TRICE: A Channel Estimation Framework for RIS-Aided Millimeter-Wave MIMO Systems

Khaled Ardah, Sepideh Gherekhloo, André L. F. de Almeida, *Senior Member, IEEE*, and Martin Haardt, *Fellow, IEEE*

***Abstract*—We consider the channel estimation problem in point-to-point reconfigurable intelligent surface (RIS)-aided millimeter-wave (mmWave) MIMO systems. By exploiting the low-rank nature of mmWave channels in the angular domains, we propose a non-iterative Two-stage RIS-aided Channel Estimation (TRICE) framework, where every stage is formulated as a multidimensional direction-of-arrival (DOA) estimation problem. As a result, our TRICE framework is very general in the sense that any efficient multidimensional DOA estimation solution can be readily used in every stage to estimate the associated channel parameters. Numerical results show that the TRICE framework has a lower training overhead and a lower computational complexity, as compared to benchmark solutions.**

***Index Terms*—Reconfigurable intelligent surface, direction of arrival estimation, compressed sensing, ESPRIT, MIMO.**

I. Introduction

RECONFIGURABLE intelligent surfaces (RISs) have been proposed recently as a cost-effective technology for reconfiguring the wireless propagation channel between transceivers [1]–[7]. In RIS-aided systems, an accurate channel state information (CSI) is required at transceivers to enable efficient signal processing techniques, e.g., beamforming and resource allocation. However, the acquisition of CSI in such systems faces several challenges. For instance, assuming a passive RIS implementation, to reduce the RIS cost and complexity, the propagation channel can only be sensed and estimated at the receiver. Furthermore, the large number of channel coefficients to be estimated limits the feasibility of CSI acquisition within a practical coherence time, since an RIS is expected to have a massive number of passive reflecting elements. Recently, channel estimation methods for RIS-aided systems have been proposed, e.g., using least-squares (LS) based methods as in [8]–[11], or minimum mean squared error based methods as in [12]. However, these works require the number of training subframes to be, at least, equal to the number of RISs reflecting elements, which is a limiting factor in practice.

In millimeter-wave (mmWave) communications [13]–[18], it was observed that the MIMO propagation channel has a low-rank structure, due to the small number of scatterers. Such a low-rank structure can be exploited to reduce the channel training overhead and complexity, as it has been shown in [19]–[24]. In these works, every channel matrix is modeled as a summation of L paths, where L is much smaller than the number of transmit and receiver antennas, and every path is completely characterized by a direction-of-departure (DOD), a direction-of-arrival (DOA), and a complex path gain. Therefore, the channel estimation is formulated as a sparse recovery problem, for which compressed sensing (CS) techniques [25] can be used to efficiently recover the channel parameters using a small training overhead. In [19], the above problem is facilitated by assuming that the RIS has a few active elements, which, however, increases the deployment cost and the energy consumption of RIS-aided systems. In [20], the authors assumed that the base station (BS)-to-RIS channel is perfectly known, while in [21], the cascade channel matrix is assumed to have a single path, i.e., $L = 1$. Differently, the authors in [24] proposed a general sparse recovery formulation for $L \geq 1$ scenarios. In most of these works, however, the channel parameters are assumed to fall perfectly on a grid, which may never be true in practice. Therefore, there exists a trade-off between the estimation accuracy and the complexity, where both increase as a function of the grid resolution. Due to the multidimensionality of the cascaded channel, a 4D sensing matrix is required by the method proposed in [24], which makes it computationally prohibitive even with low grid resolutions.

In this letter, we consider the channel estimation problem in a single-user RIS-aided mmWave MIMO communication system, similarly to [24], where the RIS has passive reflecting elements and the direct link between the BS and the mobile station (MS) is assumed to be blocked or pre-estimated by turning the RIS elements off, as in [8]. Using a structured channel training procedure, we propose a **T**wo-Stage **RI**S-aided **C**hannel **E**stimation (TRICE) framework for single-user mmWave MIMO communication systems. In the first stage, the DODs of the BS-to-RIS channel and the DOAs of the RIS-to-MS channel are first estimated. In the second stage, by using the estimated channel parameters in the first stage, the effective azimuth and elevation angles of the cascaded BS-to-RIS-to-MS channel at the RIS are estimated, one-by-one, including the effective complex path gains. In both stages, we show that the parameter estimation can be carried out *via* a multidimensional DOA estimation scheme, for which several solutions exist as in [26]–[31], among many others. Detailed simulation results are provided, showing that the proposed TRICE framework has a lower training overhead and

Manuscript received December 9, 2020; revised February 2, 2021; accepted February 6, 2021. Date of publication February 16, 2021; date of current version March 17, 2021. This work was supported in part by the German Research Foundation (DFG) under Grant HA 2239/6-2 (EXPRESS II) and Grant HA 2239/14-1 (AdAMMM), and in part by CAPES/PRINT under Grant 88887.311965/2018-00. The work of André L. F. de Almeida was supported by the CNPq under Grant 306616/2016-5. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Yunlong Cai. *(Corresponding author: Khaled Ardah.)*

Khaled Ardah, Sepideh Gherekhloo, and Martin Haardt are with the Communications Research Laboratory (CRL), Technische Universität Ilmenau, TU Ilmenau, 98693 Ilmenau, Germany (e-mail: khaled.ardah@tu-ilmenau.de; sepideh.gherekhloo@tu-ilmenau.de; martin.haardt@tu-ilmenau.de).

André L. F. de Almeida is with the Wireless Telecom Research Group (GTEL), Federal University of Ceará, Fortaleza 60020-181, Brazil (e-mail: andre@gtel.ufc.br).

Digital Object Identifier 10.1109/LSP.2021.3059363

1070-9908 © 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: University of Technology Sydney. Downloaded on May 25,2021 at 07:01:18 UTC from IEEE Xplore. Restrictions apply.

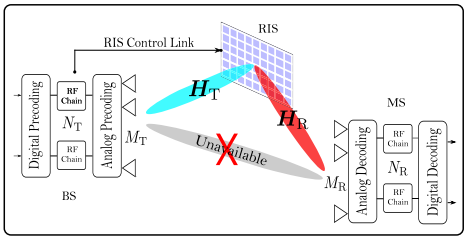

Fig. 1. An RIS-aided mmWave MIMO communication system.

a lower computational complexity, as compared to benchmark methods.

II. System and Channel Models

In this letter,[1] we consider a single-user mmWave MIMO communication system as depicted in Fig. 1, where a BS equipped with M_{T} antennas and $N_{\mathrm{T}} \leq M_{\mathrm{T}}$ RF chains is communicating with a MS that has M_{R} antennas and $N_{\mathrm{R}} \leq M_{\mathrm{R}}$ RF chains. We assume that the direct link between the BS and the MS is unavailable (e.g., due to blockage) and the indirect link is aided by an RIS composed by M_{S} phase shifters, which are arranged uniformly on a rectangular surface with $M_{\mathrm{S}}^{\mathrm{v}}$ vertical and $M_{\mathrm{S}}^{\mathrm{h}}$ horizontal elements such that $M_{\mathrm{S}} = M_{\mathrm{S}}^{\mathrm{v}} \cdot M_{\mathrm{S}}^{\mathrm{h}}$.

