psi}_i = \mathrm{diag}\{\boldsymbol{\psi}_i\} \in \mathbb{C}^{N_1 \times N_1}$ is the ith diagonal reflection matrix of RIS 1, with $\boldsymbol{\psi}_i \in \mathbb{C}^{N_1}$ and $|[\boldsymbol{\psi}_i]_{[n]}| = 1/\sqrt{N_1}, \forall n$, $\boldsymbol{\Phi}_i = \mathrm{diag}\{\boldsymbol{\phi}_i\} \in \mathbb{C}^{N_2 \times N_2}$ is the ith diagonal reflection matrix of RIS 2, with $\boldsymbol{\phi}_i \in \mathbb{C}^{N_2}$ and $|[\boldsymbol{\phi}_i]_{[m]}| = 1/\sqrt{N_2}, \forall m$, $\mathbf{f}_k \in \mathbb{C}^{M_{\mathrm{T}}}$ is the kth precoding vector at the Tx with $\|\mathbf{f}_k\| = 1, \forall k$, s_k is the kth unit-norm training signal, and $\bar{\mathbf{n}}_{i,k} \in \mathbb{C}^{M_{\mathrm{R}}}$ is the additive white Gaussian noise vector having zero-mean circularly symmetric complex Gaussian entries with variance σ^2.

Let $\mathbf{F} = [\mathbf{f}_1 s_1, \dots, \mathbf{f}_K s_K] \in \mathbb{C}^{M_{\mathrm{T}} \times K}$. Then, by staking $\{\bar{\mathbf{y}}_{i,k}\}_{k=1}^{K}$ next to each other as $\bar{\mathbf{Y}}_i = [\bar{\mathbf{y}}_{i,1}, \dots, \bar{\mathbf{y}}_{i,K}]$, the obtained measurement matrix $\bar{\mathbf{Y}}_i$ can be expressed as

$$\bar{\mathbf{Y}}_i = \mathbf{H}_{\mathrm{R}} \boldsymbol{\Phi}_i \mathbf{H}_{\mathrm{S}} \boldsymbol{\Psi}_i \mathbf{H}_{\mathrm{T}} \mathbf{F} + \bar{\mathbf{N}}_i \in \mathbb{C}^{M_{\mathrm{R}} \times K}, \tag{2}$$

where $\bar{\mathbf{N}}_i \in \mathbb{C}^{M_{\mathrm{R}} \times K}$ is expressed similarly. We assume that the training matrix $\mathbf{F}$ is designed with orthonormal rows so that $\mathbf{F}\mathbf{F}^{\mathsf{H}} = \mathbf{I}_{M_{\mathrm{T}}}$, which directly implies that $K \geq M_{\mathrm{T}}$. Then, after right filtering $\bar{\mathbf{Y}}_i$ with $\mathbf{F}^{\mathsf{H}}$, i.e., $\mathbf{Y}_i = \bar{\mathbf{Y}}_i \mathbf{F}^{\mathsf{H}}$, we can write the obtained matrix $\mathbf{Y}_i$ as

$$\mathbf{Y}_i = \mathbf{H}_{\mathrm{R}} \boldsymbol{\Phi}_i \mathbf{H}_{\mathrm{S}} \boldsymbol{\Psi}_i \mathbf{H}_{\mathrm{T}} + \mathbf{N}_i \in \mathbb{C}^{M_{\mathrm{R}} \times M_{\mathrm{T}}}, \tag{3}$$

[1]**Notation:** The conjugate, the transpose, the conjugate transpose (Hermitian), the pseudoinverse, the Kronecker product, and the Khatri-Rao product are denoted as $\mathbf{A}^*$, $\mathbf{A}^{\mathsf{T}}$, $\mathbf{A}^{\mathsf{H}}$, $\mathbf{A}^{+}$, $\otimes$, and $\diamond$, respectively. Moreover, $\mathbf{1}_N$ is the all ones vector of length N, $\mathbf{I}_N$ is the $N \times N$ identity matrix, $\mathrm{diag}\{\mathbf{a}\}$ forms a diagonal matrix $\mathbf{A}$ by putting the entries of the input vector $\mathbf{a}$ on its main diagonal, $\mathrm{vec}\{\mathbf{A}\}$ forms a vector by staking the columns of $\mathbf{A}$ over each other, $\mathrm{unvec}\{\mathbf{A}\}$ is the reverse of the vec operator, $\lceil x \rceil$ is the ceiling function, and the n-mode product of a tensor $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ with a matrix $\mathbf{B} \in \mathbb{C}^{J \times I_n}$ is denoted as $\boldsymbol{\mathcal{A}} \times_n \mathbf{B}$.. Moreover, the following properties are used: Property 1: $\mathrm{vec}\{\mathbf{ABC}\} = (\mathbf{C}^{\mathsf{T}} \otimes \mathbf{A})\mathrm{vec}\{\mathbf{B}\}$ and Property 2: $\mathrm{vec}\{\mathbf{A}\mathrm{diag}\{\mathbf{b}\}\mathbf{C}\} = (\mathbf{C}^{\mathsf{T}} \diamond \mathbf{A})\mathbf{b}$.

