Two-Stage Channel Estimation for Hybrid RIS Assisted MIMO Systems

Rafaela Schroeder, *Graduate Student Member, IEEE*, Jiguang He, *Member, IEEE*, Glauber Brante, *Senior Member, IEEE*, and Markku Juntti, *Fellow, IEEE*

***Abstract*— Reconfigurable intelligent surfaces (RISs) have been proposed as a key enabler to improve the coverage of the signals and mitigate the frequent blockages in millimeter wave (mmWave) multiple-input multiple-output (MIMO) communications. However, the channel state information (CSI) acquisition is one of the major challenges for the practical deployment of the RIS. The passive RIS without any baseband processing capabilities brings difficulty on the channel estimation (CE), since the individual channels or the cascaded one can be estimated only at base station (BS) via uplink training or mobile station (MS) via downlink training. In order to facilitate the CSI acquisition, we focus on the hybrid RIS architecture, where a small number of elements are active and able to receive and process the pilot signals at the RIS. The CE is performed in two stages by following the atomic norm minimization to recover the channel parameters, i.e., angles of departure (AoDs), angles of arrival (AoAs), and propagation path gains. Simulation results show that the proposed scheme can outperform the passive RIS CE under the same training overhead. Furthermore, we also study the theoretical performance limits in terms of mean square error (MSE) via Cramér-Rao lower bound (CRLB) analyses.**

***Index Terms*— Channel estimation, reconfigurable intelligent surface, Cramér-Rao lower bound, mmWave MIMO, hybrid RIS.**

I. Introduction

MILLIMETER-WAVE (mmWave) multiple-input multiple-output (MIMO) systems are regarded as an essential technology for the fifth generation (5G) wireless networks [1], [2]. In order to compensate for the path loss effect, large antenna arrays are required at both the transmitter and receiver. The features of the channel model in mmWave bands are tightly associated with the high frequencies, which lead to frequent blockage [3], inevitable high path loss, and inherent channel sparsity. Nevertheless, mmWave MIMO systems heavily rely on line-of-sight (LoS) to guarantee sufficient received power and spectral efficiency (SE) [4].

Reconfigurable intelligent surfaces (RISs) have been proposed as a promising solution to maintain the coverage of the signals and resolve the blockage problem in mmWave MIMO systems [5]–[10]. The RIS commonly consists of an array with low-cost discrete phase shifters. By adjusting these phase shifters, the RIS can modify their signal response to achieve a certain objective, for example, focusing the signal towards the receiver [11], [12]. The introduction of the RIS can bring other benefits, such as improved physical layer security [13], [14] and enhanced SE [5]. In addition, the RIS can also increase the accuracy of indoor and outdoor localization [15], [16].

To enable the optimal or suboptimal joint design of the beamforming vectors at the base station (BS) and mobile station (MS) [17] and RIS phase control matrix, nearly perfect channel state information (CSI) is of great essence for RIS-aided mmWave MIMO systems. The CSI generally includes the individual RIS-BS and MS-RIS channels, cascaded channel, or corresponding channel parameters. However, the CSI acquisition is already a challenging task in mmWave MIMO systems, and even more complicated in RIS-aided scenarios. For instance, in the *passive* RIS without any baseband processing capabilities, the CE can be done only at the BS or MS. Various CE methods for passive RIS have been investigated in [6], [11], [18]–[23]. In [23], Zheng and Zhang proposed a CE procedure for RIS, which is done by considering that all the elements are switched ON during the pilot training. However, despite the improvement on the performance compared to ON/OFF scheme, this method still requires a large training overhead. In our recent work [18], we proposed a two-stage CE via atomic norm minimization (ANM) for the passive RIS architecture. At the first CE stage, we aim to estimate the angles of arrival (AoAs) at the MS and angles of departure (AoDs) at the BS. Based on these estimates, we design the beam training matrix and the combing matrix for the second stage sounding. At the second CE stage, we target at recovering the remaining channel parameters, i.e., the angle difference associated with the RIS and the products of path gains from the received pilot signals. Note that we keep the beam training matrix and the combining matrix fixed, while we change the RIS phase control matrix during the second stage sounding. With this method, we can further reduce the training

Manuscript received 16 November 2021; revised 28 March 2022; accepted 11 May 2022. Date of publication 20 May 2022; date of current version 15 July 2022. This work has been financially supported in part by the Academy of Finland (EERA project), European Union's Horizon 2020 Framework Programme for Research and Innovation (ARIADNE project, under grant agreement no. 871464), Academy of Finland 6Genesis Flagship (grant 318927), CAPES (Finance Code 001), and CNPq Brazil. The associate editor coordinating the review of this article and approving it for publication was S. Jin. *(Corresponding author: Rafaela Schroeder.)*

Rafaela Schroeder and Markku Juntti are with the Centre for Wireless Communications-Radio Technologies, University of Oulu, 90014 Oulu, Finland (e-mail: rafaela.schroeder@oulu.fi; markku.juntti@oulu.fi).

Jiguang He is with the Technology Innovation Institute, Abu Dhabi, United Arab Emirates, and also with the Centre for Wireless Communications, University of Oulu, 90014 Oulu, Finland (e-mail: jiguang.he@tii.ae).

Glauber Brante is with the Graduate Program in Electrical and Computer Engineering, Federal University of Technology—Paraná, Curitiba 80230901, Brazil (e-mail: gbrante@utfpr.edu.br).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCOMM.2022.3176654.

Digital Object Identifier 10.1109/TCOMM.2022.3176654

overhead while obtaining a super-resolution estimation of the channel parameters.

Alternatively, the *hybrid* RIS with a mix of both *passive* and *active* elements has been proposed by Taha *et al.* in [24] to reduce the difficulty on CSI acquisition. The active elements are introduced to sense the received signals. With this assumption, the estimation of the entire BS-RIS-MS channel can be divided into two point-to-point (P2P) MIMO CE subproblems, which simplifies the CSI acquisition at the sacrifice of higher power consumption at the RIS. Motivated by the pioneering work in [24], the hybrid architecture has been intensively studied in [25]–[30]. For instance, Alexandropoulos *et al.* in [29] described the hardware design of the hybrid RIS and presented a full-wave proof-of-concept of the application. In our previous work [26], we compared the CE performance between the passive and hybrid RIS architectures for mmWave MIMO systems considering the two-way uplink and downlink training. Interestingly, we noticed that the passive RIS CE can outperform the hybrid RIS CE in the particular scenario of [26], where CE was totally performed at the RIS using a two-way training. Nevertheless, in order to exploit the benefits of the deployment of the hybrid RIS, i.e., mitigation of the CSI acquisition, we keep investigating methods that can further improve CE, outperforming the passive RIS.

For the reasons discussed above, we focus on the hybrid/semi-passive[1] RIS CE and extend our previous work [30] in a more comprehensive manner in this paper. In particular, in the current work we develop in detail the theoretical limit analysis via Cramér-Rao lower bound (CRLB) analysis including the combining matrices, set of active elements and the number of training sequences on the estimation performance. In addition, in the current paper we also discuss the location of the RIS and its impact on the CE performance.

By adopting the structured channel training, we develop a two-stage CE for hybrid RIS assisted mmWave MIMO systems. Different from our work in [26], we restrict the training to only one way via the uplink transmission. Despite the fact that we used the hybrid RIS architecture also in [26], the CE formulations in this paper are new. For example, in [26], the CE was performed only at the RIS with a large number of active elements. Herein, we develop the CE by taking into consideration two received signals, one at RIS and the other at BS. We also reduce the number of active RIS elements tailored for CE.

In the training procedure, the MS sends pilot signals to both the RIS and the BS. The signals are received at the RIS by the active elements, while the remaining passive elements reflect the signal towards the BS. At the first CE stage, we target at the recovery of the channel parameters in the MS-RIS channel via the ANM based on the received signals at the RIS. Subsequently, the estimates of the channel parameters are transmitted to the BS via an error-free backhaul link. Thus, the BS can reconstruct the MS-RIS channel matrix based on the feedback. In order to simplify the CE and to reduce the training overhead, we use the reconstructed channel matrix at the second CE stage, where we target at recovering the remaining channel parameters based on the received signals at the BS by resorting to the ANM. We evaluate the performance in terms of the mean square error (MSE) of the channel parameter estimation, NMSE of the channel matrices and SE. Our simulation results prove that our proposed CE method not only simplifies the CSI acquisition for RIS-aided mmWave MIMO systems with the aid of a small number of active RIS elements but also brings better CE performance than the passive RIS [18] and two-way uplink downlink training [26]. The **contributions** of this paper are summarized as follows:

- We propose a two-stage CE procedure for the hybrid RIS-assisted mmWave MIMO systems based on ANM. Since we adopt the hybrid RIS architecture, we can decouple the CE problems as to simplify the CSI acquisition.
- We provide a theoretical analysis in terms of the Cramér-Rao lower bound, which serves as the theoretical benchmark for our proposed CE method.
- We study the effect of the RIS location and path loss on the performance of hybrid RIS CE, where the BS and MS locations are fixed while the RIS location varies. The study can be used as a guidance for the RIS deployment in the current and upcoming wireless systems.

The rest of the paper is organized as follows. Section II describes the channel model, while Section III provides the sounding procedure for the RIS architectures. The proposed CE is detailed in Section IV, followed by the design of RIS phase control matrix and BS/MS beamforming vectors in the Section V. Furthermore, the CRLB is studied in the Section VI. The performance evaluation, metrics and simulation results are provided in Section VII with conclusions drawn in Section VIII.

Notation: A bold capital letter $\mathbf{A}$ denotes a matrix and a lowercase letter $\mathbf{a}$ denotes a column vector, while $()^{\mathsf{H}}$, $()^{*}$, and $()^{\mathsf{T}}$ denote the Hermitian transpose, conjugate, and transpose, respectively. $[\mathbf{a}]_m$ is the mth element of $\mathbf{a}$, $[\mathbf{A}]_{:,n}$ and $[\mathbf{A}]_{m,n}$ are the nth column and the (m,n)th entry of $\mathbf{A}$, respectively. $\otimes$ denotes the Kronecker product, $\odot$ is the Khatri-Rao product, $\mathrm{vec}(\mathbf{A})$ is the vectorization of $\mathbf{A}$, $\mathrm{diag}(\mathbf{a})$ being a square diagonal matrix with entries of $\mathbf{a}$ on its diagonal, $\|.\|_{\mathrm{F}}$ is the Frobenius norm, $\mathrm{Tr}(\mathbf{A})$ is the sum value of the diagonal elements of $\mathbf{A}$, $\mathbb{T}(\mathbf{A})$ denotes the block Toeplitz matrix constructed from the vectorized form of $\mathbf{A}$, i.e., $\mathrm{vec}(\mathbf{A})$, being its first row. $(\cdot)\dagger$ denotes the Moore–Penrose inverse, $\mathcal{A}$ is the atomic set, $\mathrm{conv}(\mathcal{A})$ denotes the convex hull of $\mathcal{A}$, and $\mathbb{E}$ is the expectation operator.

II. Channel Model

We consider the hybrid RIS-assisted mmWave MIMO systems, which include one multi-antenna BS, one hybrid multi-element RIS, and one multi-antenna MS. The number of antennas at the BS and the MS are denoted by N_{B} and N_{M}, respectively, and the number of elements at the hybrid RIS is N_{R}. We assume that the direct MS-BS channel is obstructed (e.g., due to blockage), which leads to the demand on the deployment of the RIS to maintain the connectivity between

[1]In the sequel, we use hybrid RIS to represent semi-passive/hybrid RIS for notation simplicity.

the BS and MS.[2] We adopt an uniform linear array (ULA) for the antennas/elements, but the extension to an uniform planar array (UPA) is feasible. Moreover, we focus on a CE proposal requiring only uplink training procedures.

The propagation channels consist of two tandem channels, i.e., MS-RIS and RIS-BS channels, denoted by $\mathbf{H}_{\mathrm{M,R}} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{M}}}$ and $\mathbf{H}_{\mathrm{R,B}} \in \mathbb{C}^{N_{\mathrm{B}} \times N_{\mathrm{R}}}$, respectively. We adopt a block-fading channel, which means that $\mathbf{H}_{\mathrm{M,R}}$ and $\mathbf{H}_{\mathrm{R,B}}$ stay constant during a certain period of time, known as coherence time. By following the geometric channel model, we write $\mathbf{H}_{\mathrm{M,R}}$ as

$$\begin{aligned}\mathbf{H}_{\mathrm{M,R}} &= \sum_{l=1}^{L_{\mathrm{M,R}}} [\boldsymbol{\rho}_{\mathrm{M,R}}]_l \boldsymbol{\alpha}([\boldsymbol{\phi}_{\mathrm{M,R}}]_l) \boldsymbol{\alpha}^{\mathsf{H}}([\boldsymbol{\theta}_{\mathrm{M,R}}]_l), \\ &= \mathbf{A}(\boldsymbol{\phi}_{\mathrm{M,R}}) \mathrm{diag}(\boldsymbol{\rho}_{\mathrm{M,R}}) \mathbf{A}^{\mathsf{H}}(\boldsymbol{\theta}_{\mathrm{M,R}}), \end{aligned} \tag{1}$$

where $[\boldsymbol{\rho}_{\mathrm{M,R}}]_l$ is the lth propagation path gain, $L_{\mathrm{M,R}}$ is the number of paths, $\boldsymbol{\theta}_{\mathrm{M,R}}$ and $\boldsymbol{\phi}_{\mathrm{M,R}}$, are the AoDs and AoAs of the channel, respectively. Finally, $\boldsymbol{\alpha}([\boldsymbol{\phi}_{\mathrm{M,R}}]_l)$ and $\boldsymbol{\alpha}([\boldsymbol{\theta}_{\mathrm{M,R}}]_l)$ are the array response vectors as a function of $[\boldsymbol{\phi}_{\mathrm{M,R}}]_l$ and $[\boldsymbol{\theta}_{\mathrm{M,R}}]_l$. Considering half-wavelength inter-antenna element spacing, the array response vectors are $[\boldsymbol{\alpha}([\boldsymbol{\phi}_{\mathrm{M,R}}]_l)]_n = \exp\{j\pi(n-1)\sin([\boldsymbol{\phi}_{\mathrm{M,R}}]_l)\}$, for $n = 1, \cdots, N_{\mathrm{R}}$ and $[\boldsymbol{\alpha}([\boldsymbol{\theta}_{\mathrm{M,R}}]_l)]_n = \exp\{j\pi(n-1)\sin([\boldsymbol{\theta}_{\mathrm{M,R}}]_l)\}$, for $n = 1, \cdots, N_{\mathrm{M}}$, and $j = \sqrt{-1}$. Similarly, the array response matrices $\mathbf{A}(\boldsymbol{\theta}_{\mathrm{M,R}})$ and $\mathbf{A}(\boldsymbol{\phi}_{\mathrm{M,R}})$ are formulated as

$$\mathbf{A}(\boldsymbol{\theta}_{\mathrm{M,R}}) = \Big[\boldsymbol{\alpha}([\boldsymbol{\theta}_{\mathrm{M,R}}]_1), \ldots, \boldsymbol{\alpha}([\boldsymbol{\phi}_{\mathrm{M,R}}]_{L_{\mathrm{M,R}}})\Big], \tag{2}$$

$$\mathbf{A}(\boldsymbol{\phi}_{\mathrm{M,R}}) = \Big[\boldsymbol{\alpha}([\boldsymbol{\phi}_{\mathrm{M,R}}]_1), \ldots, \boldsymbol{\alpha}([\boldsymbol{\phi}_{\mathrm{M,R}}]_{L_{\mathrm{M,R}}})\Big]. \tag{3}$$