We assume that the BS and the MS employ uniform linear arrays (ULAs).[2] Let $\boldsymbol{H}_{\mathrm{T}} \in \mathbb{C}^{M_{\mathrm{S}} \times M_{\mathrm{T}}}$ ($\boldsymbol{H}_{\mathrm{R}} \in \mathbb{C}^{M_{\mathrm{R}} \times M_{\mathrm{S}}}$) denotes the mmWave MIMO channel between the BS (RIS) and the RIS (MS). Similarly to [24], $\boldsymbol{H}_{\mathrm{T}}$ and $\boldsymbol{H}_{\mathrm{R}}$ are modeled according to the classical Saleh-Valenzuela model [32] as

$$\boldsymbol{H}_{\mathrm{T}} = \sum_{\ell=1}^{L_{\mathrm{T}}} \alpha_{\mathrm{T},\ell} \boldsymbol{v}_{\mathrm{2D}}(\mu_{\mathrm{T},\ell}^{\mathrm{v}}, \mu_{\mathrm{T},\ell}^{\mathrm{h}}) \boldsymbol{v}_{\mathrm{1D}}(\psi_{\mathrm{T},\ell})^T = \boldsymbol{B}_{\mathrm{T}} \boldsymbol{G}_{\mathrm{T}} \boldsymbol{A}_{\mathrm{T}}^T$$

$$\boldsymbol{H}_{\mathrm{R}} = \sum_{\ell=1}^{L_{\mathrm{R}}} \alpha_{\mathrm{R},\ell} \boldsymbol{v}_{\mathrm{1D}}(\psi_{\mathrm{R},\ell}) \boldsymbol{v}_{\mathrm{2D}}(\mu_{\mathrm{R},\ell}^{\mathrm{v}}, \mu_{\mathrm{R},\ell}^{\mathrm{h}})^T = \boldsymbol{A}_{\mathrm{R}} \boldsymbol{G}_{\mathrm{R}} \boldsymbol{B}_{\mathrm{R}}^T, \quad (1)$$

where $\alpha_{\mathrm{T},\ell}$ and $\alpha_{\mathrm{R},\ell}$ are the complex path gains, $\psi_{\mathrm{T},\ell}$ ($\psi_{\mathrm{R},\ell}$) is the ℓth path DOD (DOA) spatial frequency at the BS (MS), while $\mu_{\mathrm{T},\ell}^{\mathrm{h}}$ and $\mu_{\mathrm{T},\ell}^{\mathrm{v}}$ ($\mu_{\mathrm{R},\ell}^{\mathrm{h}}$ and $\mu_{\mathrm{R},\ell}^{\mathrm{v}}$) are the ℓth path azimuth and elevation DOAs (DODs) spatial frequencies at the RIS.[3] Moreover, $\boldsymbol{v}_{\mathrm{2D}}(\mu_{\mathrm{X},\ell}^{\mathrm{v}}, \mu_{\mathrm{X},\ell}^{\mathrm{h}}) = \boldsymbol{v}_{\mathrm{1D}}(\mu_{\mathrm{X},\ell}^{\mathrm{v}}) \diamond \boldsymbol{v}_{\mathrm{1D}}(\mu_{\mathrm{X},\ell}^{\mathrm{h}}) \in \mathbb{C}^{M_{\mathrm{S}}}$ and $\boldsymbol{v}_{\mathrm{1D}}(\psi_{\mathrm{X},\ell}) \in \mathbb{C}^{M_{\mathrm{X}}}$ are the functions representing the 2D and the 1D array steering vectors, respectively, where $\mathrm{X} \in \{\mathrm{T}, \mathrm{R}\}$. For a given spatial frequency ν, the steering vector $\boldsymbol{v}_{\mathrm{1D}}(\nu)$ is given as $\boldsymbol{v}_{\mathrm{1D}}(\nu) = [1, e^{j\nu}, \ldots, e^{j(M-1)\nu}]^T \in \mathbb{C}^M$. In (1), $\boldsymbol{H}_{\mathrm{T}}$ and $\boldsymbol{H}_{\mathrm{R}}$ are written in a compact form by letting $\boldsymbol{A}_{\mathrm{X}} = [\boldsymbol{v}_{\mathrm{1D}}(\psi_{\mathrm{X},1}), \ldots, \boldsymbol{v}_{\mathrm{1D}}(\psi_{\mathrm{X},L_{\mathrm{X}}})] \in \mathbb{C}^{M_{\mathrm{X}} \times L_{\mathrm{X}}}$, $\boldsymbol{B}_{\mathrm{X}} = \boldsymbol{B}_{\mathrm{X}}^{\mathrm{v}} \diamond \boldsymbol{B}_{\mathrm{X}}^{\mathrm{h}} \in \mathbb{C}^{M_{\mathrm{S}} \times L_{\mathrm{X}}}$, $\boldsymbol{B}_{\mathrm{X}}^{\mathrm{y}} = [\boldsymbol{v}_{\mathrm{1D}}(\mu_{\mathrm{X},1}^{\mathrm{y}}), \ldots, \boldsymbol{v}_{\mathrm{1D}}(\mu_{\mathrm{X},L_{\mathrm{X}}}^{\mathrm{y}})] \in \mathbb{C}^{M_{\mathrm{S}}^{\mathrm{y}} \times L_{\mathrm{X}}}$, and $\boldsymbol{G}_{\mathrm{X}} = \mathrm{diag}\{\alpha_{\mathrm{X},1}, \ldots, \alpha_{\mathrm{X},L_{\mathrm{X}}}\}$, where $\mathrm{y} \in \{\mathrm{v}, \mathrm{h}\}$.

We assume a block-fading channel, where $\boldsymbol{H}_{\mathrm{T}}$ and $\boldsymbol{H}_{\mathrm{R}}$ remain constant during each block and change from block to block. To estimate $\boldsymbol{H}_{\mathrm{T}}$ and $\boldsymbol{H}_{\mathrm{R}}$, we conduct a channel training procedure at the beginning of each block, which comprises K frames divided into $K_{\mathrm{T}} \cdot K_{\mathrm{S}}$ subframes, i.e., $K = K_{\mathrm{S}} \cdot K_{\mathrm{T}}$. At the BS, we assume that a single RF chain is used during the channel training procedure, to reduce the energy consumption, which implies that a single training vector is transmitted in every subframe. Let $\widetilde{\boldsymbol{F}} = [\widetilde{\boldsymbol{f}}_1, \ldots, \widetilde{\boldsymbol{f}}_{K_T}] \in \mathbb{C}^{M_{\mathrm{T}} \times K_{\mathrm{T}}}$ be the matrix holding the K_{T} analog training vectors of the BS, with $\left|[\widetilde{\boldsymbol{f}}_t]_{[i]}\right| = \frac{1}{\sqrt{M_{\mathrm{T}}}}, \forall t, i$, and $\widetilde{\boldsymbol{F}}^H \widetilde{\boldsymbol{F}} = \boldsymbol{I}_{K_{\mathrm{T}}}$. Moreover, let $\boldsymbol{Q} = [\boldsymbol{q}_1, \ldots, \boldsymbol{q}_{K_{\mathrm{S}}}] \in \mathbb{C}^{M_{\mathrm{S}} \times K_{\mathrm{S}}}$ be the matrix holding the K_{S} phase shift vectors of the RIS, with $\left|[\boldsymbol{q}_s]_{[j]}\right| = \frac{1}{\sqrt{M_{\mathrm{S}}}}, \forall s, j$. We propose to design $\boldsymbol{Q}$ to have a Kronecker structure as

$$\boldsymbol{Q} = \boldsymbol{Q}_{\mathrm{v}} \otimes \boldsymbol{Q}_{\mathrm{h}} \in \mathbb{C}^{M_{\mathrm{S}} \times K_{\mathrm{S}}}, \quad (2)$$

where $\boldsymbol{Q}_{\mathrm{v}} \in \mathbb{C}^{M_{\mathrm{S}}^{\mathrm{v}} \times K_{\mathrm{S}}^{\mathrm{v}}}$, $\boldsymbol{Q}_{\mathrm{h}} \in \mathbb{C}^{M_{\mathrm{S}}^{\mathrm{h}} \times K_{\mathrm{S}}^{\mathrm{h}}}$, and $K_{\mathrm{S}} = K_{\mathrm{S}}^{\mathrm{v}} \cdot K_{\mathrm{S}}^{\mathrm{h}}$. Such a design structure will be exploited in Section III to obtain a low-complexity channel estimation method.