where $\mathbf{N}_i = \bar{\mathbf{N}}_i \mathbf{F}^{\mathsf{H}}$. Given the measurement matrices $\mathbf{Y}_i, \forall i \in \{1,\dots,I\}$, our main goal in Section 2.1 is to obtain an estimate to the channel matrices $\mathbf{H}_{\mathrm{R}}$, $\mathbf{H}_{\mathrm{T}}$, and $\mathbf{H}_{\mathrm{S}}$.

2.1. Tucker2-based method for CE in D-RIS systems

By concatenating $\mathbf{Y}_1, \dots, \mathbf{Y}_I$ in (3) along the 3-mode, a 3-way tensor can be obtained as $\boldsymbol{\mathcal{Y}} = [\mathbf{Y}_1 \sqcup_3 \cdots \sqcup_3 \mathbf{Y}_I] \in \mathbb{C}^{M_{\mathrm{R}} \times M_{\mathrm{T}} \times I}$, where $\mathbf{Y}_i$ represents its ith frontal slice. Here, we note that the tensor $\boldsymbol{\mathcal{Y}}$ has a Tucker2 representation as [10]

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{S}} \times_1 \mathbf{H}_{\mathrm{R}} \times_2 \mathbf{H}_{\mathrm{T}}^{\mathsf{T}} + \boldsymbol{\mathcal{N}} \in \mathbb{C}^{M_{\mathrm{R}} \times M_{\mathrm{T}} \times I}, \tag{4}$$

where $\boldsymbol{\mathcal{N}}$ is the noise tensor and $\boldsymbol{\mathcal{S}}$ is formed by concatenating $\boldsymbol{\Phi}_i \mathbf{H}_{\mathrm{S}} \boldsymbol{\Psi}_i, \forall i \in \{1,\dots,I\}$, along the 3-mode as

$$\boldsymbol{\mathcal{S}} = [\boldsymbol{\Phi}_1 \mathbf{H}_{\mathrm{S}} \boldsymbol{\Psi}_1 \sqcup_3 \cdots \sqcup_3 \boldsymbol{\Phi}_I \mathbf{H}_{\mathrm{S}} \boldsymbol{\Psi}_I] \in \mathbb{C}^{N_2 \times N_1 \times I}. \tag{5}$$

From the above, the CE problem can be formulated as

$$\{\hat{\mathbf{H}}_{\mathrm{R}}, \hat{\mathbf{H}}_{\mathrm{T}}, \hat{\mathbf{H}}_{\mathrm{S}}\} = \underset{\mathbf{H}_{\mathrm{R}}, \mathbf{H}_{\mathrm{T}}, \mathbf{H}_{\mathrm{S}}}{\arg\min} \|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{S}} \times_1 \mathbf{H}_{\mathrm{R}} \times_2 \mathbf{H}_{\mathrm{T}}^{\mathsf{T}}\|_{\mathrm{F}}^2, \tag{6}$$

which is nonconvex due to its joint optimization. To obtain a solution, we resort to an alternating minimization approach, where we solve (6) for one variable assuming the other two are fixed. To achieve this end, we exploit the n-mode unfoldings of $\boldsymbol{\mathcal{Y}}$, i.e., $[\boldsymbol{\mathcal{Y}}]_{(n)}, n \in \{1,2,3\}$ expressed as [9, 10]

$$[\boldsymbol{\mathcal{Y}}]_{(1)} = \mathbf{H}_{\mathrm{R}} \mathbf{Z}_{\mathrm{R}}(\mathbf{H}_{\mathrm{T}}, \mathbf{H}_{\mathrm{S}}) + [\boldsymbol{\mathcal{N}}]_{(1)} \in \mathbb{C}^{M_{\mathrm{R}} \times I M_{\mathrm{T}}} \tag{7}$$

$$[\boldsymbol{\mathcal{Y}}]_{(2)} = \mathbf{H}_{\mathrm{T}}^{\mathsf{T}} \mathbf{Z}_{\mathrm{T}}(\mathbf{H}_{\mathrm{R}}, \mathbf{H}_{\mathrm{S}}) + [\boldsymbol{\mathcal{N}}]_{(2)} \in \mathbb{C}^{M_{\mathrm{T}} \times I M_{\mathrm{R}}} \tag{8}$$