Similarly, the RIS-BS channel $\mathbf{H}_{\mathrm{R,B}}$ is

$$\begin{aligned}\mathbf{H}_{\mathrm{R,B}} &= \sum_{l=1}^{L_{\mathrm{R,B}}} [\boldsymbol{\rho}_{\mathrm{R,B}}]_l \boldsymbol{\alpha}([\boldsymbol{\phi}_{\mathrm{R,B}}]_l) \boldsymbol{\alpha}^{\mathsf{H}}([\boldsymbol{\theta}_{\mathrm{R,B}}]_l), \\ &= \mathbf{A}(\boldsymbol{\phi}_{\mathrm{R,B}}) \mathrm{diag}(\boldsymbol{\rho}_{\mathrm{R,B}}) \mathbf{A}^{\mathsf{H}}(\boldsymbol{\theta}_{\mathrm{R,B}}), \end{aligned} \tag{4}$$

where $[\boldsymbol{\rho}_{\mathrm{R,B}}]_l$, $\boldsymbol{\alpha}([\boldsymbol{\phi}_{\mathrm{R,B}}]_l)$, and $\boldsymbol{\alpha}([\boldsymbol{\theta}_{\mathrm{R,B}}]_l)$ are the lth propagation path gain, and array response vectors as a function of $[\boldsymbol{\phi}_{\mathrm{R,B}}]_l$ and $[\boldsymbol{\theta}_{\mathrm{R,B}}]_l$, respectively. By applying (1) and (4), and taking RIS into consideration, we can express the complete end-to-end MS-RIS-BS channel as

$$\mathbf{H} = \mathbf{H}_{\mathrm{R,B}} \boldsymbol{\Omega} \mathbf{H}_{\mathrm{M,R}}, \tag{5}$$

where $\boldsymbol{\Omega} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{R}}}$ is the phase control matrix at the RIS with unit-modulus elements on the diagonal. By assuming that the RIS is composed by discrete phase shifters, the phase control matrix is $[\boldsymbol{\Omega}]_{k,k} = \exp(j\omega_k)$, where $\omega_k \in [0, 2\pi)$. Moreover, we the effective channel $\mathbf{G}$ is

$$\mathbf{G} = \mathrm{diag}(\boldsymbol{\rho}_{\mathrm{R,B}}) \mathbf{A}^{\mathsf{H}}(\boldsymbol{\theta}_{\mathrm{R,B}}) \boldsymbol{\Omega} \mathbf{A}(\boldsymbol{\phi}_{\mathrm{M,R}}) \mathrm{diag}(\boldsymbol{\rho}_{\mathrm{M,R}}), \tag{6}$$

which depends on the phase control matrix and angular parameters associated with the RIS and propagation path gains.

[2]The proposed algorithm can be further extended by taking into account the direct MS-BS channel. In such a case, we first estimate the direct MS-BS channel by setting the RIS in the absorption mode. After that, we can follow the proposed two-stage CE algorithm for estimating the remaining channels.

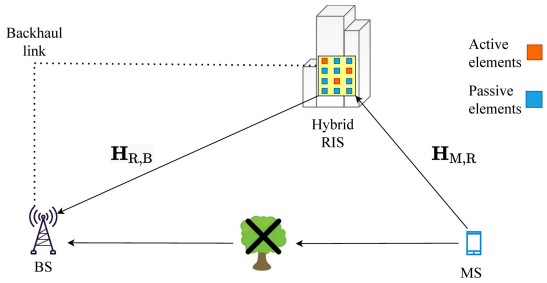


Fig. 1. Hybrid RIS with both passive and active elements.

III. Sounding Procedure for RIS Architectures

In this section, we detail the sounding procedure for the passive RIS architecture, used for comparison purposes, and the hybrid RIS, which is composed of both passive and active elements, as illustrated in Fig. 1.

A. Passive RIS Architecture

In the literature, the CE for the passive RIS architecture has been conducted in [6], [11], [19]. In this case, the CE of the individual channels or the cascaded one as in (5) can be performed only at the BS or MS. We aim at estimating the channel parameters in (1) and (4) other than the whole channel matrices by taking into consideration the inherent channel sparsity.

We assume block fading channel, so that we can divide the coherence time into two sub-intervals, the first one for CE and the second one for data transmission (DT). We further divide the CE sub-interval into K blocks. We assume the uplink pilot-based training procedure, where the MS sends a series of training matrices $\mathbf{X}_k$ during $k = 1, \cdots, K$ to the BS. The signal is reflected at the RIS by $\boldsymbol{\Omega}_k$[3] and combined at the BS by $\mathbf{W}_k$, received as $\mathbf{Y}_{\mathrm{P}k}$. In the first block of CE, we target the extraction of the AoAs at BS and AoDs at MS. The training matrix is $\mathbf{X}_0 \in \mathbb{C}^{N_{\mathrm{B}} \times N_0}$, the combining matrix at BS is $\mathbf{W}_t \in \mathbb{C}^{N_{\mathrm{M}} \times M_0}$, where N_0 is the number of training beams at the first stage of CE and M_0 is the number of columns at the combining matrix. Also, we keep the phase control matrix fixed as $\boldsymbol{\Omega}_1 \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{R}}}$. Based on the estimates of the first stage, we design the beam matrix and the combining matrix for the second stage. Thus, we have $\mathbf{X}_t \in \mathbb{C}^{N_{\mathrm{B}} \times L_{\mathrm{R,B}}}$ and $\mathbf{W}_t \in \mathbb{C}^{N_{\mathrm{M}} \times L_{\mathrm{M,R}}}$. In the second stage, we target the recovery of the angle difference and the products of path gains. We keep the training matrix and the combining matrix fixed, while the phase control matrix varies from block to block. The received signal at the BS is

$$\begin{aligned}\mathbf{Y}_{\mathrm{P}k} = \sqrt{P_{\mathrm{T}}}\beta_2 \mathbf{W}_k^{\mathsf{H}} \mathbf{H}_{\mathrm{R,B}} \boldsymbol{\Omega}_k \mathbf{H}_{\mathrm{M,R}} \mathbf{X}_k \\ + \mathbf{W}_k^{\mathsf{H}} \mathbf{Z}_k, \quad \text{for } k = 1, \cdots, K, \end{aligned} \tag{7}$$

[3]During the sounding process, the phase control matrix, the training matrices, and the analog combining matrix are constructed with random phases due to the lack of prior information on the channel parameters. However, directional pilots can be considered if prior information on the channel parameters, especially angular parameters, is available, which is left as our future work.

where P_{T} is the transmit power, $\mathbf{Z}_k$ is the additive white Gaussian noise with each entry distributed as $\mathcal{CN}(0,\sigma^2)$, and the overall path loss β_2 is given by $\beta_2 = \sqrt{\frac{1}{\beta(d_1,d_2)}}$, where d_1 and d_2 are the distance between MS and RIS and that between RIS and BS, respectively. In other words, in the entire link MS-RIS-BS, the path loss depends on both distances d_1 and d_2. In particular, the CE for the passive RIS-aided mmWave MIMO systems is addressed using the two-stage procedure via ANM, detailed in [18], which is omitted here for brevity.

B. Hybrid RIS Architecture

Thanks to the availability of measurements/received signals at the active elements illustrated in Fig. 1, we can perform the CE at RIS in the hybrid architecture. After the CE, the estimates of the channel parameters are transmitted to the BS by using the backhaul link.[4] Furthermore, by adopting the hybrid RIS architecture, we can decouple the CE problem and further simplify the CSI acquisition [24]. We assume perfect time synchronization in the time domain between MS and BS when performing CE as it is common in the literature [1].

We assume M out of N_{R} RIS elements are active with *receiver* capability and $N_{\mathrm{RF,R}} = M$ *receive* radio frequency (RF) processing chains at the RIS. We conduct uplink pilot training, where the MS sends the beam training matrix $\mathbf{X} \in \mathbb{C}^{N_{\mathrm{M}} \times T}$ to the BS and the RIS, where T is the number of training sequences. For simplicity, we ignore possible impairments due to different characteristics in the RF chains at the RIS and at the BS. Similar to passive RIS sounding, we also divide the CE sub-interval into K blocks, where each block has T channel uses.

The pilot signals are received at RIS by the M active elements, indexed by the set $\mathbb{M}$, i.e., $|\mathbb{M}| = M$, while the remaining passive[5] elements only reflect the incident signal with the phase control matrix $\mathbf{\Omega}_k$. It is worth mentioning that $\mathbf{\Omega}_k$ differs from the phase control matrix in the passive RIS since $[\mathbf{\Omega}_k]_{i,i} = 0$ for $i \in \mathbb{M}$. After the reflection at the RIS, the signal is further combined at the BS by $\mathbf{W}_{\mathrm{B}} \in \mathbb{C}^{N_{\mathrm{B}} \times N_{\mathrm{C,B}}}$, and received as $\mathbf{Y}_k$, where $N_{\mathrm{C,B}}$ is the number of columns of the combining matrix. In the training procedure, we keep the combining matrix $\mathbf{W}_{\mathrm{B}}$ and the training matrix $\mathbf{X}$ constant over all the blocks, while the phase control matrix $\mathbf{\Omega}_k$ at the RIS varies from block to block. Fig. 2 summarizes the training procedure for the hybrid RIS, occurring prior to DT. The received signal at the RIS $\mathbf{Y}_{\mathrm{H}k}$ is expressed as

$$\mathbf{Y}_{\mathrm{H}k} = \sqrt{P_{\mathrm{T}}}\beta_1 \mathbf{W}_{\mathrm{H}k}\mathbf{H}_{\mathrm{M,R}}\mathbf{X} + \mathbf{W}_{\mathrm{H}k}\mathbf{Z}_{1k}, \quad \text{for } k = 1, \cdots, K, \tag{8}$$

where $\mathbf{W}_{\mathrm{H}k}$ is a row-selection matrix containing M rows of a $N_{\mathrm{R}} \times N_{\mathrm{R}}$ identity matrix, $\mathbf{Z}_{1k} \in \mathbb{C}^{M \times T}$ is the Gaussian noise with each entry distributed as $\mathcal{CN}(0,\sigma^2)$. The term β_1 is expressed as $\beta_1 = \sqrt{\frac{1}{\beta(d_1)}}$, where $\beta(d_1)$ is the path loss associated with the MS-RIS channel. Moreover, the received signal at BS is expressed as

$$\mathbf{Y}_k = \sqrt{P_{\mathrm{T}}}\beta_2 \mathbf{W}_{\mathrm{B}}^{\mathsf{H}}\mathbf{H}_{\mathrm{R,B}}\mathbf{\Omega}_k\mathbf{H}_{\mathrm{M,R}}\mathbf{X} + \mathbf{W}_{\mathrm{B}}^{\mathsf{H}}\mathbf{Z}_{2k}, \tag{9}$$

where $\mathbf{Z}_{2k} \in \mathbb{C}^{N_{\mathrm{B}} \times T}$ is the Gaussian noise also with each entry distributed as $\mathcal{CN}(0,\sigma^2)$. After K blocks, the complete received signal at the RIS is $\mathbf{Y}_{\mathrm{H}} = [\mathbf{Y}_{\mathrm{H}1}^{\mathsf{T}}, \cdots, \mathbf{Y}_{\mathrm{H}K}^{\mathsf{T}}]^{\mathsf{T}} \in \mathbb{C}^{MK \times T}$, which can be summarized as

$$\mathbf{Y}_{\mathrm{H}} = \sqrt{P_{\mathrm{T}}}\beta_1 \mathbf{W}_{\mathrm{H}}\mathbf{H}_{\mathrm{M,R}}\mathbf{X} + \bar{\mathbf{Z}}_1, \tag{10}$$

where $\mathbf{W}_{\mathrm{H}} = [\mathbf{W}_{\mathrm{H}1}^{\mathsf{T}}, \cdots, \mathbf{W}_{\mathrm{H}K}^{\mathsf{T}}]^{\mathsf{T}} \in \mathbb{C}^{MK \times N_{\mathrm{R}}}$, $\bar{\mathbf{Z}}_1 = [(\mathbf{W}_{\mathrm{H}1}\mathbf{Z}_{11})^{\mathsf{T}}, \cdots, (\mathbf{W}_{\mathrm{H}K}\mathbf{Z}_{1K})^{\mathsf{T}}]^{\mathsf{T}} \in \mathbb{C}^{MK \times T}$. The complete received signal at the BS $\mathbf{Y} = [\mathbf{Y}_1, \cdots, \mathbf{Y}_K] \in \mathbb{C}^{N_{\mathrm{C,B}} \times TK}$ is defined as

$$\mathbf{Y} = \sqrt{P_{\mathrm{T}}}\beta_2 \mathbf{W}_{\mathrm{B}}^{\mathsf{H}}\mathbf{H}_{\mathrm{R,B}}\mathbf{U} + \bar{\mathbf{Z}}_2, \tag{11}$$

where $\mathbf{U} = [\mathbf{\Omega}_1\mathbf{H}_{\mathrm{M,R}}\mathbf{X}, \cdots, \mathbf{\Omega}_K\mathbf{H}_{\mathrm{M,R}}\mathbf{X}] \in \mathbb{C}^{N_{\mathrm{R}} \times TK}$ and $\bar{\mathbf{Z}}_2 = [\mathbf{W}_{\mathrm{B}}^{\mathsf{H}}\mathbf{Z}_{21}, \cdots, \mathbf{W}_{\mathrm{B}}^{\mathsf{H}}\mathbf{Z}_{2K}] \in \mathbb{C}^{N_{\mathrm{C,B}} \times TK}$.