The received signal at the MS at the (s,t)th subframe, $s \in \{1, \ldots, K_{\mathrm{S}}\}$, $t \in \{1, \ldots, K_{\mathrm{T}}\}$, is given as

$$\boldsymbol{y}_{s,t} = \boldsymbol{W}^T \boldsymbol{H}_{\mathrm{R}} \mathrm{diag}\{\boldsymbol{q}_s\} \boldsymbol{H}_{\mathrm{T}} \widetilde{\boldsymbol{f}}_t p_t + \boldsymbol{z}_{s,t} \in \mathbb{C}^{N_{\mathrm{R}}}, \quad (3)$$

where $\boldsymbol{W} \in \mathbb{C}^{M_{\mathrm{R}} \times N_{\mathrm{R}}}$ is the decoding matrix, $p_t \in \mathbb{C}$ is the unit-power pilot signal, and $\boldsymbol{z}_{s,t} \in \mathbb{C}^{N_{\mathrm{R}}}$ is the additive white Gaussian noise vector having zero-mean circularly symmetric complex-valued entries with variance σ^2. Let $\boldsymbol{F} = [\widetilde{\boldsymbol{f}}_1 p_1, \ldots, \widetilde{\boldsymbol{f}}_{K_T} p_{K_{\mathrm{T}}}]$. Then, by stacking $\boldsymbol{y}_{s,t}, \forall t$, on top of each other as $\boldsymbol{y}_s = [\boldsymbol{y}_{s,1}^T, \ldots, \boldsymbol{y}_{s,K_{\mathrm{T}}}^T]^T$, we have

$$\boldsymbol{y}_s = (\boldsymbol{F}^T \boldsymbol{H}_{\mathrm{T}}^T \diamond \boldsymbol{W}^T \boldsymbol{H}_{\mathrm{R}}) \boldsymbol{q}_s + \boldsymbol{z}_s \in \mathbb{C}^{N_{\mathrm{R}} K_{\mathrm{T}}}, \quad (4)$$

where $\boldsymbol{z}_s = [\boldsymbol{z}_{s,1}^T, \ldots, \boldsymbol{z}_{s,K_{\mathrm{T}}}^T]^T$. Let $\boldsymbol{H} = (\boldsymbol{H}_{\mathrm{T}}^T \diamond \boldsymbol{H}_{\mathrm{R}}) \in \mathbb{C}^{M_{\mathrm{R}} M_{\mathrm{T}} \times M_{\mathrm{S}}}$ and $\boldsymbol{Z} = [\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{K_{\mathrm{S}}}]$. Then, by stacking $\boldsymbol{y}_s, \forall s$, as $\boldsymbol{Y} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{K_{\mathrm{S}}}]$ and applying Property 2, we have

$$\boldsymbol{Y} = (\boldsymbol{F}^T \otimes \boldsymbol{W}^T) \boldsymbol{H} \boldsymbol{Q} + \boldsymbol{Z} \in \mathbb{C}^{N_{\mathrm{R}} K_{\mathrm{T}} \times K_{\mathrm{S}}}. \quad (5)$$

Our main goal is to estimate $\boldsymbol{H}$ from (5). One direct solution is to use the LS-based method. By applying Property 1, the vectorized form of (5) can be written as $\boldsymbol{y} = \boldsymbol{\Upsilon} \boldsymbol{h} + \boldsymbol{z}$, where $\boldsymbol{\Upsilon} = (\boldsymbol{Q}^T \otimes \boldsymbol{F}^T \otimes \boldsymbol{W}^T) \in \mathbb{C}^{N_{\mathrm{R}} K_{\mathrm{T}} K_{\mathrm{S}} \times M_{\mathrm{R}} M_{\mathrm{T}} M_{\mathrm{S}}}$, $\boldsymbol{h} = \mathrm{vec}\{\boldsymbol{H}\}$, and $\boldsymbol{z} = \mathrm{vec}\{\boldsymbol{Z}\}$. Therefore, an estimate to the channel vector $\boldsymbol{h}$ can be obtained as $\widehat{\boldsymbol{h}}_{\mathrm{LS}} = \boldsymbol{\Upsilon}^+ \boldsymbol{y}$, which requires $K = K_{\mathrm{T}} K_{\mathrm{S}} \geq \frac{M_{\mathrm{R}} M_{\mathrm{T}} M_{\mathrm{S}}}{N_{\mathrm{R}}}$ to have an accurate channel estimate. Such an approach, however, becomes impractical in a massive MIMO setup, since it requires a large number of training subframes K and a long channel coherence time.

III. Proposed TRICE Framework

From (1), the cascaded channel matrix $\boldsymbol{H}$ can be written as

$$\boldsymbol{H} = (\boldsymbol{A}_{\mathrm{T}} \boldsymbol{G}_{\mathrm{T}} \boldsymbol{B}_{\mathrm{T}}^T \diamond \boldsymbol{A}_{\mathrm{R}} \boldsymbol{G}_{\mathrm{R}} \boldsymbol{B}_{\mathrm{R}}^T) \overset{(a)}{=} (\boldsymbol{A}_{\mathrm{T}} \otimes \boldsymbol{A}_{\mathrm{R}}) \boldsymbol{G} \boldsymbol{B}, \quad (6)$$

where $\boldsymbol{G} = (\boldsymbol{G}_{\mathrm{T}} \otimes \boldsymbol{G}_{\mathrm{R}}) \in \mathbb{C}^{L \times L}$, $\boldsymbol{B} = (\boldsymbol{B}_{\mathrm{T}}^T \diamond \boldsymbol{B}_{\mathrm{R}}^T) \in \mathbb{C}^{L \times M_{\mathrm{S}}}$, $L = L_{\mathrm{R}} L_{\mathrm{T}}$, and $\overset{(a)}{=}$ is obtained from Property 2.

[1] *Notation:* Matrices (vectors) are represented by boldface capital (lowercase) letters, $\boldsymbol{A}^T$, $\boldsymbol{A}^+$, $\otimes$, $\diamond$, and $\odot$ denote the transpose, the Moore-Penrose pseudo-inverse, the Kronecker, the Khatri-Rao, and the Hadamard products, respectively, $\mathrm{diag}\{\boldsymbol{a}\}$ forms a matrix by placing $\boldsymbol{a}$ on its main diagonal, and $\mathrm{vec}\{\boldsymbol{A}\}$ vectorizes $\boldsymbol{A}$ by arranging its columns on top of each other. We define $[\boldsymbol{a}]_{[n]}$ as the nth entry of vector $\boldsymbol{a}$, $\boldsymbol{1}_N$ as the all ones vector of length N, $\boldsymbol{I}_N$ as the $N \times N$ identity matrix, $\mathcal{CN}(\boldsymbol{0}, \boldsymbol{R})$ as the circularly symmetric complex Gaussian distribution with zeros mean and covariance matrix $\boldsymbol{R}$, and $\mathcal{U}(a_1, a_2)$ as the uniform distribution within the interval $[a_1, a_2]$. Moreover, the following properties are used: Property 1: $\mathrm{vec}\{\boldsymbol{ABC}\} = (\boldsymbol{C}^T \otimes \boldsymbol{A})\mathrm{vec}\{\boldsymbol{B}\}$. Property 2: $(\boldsymbol{AB} \diamond \boldsymbol{CD}) = (\boldsymbol{A} \otimes \boldsymbol{C})(\boldsymbol{B} \diamond \boldsymbol{D})$. Property 3: $(\boldsymbol{A} \otimes \boldsymbol{C})(\boldsymbol{B} \otimes \boldsymbol{D}) = (\boldsymbol{AB} \otimes \boldsymbol{CD})$.

[2] The extension of the proposed TRICE framework to scenarios where the BS and/or the MS are equipped with URAs is straightforward.