$$[\boldsymbol{\mathcal{Y}}]_{(3)} = [\boldsymbol{\mathcal{S}}]_{(3)} (\mathbf{H}_{\mathrm{T}}^{\mathsf{T}} \otimes \mathbf{H}_{\mathrm{R}})^{\mathsf{T}} + [\boldsymbol{\mathcal{N}}]_{(3)} \in \mathbb{C}^{I \times M_{\mathrm{T}} M_{\mathrm{R}}}, \tag{9}$$

where $\mathbf{Z}_{\mathrm{R}}(\mathbf{H}_{\mathrm{T}}, \mathbf{H}_{\mathrm{S}}) = [\boldsymbol{\mathcal{S}}]_{(1)} (\mathbf{I}_I \otimes \mathbf{H}_{\mathrm{T}}^{\mathsf{T}})^{\mathsf{T}} \in \mathbb{C}^{N_2 \times I M_{\mathrm{T}}}$ and $\mathbf{Z}_{\mathrm{T}}(\mathbf{H}_{\mathrm{R}}, \mathbf{H}_{\mathrm{S}}) = [\boldsymbol{\mathcal{S}}]_{(2)} (\mathbf{I}_I \otimes \mathbf{H}_{\mathrm{R}})^{\mathsf{T}} \in \mathbb{C}^{N_1 \times I M_{\mathrm{R}}}$. Note that, according to the definition of n-mode unfoldings [9], $[\boldsymbol{\mathcal{S}}]_{(3)} \in \mathbb{C}^{I \times N_1 N_2}$ can be expressed as $[\boldsymbol{\mathcal{S}}]_{(3)} = (\boldsymbol{\Psi} \diamond \boldsymbol{\Phi})^{\mathsf{T}} \mathrm{diag}\{\mathbf{h}_{\mathrm{S}}\}$, where $\mathbf{h}_{\mathrm{S}} = \mathrm{vec}\{\mathbf{H}_{\mathrm{S}}\} \in \mathbb{C}^{N_1 N_2}$, $\boldsymbol{\Psi} = [\boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_I] \in \mathbb{C}^{N_1 \times I}$, and $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \dots, \boldsymbol{\phi}_I] \in \mathbb{C}^{N_2 \times I}$. Therefore,

$$[\boldsymbol{\mathcal{Y}}]_{(3)} = (\boldsymbol{\Psi} \diamond \boldsymbol{\Phi})^{\mathsf{T}} \mathrm{diag}\{\mathbf{h}_{\mathrm{S}}\} (\mathbf{H}_{\mathrm{T}}^{\mathsf{T}} \otimes \mathbf{H}_{\mathrm{R}})^{\mathsf{T}} + [\boldsymbol{\mathcal{N}}]_{(3)}. \tag{10}$$

The vectorized form of $[\boldsymbol{\mathcal{Y}}]_{(3)}$, i.e., $\mathbf{y}_{(3)} = \mathrm{vec}\{[\boldsymbol{\mathcal{Y}}]_{(3)}\}$ can be expressed as

$$\mathbf{y}_{(3)} = \mathbf{Z}_{\mathrm{S}}(\mathbf{H}_{\mathrm{T}}, \mathbf{H}_{\mathrm{R}})\, \mathbf{h}_{\mathrm{S}} + \mathbf{n}_{(3)} \in \mathbb{C}^{I M_{\mathrm{T}} M_{\mathrm{R}}}, \tag{11}$$

where $\mathbf{n}_{(3)} = \mathrm{vec}\{[\boldsymbol{\mathcal{N}}]_{(3)}\}$, $\mathbf{Z}_{\mathrm{S}}(\mathbf{H}_{\mathrm{T}}, \mathbf{H}_{\mathrm{R}}) = \Big((\mathbf{H}_{\mathrm{T}}^{\mathsf{T}} \otimes \mathbf{H}_{\mathrm{R}}) \diamond (\boldsymbol{\Psi} \diamond \boldsymbol{\Phi})^{\mathsf{T}}\Big) \in \mathbb{C}^{I M_{\mathrm{R}} M_{\mathrm{T}} \times N_1 N_2}$. By exploiting (7), (8), and (11), an estimate of $\mathbf{H}_{\mathrm{R}}$, $\mathbf{H}_{\mathrm{T}}$, and $\mathbf{h}_{\mathrm{S}}$ can be obtained as

$$\hat{\mathbf{H}}_{\mathrm{R}} = \underset{\mathbf{H}_{\mathrm{R}}}{\arg\min} \|[\boldsymbol{\mathcal{Y}}]_{(1)} - \mathbf{H}_{\mathrm{R}} \mathbf{Z}_{\mathrm{R}}(\mathbf{H}_{\mathrm{T}}, \mathbf{H}_{\mathrm{S}})\|_{\mathrm{F}}^2 \tag{12}$$