We propose a two-stage CE to recover the channel parameters from the received signals in (10) and (11). At the first stage, we recover the channel parameters in $\mathbf{H}_{\mathrm{M,R}}$ based on the received signal at the RIS in (10). After the estimation of $\mathbf{H}_{\mathrm{M,R}}$, the RIS sends the estimates of the channel parameters to the BS via the backhaul link. The BS reconstructs the channel $\mathbf{H}_{\mathrm{M,R}}$ based on the estimates of the channel parameters, denoted as $\hat{\mathbf{H}}_{\mathrm{M,R}}$. At the second stage, we assume $\hat{\mathbf{H}}_{\mathrm{M,R}} \approx \mathbf{H}_{\mathrm{M,R}}$ and target at the recovery of the channel parameters in $\mathbf{H}_{\mathrm{R,B}}$ from the received signal at the BS in (11).

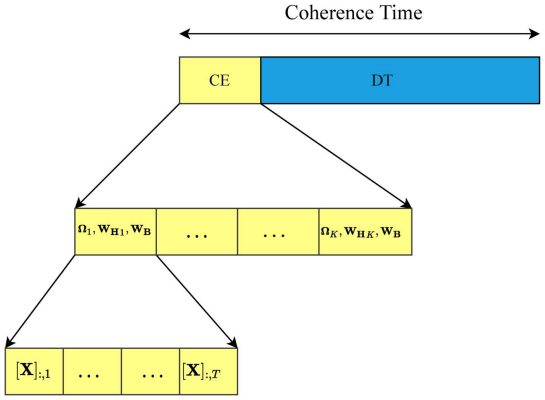


Fig. 2. Uplink training procedure for the hybrid RIS. We divide the coherence time in two sub-intervals, one for CE and the other for DT. We further divide the CE sub-interval into K blocks with each block containing T channel uses.

IV. RIS Channel Estimation

Channel estimation for mmWave MIMO systems has been investigated by applying compressive sensing (CS) methods. By leveraging the sparsity property of mmWave MIMO channels, the CS methods can reduce significantly the training overhead compared to least-squares (LS) [31] and minimum mean squared error (MMSE) [32] based methods. CS methods can be generally classified into two categories, i.e., on-the-grid methods [33] as well as more advanced off-the-grid methods, such as ANM [34], [35]. ANM is a well-known technique that can reconstruct the sparse signals without discretizing

[4]Alternatively, the RIS can also use the backhaul link to share its received signals to the BS in order to offload the computation to the BS. In this sense, we can keep a simple architecture for the RIS.

[5]Since the passive RIS elements are not connected to the RF chains and baseband processing units, they cause no interference to the received signal at the RIS.

the entire angle spaces for the angle of arrival and angle of departure into grids. By using ANM, we can successfully avoid the basis mismatch that degrades the performance, commonly seen in the on-the-grid methods [36].

A. Atomic Norm Minimization

We resort to the ANM for the recovery of the angular parameters or equivalent spatial frequencies. We substitute the angular parameters with spatial frequencies in the following. Thus, we can express (1) as

$$\mathbf{H}_{\mathrm{M,R}} = \sum_{l=1}^{L_{\mathrm{M,R}}} [\boldsymbol{\rho}_{\mathrm{M,R}}]_l \boldsymbol{\alpha}([\mathbf{g}_1]_l)\boldsymbol{\alpha}^{\mathsf{H}}([\mathbf{f}_1]_l), \tag{12}$$

where the spatial frequencies $\mathbf{f}_1 = \sin(\boldsymbol{\theta}_{\mathrm{M,R}})$, $\mathbf{g}_1 = \sin(\boldsymbol{\phi}_{\mathrm{M,R}})$. With the abuse of $\boldsymbol{\alpha}(\cdot)$, we redefine $[\boldsymbol{\alpha}([\mathbf{g}_1]_l)]_n = \exp\{j\pi(n-1)[\mathbf{g}_1]_l\}$. Similarly, we express (4) as

$$\mathbf{H}_{\mathrm{R,B}} = \sum_{l=1}^{L_{\mathrm{R,B}}} [\boldsymbol{\rho}_{\mathrm{R,B}}]_l \boldsymbol{\alpha}([\mathbf{g}_2]_l)\boldsymbol{\alpha}^{\mathsf{H}}([\mathbf{f}_2]_l), \tag{13}$$

where the spatial frequencies $\mathbf{f}_2 = \sin(\boldsymbol{\theta}_{\mathrm{R,B}})$, $\mathbf{g}_2 = \sin(\boldsymbol{\phi}_{\mathrm{R,B}})$. These spatial frequencies are within $[0, 1)$ by assuming the angles are within $[0, \pi)$. The atomic set of $\mathbf{H}_{\mathrm{M,R}}$, denoted by $\mathcal{A}_{\mathrm{M}}$, is expressed as

$$\mathcal{A}_{\mathrm{M}} = \left\{ \mathbf{Q}_1(f_1, g_1) : f_1 \in \left[0, 1\right), g_1 \in \left[0, 1\right) \right\}, \tag{14}$$

where $\mathbf{Q}_1(f_1, g_1) = \boldsymbol{\alpha}(f_1)\boldsymbol{\alpha}^{\mathsf{H}}(g_1)$ is the matrix atom. Similarly, the atomic set of $\mathbf{H}_{\mathrm{R,B}}$ is given by

$$\mathcal{A}_{\mathrm{N}} = \left\{ \mathbf{Q}_2(f_2, g_2) : f_2 \in \left[0, 1\right), g_2 \in \left[0, 1\right) \right\}, \tag{15}$$

where $\mathbf{Q}_2(f_2, g_2) = \boldsymbol{\alpha}(f_2)\boldsymbol{\alpha}^{\mathsf{H}}(g_2)$ denotes the matrix atom.

B. First Stage of CE for Hybrid RIS

At the first stage, we aim at the recovery of the channel parameters in $\mathbf{H}_{\mathrm{M,R}}$. For this reason, we formulate the atomic norm with respect to the atomic set $\mathcal{A}_{\mathrm{M}}$ as

$$\|\mathbf{H}_{\mathrm{M,R}}\|_{\mathcal{A}_{\mathrm{M}}} = \inf\{q : \mathbf{H}_{\mathrm{M,R}} \in \mathrm{conv}(\mathcal{A}_{\mathrm{M}})\}. \tag{16}$$

The equivalent form as a semidefinite programming (SDP) problem is [36]

$$\begin{aligned} \|\mathbf{H}_{\mathrm{M,R}}\|_{\mathcal{A}_{\mathrm{M}}} = \inf_{\{\mathbf{C},\mathbf{V}\}} & \Big\{ \frac{1}{2N_{\mathrm{M}}}\mathrm{Tr}(\mathbb{T}(\mathbf{C})) + \frac{1}{2N_{\mathrm{R}}}\mathrm{Tr}(\mathbb{T}(\mathbf{V})) \Big\} \\ & \text{s.t.} \begin{bmatrix} \mathbb{T}(\mathbf{C}) & \mathbf{H}_{\mathrm{M,R}} \\ \mathbf{H}_{\mathrm{M,R}}^{\mathsf{H}} & \mathbb{T}(\mathbf{V}) \end{bmatrix} \succeq 0, \end{aligned} \tag{17}$$

where $\mathbb{T}(\mathbf{C})$ and $\mathbb{T}(\mathbf{v})$ are 2-level Toeplitz matrices. The recovery of the angles $\boldsymbol{\theta}_{\mathrm{M,R}}$ and $\boldsymbol{\phi}_{\mathrm{M,R}}$ (or equivalently $\mathbf{f}_1$ and $\mathbf{g}_1$) can be done by addressing the following convex problem

$$\hat{\mathbf{H}}_{\mathrm{M,R}} = \arg\min_{\mathbf{H}_{\mathrm{M,R}}} \tau\|\mathbf{H}_{\mathrm{M,R}}\|_{\mathcal{A}_{\mathrm{M}}} + \frac{1}{2}\|\sqrt{P_{\mathrm{T}}}\beta_{\mathrm{l}}\mathbf{W}_{\mathrm{H}}\mathbf{H}_{\mathrm{M,R}}\mathbf{X} - \mathbf{Y}_{\mathrm{H}}\|_{\mathrm{F}}^2, \tag{18}$$

where τ is the regularization parameter set as $\tau \propto \sigma\sqrt{N_{\mathrm{R}}N_{\mathrm{M}}\log(N_{\mathrm{R}}N_{\mathrm{M}})}$. Using the SDP formulation, we further define the problem as

$$\begin{aligned} \hat{\mathbf{H}}_{\mathrm{M,R}} = \arg \min_{\mathbf{H}_{\mathrm{M,R}},\mathbf{C},\mathbf{V}} & \frac{\tau}{2N_{\mathrm{M}}}\mathrm{Tr}(\mathbb{T}(\mathbf{C})) + \frac{\tau}{2N_{\mathrm{R}}}\mathrm{Tr}(\mathbb{T}(\mathbf{V})) \\ & + \frac{1}{2}\|\sqrt{P_{\mathrm{T}}}\beta_{\mathrm{l}}\mathbf{W}_{\mathrm{H}}\mathbf{H}_{\mathrm{M,R}}\mathbf{X} - \mathbf{Y}_{\mathrm{H}}\|_{\mathrm{F}}^2 \\ & \text{s.t.} \begin{bmatrix} \mathbb{T}(\mathbf{C}) & \mathbf{H}_{\mathrm{M,R}} \\ \mathbf{H}_{\mathrm{M,R}}^{\mathsf{H}} & \mathbb{T}(\mathbf{V}) \end{bmatrix} \succeq 0. \end{aligned} \tag{19}$$

The solutions of the Toeplitz matrices $\mathbb{T}(\mathbf{C})$ and $\mathbb{T}(\mathbf{V})$ lead us to the recovery of the angles $\boldsymbol{\theta}_{\mathrm{M,R}}$ and $\boldsymbol{\phi}_{\mathrm{M,R}}$, respectively, by applying the ROOTMUSIC algorithm [37].[6] We assume the order information of the angles and the number of paths as known prior information. In practice, this information can be obtained by using off-line channel measurements, CS-based recovery algorithms [38], or ray tracing based network planning tools. We estimate the path gain vector $\boldsymbol{\rho}_{\mathrm{M,R}}$ by applying the least squares (LS), which results in

$$\hat{\boldsymbol{\rho}}_{\mathrm{M,R}} = \Big[\sqrt{P_{\mathrm{T}}}\beta_{\mathrm{l}}(\mathbf{X}^{\mathsf{T}} \otimes \mathbf{W}_{\mathrm{H}})\big((\mathbf{A}^*(\hat{\boldsymbol{\theta}}_{\mathrm{M,R}}) \odot \mathbf{A}(\hat{\boldsymbol{\phi}}_{\mathrm{M,R}})\big)\Big]^{\dagger}\mathbf{y}_{\mathrm{H}}, \tag{20}$$

where $\mathbf{y}_{\mathrm{H}}$ is defined as $\mathbf{y}_{\mathrm{H}} = \mathrm{vec}(\mathbf{Y}_{\mathrm{H}})$,$\hat{\boldsymbol{\theta}}_{\mathrm{M,R}}$ and $\hat{\boldsymbol{\phi}}_{\mathrm{M,R}}$ are the estimates of $\boldsymbol{\theta}_{\mathrm{M,R}}$ and $\boldsymbol{\phi}_{\mathrm{M,R}}$, respectively.

C. Second Stage of CE for Hybrid RIS

At the second stage, we aim at extracting the channel parameters from the received signals at BS.[7] In order to simplify the CE, we use the estimate of $\hat{\mathbf{H}}_{\mathrm{M,R}}$ from the first stage,which results in

$$\hat{\mathbf{U}} = [\boldsymbol{\Omega}_1\hat{\mathbf{H}}_{\mathrm{M,R}}\mathbf{X}, \cdots, \boldsymbol{\Omega}_K\hat{\mathbf{H}}_{\mathrm{M,R}}\mathbf{X}] \in \mathbb{C}^{N_{\mathrm{R}} \times TK}. \tag{21}$$

We formulate the atomic norm of $\mathbf{H}_{\mathrm{R,B}}$ as

$$\|\mathbf{H}_{\mathrm{R,B}}\|_{\mathcal{A}_{\mathrm{N}}} = \inf\{q : \mathbf{H}_{\mathrm{R,B}} \in \mathrm{conv}(\mathcal{A}_{\mathrm{N}})\}. \tag{22}$$

By following the SDP formulation, we can write

$$\begin{aligned} \|\mathbf{H}_{\mathrm{R,B}}\|_{\mathcal{A}_{\mathrm{N}}} = \inf_{\{\mathbf{S},\mathbf{O}\}} & \Big\{ \frac{1}{2N_{\mathrm{R}}}\mathrm{Tr}(\mathbb{T}(\mathbf{S})) + \frac{1}{2N_{\mathrm{B}}}\mathrm{Tr}(\mathbb{T}(\mathbf{O})) \Big\} \\ & \text{s.t.} \begin{bmatrix} \mathbb{T}(\mathbf{S}) & \mathbf{H}_{\mathrm{R,B}} \\ \mathbf{H}_{\mathrm{R,B}}^{\mathsf{H}} & \mathbb{T}(\mathbf{O}) \end{bmatrix} \succeq 0. \end{aligned} \tag{23}$$

where $\mathbb{T}(\mathbf{S})$ and $\mathbb{T}(\mathbf{O})$ are 2-level Toeplitz matrices. In order to recover the angles $\boldsymbol{\theta}_{\mathrm{R,B}}$ and $\boldsymbol{\phi}_{\mathrm{R,B}}$ in $\mathbf{H}_{\mathrm{R,B}}$,

[6]The ROOTMUSIC algorithm can be easily implemented by using the function *rootmusic* in MATLAB.