[3] Let d denotes the antenna spacing and λ be the signal wavelength. Then, the spatial frequencies are defined as $\psi_{\mathrm{X}\ell} = 2\pi \frac{d}{\lambda} \cos(\phi_{\mathrm{X},\ell})$, $\mu_{\mathrm{X},\ell}^{\mathrm{h}} = 2\pi \frac{d}{\lambda} \cos(\theta_{\mathrm{X},\ell}^{\mathrm{h}})$, and $\mu_{\mathrm{X},\ell}^{\mathrm{v}} = 2\pi \frac{d}{\lambda} \sin(\theta_{\mathrm{X},\ell}^{\mathrm{h}}) \cos(\theta_{\mathrm{X},\ell}^{\mathrm{v}})$, where $\phi_{\mathrm{X},\ell} \in [-180°, 180°]$ is the ℓth path angle in the angular domain, while $\theta_{\mathrm{X},\ell}^{\mathrm{h}} \in [-180°, 180°]$ ($\theta_{\mathrm{X},\ell}^{\mathrm{v}} \in [-90°, 90°]$) is the ℓth path azimuth (elevation) angle at the RIS in the angular domain.

Authorized licensed use limited to: University of Technology Sydney. Downloaded on May 25,2021 at 07:01:18 UTC from IEEE Xplore. Restrictions apply.

Using (6) and applying Property 3, we have

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{Z} \in \mathbb{C}^{N_{\mathrm{R}}K_{\mathrm{T}}\times K_{\mathrm{S}}}, \tag{7}$$

where $\boldsymbol{A} = (\boldsymbol{F}^T\boldsymbol{A}_{\mathrm{T}} \otimes \boldsymbol{W}^T\boldsymbol{A}_{\mathrm{R}})$ and $\boldsymbol{X} = \boldsymbol{G}\boldsymbol{B}\boldsymbol{Q}$. Observing (7), we can see that $\boldsymbol{A}$ is completely characterized by the frequency vectors defined as $\boldsymbol{\psi}_{\mathrm{T}} = [\psi_{\mathrm{T},1}, \ldots, \psi_{\mathrm{T},L_{\mathrm{T}}}]^T$ and $\boldsymbol{\psi}_{\mathrm{R}} = [\psi_{\mathrm{R},1}, \ldots, \psi_{\mathrm{R},L_{\mathrm{R}}}]^T$. Therefore, estimating $\boldsymbol{\psi}_{\mathrm{T}}$ and $\boldsymbol{\psi}_{\mathrm{R}}$ from (7) is, in fact, a 2D DOA estimation problem, where several methods exist in the literature, such as in [26]–[31], among many others. For instance, the DFT-beamspace ESPRIT methods of [26], [27] can be readily applied to estimate $\boldsymbol{\psi}_{\mathrm{T}}$ and $\boldsymbol{\psi}_{\mathrm{R}}$ in a closed form with guaranteed automatic pairing [33]. While subspace-based methods perform asymptotically optimal, they suffer from a performance degradation in the case of difficult scenarios such as high noise power and small number of measurement vectors. Alternatively, CS techniques [28]–[31] have been shown to provide an attractive alternative to subspace-based methods, yielding good estimation performance even in difficult scenarios. To show this, we note that (7) can be written in a sparse form as

$$\boldsymbol{Y} \approx (\boldsymbol{F}^T\overline{\boldsymbol{A}}_{\mathrm{T}} \otimes \boldsymbol{W}^T\overline{\boldsymbol{A}}_{\mathrm{R}})\overline{\boldsymbol{X}} + \boldsymbol{Z} \in \mathbb{C}^{N_{\mathrm{R}}K_{\mathrm{T}}\times K_{\mathrm{S}}}, \tag{8}$$

where $\overline{\boldsymbol{A}}_{\mathrm{T}} \in \mathbb{C}^{M_{\mathrm{T}}\times \bar{L}_{\mathrm{T}}}$ and $\overline{\boldsymbol{A}}_{\mathrm{R}} \in \mathbb{C}^{M_{\mathrm{R}}\times \bar{L}_{\mathrm{R}}}$ represent two dictionary matrices, in which $\bar{L}_{\mathrm{T}} \gg L_{\mathrm{T}}$ and $\bar{L}_{\mathrm{R}} \gg L_{\mathrm{R}}$ define the number of grid points or, in other words, the grid resolution, while $\overline{\boldsymbol{X}} \in \mathbb{C}^{\bar{L}_{\mathrm{T}}\bar{L}_{\mathrm{R}}\times K_{\mathrm{S}}}$ is an L row-sparse matrix [28]. Here, (8) can be written with equality if, and only if, the true angles $\boldsymbol{\psi}_{\mathrm{T}}$ and $\boldsymbol{\psi}_{\mathrm{R}}$ fall perfectly on the grid points. In this latter case, the kth nonzero row of $\overline{\boldsymbol{X}}$ equals to the kth row of $\boldsymbol{X}$. Note that (8) corresponds to a sparse recovery problem. Therefore, known CS techniques, e.g., [28]–[31], including the OMP method [34] can readily be applied to estimate $\overline{\boldsymbol{X}}$, as well as, $\boldsymbol{\psi}_{\mathrm{T}}$ and $\boldsymbol{\psi}_{\mathrm{R}}$, with automatic pairing. Since $L \ll M_{\mathrm{R}}M_{\mathrm{T}}$, due to the low-rank nature of the mmWave channels, only a few measurements (training overhead) are required, i.e., $N_{\mathrm{R}}K_{\mathrm{T}} \approx \mathcal{O}(L\log(\bar{L}_{\mathrm{R}}\bar{L}_{\mathrm{T}}/L)) \ll M_{\mathrm{R}}M_{\mathrm{T}}$ [35].[4]

To proceed, let $\widehat{\boldsymbol{\psi}}_{\mathrm{T}}$ and $\widehat{\boldsymbol{\psi}}_{\mathrm{R}}$ denote the estimated frequency vectors of $\boldsymbol{\psi}_{\mathrm{T}}$ and $\boldsymbol{\psi}_{\mathrm{R}}$. Then, we construct $\widehat{\boldsymbol{A}}_{\mathrm{T}}$, $\widehat{\boldsymbol{A}}_{\mathrm{R}}$, and $\widehat{\boldsymbol{A}} = (\boldsymbol{F}^T\widehat{\boldsymbol{A}}_{\mathrm{T}} \otimes \boldsymbol{W}^T\widehat{\boldsymbol{A}}_{\mathrm{R}})$. Therefore, to estimate $\boldsymbol{H}$ in (6), an estimate of $\boldsymbol{G}$ and $\boldsymbol{B}$ is required. Let us assume that $\boldsymbol{\psi}_{\mathrm{T}}$ and $\boldsymbol{\psi}_{\mathrm{R}}$ are estimated perfectly and that the rank$\{\widehat{\boldsymbol{A}}\} \geq L$. Then, multiplying (7) by $\widehat{\boldsymbol{A}}^+$ from the left-hand-side we get

$$\underline{\boldsymbol{Y}} = \widehat{\boldsymbol{A}}^+\boldsymbol{Y} = \boldsymbol{G}\boldsymbol{B}\boldsymbol{Q} + \underline{\boldsymbol{Z}} \in \mathbb{C}^{L\times K_{\mathrm{S}}}, \tag{9}$$

where $\underline{\boldsymbol{Z}} = \widehat{\boldsymbol{A}}^+\boldsymbol{Z} \in \mathbb{C}^{L\times K_{\mathrm{S}}}$ is the filtered noise. Since $\boldsymbol{G} = \boldsymbol{G}^T$, due to its diagonal structure, we can write $\underline{\boldsymbol{Y}}^T$ as

$$\underline{\boldsymbol{Y}}^T = \boldsymbol{Q}^T\boldsymbol{B}^T\boldsymbol{G} + \underline{\boldsymbol{Z}}^T \in \mathbb{C}^{K_{\mathrm{S}}\times L}. \tag{10}$$