$$\hat{\mathbf{H}}_{\mathrm{T}} = \underset{\mathbf{H}_{\mathrm{T}}}{\arg\min} \|[\boldsymbol{\mathcal{Y}}]_{(2)} - \mathbf{H}_{\mathrm{T}}^{\mathsf{T}} \mathbf{Z}_{\mathrm{T}}(\mathbf{H}_{\mathrm{R}}, \mathbf{H}_{\mathrm{S}})\|_{\mathrm{F}}^2 \tag{13}$$

$$\hat{\mathbf{h}}_{\mathrm{S}} = \underset{\mathbf{h}_{\mathrm{S}}}{\arg\min} \|\mathbf{y}_{(3)} - \mathbf{Z}_{\mathrm{S}}(\mathbf{H}_{\mathrm{T}}, \mathbf{H}_{\mathrm{R}})\, \mathbf{h}_{\mathrm{S}}\|_2^2. \tag{14}$$

The above problems are convex and can be solved using the ALS method, as summarized in Algorithm 1, which is guaranteed to converge monotonically to, at least, a locally optimal solution [10].

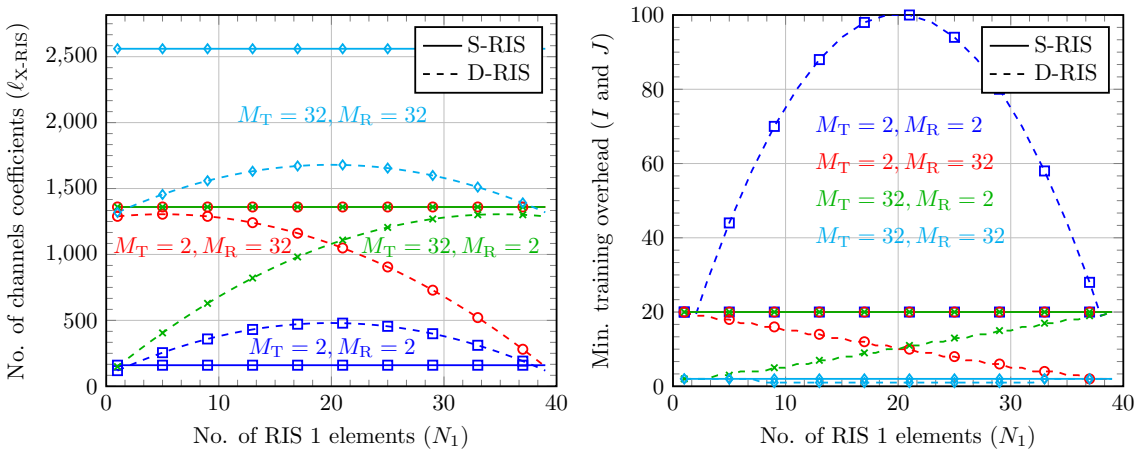

Fig. 2. Number of channels coefficients ($\ell_{\text{X-RIS}}$) and minimum training overhead (I and J) [$N = 40$, $N_2 = N - N_1$].

Algorithm 1 ALS method for CE in D-RIS MIMO systems

1: Input: Measurement tensor $\boldsymbol{\mathcal{Y}} \in \mathbb{C}^{M_{\text{R}} \times M_{\text{T}} \times I}$ as in (4)
2: Initialize: $\mathbf{H}_{\text{T}}^{(0)}$ and $\mathbf{H}_{\text{S}}^{(0)}$ and select $t_{\max}$.
3: **for** $t = 1$ to $t_{\max}$ **do**
4: $\quad \mathbf{H}_{\text{R}}^{(t)} = [\boldsymbol{\mathcal{Y}}]_{(1)} \{\mathbf{Z}_{\text{R}}(\mathbf{H}_{\text{T}}^{(t-1)}, \mathbf{H}_{\text{S}}^{(t-1)})\}^{+}$
5: $\quad \hat{\mathbf{H}}_{\text{T}}^{(t)} = \left([\boldsymbol{\mathcal{Y}}]_{(2)} \{\mathbf{Z}_{\text{T}}(\mathbf{H}_{\text{R}}^{(t)}, \mathbf{H}_{\text{S}}^{(t-1)})\}^{+}\right)^{\mathsf{T}}$
6: $\quad \hat{\mathbf{H}}_{\text{S}}^{(t)} = \text{unvec}\left\{ \{\mathbf{Z}_{\text{S}}(\mathbf{H}_{\text{T}}^{(t)}, \mathbf{H}_{\text{R}}^{(t)})\}^{+} \mathbf{y}_{(3)} \right\}$
7: **end for**

3. COMPARISON WITH S-RIS-AIDED SYSTEMS

In S-RIS-aided systems, on the other hand, the communication between the Tx and the Rx with M_{T} and M_{R} antennas, respectively, is aided via a single RIS with N elements, as depicted on the right-side of Fig. 1. Let $\mathbf{G}_{\text{T}} \in \mathbb{C}^{N \times M_{\text{T}}}$ be the Tx to RIS channel and $\mathbf{G}_{\text{R}} \in \mathbb{C}^{M_{\text{R}} \times N}$ be the RIS to Rx channel. Then, it was shown in [7] that the received signals at the Rx can be arranged in a 3-way tensor admitting a canonical polyadic (CP) decomposition given as