[7]Another alternative is to apply joint processing on the received signals at both the BS and the RIS, which is left as our future work.

we formulate the problem as

$$\hat{\mathbf{H}}_{\mathrm{R,B}} = \arg\min_{\mathbf{H}_{\mathrm{R,B}}} \nu\|\mathbf{H}_{\mathrm{R,B}}\|_{\mathcal{A}_{\mathrm{N}}} + \frac{1}{2}\|\sqrt{P_{\mathrm{T}}}\beta_2\mathbf{W}_{\mathrm{B}}^{\mathsf{H}}\mathbf{H}_{\mathrm{R,B}}\hat{\mathbf{U}} - \mathbf{Y}\|_{\mathrm{F}}^2, \tag{24}$$

where the regularization parameter ν is set as $\nu \propto \sigma\sqrt{N_{\mathrm{R}}N_{\mathrm{B}}\log(N_{\mathrm{R}}N_{\mathrm{B}})}$. By following the SDP formulation, we define the problem as

$$\begin{aligned}\hat{\mathbf{H}}_{\mathrm{R,B}} = \arg\min_{\mathbf{H}_{\mathrm{R,B}},\mathbf{S},\mathbf{O}} \ & \frac{\nu}{2N_{\mathrm{R}}}\mathrm{Tr}(\mathbb{T}(\mathbf{S})) + \frac{\nu}{2N_{\mathrm{B}}}\mathrm{Tr}(\mathbb{T}(\mathbf{O})) \\ & + \frac{1}{2}\|\sqrt{P_{\mathrm{T}}}\beta_2\mathbf{W}_{\mathrm{B}}^{\mathsf{H}}\mathbf{H}_{\mathrm{R,B}}\hat{\mathbf{U}} - \mathbf{Y}\|_{\mathrm{F}}^2 \\ \text{s.t.} \ & \begin{bmatrix}\mathbb{T}(\mathbf{S}) & \mathbf{H}_{\mathrm{R,B}} \\ \mathbf{H}_{\mathrm{R,B}}^{\mathsf{H}} & \mathbb{T}(\mathbf{O})\end{bmatrix} \succeq 0. \end{aligned} \tag{25}$$

Similarly, the angles $\boldsymbol{\theta}_{\mathrm{R,B}}$ and $\boldsymbol{\phi}_{\mathrm{R,B}}$ can be recovered based on the solutions of $\mathbb{T}(\mathbf{S})$ and $\mathbb{T}(\mathbf{O})$, respectively, which can be solved by ROOTMUSIC algorithm [37]. The estimate of the path gain vector is addressed by LS as

$$\hat{\boldsymbol{\rho}}_{\mathrm{R,B}} = \Big[\sqrt{P_{\mathrm{T}}}\beta_2(\hat{\mathbf{U}}^{\mathsf{T}}\otimes\mathbf{W}_{\mathrm{B}}^{\mathsf{H}})\big((\mathbf{A}^*(\hat{\boldsymbol{\theta}}_{\mathrm{R,B}})\odot\mathbf{A}(\hat{\boldsymbol{\phi}}_{\mathrm{R,B}})\big)\Big]^{\dagger}\mathbf{y}, \tag{26}$$

where $\mathbf{y} = \mathrm{vec}(\mathbf{Y})$, $\hat{\boldsymbol{\theta}}_{\mathrm{R,B}}$ and $\hat{\boldsymbol{\phi}}_{\mathrm{R,B}}$ are the estimates of $\boldsymbol{\theta}_{\mathrm{R,B}}$ and $\boldsymbol{\phi}_{\mathrm{R,B}}$, respectively. The proposed two-stage CE method is summarized in the **Algorithm 1**.

With the estimation of all the channel parameters, we can calculate the angle differences associated with the RIS (written as a matrix) and the products of path gains (written as a vector). The angle differences are functions of the estimates of $\boldsymbol{\phi}_{\mathrm{R,B}}$ and $\boldsymbol{\theta}_{\mathrm{M,R}}$, expressed as

$$[\hat{\boldsymbol{\Delta}}]_{lp} = \mathrm{asin}\big[\sin([\hat{\boldsymbol{\phi}}_{\mathrm{M,R}}]_l) - \sin([\hat{\boldsymbol{\theta}}_{\mathrm{R,B}}]_p)\big], \quad \text{for } l = 1,\cdots,L_{\mathrm{M,R}}, p = 1,\cdots,L_{\mathrm{R,B}}. \tag{27}$$

Moreover, the estimated products of path gains $\hat{\boldsymbol{\rho}} \in \mathbb{C}^{L_{\mathrm{R,B}}L_{\mathrm{M,R}}\times 1}$ are defined as follows

$$\hat{\boldsymbol{\rho}} = \hat{\boldsymbol{\rho}}_{\mathrm{R,B}} \otimes \hat{\boldsymbol{\rho}}_{\mathrm{M,R}}. \tag{28}$$

The training overhead for the proposed CE for hybrid RIS-aided mmWave MIMO systems is given as

$$T_{\mathrm{H}} = KT\Big\lceil\frac{N_{\mathrm{C,B}}}{N_{\mathrm{RF,B}}}\Big\rceil\Big\lceil\frac{M}{N_{\mathrm{RF,R}}}\Big\rceil, \tag{29}$$

where $N_{\mathrm{RF,B}}$ is the number of RF chains at BS. The complexity of the algorithm depends on the size of the PSD (positive semidefinite) matrix in (19) and in (25), given respectively by $O(N_{\mathrm{R}}+N_{\mathrm{B}})^{3.5}$ and $O(N_{\mathrm{M}}+N_{\mathrm{R}})^{3.5}$ [39]. Note also that the complexity order of the ANM is much higher than that of the LS algorithm in (20) and (26). For this reason, the overall complexity order is $\max\{O(N_{\mathrm{R}}+N_{\mathrm{B}})^{3.5}, O(N_{\mathrm{M}}+N_{\mathrm{R}})^{3.5}\}$.

Algorithm 1 The Proposed Two-Stage CE for Hybrid RIS

Input : $\mathbf{f}_1$, $\mathbf{g}_1$, $\mathbf{f}_2$, $\mathbf{g}_2$, $\boldsymbol{\rho}_{\mathrm{R,B}}$, $\boldsymbol{\rho}_{\mathrm{M,R}}$, M, K
Define $\mathbf{H}_{\mathrm{M,R}}$, $\mathbf{H}_{\mathrm{R,B}}$, $\mathbf{W}_{\mathrm{B}}$, and $\mathbf{X}$;
▷ `Uplink training`
for $k \leftarrow 1$ **to** K **do**
  Define $|\mathbb{M}| = M$;
  Get $\mathbf{Y}_{\mathrm{H}k}$ accordingly to (8);
  Set $[\boldsymbol{\Omega}_k]_{i,i} = 0$, for $i \in \mathbb{M}$;
  Get $\mathbf{Y}_k$ by following (9);
end
Collect the received signals $\mathbf{Y}_{\mathrm{H}} = [\mathbf{Y}_{\mathrm{H}1}^{\mathsf{T}},\cdots,\mathbf{Y}_{\mathrm{H}K}^{\mathsf{T}}]^{\mathsf{T}} \in \mathbb{C}^{MK\times T}$ and $\mathbf{Y} = [\mathbf{Y}_1,\cdots,\mathbf{Y}_K] \in \mathbb{C}^{N_{\mathrm{C,B}}\times TK}$;
▷ `First stage of CE`
Estimate $\boldsymbol{\theta}_{\mathrm{M,R}}$ and $\boldsymbol{\phi}_{\mathrm{M,R}}$ by following (19);
Estimate $\boldsymbol{\rho}_{\mathrm{M,R}}$ by using LS (20);
▷ `Second stage of CE`
Reconstruct $\hat{\mathbf{H}}_{\mathrm{M,R}}$ at the BS;
Get $\hat{\mathbf{U}}$ by following (21);
Estimate $\boldsymbol{\theta}_{\mathrm{R,B}}$ and $\boldsymbol{\phi}_{\mathrm{R,B}}$ by following (25);
Estimate $\boldsymbol{\rho}_{\mathrm{R,B}}$ by applying (26);
Reconstruct $\hat{\mathbf{H}}_{\mathrm{R,B}}$;
Output: $\hat{\mathbf{H}}_{\mathrm{M,R}}$, $\hat{\mathbf{H}}_{\mathrm{R,B}}$, $\hat{\boldsymbol{\rho}}_{\mathrm{M,R}}$, $\hat{\boldsymbol{\theta}}_{\mathrm{M,R}}$, $\hat{\boldsymbol{\phi}}_{\mathrm{M,R}}$, $\hat{\boldsymbol{\rho}}_{\mathrm{R,B}}$, $\hat{\boldsymbol{\theta}}_{\mathrm{R,B}}$, and $\hat{\boldsymbol{\phi}}_{\mathrm{R,B}}$

V. Design of RIS Phase Control Matrix and Beamforming Vectors

In this section, we present the design of the phase control matrix at the RIS and the beamforming vectors at BS and MS.

A. RIS Phase Control Matrix

We design the RIS phase control matrix at the RIS based on the maximization of the power of the effective channel, as defined in (6). Based on this criterion, the optimal phase control matrix is

$$\boldsymbol{\Omega}^* = \arg\max_{\boldsymbol{\Omega}} \|\mathbf{G}\|_{\mathrm{F}}^2, \tag{30}$$

where $\|\mathbf{G}\|_{\mathrm{F}}^2 = \|\mathrm{diag}(\hat{\boldsymbol{\rho}}_{\mathrm{R,B}})\mathbf{A}^{\mathsf{H}}(\hat{\boldsymbol{\theta}}_{\mathrm{R,B}})\boldsymbol{\Omega}\mathbf{A}(\hat{\boldsymbol{\phi}}_{\mathrm{M,R}})\mathrm{diag}(\hat{\boldsymbol{\rho}}_{\mathrm{M,R}})\|_{\mathrm{F}}^2$, whose (l,p)th entry can be written as

$$[\mathbf{G}]_{l,p} = [\hat{\boldsymbol{\rho}}_{\mathrm{R,B}}]_p\boldsymbol{\omega}^{\mathsf{T}}\boldsymbol{\alpha}([\hat{\boldsymbol{\Delta}}]_{l,p})[\hat{\boldsymbol{\rho}}_{\mathrm{R,M}}]_l, \quad \text{for } l = 1,\cdots,L_{\mathrm{R,B}},\ p = 1,\cdots,L_{\mathrm{R,M}}, \tag{31}$$

where $\boldsymbol{\Omega} = \mathrm{diag}(\boldsymbol{\omega})$. We further define the vectorization of $\mathbf{G}$ as $\mathbf{g} = \mathrm{vec}(\mathbf{G})$, and the ith element of $\mathbf{g}$ is expressed as

$$[\mathbf{g}]_i = [\hat{\boldsymbol{\rho}}]_i\boldsymbol{\omega}^{\mathsf{T}}\boldsymbol{\alpha}([\hat{\boldsymbol{\delta}}]_i), \text{ for } i = 1,\cdots,L_{\mathrm{R,B}}L_{\mathrm{M,R}}, \tag{32}$$

where $\hat{\boldsymbol{\delta}} = \mathrm{vec}(\hat{\boldsymbol{\Delta}})$.

By following (31) and (32), we can further express $\|\mathbf{G}\|_{\mathrm{F}}^2$ as

$$\|\mathbf{G}\|_{\mathrm{F}}^2 = \sum_{i=1}^{L_{\mathrm{R,B}}L_{\mathrm{M,R}}} |[\hat{\boldsymbol{\rho}}]_i\boldsymbol{\omega}^{\mathsf{T}}\boldsymbol{\alpha}([\hat{\boldsymbol{\delta}}]_i)|^2. \tag{33}$$

Substituting (33) in (30), we define the optimal $\boldsymbol{\omega}^{*}$[8] as

$$\boldsymbol{\omega}^{*} = \arg\max_{\boldsymbol{\omega}} \sum_{i=1}^{L_{\mathrm{R,B}}L_{\mathrm{M,R}}} |[\hat{\boldsymbol{\rho}}]_i \boldsymbol{\omega}^{\mathsf{T}} \boldsymbol{\alpha}([\hat{\boldsymbol{\delta}}]_i)|^2 = \arg\max_{\boldsymbol{\omega}} \boldsymbol{\omega}^{\mathsf{T}} \mathbf{C}\mathbf{C}^{\mathsf{H}} \boldsymbol{\omega}^{*}, \tag{34}$$

$$\mathbf{C} = \left[\boldsymbol{\alpha}([\hat{\boldsymbol{\delta}}]_1), \ldots, \boldsymbol{\alpha}([\hat{\boldsymbol{\delta}}]_{L_{\mathrm{R,B}}L_{\mathrm{M,R}}})\right] \mathrm{diag}\left(\left[[\hat{\boldsymbol{\rho}}]_1, \ldots, [\hat{\boldsymbol{\rho}}]_{L_{\mathrm{R,B}}L_{\mathrm{M,R}}}\right]\right). \tag{35}$$

We apply singular value decomposition (SVD) on $\mathbf{C}\mathbf{C}^{\mathsf{H}}$, resulting in $\mathbf{C}\mathbf{C}^{\mathsf{H}} = \mathbf{E}\mathbf{D}\mathbf{E}^{\mathsf{H}}$. Then, in order to obtain the optimal $\boldsymbol{\omega}^{*}$, we select the first column of $\mathbf{E}$ and further project it to the unit-modulus vector space, resulting in $\boldsymbol{\omega}^{*} = \exp(-j\mathrm{phase}([\mathbf{E}]_{:,1}))$, where $\mathrm{phase}()$ means the element-wise operation of extracting the phase of the argument.

B. Beamforming Vectors

The design of the beamforming vectors at BS and MS is based on the estimates of the channel matrices and the already-designed phase control matrix, i.e., the reconstruction of the composite channel $\hat{\mathbf{H}} = \hat{\mathbf{H}}_{\mathrm{R,B}}\mathrm{diag}(\boldsymbol{\omega}^{*})\hat{\mathbf{H}}_{\mathrm{M,R}}$. We conduct SVD on $\hat{\mathbf{H}}$, which results in $\hat{\mathbf{H}} = \mathbf{U}\mathbf{S}\mathbf{V}$, so that we select the first left-singular vector and right-singular vector as beamforming vectors, i.e., $\mathbf{f} = [\mathbf{V}]_{:,1}$ and $\mathbf{w} = [\mathbf{U}]_{:,1}$.