Note that, $\boldsymbol{B}^T \in \mathbb{C}^{M_{\mathrm{S}}\times L}$ can be written as

$$\begin{aligned}\boldsymbol{B}^T &= \left[(\boldsymbol{b}_{\mathrm{T},1}^T \diamond \boldsymbol{b}_{\mathrm{R},1}^T)^T, \ldots, (\boldsymbol{b}_{\mathrm{T},1}^T \diamond \boldsymbol{b}_{\mathrm{R},L_{\mathrm{R}}}^T)^T, \ldots, (\boldsymbol{b}_{\mathrm{T},L_{\mathrm{T}}}^T \diamond \boldsymbol{b}_{\mathrm{R},L_{\mathrm{R}}}^T)^T\right] \\ &= \left[(\boldsymbol{b}_{\mathrm{T},1} \odot \boldsymbol{b}_{\mathrm{R},1}), \ldots, (\boldsymbol{b}_{\mathrm{T},1} \odot \boldsymbol{b}_{\mathrm{R},L_{\mathrm{R}}}), \ldots, (\boldsymbol{b}_{\mathrm{T},L_{\mathrm{T}}} \odot \boldsymbol{b}_{\mathrm{R},L_{\mathrm{R}}})\right],\end{aligned} \tag{11}$$

where $\boldsymbol{b}_{\mathrm{T},\ell} = \boldsymbol{v}_{1\mathrm{D}}(\mu_{\mathrm{T},\ell}^{\mathrm{v}}) \diamond \boldsymbol{v}_{1\mathrm{D}}(\mu_{\mathrm{T},\ell}^{\mathrm{h}})$ and $\boldsymbol{b}_{\mathrm{R},k} = \boldsymbol{v}_{1\mathrm{D}}(\mu_{\mathrm{R},k}^{\mathrm{v}}) \diamond \boldsymbol{v}_{1\mathrm{D}}(\mu_{\mathrm{R},k}^{\mathrm{h}})$ are the ℓth and the kth column vectors of $\boldsymbol{B}_{\mathrm{T}}$ and $\boldsymbol{B}_{\mathrm{R}}$, respectively, $\ell \in \{1, \ldots, L_{\mathrm{T}}\}$, $k \in \{1, \ldots, L_{\mathrm{R}}\}$, i.e.,

$$\boldsymbol{b}_{\mathrm{T},\ell} = \begin{bmatrix} 1 \cdot \boldsymbol{v}_{1\mathrm{D}}(\mu_{\mathrm{T},\ell}^{\mathrm{h}}) \\ e^{j\mu_{\mathrm{T},\ell}^{\mathrm{v}}} \cdot \boldsymbol{v}_{1\mathrm{D}}(\mu_{\mathrm{T},\ell}^{\mathrm{h}}) \\ \vdots \\ e^{j(M_{\mathrm{S}}^{\mathrm{v}}-1)\mu_{\mathrm{T},\ell}^{\mathrm{v}}} \cdot \boldsymbol{v}_{1\mathrm{D}}(\mu_{\mathrm{T},\ell}^{\mathrm{h}}) \end{bmatrix},$$

$$\boldsymbol{b}_{\mathrm{R},k} = \begin{bmatrix} 1 \cdot \boldsymbol{v}_{1\mathrm{D}}(\mu_{\mathrm{R},k}^{\mathrm{h}}) \\ e^{j\mu_{\mathrm{R},k}^{\mathrm{v}}} \cdot \boldsymbol{v}_{1\mathrm{D}}(\mu_{\mathrm{R},k}^{\mathrm{h}}) \\ \vdots \\ e^{j(M_{\mathrm{S}}^{\mathrm{v}}-1)\mu_{\mathrm{R},k}^{\mathrm{v}}} \cdot \boldsymbol{v}_{1\mathrm{D}}(\mu_{\mathrm{R},k}^{\mathrm{h}}) \end{bmatrix}.$$

Therefore, the nth column of $\boldsymbol{B}^T$, i.e., $\boldsymbol{b}_n = (\boldsymbol{b}_{\mathrm{T},\ell} \odot \boldsymbol{b}_{\mathrm{R},k})$ has a Khatri-Rao structure given as

$$\boldsymbol{b}_n = \begin{bmatrix} 1 & \cdot & \boldsymbol{v}_{1\mathrm{D}}(\mu_{\mathrm{T},\ell}^{\mathrm{h}} + \mu_{\mathrm{R},k}^{\mathrm{h}}) \\ e^{j(\mu_{\mathrm{T},\ell}^{\mathrm{v}} + \mu_{\mathrm{R},k}^{\mathrm{v}})} & \cdot & \boldsymbol{v}_{1\mathrm{D}}(\mu_{\mathrm{T},\ell}^{\mathrm{h}} + \mu_{\mathrm{R},k}^{\mathrm{h}}) \\ & \vdots & \\ e^{j(M_{\mathrm{S}}^{\mathrm{v}}-1)(\mu_{\mathrm{T},\ell}^{\mathrm{v}} + \mu_{\mathrm{R},k}^{\mathrm{v}})} & \cdot & \boldsymbol{v}_{1\mathrm{D}}(\mu_{\mathrm{T},\ell}^{\mathrm{h}} + \mu_{\mathrm{R},k}^{\mathrm{h}}) \end{bmatrix},$$

where $n = (\ell - 1) \cdot L_{\mathrm{R}} + k \in \{1, \ldots, L\}$. Let $\mu_n^{\mathrm{v}} = \mu_{\mathrm{T},\ell}^{\mathrm{v}} + \mu_{\mathrm{R},k}^{\mathrm{v}}$ and $\mu_n^{\mathrm{h}} = \mu_{\mathrm{T},\ell}^{\mathrm{h}} + \mu_{\mathrm{R},k}^{\mathrm{h}}$. Then, we have

$$\boldsymbol{b}_n = \boldsymbol{v}_{1\mathrm{D}}(\mu_n^{\mathrm{v}}) \diamond \boldsymbol{v}_{1\mathrm{D}}(\mu_n^{\mathrm{h}}) \in \mathbb{C}^{M_{\mathrm{S}}},$$

where $\boldsymbol{v}_{1\mathrm{D}}(\mu_n^{\mathrm{v}}) \in \mathbb{C}^{M_{\mathrm{S}}^{\mathrm{v}}}$ and $\boldsymbol{v}_{1\mathrm{D}}(\mu_n^{\mathrm{h}}) \in \mathbb{C}^{M_{\mathrm{S}}^{\mathrm{h}}}$. Accordingly, $\boldsymbol{B}^T = (\boldsymbol{B}^{\mathrm{v}} \diamond \boldsymbol{B}^{\mathrm{h}})$, in which $\boldsymbol{B}^{\mathrm{v}} = [\boldsymbol{v}_{1\mathrm{D}}(\mu_1^{\mathrm{v}}), \ldots, \boldsymbol{v}_{1\mathrm{D}}(\mu_L^{\mathrm{v}})]$, $\boldsymbol{B}^{\mathrm{h}} = [\boldsymbol{v}_{1\mathrm{D}}(\mu_1^{\mathrm{h}}), \ldots, \boldsymbol{v}_{1\mathrm{D}}(\mu_L^{\mathrm{h}})]$, and (10) can be written as

$$\underline{\boldsymbol{Y}}^T \overset{(a)}{=} (\boldsymbol{Q}_{\mathrm{v}}^T\boldsymbol{B}^{\mathrm{v}} \diamond \boldsymbol{Q}_{\mathrm{h}}^T\boldsymbol{B}^{\mathrm{h}})\boldsymbol{G} + \underline{\boldsymbol{Z}}^T \in \mathbb{C}^{K_{\mathrm{S}}\times L}, \tag{12}$$