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{I}}_{3,N} \times_1 \mathbf{G}_{\text{R}} \times_2 \mathbf{G}_{\text{T}}^{\mathsf{T}} \times_3 \boldsymbol{\Omega}^{\mathsf{T}} + \boldsymbol{\mathcal{E}} \in \mathbb{C}^{M_{\text{R}} \times M_{\text{T}} \times J}, \quad (15)$$

where $\boldsymbol{\mathcal{I}}_{3,N} \in \mathbb{Z}^{N \times N \times N}$ is the super-diagonal tensor, $\boldsymbol{\mathcal{E}}$ is the noise tensor, $\boldsymbol{\Omega} = [\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_J] \in \mathbb{C}^{N \times J}$ is the RIS training matrix with J training beams and $|[\boldsymbol{\omega}_j]_{[\ell]}| = 1/\sqrt{N}, \forall \ell$. The n-mode unfoldings of $\boldsymbol{\mathcal{X}}$, $n \in \{1, 2\}$, can be expressed as

$$[\boldsymbol{\mathcal{X}}]_{(1)} = \mathbf{G}_{\text{R}} \, \mathbf{V}_{\text{R}}(\mathbf{G}_{\text{T}}) + [\boldsymbol{\mathcal{E}}]_{(1)} \in \mathbb{C}^{M_{\text{R}} \times J M_{\text{T}}} \quad (16)$$

$$[\boldsymbol{\mathcal{X}}]_{(2)} = \mathbf{G}_{\text{T}}^{\mathsf{T}} \, \mathbf{V}_{\text{T}}(\mathbf{G}_{\text{R}}) + [\boldsymbol{\mathcal{E}}]_{(2)} \in \mathbb{C}^{M_{\text{T}} \times J M_{\text{R}}}, \quad (17)$$

where $\mathbf{V}_{\text{R}}(\mathbf{G}_{\text{T}}) = (\boldsymbol{\Omega}^{\mathsf{T}} \diamond \mathbf{G}_{\text{T}}^{\mathsf{T}})^{\mathsf{T}} \in \mathbb{C}^{N \times J M_{\text{T}}}$ and $\mathbf{V}_{\text{T}}(\mathbf{G}_{\text{R}}) = (\boldsymbol{\Omega}^{\mathsf{T}} \diamond \mathbf{G}_{\text{R}})^{\mathsf{T}} \in \mathbb{C}^{N \times J M_{\text{R}}}$. Therefore, an ALS-based method, similarly to Algorithm 1, has been proposed in [7] to obtain an estimate of $\mathbf{G}_{\text{R}}$ and $\mathbf{G}_{\text{T}}$, as summarized in Algorithm 2.

Identifiablity, in the LS sense, can be obtained by noting that $\mathbf{Z}_{\text{R}}$, $\mathbf{Z}_{\text{T}}$, $\mathbf{V}_{\text{R}}$, and $\mathbf{V}_{\text{T}}$ need to have full column-rank, while $\mathbf{Z}_{\text{S}}$ needs to have full row-rank [10]. This leads to the following conditions: $IM_{\text{T}} \geq N_2$ (for $\mathbf{Z}_{\text{R}}$), $IM_{\text{R}} \geq N_1$ (for $\mathbf{Z}_{\text{T}}$), $JM_{\text{T}} \geq N$ (for $\mathbf{V}_{\text{R}}$), $JM_{\text{R}} \geq N$ (for $\mathbf{V}_{\text{T}}$), and $IM_{\text{R}}M_{\text{T}} \geq N_1 N_2$ (for $\mathbf{Z}_{\text{S}}$). Therefore, we conclude that

$$I \geq \max\left\{\lceil N_2/M_{\text{T}} \rceil, \ \lceil N_1/M_{\text{R}} \rceil, \ \lceil N_1 N_2/M_{\text{R}} M_{\text{T}} \rceil\right\} \quad (18)$$

$$J \geq \max\left\{\lceil N/M_{\text{T}} \rceil, \ \lceil N/M_{\text{R}} \rceil\right\}, \quad (19)$$

where (18) is for the D-RIS systems and (19) is for the S-RIS systems. Let $\ell_{\text{D-RIS}}$ and $\ell_{\text{S-RIS}}$ denote the total number of channel coefficients in the D-RIS and the S-RIS communication scenarios, respectively, which are given as

$$\ell_{\text{D-RIS}} = M_{\text{T}} N_1 + N_1 N_2 + M_{\text{R}} N_2 \quad (20)$$