VI. CRLB Analyses

The CRLB is the lower bound of any unbiased estimator [41], so that it serves as a performance indicator of our proposed CE method. Thus, in this section, we develop the CRLB for the channel parameter estimation based on [41], [42]. Through our analyses, the observation noise in (10) and (11) are assumed to be Gaussian. In our proposed two-stage CE, we first estimate the channel $\mathbf{H}_{\mathrm{M,R}}$ based on the received signals at RIS, and then estimate the channel $\mathbf{H}_{\mathrm{R,B}}$ based on the received signals at BS. In order to align with our proposed two-stage CE procedure, we also divide the calculations of the CRLB into two steps.

At the first stage of the proposed CE, the parameters $\boldsymbol{\zeta} = \{\boldsymbol{\theta}_{\mathrm{M,R}}, \boldsymbol{\phi}_{\mathrm{M,R}}, \boldsymbol{\rho}_{\mathrm{M,R}}\}$ are estimated, which are related to the MS-RIS channel. The MSEs of the channel parameters in $\boldsymbol{\zeta}$ are lower bounded by $\mathrm{CRLB}(\boldsymbol{\zeta}) = \mathbf{J}^{-1}(\boldsymbol{\zeta})$, where $\mathbf{J}(\boldsymbol{\zeta})$ is the fisher information matrix (FIM) of $\boldsymbol{\zeta}$. The details of the derivations of the FIM are provided in Appendix A.

At the second stage of the proposed CE procedure, the parameters related to the RIS-BS channel[9] $\boldsymbol{\eta} = \{\boldsymbol{\theta}_{\mathrm{R,B}}, \boldsymbol{\phi}_{\mathrm{R,B}}, \boldsymbol{\rho}_{\mathrm{R,B}}\}$ are estimated. For simplicity, we calculate the CRLB for $\boldsymbol{\eta}$ by considering the following approximation,

$$\hat{\mathbf{H}}_{\mathrm{M,R}} \approx \mathbf{H}_{\mathrm{M,R}}, \tag{36}$$

which leads to

$$\tilde{\mathbf{U}} = [\boldsymbol{\Omega}_1 \mathbf{H}_{\mathrm{M,R}}\mathbf{X}, \cdots, \boldsymbol{\Omega}_K \mathbf{H}_{\mathrm{M,R}}\mathbf{X}]. \tag{37}$$

[8]The design of the phase control matrix can be done by following other algorithms, such as element iteration algorithm and majorization-minimization (MM) method [40]. Other methods to improve the design of BS beamforming and RIS phase control matrix are left for future investigation.

[9]Since the BS knows the estimates of the channel parameters in $\mathbf{H}_{\mathrm{M,R}}$, it is reasonable to also ignore those parameters in the CRLB analysis at the second stage.

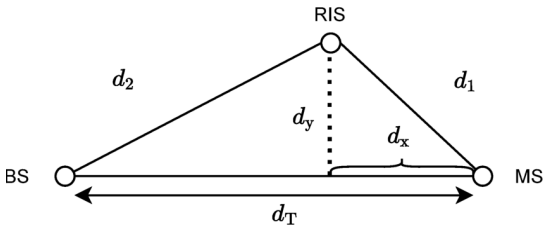


Fig. 3. Network topology.

TABLE I
PARAMETER SETUP

Symbol	Parameter	Value
γ	Path loss exponent	3
f_c	Carrier frequency	28 [MHz]
B	Bandwidth	100 [MHz]
N_0	Noise power density	-173 [dBm/Hz]
P_{T}	Transmit power	$0:5:20$ [dBm]
d_0	Reference distance	1 [m]

Finally, the details of the derivations of $\mathrm{CRLB}(\boldsymbol{\eta}) = \mathbf{J}^{-1}(\boldsymbol{\eta})$ are included in Appendix B.

VII. Performance Evaluation

In this section, we describe the parameter setup and the performance metrics to evaluate the proposed two-stage CE approach. The propagation path gains are distributed as $\mathcal{CN}(0,1)$. We assume $N_{\mathrm{B}} = 16$, $N_{\mathrm{R}} \in \{32, 64\}$, $N_{\mathrm{M}} = 16$, $L_{\mathrm{R,B}} = L_{\mathrm{M,R}} = 2$. We perform 1000 trials to average out the results. As shown in Fig. 3, we consider the BS at a fixed location (0,0), the RIS at the location (x,y) and the MS at (x_T,0), where $x = d_{\mathrm{T}} - d_{\mathrm{x}}$, $y = d_{\mathrm{y}}$ and $x_T = d_{\mathrm{T}}$. As a consequence, the distance between MS and RIS is $d_1 = \sqrt{d_{\mathrm{x}}^2 + d_{\mathrm{y}}^2}$, while the distance between BS and RIS is $d_2 = \sqrt{(d_{\mathrm{T}} - d_{\mathrm{x}})^2 + d_{\mathrm{y}}^2}$. The path loss as a function of distance is given by [5], so that $\boldsymbol{\beta}(d_1) = \beta_0\left(\frac{d_0}{d_1}\right)^{\gamma}$ and $\beta(d_2) = \beta_0\left(\frac{d_0}{d_1 d_2}\right)^{\gamma}$, where $\beta_0 = (\frac{\lambda}{4\pi d_0})^2$ is the path loss at the reference distance d_0 and $\lambda = \frac{c}{f_c}$ denotes the wavelength, with $c = 3 \times 10^8$ [m/s] being the speed of the light and f_c the carrier frequency in Hz. We set $d_0 = 1$ [m], the path loss exponent as $\gamma = 3$, $f_c = 28$ [MHz], the noise power density as $N_0 = -173$ [dBm/Hz], the bandwidth as $B = 100$ [MHz], and the transmit power as $P_{\mathrm{T}} = 0:5:20$ [dBm]. Table I summarizes the parameters of the system model.

We evaluate the performance considering different numbers of active elements at the RIS. Table II summarizes the parameters setup for the hybrid RIS architecture. We denote Setups 1 and 2 as the *small* RIS, where $N_{\mathrm{R}} = 32$ for both cases, but which differ from each other in the number of active elements, with $M \in \{2, 4\}$. Setups 3 and 4 are denoted as the *large* RIS, in which we increase the number of elements at the RIS to $N_{\mathrm{R}} = 64$, differing from each other in $M \in \{6, 8\}$. Also, due to the higher number of elements at the large RIS, we further increase the training overhead compared to the

TABLE II
PARAMETERS OF THE HYBRID RIS

	N_{R}	M	$N_{\mathrm{RF,R}}$	K	T	$N_{\mathrm{C,B}}$	$N_{\mathrm{RF,B}}$	T_{H}
Setup 1	32	4	4	5	8	8	8	40
Setup 2	32	2	2	5	8	8	8	40
Setup 3	64	8	8	7	8	8	8	56
Setup 4	64	6	6	7	8	8	8	56

small RIS. We also consider $N_{\mathrm{RF,R}} = M$ and $N_{\mathrm{C,B}} = N_{\mathrm{RF,B}}$ for all the hybrid setups in order to reduce the training overhead.

In addition, since we consider ANM in our proposed CE method, we assume the spatial frequencies are separated at least by $\left(\frac{4}{N_{\mathrm{B}}}\right)$, $\left(\frac{4}{N_{\mathrm{R}}}\right)$, $\left(\frac{4}{N_{\mathrm{M}}}\right)$ as to guarantee the super-resolution estimation.[10]

A. Performance Metrics

We evaluate the performance of MSEs[11] of the estimates of AoDs, AoAs, angle differences and propagation path gains, which are respectively given as

$$\mathrm{MSE}(\boldsymbol{\phi}_{\mathrm{M,R}}) = \mathbb{E}\left[\frac{\|\boldsymbol{\phi}_{\mathrm{M,R}} - \hat{\boldsymbol{\phi}}_{\mathrm{M,R}}\|_2^2}{L_{\mathrm{M,R}}}\right], \tag{38}$$

$$\mathrm{MSE}(\boldsymbol{\theta}_{\mathrm{M,R}}) = \mathbb{E}\left[\frac{\|\boldsymbol{\theta}_{\mathrm{M,R}} - \hat{\boldsymbol{\theta}}_{\mathrm{M,R}}\|_2^2}{L_{\mathrm{M,R}}}\right], \tag{39}$$

$$\mathrm{MSE}(\boldsymbol{\rho}_{\mathrm{M,R}}) = \mathbb{E}\left[\frac{\|\boldsymbol{\rho}_{\mathrm{M,R}} - \hat{\boldsymbol{\rho}}_{\mathrm{M,R}}\|_2^2}{L_{\mathrm{M,R}}}\right], \tag{40}$$

$$\mathrm{MSE}(\boldsymbol{\delta}) = \mathbb{E}\left[\frac{\|\boldsymbol{\delta} - \hat{\boldsymbol{\delta}}\|_2^2}{L_{\mathrm{R,B}} L_{\mathrm{M,R}}}\right]. \tag{41}$$

The MSEs of $\boldsymbol{\phi}_{\mathrm{R,B}}$, $\boldsymbol{\theta}_{\mathrm{R,B}}$, $\boldsymbol{\rho}_{\mathrm{R,B}}$ are defined in the same manner. Moreover, we adopt the NMSE of the channel matrices as a performance metric. The NMSE of $\mathbf{H}_{\mathrm{M,R}}$ is defined as,

$$\mathrm{NMSE} = \mathbb{E}\left[\frac{\|\mathbf{H}_{\mathrm{M,R}} - \hat{\mathbf{H}}_{\mathrm{M,R}}\|_{\mathrm{F}}^2}{\|\mathbf{H}_{\mathrm{M,R}}\|_{\mathrm{F}}^2}\right] \tag{42}$$

where $\hat{\mathbf{H}}_{\mathrm{M,R}}$ is the estimates of the channel matrix $\mathbf{H}_{\mathrm{M,R}}$. The NMSE of $\mathbf{H}_{\mathrm{R,B}}$ can be similarly defined. We also evaluate the performance in terms of average effective SE in [bits/s/Hz] defined as [43]

$$R = \mathbb{E}\left[\frac{T_{\mathrm{c}} - T_{\mathrm{H}}}{T_{\mathrm{c}}} \log_2\left(1 + \frac{|\mathbf{w}^{\mathsf{H}}\hat{\mathbf{H}}\mathbf{f}|^2}{\sigma^2 + \mathrm{var}(\mathbf{w}^{\mathsf{H}}\mathbf{H}_{\mathrm{e}}\mathbf{f})}\right)\right], \tag{43}$$

where T_{c} is the number of time slots in a coherence time interval and $\mathbf{H}_{\mathrm{e}}$ denotes the channel estimation error, i.e., $\mathbf{H}_{\mathrm{e}} = \mathbf{H}(\boldsymbol{\omega}^*) - \hat{\mathbf{H}}$. We assume that the coherence time has 500 channel uses, i.e., $T_{\mathrm{c}} = 500$.

B. Comparison With the Passive RIS

The simulation results for the MSEs of channel parameters $\boldsymbol{\theta}_{\mathrm{M,R}}$, $\boldsymbol{\phi}_{\mathrm{R,B}}$ and $\boldsymbol{\delta}$ are shown in Fig. 4 for the small RIS.

[10]Accordingly with [39, Thrm. 1], the convergence of the algorithm is guaranteed if the signal of interested is composed of adequately separated frequencies. Otherwise, calculating the atomic norm is an infinite programming problem over all feasible set of frequencies, where convergence is not guaranteed.

[11]The MSE can be also formulated by considering the sine of the angles/spatial frequencies. We clarify that the formulations are equivalent.

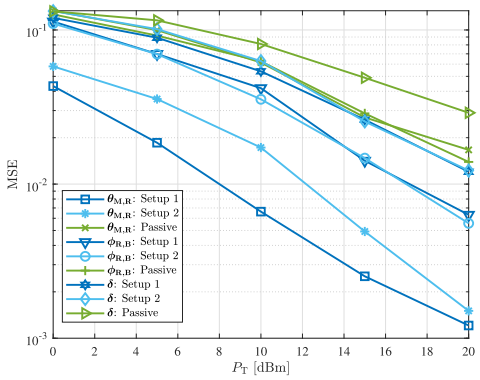


Fig. 4. MSEs of the angular parameters for the RIS with $N_{\mathrm{R}} = 32$ elements.

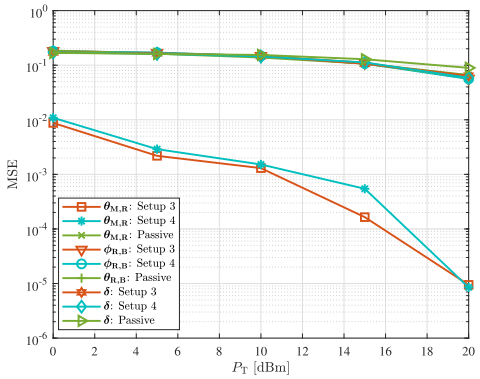


Fig. 5. MSEs of the angular parameters for the RIS with $N_{\mathrm{R}} = 64$ elements.

We consider $d_{\mathrm{T}} = 22$ [m], $d_{\mathrm{x}} = 15$ [m], $d_{\mathrm{y}} = 2$ [m] and compare our results with the benchmark Thanks to the super-resolution of the channel parameters at the first stage, the error propagation to the second stage is relatively small.scheme, i.e., the passive RIS detailed in [18]. For the estimation of $\boldsymbol{\theta}_{\mathrm{M,R}}$, all the setups of the hybrid RIS outperform the passive RIS. Setup 1 brings better performance than Setup 2, while Setup 2 has relatively lower power consumption due to the reduced number of RF chains at the RIS. The simulation results can be well explained by taking into account the path loss effect on the CE. In our proposed CE method, we perform CE at the RIS, where the path loss is proportional to the distance d_1.

However, when the CE is performed at the BS, the path loss is proportional to $d_1 d_2$. That's why the proposed hybrid RIS CE can in general outperform the passive RIS CE. Regarding the MSE of $\boldsymbol{\delta}$, the hybrid RIS also has better performance than the passive RIS. For this parameter $\boldsymbol{\delta}$, the performance depends on the estimates from both the first and second stage. At the first stage, we can obtain better estimation of the channel parameters due to the lower path loss. However, at the second stage, the path loss is larger, which affects the estimation of $\boldsymbol{\rho}_{\mathrm{R,B}}$. The overall performance heavily relies on the worse estimation, so the two setups do not have big performance gap, as shown in Fig. 4.