where $\overset{(a)}{=}$ is obtained by utilizing the structure of $\boldsymbol{Q}$ in (2) and Property 2. Similarly to (7), the first term on the right-hand-side of (12) is completely characterized by the frequency vectors $\boldsymbol{\mu}^{\mathrm{v}} = [\mu_1^{\mathrm{v}}, \ldots, \mu_L^{\mathrm{v}}]$ and $\boldsymbol{\mu}^{\mathrm{h}} = [\mu_1^{\mathrm{h}}, \ldots, \mu_L^{\mathrm{h}}]$. Therefore, $\boldsymbol{\mu}^{\mathrm{v}}$ and $\boldsymbol{\mu}^{\mathrm{h}}$ can be estimated using the same methods discussed above. However, it should be noted that the joint estimation of $\boldsymbol{\mu}^{\mathrm{v}}$ and $\boldsymbol{\mu}^{\mathrm{h}}$ does not guarantee the automatic pairing with the pre-estimated frequency vectors $\boldsymbol{\psi}_{\mathrm{T}}$ and $\boldsymbol{\psi}_{\mathrm{R}}$. To overcome this issue, we utilize the diagonal structure of the $\boldsymbol{G}$ matrix in (12) and propose to estimate $\boldsymbol{\mu}^{\mathrm{v}}$ and $\boldsymbol{\mu}^{\mathrm{h}}$ sequentially, where the nth entries μ_n^{v} and μ_n^{h} can be jointly estimated from the nth column vector of $\underline{\boldsymbol{Y}}^T$ in (12), i.e., $\underline{\boldsymbol{y}}_n$ that is given as

$$\underline{\boldsymbol{y}}_n = (\boldsymbol{Q}_{\mathrm{v}}^T\boldsymbol{v}_{1\mathrm{D}}(\mu_n^{\mathrm{v}}) \diamond \boldsymbol{Q}_{\mathrm{h}}^T\boldsymbol{v}_{1\mathrm{D}}(\mu_n^{\mathrm{h}}))\alpha_n + \underline{\boldsymbol{z}}_n \in \mathbb{C}^{K_{\mathrm{S}}}, \tag{13}$$

where α_n is the nth diagonal entry of $\boldsymbol{G}$ and $\underline{\boldsymbol{z}}_n$ is the nth column vector of $\underline{\boldsymbol{Z}}^T$. Note that, due to the Kronecker structure of $\boldsymbol{Q}$ in (2), it is possible to apply the DFT-beamspace ESPRIT method of [26] on (13) to obtain closed form estimates of μ_n^{v} and μ_n^{h}. Next, for given $\hat{\mu}_n^{\mathrm{h}}$ and $\hat{\mu}_n^{\mathrm{v}}$, the nth path gain α_n can be estimated from (13) using LS as

$$\hat{\alpha}_n = \left(\boldsymbol{Q}_{\mathrm{v}}^T\boldsymbol{v}_{1\mathrm{D}}(\hat{\mu}_n^{\mathrm{v}}) \diamond \boldsymbol{Q}_{\mathrm{h}}^T\boldsymbol{v}_{1\mathrm{D}}(\hat{\mu}_n^{\mathrm{h}})\right)^+ \underline{\boldsymbol{y}}_n. \tag{14}$$

Finally, the $\boldsymbol{B}$ matrix in (6) can be reconstructed as $\widehat{\boldsymbol{B}} = (\widehat{\boldsymbol{B}}^{\mathrm{v}} \diamond \widehat{\boldsymbol{B}}^{\mathrm{h}})^T \in \mathbb{C}^{L\times M_{\mathrm{S}}}$. In summary, the proposed TRICE framework is given by Algorithm 1, where in Step 10, an estimate of $\boldsymbol{H}_{\mathrm{T}}$ and $\boldsymbol{H}_{\mathrm{R}}$, up to trivial scaling factors, can be obtained from $\widehat{\boldsymbol{H}}$ using the LS Khatri-Rao factorization (LSKRF) algorithm proposed in [11], [38]. Please note that Algorithm 1 is very general in the sense that any other efficient 2D parameter estimation method can be readily used in Steps 2 and 6, e.g., the methods proposed in [28]–[31].

[4]Note that the recoverability guarantee of $\overline{\boldsymbol{X}}$ in (8) can be improved by probably designing the sensing matrix $\overline{\boldsymbol{A}} = (\boldsymbol{F}^T\overline{\boldsymbol{A}}_{\mathrm{T}} \otimes \boldsymbol{W}^T\overline{\boldsymbol{A}}_{\mathrm{R}})$, as in [36], [37], which is out of the scope of this letter.

Authorized licensed use limited to: University of Technology Sydney. Downloaded on May 25,2021 at 07:01:18 UTC from IEEE Xplore. Restrictions apply.

TABLE I
TRAINING OVERHEAD AND COMPUTATIONAL COMPLEXITY ANALYSIS

Method	Training overhead	Computational complexity
TRICE-BES	$K_\mathrm{S} \geq L \geq 4$, $N_\mathrm{R} \geq L_\mathrm{R}+1$, $K_\mathrm{T} \geq L_\mathrm{T}+1$, $(K_\mathrm{T}-1)N_\mathrm{R} \geq L$, $(N_\mathrm{R}-1)K_\mathrm{T} \geq L$	$\mathcal{O}\big((N_\mathrm{R}K_\mathrm{T})^2K_\mathrm{S}+K_\mathrm{S}^3+3L^3+L\big)$
TRICE-CS	$N_\mathrm{R}K_\mathrm{T} \approx \mathcal{O}(L\log(\bar{L}_\mathrm{R}\bar{L}_\mathrm{T}/L))$, $K_\mathrm{S} \approx \mathcal{O}(\log(\bar{L}_\mathrm{S}^\mathrm{v}\bar{L}_\mathrm{S}^\mathrm{h}))$	$\mathcal{O}(L(N_\mathrm{R}K_\mathrm{T}(\bar{L}_\mathrm{T}\bar{L}_\mathrm{R}+L+L^2))+2L^3+LK_\mathrm{S}\bar{L}_\mathrm{S}^\mathrm{v}\bar{L}_\mathrm{S}^\mathrm{h})$
Joint-CS [24]	$N_\mathrm{R}K_\mathrm{T}K_\mathrm{S} \approx \mathcal{O}(L\log(\bar{L}_\mathrm{S}^\mathrm{v}\bar{L}_\mathrm{S}^\mathrm{h}\bar{L}_\mathrm{R}\bar{L}_\mathrm{T}/L))$	$\mathcal{O}(L(N_\mathrm{R}K_\mathrm{T}K_\mathrm{S}(\bar{L}_\mathrm{S}^\mathrm{v}\bar{L}_\mathrm{S}^\mathrm{h}\bar{L}_\mathrm{T}\bar{L}_\mathrm{R}+L+L^2))+L^3)$

Algorithm 1: Two-Stage RIS-Aided MIMO Channel Estimation (TRICE).