$$\ell_{\text{S-RIS}} = M_{\text{T}} N + M_{\text{R}} N. \quad (21)$$

Algorithm 2 ALS method for CE in S-RIS MIMO systems

1: Input: Measurement tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{M_{\text{R}} \times M_{\text{T}} \times J}$ as in (15)
2: Initialize: $\mathbf{G}_{\text{T}}^{(0)}$ and select $t_{\max}$.
3: **for** $t = 1$ to $t_{\max}$ **do**
4: $\quad \mathbf{G}_{\text{R}}^{(t)} = [\boldsymbol{\mathcal{X}}]_{(1)} \{\mathbf{V}_{\text{R}}(\mathbf{G}_{\text{T}}^{(t-1)})\}^{+}$
5: $\quad \mathbf{G}_{\text{T}}^{(t)} = \left([\boldsymbol{\mathcal{X}}]_{(2)} \{\mathbf{V}_{\text{T}}(\mathbf{G}_{\text{R}}^{(t)})\}^{+}\right)^{\mathsf{T}}$
6: **end for**

Let us assume that the S-RIS elements N are distributed between RIS 1 and RIS 2 in the D-RIS scenario such that $N = N_1 + N_2$. In Fig. 2 we plot the results of (18), (19), (20), and (21) for different M_{R} and M_{T} values assuming $N = 40$. Note that along the x-axis we vary N_1 so that $N_2 = N - N_1$. From Fig. 2, we have the following remarks:

***Remark 1**: If $N \gg \max\{M_R, M_T\}$ and $M_R \approx M_T$, then S-RIS, i.e., Algorithm 2 requires less training overhead compared to D-RIS, i.e., Algorithm 1, in most of the N_1 and N_2 distribution scenarios. This comes from the fact that the number of channel coefficients that D-RIS needs to estimate, i.e., $\ell_{\text{D-RIS}}$ is much larger than that of S-RIS, i.e., $\ell_{\text{S-RIS}}$.*

***Remark 2**: If $N \approx \max\{M_R, M_T\}$, then D-RIS requires less training overhead compared to S-RIS for the same reason mentioned in Remark 1, i.e., $\ell_{\text{D-RIS}} < \ell_{\text{S-RIS}}$.*

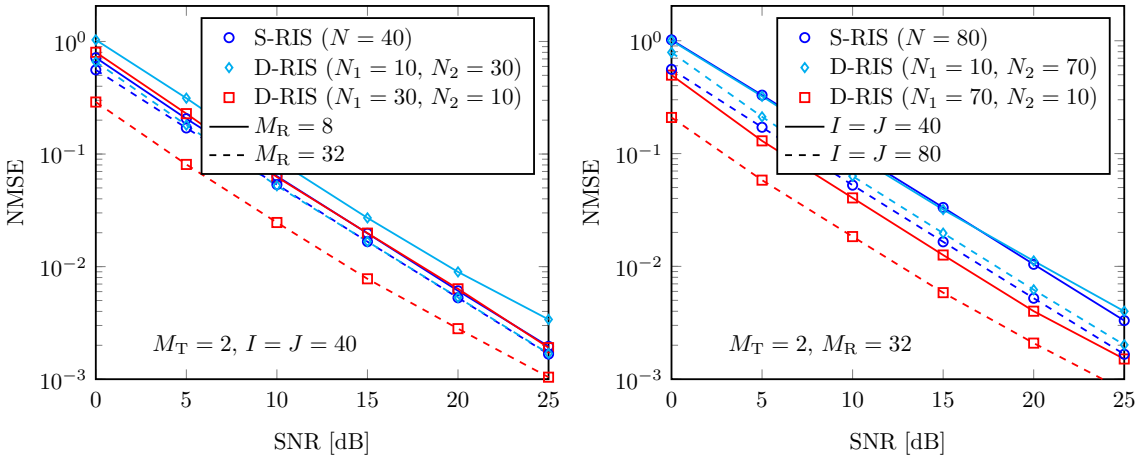

Fig. 3. NMSE versus SNR comparing the D-RIS against the S-RIS systems for different system settings [$M_{\mathrm{T}} = 2$].

***Remark 3**: In the D-RIS systems, the careful distribution of the N elements between RIS 1 and RIS 2 (i.e., N_1 and N_2) can reduce the training overhead of Algorithm 1. From Fig. 2, we can note that the best distribution depends on the M_R and the M_T values as: if $M_R > M_T$, then it is more beneficial to allocate more elements to RIS 1 than RIS 2, i.e., $N_1 > N_2$. This observation is reversed if $M_R < M_T$, i.e., more elements should be allocated to RIS 2 than RIS 1 as $N_1 < N_2$.*

Computational complexity: Assuming that the conditions in (18) and (19) are satisfied, the complexities[2] of Algorithm 1 and Algorithm 2 are on the order of $\mathcal{O}\big(t_{\max} \cdot (N_2^3 + N_1^3 + (N_1 \cdot N_2)^3)\big)$ and $\mathcal{O}\big(t_{\max} \cdot 2N^3\big)$, respectively.