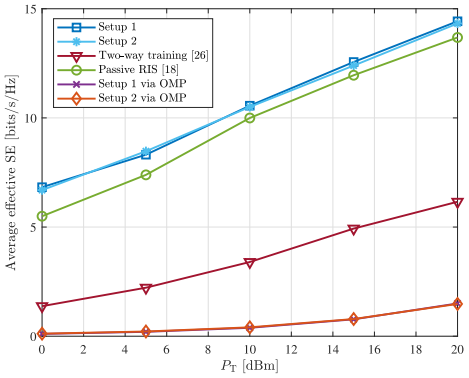


Fig. 6. Overall analyses in terms of SE of our proposed method in comparison with the OMP, passive RIS [18] and hybrid RIS via two-way *uplink downlink* training [26].

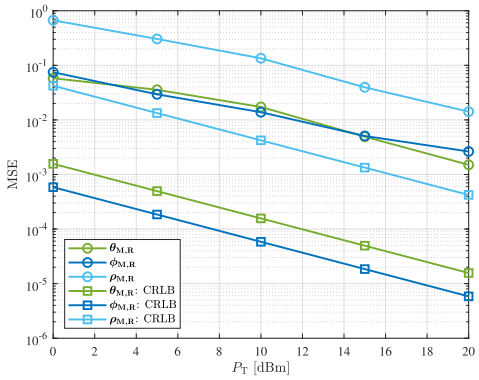


Fig. 7. MSEs of the channel parameters in $\mathbf{H}_{\mathrm{M,R}}$ vs. CRLBs.

We further evaluate the MSEs of the angular parameters for the large RIS with 64 elements, i.e., Setups 3 and 4, in Fig. 5. We consider $d_{\mathrm{T}} = 50$ [m], $d_{\mathrm{x}} = 10$ [m], $d_{\mathrm{y}} = 2$ [m]. Since we consider a larger distance in the LoS link, the path loss of the MS-RIS-BS link brings more impact on the CE performance. Herein, we can clearly see that the estimate of $\boldsymbol{\theta}_{\mathrm{M,R}}$ is better for the hybrid setups in Fig. 5. The MSEs of $\boldsymbol{\phi}_{\mathrm{R,B}}$ and $\boldsymbol{\delta}$ show similar results as the small RIS in Fig. 4. That is, the performance of both architectures is strongly affected by the path loss.

C. Overall Analyses of the Hybrid RIS

We evaluate the overall performance in terms of SE of the Setups 1 and 2, and passive RIS via ANM [18] and that of the hybrid RIS via two-way *uplink and downlink* training [26]. We also evaluate the performance of Setups 1 and 2 via OMP. In particular, the OMP algorithm employs a overcomplete dictionary constructed by array response vectors with quantized angle grids [33]. The implementation of the OMP follows the two-stage procedure. At the first stage, we aim to recover the channel matrix $\mathbf{H}_{\mathrm{M,R}}$ from the received signal at the RIS (10). For this purpose, we constructed the dictionary matrices at the MS and RIS by considering the grid sizes $G_M = N_{\mathrm{M}}$ and $G_R = N_{\mathrm{R}}$, respectively. At the second stage, we target the recovery of $\mathbf{H}_{\mathrm{R,B}}$ based on the received signal at the BS (11). Again, we constructed the dictionary matrices at RIS and BS by quantizing the angle domain. The grid sizes at BS and RIS are set as $G_{\mathrm{B}} = N_{\mathrm{B}}$ and $G_{\mathrm{R}} = N_{\mathrm{R}}$, respectively. We clarify that it is possible to employ a larger grid size. However, due to the inevitable quantization error and increased correlation among the atoms, this may not further improve the overall performance [44].

We assume $M = N_{\mathrm{RF,R}} = 10$ for the two-way training. For simplicity, we focus on the small RIS with $N_{\mathrm{R}} = 32$ elements. The results of the average effective SE are provided in Fig. 6. From the figure, we observe that results of the hybrid setups and the passive RIS are aligned with the MSEs in Fig. 4. Also, we can notice that the performance of our proposed method via ANM can significantly outperform that of the OMP. Moreover, we observe that the overall performance of the Setups 1 and 2 are very similar, which means that we can use a reduced number of active elements and still achieve similar SE. The two-stage CE for hybrid RIS offers the possibility to improve the channel estimation of one of the individual channels, while the other channel still suffer from the effect of the path loss at the receiver. Thanks to the combination of efficient training and availability of measurements at the RIS, the overall performance of our proposed method is better than the passive RIS and the hybrid RIS via two-way uplink downlink training.

1) CRLB: We now examine the performance of our proposed method and the CRLB (More details can be found in Section VI with calculations/derivations drawn in the Appendices A and B). Fig. 7 shows the MSEs of the channel parameters of the Setup 2, where $d_{\mathrm{T}} = 22$ [m], $d_{\mathrm{x}} = 15$ [m], $d_{\mathrm{y}} = 2$ [m]. Despite the better performance obtained in comparison with the literature as show in Fig. 6, the gap between the MSEs of the proposed estimator and the CRLB shows that our method can still be further improved. Nevertheless, this improvement would inevitably come at the expense of more complexity, energy or time consumption, which is not desirable in this context. It is worth mentioning that the gap between the CRLB and the ANM may result from different factors. For instance, the regularization parameter may not be optimal (and finding the optimal one is still an open problem), which affects the performance of the ANM. Nevertheless, CRLB is helpful in providing the guidance for the CE algorithm development in this paper.

In Fig. 8, we evaluate the MSEs of the channel parameters $\boldsymbol{\theta}_{\mathrm{R,B}}$, $\boldsymbol{\phi}_{\mathrm{R,B}}$, $\boldsymbol{\rho}_{\mathrm{R,B}}$ and make comparisons to their CRLBs. Similar to the previous case, we can see that the MSEs of the angles has a clear gap to the CRLBs, which may come from the assumption of perfect recovery of $\mathbf{H}_{\mathrm{M,R}}$ in the CRLB analyses. However, the MSE of the propagation path gain has a reduced gap to the lower bound. Comparing the results in Figs. 7 and 8, we can see that the MSEs of the channel parameters are better at the first stage. This can be understood due to the reduced path loss at the hybrid RIS. Moreover, the estimation of the channel parameters $\boldsymbol{\theta}_{\mathrm{R,B}}$, $\boldsymbol{\phi}_{\mathrm{R,B}}$, $\boldsymbol{\rho}_{\mathrm{R,B}}$ depends on the estimates of the previous stage.

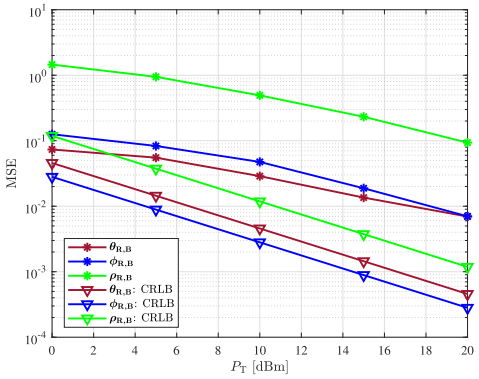


Fig. 8. MSEs of the channel parameters in $\mathbf{H}_{R,B}$ vs. CRLBs.

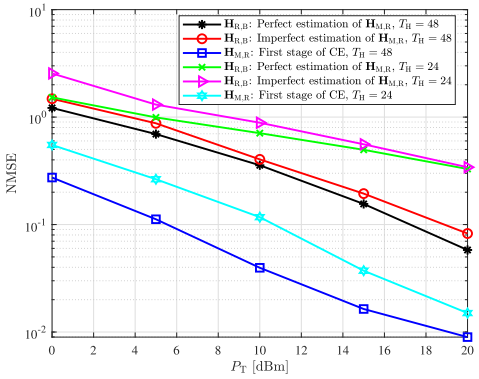


Fig. 9. NMSE of $\mathbf{H}_{R,B}$ and $\mathbf{H}_{M,R}$.

Again, note that we develop the CRLB for the second stage of CE by adopting $\hat{\mathbf{H}}_{M,R} \approx \mathbf{H}_{M,R}$. In practice, the estimates of $\mathbf{H}_{M,R}$ depend on the training overhead, the number of RF chains at the RIS, and the level of SNR. As a consequence, $\hat{\mathbf{H}}_{M,R}$ may not be approximate to $\mathbf{H}_{M,R}$. However, characterizing the channel (parameter) estimation errors at the first stage may yield a non-closed-form solution of the CRLB for the second stage of CE. A more rigorous study on the CRLB analysis for the second stage of CE is left as future investigation.

In order to clarify the effect of error propagation from the first stage to the second stage, we extend the discussion by only focusing on Setup 1. We consider the training overhead as $T_H = \{24, 48\}$. In particular, we set the number of blocks as $K = \{3, 6\}$, while the number of training beams is set to be $T = 8$. In Fig. 9, we evaluate the NMSEs of the channel matrices. First, we focus on the analyses of the error propagation associated with the training overhead $T_H = 48$. From Fig. 9, we can see that the NMSE of $\mathbf{H}_{M,R}$ is low. For instance, when the transmit power is high, such as $P_T = 20$ [dBm], the NMSE reaches the level of 10^{-2}. This result is beneficial from the efficient training and higher received power at the RIS, which leads to a super-resolution estimation. The simulation results for the estimate of $\mathbf{H}_{R,B}$ take two scenarios into account: perfect and imperfect estimation at the first stage of CE. As excepted, the perfect estimation at the first stage yields better performance at the second stage. However, we would like to highlight that the gap between the NMSEs of $\mathbf{H}_{R,B}$ is small for the two scenarios. Moreover, when the training overhead is reduced, the gap between the NMSEs of $\mathbf{H}_{R,B}$ for the two scenarios is also small. We observe that for high transmit power, the NMSEs of $\mathbf{H}_{R,B}$ have similar values. Thus, we can conclude that the error propagation from the first stage has a negligible effect on the performance at the second stage.

We further extend the discussion to the impact of the training overhead on the CE. From Fig. 9, we can see that by increasing the training overhead, better estimation accuracy of $\mathbf{H}_{M,R}$ and $\mathbf{H}_{R,B}$ is achieved. This is reasonable and also intuitive. Also, we observe that the performance for a lower training overhead only degrades slightly. For a high transmit power, such as $P_T = 20$ [dBm], the gap between the NMSEs of $\mathbf{H}_{M,R}$ is equal to 0.0060. Similar conclusions can be drawn for the NMSEs of $\mathbf{H}_{R,B}$. Overall, we can see that our proposed algorithm can achieve effective performance even with reduced training overhead.

2) Location of the RIS: We also evaluate the effect of the position of the RIS on the CE of the first stage (i.e., recovery of $\mathbf{H}_{M,R}$). Considering the Setup 2, the MSEs of the channel parameters $\boldsymbol{\theta}_{M,R}$, $\boldsymbol{\phi}_{M,R}$, $\boldsymbol{\rho}_{M,R}$ are shown in Fig. 10. We consider $d_T = 22$ [m], $d_x \in \{10, 15\}$ [m]. We can see that the estimation of the channel parameters for $d_x = 10$ [m] has better performance. We extend our study for the Setup 3, where the RIS has $N_R = 64$ elements. We evaluate the performance by considering $d_T = 50$ [m], $d_y = 2$ [m], and $d_x \in \{20, 40\}$ [m]. For the Setup 3, the distance between MS and RIS also affects the performance of the CE. We can see in Fig. 11 that the gap between the MSEs of the channel parameters for $d_x = 20$ [m] and $d_x = 40$ [m] is significant.

Next, in Fig. 12 we further evaluate the MSEs of the Setups 3 and 4 for a larger range of distance. We consider the transmission power fixed as $P_T = 10$ [dBm], $d_T = 100$ [m] and the distance $d_x = \{20 : 15 : 80\}$ [m]. We can observe that the MSEs of the channel parameters increase significantly when the distance between the RIS and the MS (d_x) also increases. Moreover, the performance of the Setups 3 and 4 are quite similar to each other, which is aligned with the findings in Figs. 4 and 5. The location of the RIS and how it affects the MSEs of the channel parameters depend on the number of elements/antennas at the BS, RIS, and MS and also the distances among them. The overall performance may not be significantly affected by the position of the RIS. However, we clarify that finding the optimal location of the RIS can bring benefits for its practical deployment.

3) Discussions: We assume that the RIS feeds back the channel parameters to the BS by using the error-free backhaul link. Since we consider structured channel, the BS can reconstruct the MS-RIS channel matrix, i.e., $\hat{\mathbf{H}}_{M,R}$, based on the estimates of the channel parameters. An alternative way is to feed back the entire channel matrix obtained from (19). However, this assumption increases the volume of the feedback due to the higher dimensions of the channel matrix than these of channel parameters. Besides, by using the structured channel, we can exploit the inherent sparsity of the channel and reduce significantly the training overhead.

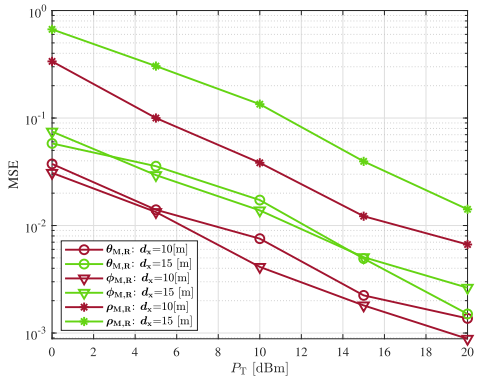


Fig. 10. The effect of the location of the RIS for the channel parameter estimation of the Setup 2.

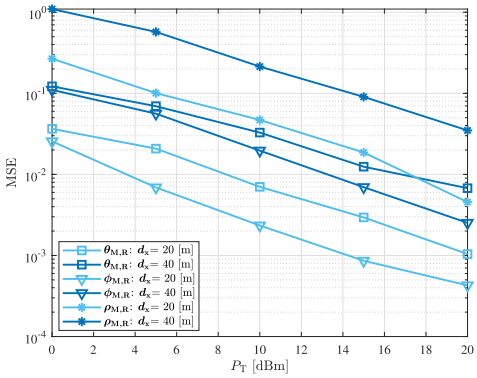


Fig. 11. The effect of the location of the RIS for the channel parameter estimation considering the Setup 3.