Inputs: Measurement matrix $\boldsymbol{Y}$ in (5)
Stage 1: Get $\widehat{\boldsymbol{\psi}}_\mathrm{T}$, $\widehat{\boldsymbol{\psi}}_\mathrm{R}$ using, e.g., OMP or method in [26]
Stage 2: Assuming knowledge of $\widehat{\boldsymbol{\psi}}_\mathrm{T}$ and $\widehat{\boldsymbol{\psi}}_\mathrm{R}$ **do**
Get $\underline{\boldsymbol{Y}}^T = [\underline{\boldsymbol{y}}_1, \ldots, \underline{\boldsymbol{y}}_L] \in \mathbb{C}^{M_\mathrm{S}\times L}$ from (9)
For $n = 1$ to L
  Get $\hat{\mu}_n^\mathrm{h}$ and $\hat{\mu}_n^\mathrm{v}$ using, e.g., OMP or method in [26]
  Get nth diagonal entry of $\widehat{\boldsymbol{G}}$, i.e., $\hat{\alpha}_n$ using (14)
EndFor
Construct $\widehat{\boldsymbol{H}} = (\widehat{\boldsymbol{A}}_\mathrm{T} \otimes \widehat{\boldsymbol{A}}_\mathrm{R})\widehat{\boldsymbol{G}}\widehat{\boldsymbol{B}}$ (according to (6))
Estimate $\widehat{\boldsymbol{H}}_\mathrm{T}$ and $\widehat{\boldsymbol{H}}_\mathrm{R}$ from $\widehat{\boldsymbol{H}}$ using [11, Algorithm 1]

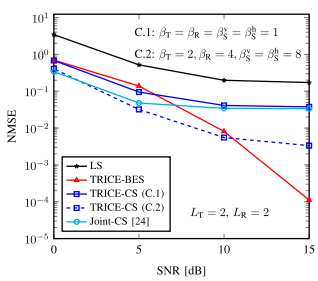

Fig. 2. NMSE vs. SNR. $N_\mathrm{R}=8$, $K_\mathrm{T}=8$, and $K_\mathrm{S}=16[4\times 4]$.

IV. NUMERICAL RESULTS

In this section, we show simulation results assuming that the TRICE framework employs, at both stages, (i) the 2D DFT-beamspace ESPRIT method from [26], denoted as TRICE-BES, and (ii) the on-grid CS method, denoted as TRICE-CS. For comparison, we also included the simulation results of the on-grid CS method proposed in [24], denoted as Joint-CS. For the CS-based methods, the estimation is performed using the classical OMP technique [34]. To comply with the DFT-beamspace ESPRIT method requirements as discussed in [26, Lemma 1], we assume that $\psi_{\mathrm{R},\ell} \sim \mathcal{U}(0, 2\pi(N_\mathrm{R}-1)/M_\mathrm{R})$, $\psi_{\mathrm{T},\ell} \sim \mathcal{U}(0, 2\pi(K_\mathrm{T}-1)/M_\mathrm{T})$, $\mu_\ell^\mathrm{h} \sim \mathcal{U}(0, 2\pi(K_\mathrm{S}^\mathrm{h}-1)/M_\mathrm{S}^\mathrm{h})$, $\mu_\ell^\mathrm{v} \sim \mathcal{U}(0, 2\pi(K_\mathrm{S}^\mathrm{v}-1)/M_\mathrm{S}^\mathrm{v})$, and the training matrices are chosen as $\boldsymbol{W}^T = [\boldsymbol{U}_{M_\mathrm{R}}]_{[1:N_\mathrm{R},:]}$, $\boldsymbol{F}^T = [\boldsymbol{U}_{M_\mathrm{T}}]_{[1:K_\mathrm{T},:]}$, $\boldsymbol{Q}_\mathrm{h}^T = [\boldsymbol{U}_{M_\mathrm{S}^\mathrm{h}}]_{[1:K_\mathrm{S}^\mathrm{h},:]}$, and $\boldsymbol{Q}_\mathrm{v}^T = [\boldsymbol{U}_{M_\mathrm{S}^\mathrm{v}}]_{[1:K_\mathrm{S}^\mathrm{v},:]}$, where $\boldsymbol{U}_M$ denotes the normalized $M\times M$ DFT-matrix. Moreover, we assume that $\alpha_\ell \sim \mathcal{CN}(0,1)$ and define the SNR $= \mathbb{E}\{\|\boldsymbol{Y}-\boldsymbol{Z}\|_\mathrm{F}^2/\|\boldsymbol{Z}\|_\mathrm{F}^2\}$ and the NMSE $= \mathbb{E}[\|\boldsymbol{H}-\widehat{\boldsymbol{H}}\|_\mathrm{F}^2/\|\boldsymbol{H}\|_\mathrm{F}^2]$. Table I summarizes the training overhead and the complexity of the simulated algorithms [39]. Note that the major difference between TRICE-CS and Joint-CS is that the former decouples the channel parameter estimation into two stages, while the latter jointly estimates them. Therefore, TRICE-CS requires a 2D dictionary in every stage, while Joint-CS requires a single 4D dictionary.

In Figs. 2, 3, and 4, we assume that $M_\mathrm{T}=64$, $M_\mathrm{R}=32$, and $M_\mathrm{S}=256[16\times 16]$. The 4D dictionary for Joint-CS is formed by using $64\times 32\times 16\times 16$ grid points, i.e., it has 524 288 atoms. On the other hand, for TRICE-CS, the first stage 2D dictionary is formed by using $\beta_\mathrm{T}M_\mathrm{T}\times\beta_\mathrm{R}M_\mathrm{R}$ grid points ($\bar{L}_\mathrm{T}=\beta_\mathrm{T}M_\mathrm{T}$, $\bar{L}_\mathrm{R}=\beta_\mathrm{R}M_\mathrm{R}$), while the second stage 2D dictionary is formed by using $\beta_\mathrm{S}^\mathrm{v}M_\mathrm{S}^\mathrm{v}\times\beta_\mathrm{S}^\mathrm{h}M_\mathrm{S}^\mathrm{h}$ grid points ($\bar{L}_\mathrm{S}^\mathrm{v}=\beta_\mathrm{S}^\mathrm{v}M_\mathrm{S}^\mathrm{v}$, $\bar{L}_\mathrm{S}^\mathrm{h}=\beta_\mathrm{S}^\mathrm{h}M_\mathrm{S}^\mathrm{h}$), where $\{\beta_\mathrm{T},\beta_\mathrm{R},\beta_\mathrm{S}^\mathrm{v},\beta_\mathrm{S}^\mathrm{h}\}\in\{1,2,\ldots\}$.

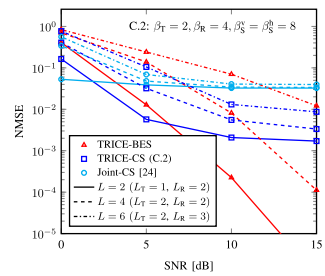

Fig. 3. NMSE vs. SNR. $N_\mathrm{R}=8$, $K_\mathrm{T}=8$, and $K_\mathrm{S}=16[4\times 4]$.

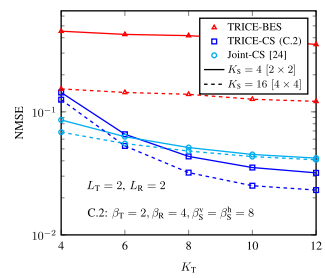

Fig. 4. NMSE vs. K_T and K_S. SNR $=5$ dB.

From Fig. 2, in case of C.1, we can see that TRICE-CS approaches the Joint-CS performance as the SNR increases, since in this case both methods have the same grid resolution, while Joint-CS outperforms TRICE-CS in the low SNR regime, due to its joint estimation. However, by increasing the grid resolutions as in C.2, TRICE-CS outperforms Joint-CS even in the low SNR regime. Note that, using the C.2 case, TRICE-CS has a much lower complexity when compared to Joint-CS, since it has $\approx 94\%$ less atoms. By its turn, TRICE-BES has a good performance in the medium and the high SNR regimes, where Fig. 3 shows that TRICE-BES provides a satisfactory performance in case of very sparse channels, cf. the $L=2$ case. Further, Fig. 4 shows that the estimation accuracy can be improved by increasing K_T and/or K_S.

V. CONCLUSIONS

The proposed TRICE framework is a two-stage channel parameter estimation scheme for single-user RIS-aided MIMO mmWave systems. By exploiting the low-rank nature of mmWave channels and by decoupling the channel parameter estimation problem into two stages, we have shown that TRICE not only has a high estimation performance, but also affords a low training overhead and has a low computational complexity, which makes it appealing in practical applications.