Ambiguities: Assuming that the conditions in (18) and (19) are satisfied, then the estimated MIMO channels by Algorithm 1 and Algorithm 2 are unique up to scalar ambiguities[7, 10]. In particular, the estimated channels are related to the perfect (true) channels as: $\hat{\mathbf{H}}_{\mathrm{R}} \approx \mathbf{H}_{\mathrm{R}}\boldsymbol{\Delta}_{\mathrm{R}}$, $\hat{\mathbf{H}}_{\mathrm{T}} \approx \boldsymbol{\Delta}_{\mathrm{T}}\mathbf{H}_{\mathrm{T}}$, $\mathbf{H}_{\mathrm{S}} \approx \boldsymbol{\Delta}_{\mathrm{R}}^{-1}\hat{\mathbf{H}}_{\mathrm{S}}\boldsymbol{\Delta}_{\mathrm{T}}^{-1}$, $\hat{\mathbf{G}}_{\mathrm{R}} \approx \mathbf{G}_{\mathrm{R}}\boldsymbol{\Lambda}$, and $\hat{\mathbf{G}}_{\mathrm{T}} \approx \boldsymbol{\Lambda}^{-1}\hat{\mathbf{G}}_{\mathrm{T}}$, where $\boldsymbol{\Delta}$ and $\boldsymbol{\Lambda}$ are diagonal matrices holding the scaling ambiguities. However, these ambiguities disappear when reconstructing an estimate of the effective end-to-end channels $\hat{\mathbf{H}}_{\mathrm{e}} = \hat{\mathbf{H}}_{\mathrm{R}}\hat{\mathbf{H}}_{\mathrm{S}}\hat{\mathbf{H}}_{\mathrm{T}}$ and $\hat{\mathbf{G}}_{\mathrm{e}} = \hat{\mathbf{G}}_{\mathrm{R}}\hat{\mathbf{G}}_{\mathrm{T}}$. Moreover, note that, due to the knowledge of the RIS reflection matrices $\boldsymbol{\Psi}$, $\boldsymbol{\Phi}$, and $\boldsymbol{\Omega}$ at the Rx, the permutation ambiguities do not exist [7].

4. SIMULATION RESULTS

We assume that the entries of the channel matrices $\mathbf{H}_{\mathrm{R}}$, $\mathbf{H}_{\mathrm{T}}$, $\mathbf{H}_{\mathrm{S}}$, $\mathbf{G}_{\mathrm{R}}$, and $\mathbf{G}_{\mathrm{T}}$ follow a Rayleigh fading distribution. We show results in terms of the normalized-mean-square-error (NMSE) of the effective channels defined as $\mathrm{NMSE} = \mathbb{E}\{\|\mathbf{H}_{\mathrm{e}} - \hat{\mathbf{H}}_{\mathrm{e}}\|_{\mathrm{F}}^2\}/\mathbb{E}\{\|\mathbf{H}_{\mathrm{e}}\|_{\mathrm{F}}^2\}$, for the D-RIS, and $\mathrm{NMSE} = \mathbb{E}\{\|\mathbf{G}_{\mathrm{e}} - \hat{\mathbf{G}}_{\mathrm{e}}\|_{\mathrm{F}}^2\}/\mathbb{E}\{\|\mathbf{G}_{\mathrm{e}}\|_{\mathrm{F}}^2\}$, for the S-RIS. We define the signal-to-noise ratio (SNR) as $\mathrm{SNR} = \mathbb{E}\{\|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{N}}\|_{\mathrm{F}}^2\}/\mathbb{E}\{\|\boldsymbol{\mathcal{N}}\|_{\mathrm{F}}^2\}$, for the D-RIS, and $\mathrm{SNR} = \mathbb{E}\{\|\boldsymbol{\mathcal{X}} - \boldsymbol{\mathcal{E}}\|_{\mathrm{F}}^2\}/\mathbb{E}\{\|\boldsymbol{\mathcal{E}}\|_{\mathrm{F}}^2\}$, for the S-RIS. Moreover, assuming that $I \leq N_1N_2$ and $J \leq N$, the training matrices $\boldsymbol{\Psi}$, $\boldsymbol{\Phi}$, and $\boldsymbol{\Omega}$ are designed using a DFT-based approach as: $\boldsymbol{\Phi} = [\mathbf{W}_{N_2} \otimes \mathbf{1}_{\bar{I}_2}^{\mathsf{T}}]_{[:,1:I]}$, $\boldsymbol{\Psi} = [\mathbf{1}_{\bar{I}_1}^{\mathsf{T}} \otimes \mathbf{W}_{N_1}]_{[:,1:I]}$, and $\boldsymbol{\Omega} = [\mathbf{W}_N]_{[:,1:J]}$, where $\bar{I}_1 = \left\lceil \frac{I}{N_1} \right\rceil$, $\bar{I}_2 = \left\lceil \frac{I}{N_2} \right\rceil$, and $\mathbf{W}_K$ is the normalized $K \times K$ DFT matrix such that $\boldsymbol{\Omega}^{\mathsf{H}}\boldsymbol{\Omega} = \mathbf{I}_J$ and $\boldsymbol{\Upsilon}^{\mathsf{H}}\boldsymbol{\Upsilon} = \mathbf{I}_I$, where $\boldsymbol{\Upsilon} \overset{\text{def}}{=} \boldsymbol{\Psi} \diamond \boldsymbol{\Phi}$.