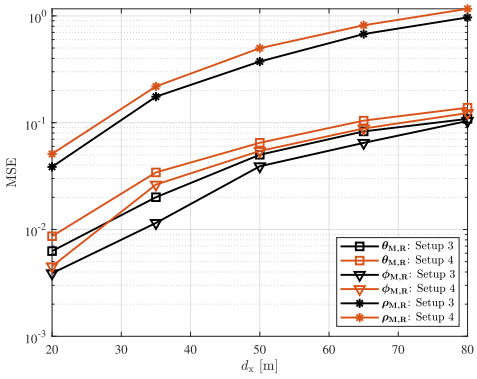


Fig. 12. MSEs of the channel parameters for fixed $P_T = 10$ [dBm] and varying the MS-RIS distance (d_x).

Furthermore, in this work we assume that the active elements are used to collect signal observations at RIS. We find that their positions do not affect the accuracy of the proposed CE algorithm as long as the total number of active elements is fixed. Nevertheless, the performance of CE can be further improved by increasing the training overhead and the number of RF chains and active elements.

VIII. Conclusion and Future Works

In this paper, we have studied the channel estimation via hybrid RIS for mmWave MIMO systems and proposed a two-stage CE procedure via ANM. Our results have shown that our proposed CE method could bring better performance than the passive RIS with the help of a small number of active elements. We have developed the analytical expressions of the CRLB for our proposed method, which provide another baselines in addition to passive RIS CE. It has been verified that the proposed hybrid RIS CE can outperform passive RIS CE under the same assumption of training overhead. We have shown that the location of the RIS can also affect the performance of the proposed hybrid RIS CE.

The study of the minimum training overhead could provide guidance to the design of CE, which is left as our future work. Future research directions also include the investigation of the CE performance for multi-user MIMO systems. In addition, the trade-off between energy efficiency and hardware cost for the RIS architectures is left as future investigation. The assumption on knowing the exact number of paths and their order information as known *prior* information should be relaxed in order to make the proposed method more applicable to the practical systems.

Appendix A
Calculation of the CRLB: Stage 1 of CE

In the following, we describe the calculations of the CRLB for the proposed CE method. At the first stage of CE, we estimate the parameters in the MS-RIS channel, defined as $\boldsymbol{\zeta} = \{\boldsymbol{\theta}_{\mathrm{M,R}}, \boldsymbol{\phi}_{\mathrm{M,R}}, \boldsymbol{\rho}_{\mathrm{M,R}}\}$. The MSEs of the channel parameters in $\boldsymbol{\zeta}$ are lower bounded by CRLB($\boldsymbol{\zeta}$), which is formulated as

$$\mathrm{CRLB}(\boldsymbol{\zeta}) \geq \mathbf{J}(\boldsymbol{\zeta})^{-1}, \tag{44}$$

with $\mathbf{J}(\boldsymbol{\zeta})$ being the FIM. For instance, the observation vector $\mathbf{y}_{\mathrm{H}} = \mathrm{vec}(\mathbf{Y}_{\mathrm{H}})$ follows Gaussian distribution with $\mathcal{CN}(\boldsymbol{\mu}_1, \sigma^2\mathbf{I}_{MKT})$, where the mean $\boldsymbol{\mu}_1$ is expressed by

$$\boldsymbol{\mu}_1 = \sum_{l=1}^{L_{\mathrm{M,R}}} [\boldsymbol{\rho}_{\mathrm{M,R}}]_l (\mathbf{X}^{\mathsf{T}} \otimes \mathbf{W}_{\mathrm{H}}) \left[\boldsymbol{\alpha}^*([\boldsymbol{\theta}_{\mathrm{M,R}}]_l) \otimes \boldsymbol{\alpha}([\boldsymbol{\phi}_{\mathrm{M,R}}]_l)\right]. \tag{45}$$

With this assumption, we can express the (l, m)th entry of the FIM as

$$[\mathbf{J}(\boldsymbol{\zeta})]_{l,m} = \frac{2}{\sigma^2} \Re\left\{ \left(\frac{\partial \boldsymbol{\mu}_1}{\partial [\boldsymbol{\zeta}]_l}\right)^{\mathsf{H}} \left(\frac{\partial \boldsymbol{\mu}_1}{\partial [\boldsymbol{\zeta}]_m}\right) \right\}. \tag{46}$$

In the next step, we describe the derivatives of $\boldsymbol{\mu}_1$ with respect to the channel parameters in $\boldsymbol{\zeta}$ and further derive the entries of the FIM.

A. Partial Derivatives

First, we develop the partial derivatives of $\boldsymbol{\mu}_1$ with respect to the channel parameter $\boldsymbol{\theta}_{\mathrm{M,R}}$.

$$\left(\frac{\partial \boldsymbol{\mu}_1}{\partial [\boldsymbol{\theta}_{\mathrm{M,R}}]_l}\right) = (\mathbf{X}^{\mathsf{T}} \otimes \mathbf{W}_{\mathrm{H}})([\boldsymbol{\rho}_{\mathrm{M,R}}]_l) j\pi \cos([\boldsymbol{\theta}_{\mathrm{M,R}}]_l) \left[\tilde{\boldsymbol{\alpha}}^*([\boldsymbol{\theta}_{\mathrm{M,R}}]_l) \otimes \boldsymbol{\alpha}([\boldsymbol{\phi}_{\mathrm{M,R}}]_l)\right], \tag{47}$$

where $[\tilde{\boldsymbol{\alpha}}([\boldsymbol{\theta}_{\mathrm{M,R}}]_l)]_n = (n-1)\exp\{j\pi(n-1)\sin([\boldsymbol{\theta}_{\mathrm{M,R}}]_l)\}$ for $n = 1, \cdots, N_{\mathrm{M}}$. Similarly, we obtain the partial derivatives with respect to the remaining parameters as follows

$$\left(\frac{\partial \boldsymbol{\mu}_1}{\partial [\boldsymbol{\phi}_{\mathrm{M,R}}]_l}\right) = (\mathbf{X}^{\mathsf{T}} \otimes \mathbf{W}_{\mathrm{H}})([\boldsymbol{\rho}_{\mathrm{M,R}}]_l) j\pi \cos([\boldsymbol{\phi}_{\mathrm{M,R}}]_l) \big[\boldsymbol{\alpha}^*([\boldsymbol{\theta}_{\mathrm{M,R}}]_l) \otimes \tilde{\boldsymbol{\alpha}}([\boldsymbol{\phi}_{\mathrm{M,R}}]_l)\big], \tag{48}$$

$$\left(\frac{\partial \boldsymbol{\mu}_1}{\partial [\boldsymbol{\rho}_{\mathrm{M,R}}]_l}\right) = (\mathbf{X}^{\mathsf{T}} \otimes \mathbf{W}_{\mathrm{H}})\big[\boldsymbol{\alpha}^*([\boldsymbol{\theta}_{\mathrm{M,R}}]_l) \otimes \boldsymbol{\alpha}([\boldsymbol{\phi}_{\mathrm{M,R}}]_l)\big], \tag{49}$$

where $[\tilde{\boldsymbol{\alpha}}([\boldsymbol{\phi}_{\mathrm{M,R}}]_l)]_n = (n-1)\exp\{j\pi(n-1)\sin([\boldsymbol{\phi}_{\mathrm{M,R}}]_l)\}$ for $n = 1, \cdots, N_{\mathrm{R}}$. The derivatives $\left(\frac{\partial \boldsymbol{\mu}_1}{\partial [\boldsymbol{\theta}_{\mathrm{M,R}}]_l}\right)^{\mathsf{H}}$, $\left(\frac{\partial \boldsymbol{\mu}_1}{\partial [\boldsymbol{\phi}_{\mathrm{M,R}}]_l}\right)^{\mathsf{H}}$, and $\left(\frac{\partial \boldsymbol{\mu}_1}{\partial [\boldsymbol{\rho}_{\mathrm{M,R}}]_l}\right)^{\mathsf{H}}$ can be obtained by following the equations developed above and by applying the conjugate transpose.

B. Calculation of Fisher Information Matrix

We first calculate the entries of the elements in the principal diagonal. For instance, the (l,l)th entry of $\mathbf{J}(\boldsymbol{\zeta})$ is given by

$$\mathbf{J}([\boldsymbol{\theta}_{\mathrm{M,R}}]_l, [\boldsymbol{\theta}_{\mathrm{M,R}}]_l) = \frac{2}{\sigma^2}\left\{\left(\frac{\partial \boldsymbol{\mu}_1}{\partial [\boldsymbol{\theta}_{\mathrm{M,R}}]_l}\right)^{\mathsf{H}}\left(\frac{\partial \boldsymbol{\mu}_1}{\partial [\boldsymbol{\theta}_{\mathrm{M,R}}]_l}\right)\right\}. \tag{50}$$

By substituting the derivatives of $\boldsymbol{\mu}_1$ with respect to $\boldsymbol{\theta}_{\mathrm{M,R}}$, we obtain

$$\mathbf{J}([\boldsymbol{\theta}_{\mathrm{M,R}}]_l, [\boldsymbol{\theta}_{\mathrm{M,R}}]_l) = \frac{2}{\sigma^2}\left\{([\boldsymbol{\rho}_{\mathrm{M,R}}]_l)^2\pi^2\cos^2([\boldsymbol{\theta}_{\mathrm{M,R}}]_l)\big[\boldsymbol{\chi}_1\boldsymbol{\xi}_1\boldsymbol{\kappa}_1\big]\right\}, \tag{51}$$

where $\boldsymbol{\chi}_1 = \big[\tilde{\boldsymbol{\alpha}}^{\mathsf{T}}([\boldsymbol{\theta}_{\mathrm{M,R}}]_l) \otimes \boldsymbol{\alpha}^{\mathsf{H}}([\boldsymbol{\phi}_{\mathrm{M,R}}]_l)\big]$, $\boldsymbol{\xi}_1 = (\mathbf{X}^{\mathsf{T}} \otimes \mathbf{W}_{\mathrm{H}})^{\mathsf{H}}(\mathbf{X}^{\mathsf{T}} \otimes \mathbf{W}_{\mathrm{H}})$, and $\boldsymbol{\kappa}_1 = \big[\tilde{\boldsymbol{\alpha}}^*([\boldsymbol{\theta}_{\mathrm{M,R}}]_l) \otimes \boldsymbol{\alpha}([\boldsymbol{\phi}_{\mathrm{M,R}}]_l)\big]$. Similarly, we can obtain

$$\begin{aligned}\mathbf{J}([\boldsymbol{\phi}_{\mathrm{M,R}}]_l, [\boldsymbol{\phi}_{\mathrm{M,R}}]_l) &= \frac{2}{\sigma^2}\left\{\left(\frac{\partial \boldsymbol{\mu}_1}{\partial [\boldsymbol{\phi}_{\mathrm{M,R}}]_l}\right)^{\mathsf{H}}\left(\frac{\partial \boldsymbol{\mu}_1}{\partial [\boldsymbol{\phi}_{\mathrm{M,R}}]_l}\right)\right\} \\ &= \frac{2}{\sigma^2}([\boldsymbol{\rho}_{\mathrm{M,R}}]_l)^2\pi^2\cos^2([\boldsymbol{\phi}_{\mathrm{M,R}}]_l)\big[\boldsymbol{\chi}_2\boldsymbol{\xi}_1\boldsymbol{\kappa}_2\big],\end{aligned} \tag{52}$$

$$\begin{aligned}\mathbf{J}([\boldsymbol{\rho}_{\mathrm{M,R}}]_l, [\boldsymbol{\rho}_{\mathrm{M,R}}]_l) &= \frac{2}{\sigma^2}\left\{\left(\frac{\partial \boldsymbol{\mu}_1}{\partial [\boldsymbol{\rho}_{\mathrm{M,R}}]_l}\right)^{\mathsf{H}}\left(\frac{\partial \boldsymbol{\mu}_1}{\partial [\boldsymbol{\rho}_{\mathrm{M,R}}]_l}\right)\right\} \\ &= \frac{2}{\sigma^2}\big[\boldsymbol{\chi}_3\boldsymbol{\xi}_1\boldsymbol{\kappa}_3\big],\end{aligned} \tag{53}$$

where $\boldsymbol{\chi}_2 = \big[\boldsymbol{\alpha}^{\mathsf{T}}([\boldsymbol{\theta}_{\mathrm{M,R}}]_l) \otimes \tilde{\boldsymbol{\alpha}}^{\mathsf{H}}([\boldsymbol{\phi}_{\mathrm{M,R}}]_l)\big]$, $\boldsymbol{\kappa}_2 = \big[\boldsymbol{\alpha}^*([\boldsymbol{\theta}_{\mathrm{M,R}}]_l) \otimes \tilde{\boldsymbol{\alpha}}([\boldsymbol{\phi}_{\mathrm{M,R}}]_l)\big]$, $\boldsymbol{\chi}_3 = \big[\boldsymbol{\alpha}^{\mathsf{T}}([\boldsymbol{\theta}_{\mathrm{M,R}}]_l) \otimes \boldsymbol{\alpha}^{\mathsf{H}}([\boldsymbol{\phi}_{\mathrm{M,R}}]_l)\big]$, and $\boldsymbol{\kappa}_3 = \big[\boldsymbol{\alpha}^*([\boldsymbol{\theta}_{\mathrm{M,R}}]_l) \otimes \boldsymbol{\alpha}([\boldsymbol{\phi}_{\mathrm{M,R}}]_l)\big]$. Note that the calculations of the (l,m)th off-diagonal entries follow the same procedure.