Authorized licensed use limited to: University of Technology Sydney. Downloaded on May 25,2021 at 07:01:18 UTC from IEEE Xplore. Restrictions apply.

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, Sep. 2018.

[2] C. Huang *et al.*, "Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges, and trends," *IEEE Wireless Commun.*, vol. 27, no. 5, pp. 118–125, Oct. 2020.

[3] M. Di Renzo *et al.*, "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, Nov. 2020.

[4] Ö. Özdogan, E. Björnson, and E. G. Larsson, "Intelligent reflecting surfaces: Physics, propagation, and pathloss modeling," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 581–585, May 2020.

[5] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, 2019.

[6] S. Gong *et al.*, "Towards smart wireless communications via intelligent reflecting surfaces: A contemporary survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2283–2314, Jun. 2020.

[7] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface aided wireless communications: A tutorial," *IEEE Trans. Wireless Commun.*, early access, Jan. 18, 2021, doi: 10.1109/TCOMM.2021.3051897.

[8] D. Mishra and H. Johansson, "Channel estimation and low-complexity beamforming design for passive intelligent surface assisted MISO wireless energy transfer," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2019, pp. 4659–4663.

[9] T. L. Jensen and E. De Carvalho, "An optimal channel estimation scheme for intelligent reflecting surfaces based on a minimum variance unbiased estimator," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2020, pp. 5000–5004.

[10] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization," *IEEE Wireless Commun. Lett.*, vol. 9, no. 4, pp. 518–522, Apr. 2020.

[11] G. T. de Araújo and A. L. F. de Almeida, "PARAFAC-based channel estimation for intelligent reflective surface assisted MIMO system," in *Proc. IEEE 11th Sensor Array Multichannel Signal Process. Workshop*, Jan. 2020, pp. 1–5.

[12] Q. Nadeem, H. Alwazani, A. Kammoun, A. Chaaban, M. Debbah, and M. Alouini, "Intelligent reflecting surface-assisted multi-user MISO communication: Channel estimation and beamforming design," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 661–680, 2020.

[13] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

[14] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of millimeter wave communications for fifth-generation (5G) wireless networks-with a focus on propagation models," *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6213–6230, Dec. 2017.

[15] K. Ardah, G. Fodor, Y. C. B. Silva, W. C. Freitas, and A. L. F. de Almeida, "Hybrid analog-digital beamforming design for SE and EE maximization in massive MIMO networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 377–389, Jan. 2020.

[16] S. Gherekhloo, K. Ardah, and M. Haardt, "Hybrid beamforming design for downlink MU-MIMO-OFDM millimeter-wave systems," in *Proc. IEEE 11th Sensor Array Multichannel Signal Process. Workshop*, Jun. 2020, pp. 1–5.

[17] J. Zhang, A. Wiesel, and M. Haardt, "Low rank approximation based hybrid precoding schemes for multi-carrier single-user massive MIMO systems," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2016, pp. 3281–3285.

[18] K. Ardah, G. Fodor, Y. C. B. Silva, W. C. Freitas, and F. R. P. Cavalcanti, "A unifying design of hybrid beamforming architectures employing phase shifters or switches," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 11 243–11 247, Nov. 2018.

[19] A. Taha, M. Alrabeiah, and A. Alkhateeb, "Enabling large intelligent surfaces with compressive sensing and deep learning," Apr. 2019, *arXiv:1904.10136*.

[20] Z. Wan, Z. Gao, and M. Alouini, "Broadband channel estimation for intelligent reflecting surface aided mmwave massive MIMO systems," in *Proc. IEEE Int. Conf. Commun.*, Feb. 2020, pp. 1–6.

[21] J. He, M. Leinonen, H. Wymeersch, and M. Juntti, "Channel estimation for RIS-aided mmwave MIMO channels," in *Proc. IEEE Global Commun. Conf.*, Dec. 2020, pp. 1–6.

[22] Z. He and X. Yuan, "Channel estimation for RIS-aided mmwave MIMO systems," in *Proc. IEEE Global Commun. Conf.*, Dec. 2020, pp. 1–6.

[23] J. Chen, Y.-C. Liang, H. V. Cheng, and W. Yu, "Channel estimation for reconfigurable intelligent surface aided multi-user MIMO systems," Dec. 2019, *arXiv:1912.03619*.

[24] P. Wang, J. Fang, H. Duan, and H. Li, "Compressed channel estimation for intelligent reflecting surface-assisted millimeter wave systems," *IEEE Signal Process. Lett.*, vol. 27, pp. 905–909, 2020.

[25] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[26] J. Zhang and M. Haardt, "Channel estimation and training design for hybrid multi-carrier mmwave massive MIMO systems: The beamspace ESPRIT approach," in *Proc. 25th Eur. Signal Process. Conf.*, Sep. 2017, pp. 385–389.

[27] J. Zhang and M. Haardt, "Channel estimation for hybrid multi-carrier mmwave MIMO systems using three-dimensional unitary ESPRIT in DFT beamspace," in *Proc. IEEE 7th Int. Workshop Comput. Adv. Multi-Sensor Adaptive Process.*, Dec. 2017, pp. 1–5.

[28] C. Steffens, M. Pesavento, and M. E. Pfetsch, "A compact formulation for the $\ell_{2,1}$ mixed-norm minimization problem," *IEEE Trans. Signal Process.*, vol. 66, no. 6, pp. 1483–1497, Mar. 2018.

[29] K. Ardah, A. L. F. de Almeida, and M. Haardt, "A gridless CS approach for channel estimation in hybrid massive MIMO systems," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2019, pp. 4160–4164.

[30] B. Mamandipoor, D. Ramasamy, and U. Madhow, "Newtonized orthogonal matching pursuit: Frequency estimation over the continuum," *IEEE Trans. Signal Process.*, vol. 64, no. 19, pp. 5066–5081, Jun. 2016.

[31] M. Cao, X. Mao, X. Long, and L. Huang, "Direction-of-arrival estimation for uniform rectangular array: A multilinear projection approach," in *Proc. 26th Eur. Signal Process. Conf.*, Sep. 2018, pp. 1237–1241.

[32] A. A. M. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE J. Sel. Areas Commun.*, vol. 5, no. 2, pp. 128–137, Feb. 1987.

[33] M. Haardt and J. A. Nossek, "Simultaneous schur decomposition of several nonsymmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems," *IEEE Trans. Signal Process.*, vol. 46, no. 1, pp. 161–169, Jan. 1998.

[34] B. L. Sturm and M. G. Christensen, "Comparison of orthogonal matching pursuit implementations," in *Proc. of the 20th Eur. Signal Process. Conf.*, Aug. 2012, pp. 220–224.

[35] H. Reboredo, F. Renna, R. Calderbank, and M. R. D. Rodrigues, "Bounds on the number of measurements for reliable compressive classification," *IEEE Trans. Signal Process.*, vol. 64, no. 22, pp. 5778–5793, Nov. 2016.

[36] K. Ardah, B. Sokal, A. L. F. de Almeida, and M. Haardt, "Compressed sensing based channel estimation and open-loop training design for hybrid analog-digital massive MIMO systems," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2020, pp. 4597–4601.

[37] K. Ardah, M. Pesavento, and M. Haardt, "A novel sensing matrix design for compressed sensing via mutual coherence minimization," in *Proc. IEEE 8th Int. Workshop Comput. Adv. Multi-Sensor Adaptive Process.*, Dec. 2019, pp. 66–70.

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

[39] G. H. Golub and C. F. van Loan, *Matrix Computations*, 4th ed. Baltimore, MD, USA: Johns Hopkins Univ. Press, Feb. 2013.

Authorized licensed use limited to: University of Technology Sydney. Downloaded on May 25,2021 at 07:01:18 UTC from IEEE Xplore. Restrictions apply.