Fig. 3 shows the NMSE versus SNR results for different system settings. From the left-side figure, we can see that when $M_{\mathrm{R}} = 8$, the D-RIS, i.e., Algorithm 1 has a worse NMSE performance compared to the S-RIS, i.e., Algorithm 2 especially with the $[N_1, N_2] = [10, 30]$ distribution scenario. This can be explained from Fig. 2 and Remarks 1 and 3. Note that in a such system setting, the D-RIS has a larger number of channels coefficients $\ell_{\text{D-RIS}} = 560$ compared to the S-RIS $\ell_{\text{S-RIS}} = 400$. Moreover, as we have highlighted in Remark 3, we can see that the $[N_1, N_2] = [30, 10]$ distribution scenario has a better NMSE performance than $[N_1, N_2] = [10, 30]$, since $M_{\mathrm{T}} < M_{\mathrm{R}}$. On the other hand, when $M_{\mathrm{R}} = 32$, we can see that the D-RIS has a much better NMSE performance compared to the S-RIS, especially with the $[N_1, N_2] = [30, 10]$ distribution scenario. This can be explained in the same way from Fig. 2 and Remarks 2 and 3. From the right-side figure, we can see that the same observations hold true when we increase N from 40 to 80 or when we increase the training overhead I and J from 40 to 80.

5. CONCLUSIONS

In this paper, we have shown that D-RIS MIMO systems can be used to reduce the training overhead and to improve the channel estimation accuracy compared to S-RIS aided systems. This comes from the observation that if the RIS elements in the S-RIS system are distributed carefully between the two RISs in the D-RIS system, the number of channel coefficients in the D-RIS system that need to be estimated reduces significantly compared to the S-RIS system. Therefore, D-RIS systems can be seen as an appealing approach to further increase the coverage, capacity, and efficiency of wireless networks compared to S-RIS systems.

[2]Here, we have assumed that the complexity of calculating the Moore-Penrose inverse of a $n \times m$ matrix is on the order of $\mathcal{O}(\min\{n, m\}^3)$.

6. REFERENCES

[1] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "A new wireless communication paradigm through software-controlled metasurfaces," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 162–169, 2018.

[2] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, 2020.

[3] S. Zhang and R. Zhang, "Capacity characterization for intelligent reflecting surface aided MIMO communication," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1823–1838, Aug. 2020.

[4] K. Ardah, S. Gherekhloo, A. L. F. de Almeida, and M. Haardt, "TRICE: A channel estimation framework for RIS-aided millimeter-wave MIMO systems," *IEEE Signal Process. Lett.*, vol. 28, pp. 513–517, Feb. 2021.

[5] S. Gherekhloo, K. Ardah, A. L. F. de Almeida, and M. Haardt, "Tensor-based channel estimation and reflection design for RIS-aided millimeter-wave MIMO communication systems," in *Proc. of 55th Asilomar Conf. on Signals, Systems, and Computers*, Nov. 2021.

[6] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, 2019.

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

[8] E. Björnson and L. Sanguinetti, "Power scaling laws and near-field behaviors of massive MIMO and intelligent reflecting surfaces," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 1306–1324, 2020.

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

[10] P. Comon, X. Luciani, and A. L. F. de Almeida, "Tensor decompositions, alternating least squares and other tales," *Journal of Chemometrics*, vol. 23, no. 7-8, pp. 393–405, 2009.

[11] K. Ardah, A. L. F. de Almeida, and M. Haardt, "Low-complexity millimeter wave CSI estimation in MIMO-OFDM hybrid beamforming systems," in *Proc. 23rd International ITG Workshop on Smart Antennas (WSA)*, Apr. 2019, pp. 1–5.

[12] V. Jamali, A. M. Tulino, G. Fischer, R. R. Müller, and R. Schober, "Intelligent surface-aided transmitter architectures for millimeter-wave ultra massive MIMO systems," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 144–167, 2021.