Appendix B
Calculation of the CRLB: Stage 2 of CE

At the second CE stage, we target at recovering the parameters of the RIS-BS channel, defined as $\boldsymbol{\eta} = \{\boldsymbol{\theta}_{\mathrm{R,B}}, \boldsymbol{\phi}_{\mathrm{R,B}}, \boldsymbol{\rho}_{\mathrm{R,B}}\}$. Note that the procedure to obtain the elements of the FIM and the corresponding CRLB is similar to the analysis in Appendice A. We assume that the observation vector $\mathbf{y} = \mathrm{vec}(\mathbf{Y})$ follows Gaussian distribution with $\mathcal{CN}(\boldsymbol{\mu}_2, \sigma^2\mathbf{I}_{TKN_{\mathrm{C,B}}})$, where $\boldsymbol{\mu}_2$ is defined by

$$\boldsymbol{\mu}_2 = \sum_{l=1}^{L_{\mathrm{R,B}}}[\boldsymbol{\rho}_{\mathrm{R,B}}]_l(\tilde{\mathbf{U}}^{\mathsf{T}} \otimes \mathbf{W}_{\mathrm{B}}^{\mathsf{H}})\big(\boldsymbol{\alpha}^*([\boldsymbol{\theta}_{\mathrm{R,B}}]_l) \otimes \boldsymbol{\alpha}([\boldsymbol{\phi}_{\mathrm{R,B}}]_l)\big). \tag{54}$$

By adopting this assumption, we can formulate the FIM $\mathbf{J}(\boldsymbol{\eta})$ as

$$[\mathbf{J}(\boldsymbol{\eta})]_{l_1,l_2} = \frac{2}{\sigma^2}\Re\left\{\left(\frac{\partial \boldsymbol{\mu}_2}{\partial [\boldsymbol{\eta}]_{l_1}}\right)^{\mathsf{H}}\left(\frac{\partial \boldsymbol{\mu}_2}{\partial [\boldsymbol{\eta}]_{l_2}}\right)\right\}. \tag{55}$$

In the following, we develop the derivatives of $\boldsymbol{\mu}_2$ with respect to $\boldsymbol{\eta}$ in detail.

A. Partial Derivatives

We describe the derivatives of $\boldsymbol{\mu}_2$ with respect to the channel parameters in $\boldsymbol{\eta}$.

$$\left(\frac{\partial \boldsymbol{\mu}_2}{\partial [\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}}\right) = (\tilde{\mathbf{U}}^{\mathsf{T}} \otimes \mathbf{W}_{\mathrm{B}}^{\mathsf{H}})([\boldsymbol{\rho}_{\mathrm{R,B}}]_l) j\pi \cos([\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}) \big[\tilde{\boldsymbol{\alpha}}^*([\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}) \otimes \boldsymbol{\alpha}([\boldsymbol{\phi}_{\mathrm{R,B}}]_{l_1})\big], \tag{56, 57}$$

$$\left(\frac{\partial \boldsymbol{\mu}_2}{\partial [\boldsymbol{\phi}_{\mathrm{R,B}}]_{l_1}}\right) = (\tilde{\mathbf{U}}^{\mathsf{T}} \otimes \mathbf{W}_{\mathrm{B}}^{\mathsf{H}})([\boldsymbol{\rho}_{\mathrm{R,B}}]_{l_1}) j\pi \cos([\boldsymbol{\phi}_{\mathrm{R,B}}]_{l_1}) \big[\boldsymbol{\alpha}^*([\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}) \otimes \tilde{\boldsymbol{\alpha}}([\boldsymbol{\phi}_{\mathrm{R,B}}]_{l_1})\big], \tag{58, 59}$$

$$\left(\frac{\partial \boldsymbol{\mu}_2}{\partial [\boldsymbol{\rho}_{\mathrm{R,B}}]_{l_1}}\right) = (\tilde{\mathbf{U}}^{\mathsf{T}} \otimes \mathbf{W}_{\mathrm{B}}^{\mathsf{H}})\big[\boldsymbol{\alpha}^*([\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}) \otimes \boldsymbol{\alpha}([\boldsymbol{\phi}_{\mathrm{R,B}}]_{l_1})\big], \tag{60}$$

where $[\tilde{\boldsymbol{\alpha}}([\boldsymbol{\phi}_{\mathrm{R,B}}]_{l_1})]_n = (n-1)\exp\{j\pi(n-1)\sin([\boldsymbol{\phi}_{\mathrm{R,B}}]_{l_1})\}$ for $n = 1, \cdots, N_{\mathrm{B}}$ and $[\tilde{\boldsymbol{\alpha}}([\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1})]_n = (n-1)\exp\{j\pi(n-1)\sin([\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1})\}$ for $n = 1, \cdots, N_{\mathrm{R}}$.

B. Calculation of Fisher Information Matrix

For sake of brevity, we only define the entries of the elements on the main diagonal of the FIM as

$$\mathbf{J}([\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}, [\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}) = \frac{2}{\sigma^2}\left\{\left(\frac{\partial \boldsymbol{\mu}_2}{\partial [\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}}\right)^{\mathsf{H}}\left(\frac{\partial \boldsymbol{\mu}_2}{\partial [\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}}\right)\right\}, \tag{61}$$

$$\mathbf{J}([\boldsymbol{\phi}_{\mathrm{R,B}}]_{l_1}, [\boldsymbol{\phi}_{\mathrm{R,B}}]_{l_1}) = \frac{2}{\sigma^2}\left\{\left(\frac{\partial \boldsymbol{\mu}_2}{\partial [\boldsymbol{\phi}_{\mathrm{R,B}}]_{l_1}}\right)^{\mathsf{H}}\left(\frac{\partial \boldsymbol{\mu}_2}{\partial [\boldsymbol{\phi}_{\mathrm{R,B}}]_{l_1}}\right)\right\}, \tag{62}$$

$$\mathbf{J}([\boldsymbol{\rho}_{\mathrm{R,B}}]_{l_1}, [\boldsymbol{\rho}_{\mathrm{R,B}}]_{l_1}) = \frac{2}{\sigma^2}\left\{\left(\frac{\partial \boldsymbol{\mu}_2}{\partial [\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}}\right)^{\mathsf{H}}\left(\frac{\partial \boldsymbol{\mu}_2}{\partial [\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}}\right)\right\}. \tag{63}$$

By applying the derivatives of $\boldsymbol{\mu}_2$ with respect to $\boldsymbol{\eta}$, we obtain

$$\begin{aligned}&\mathbf{J}([\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}, [\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}) \\ &\quad= \frac{2}{\sigma^2}\left\{([\boldsymbol{\rho}_{\mathrm{R,B}}]_{l_1})^2\pi^2\cos^2([\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1})\big[\boldsymbol{\chi}_4\boldsymbol{\xi}_2\boldsymbol{\kappa}_4\big]\right\},\end{aligned} \tag{64}$$

where $\boldsymbol{\chi}_4 = \big[\boldsymbol{\alpha}^{\mathsf{T}}([\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}) \otimes \tilde{\boldsymbol{\alpha}}^{\mathsf{H}}([\boldsymbol{\phi}_{\mathrm{R,B}}]_{l_1})\big]$, $\boldsymbol{\xi}_2 = (\tilde{\mathbf{U}}^{\mathsf{T}} \otimes \mathbf{W}_{\mathrm{B}}^{\mathsf{H}})^{\mathsf{H}}(\tilde{\mathbf{U}}^{\mathsf{T}} \otimes \mathbf{W}_{\mathrm{B}}^{\mathsf{H}})$, and $\boldsymbol{\kappa}_4 = \big[\boldsymbol{\alpha}^*([\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}) \otimes \tilde{\boldsymbol{\alpha}}([\boldsymbol{\phi}_{\mathrm{R,B}}]_{l_1})\big]$.

In a similar procedure, we can re-write (62) and (63), respectively, as

$$\mathbf{J}([\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1},[\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}) = \frac{2}{\sigma^2}([\boldsymbol{\rho}_{\mathrm{R,B}}]_{l_1})^2\pi^2\cos^2([\boldsymbol{\phi}_{\mathrm{R,B}}]_{l_1})\left[\boldsymbol{\chi}_5\boldsymbol{\xi}_2\boldsymbol{\kappa}_5\right], \tag{65}$$

$$\mathbf{J}([\boldsymbol{\rho}_{\mathrm{R,B}}]_{l_1},[\boldsymbol{\rho}_{\mathrm{R,B}}]_{l_1}) = \frac{2}{\sigma^2}\left[\boldsymbol{\chi}_6\boldsymbol{\xi}_2\boldsymbol{\kappa}_6\right], \tag{66}$$

where $\boldsymbol{\chi}_5 = \left[\boldsymbol{\alpha}^{\mathsf{T}}([\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}) \otimes \tilde{\boldsymbol{\alpha}}^{\mathsf{H}}([\boldsymbol{\phi}_{\mathrm{R,B}}]_{l_1})\right]$, $\boldsymbol{\kappa}_5 = \left[\boldsymbol{\alpha}^{*}([\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}) \otimes \tilde{\boldsymbol{\alpha}}([\boldsymbol{\phi}_{\mathrm{R,B}}]_{l_1})\right]$, $\boldsymbol{\chi}_6 = \left[\boldsymbol{\alpha}^{\mathsf{T}}([\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}) \otimes \boldsymbol{\alpha}^{\mathsf{H}}([\boldsymbol{\phi}_{\mathrm{R,B}}]_{l_1})\right]$, and $\boldsymbol{\kappa}_6 = \left[\boldsymbol{\alpha}^{*}([\boldsymbol{\theta}_{\mathrm{R,B}}]_{l_1}) \otimes \boldsymbol{\alpha}([\boldsymbol{\phi}_{\mathrm{R,B}}]_{l_1})\right]$. The remaining entries of the FIM can be derived by following the same procedure.

REFERENCES

[1] R. W. Heath, Jr., 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. 2017.

[2] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[3] V. Raghavan *et al.*, "Statistical blockage modeling and robustness of beamforming in millimeter-wave systems," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 7, pp. 3010–3024, Jul. 2019.

[4] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, Jr., "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[5] 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.

[6] J. He, M. Leinonen, H. Wymeersch, and M. Juntti, "Channel estimation for RIS-aided mmWave MIMO systems," in *Proc. IEEE GLOBECOM*, Dec. 2020, pp. 1–6.

[7] 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, Nov. 2019.

[8] B. Li, Z. Zhang, Z. Hu, and Y. Chen, "Joint array diagnosis and channel estimation for RIS-aided mmWave MIMO system," *IEEE Access*, vol. 8, pp. 193992–194006, 2020.

[9] 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. 116753–116773, 2019.

[10] 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.

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

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

[13] L. Dong and H.-M. Wang, "Secure MIMO transmission via intelligent reflecting surface," *IEEE Wireless Commun. Lett.* vol. 9, no. 6, pp. 787–790, Jun. 2020.

[14] Z. Chu, W. Hao, P. Xiao, and J. Shi, "Intelligent reflecting surface aided multi-antenna secure transmission," *IEEE Wireless Commun. Lett.*, vol. 9, no. 1, pp. 108–112, Jan. 2020.

[15] H. Wymeersch, J. He, B. Denis, A. Clemente, and M. Juntti, "Radio localization and mapping with reconfigurable intelligent surfaces: Challenges, opportunities, and research directions," *IEEE Veh. Technol. Mag.*, vol. 15, no. 4, pp. 52–61, Dec. 2020.

[16] J. He, H. Wymeersch, T. Sanguanpuak, O. Silven, and M. Juntti, "Adaptive beamforming design for mmWave RIS-aided joint localization and communication," in *Proc. IEEE WCNCW*, Apr. 2020, pp. 1–6.

[17] S. Abeywickrama, R. Zhang, Q. Wu, and C. Yuen, "Intelligent reflecting surface: Practical phase shift model and beamforming optimization," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5849–5863, Sep. 2020.

[18] J. He, H. Wymeersch, and M. Juntti, "Channel estimation for RIS-aided mmWave MIMO systems via atomic norm minimization," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 5786–5797, Sep. 2021.

[19] Z.-Q. He and X. Yuan, "Cascaded channel estimation for large intelligent metasurface assisted massive MIMO," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 210–214, Feb. 2020.

[20] J. Mirza and B. Ali, "Channel estimation method and phase shift design for reconfigurable intelligent surface assisted MIMO networks," *IEEE Trans. Cogn. Commun. Netw.*, vol. 7, no. 2, pp. 441–451, Jun. 2021.

[21] 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, Apr. 2021.

[22] C. You, B. Zheng, and R. Zhang, "Intelligent reflecting surface with discrete phase shifts: Channel estimation and passive beamforming," in *Proc. IEEE ICC*, Jun. 2020, pp. 1–6.

[23] 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.

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

[25] G. C. Alexandropoulos and E. Vlachos, "A hardware architecture for reconfigurable intelligent surfaces with minimal active elements for explicit channel estimation," in *Proc. IEEE ICASSP*, May 2020, pp. 9175–9179.

[26] R. Schroeder, J. He, and M. Juntti, "Passive RIS vs. hybrid RIS: A comparative study on channel estimation," in *Proc. IEEE VTC*, Apr. 2021, pp. 1–7.

[27] N. T. Nguyen, D. Vu, K. Lee, and M. Juntti, "Hybrid relay-reflecting intelligent surface-assisted wireless communications," *IEEE Trans. Veh. Technol.*, early access, Mar. 11, 2022, doi: 10.1109/TVT.2022.3158686.

[28] Y. Lin, S. Jin, M. Matthaiou, and X. You, "Tensor-based algebraic channel estimation for hybrid IRS-assisted MIMO-OFDM," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3770–3784, Jun. 2021.

[29] G. C. Alexandropoulos, N. Shlezinger, I. Alamzadeh, M. F. Imani, H. Zhang, and Y. C. Eldar, "Hybrid reconfigurable intelligent metasurfaces: Enabling simultaneous tunable reflections and sensing for 6G wireless communications," 2021, *arXiv:2104.04690*.

[30] R. Schroeder, J. He, and M. Juntti, "Channel estimation for hybrid RIS aided MIMO communications via atomic norm minimization," 2021, *arXiv:2106.10909*.

[31] 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 ICASSP*, May 2020, pp. 5000–5004.

[32] Q.-U.-A. Nadeem, H. Alwazani, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. 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.

[33] J. Lee, G.-T. Gil, and Y. H. Lee, "Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2370–2386, Jun. 2016.

[34] Y. Chi and M. F. Da Costa, "Harnessing sparsity over the continuum: Atomic norm minimization for superresolution," *IEEE Signal Process. Mag.*, vol. 37, no. 2, pp. 39–57, Mar. 2020.

[35] Z. Yang and L. Xie, "On gridless sparse methods for line spectral estimation from complete and incomplete data," *IEEE Trans. Signal Process.*, vol. 63, no. 12, pp. 3139–3153, Jun. 2015.

[36] Y. Tsai, L. Zheng, and X. Wang, "Millimeter-wave beamformed full-dimensional MIMO channel estimation based on atomic norm minimization," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6150–6163, Dec. 2018.

[37] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 276–280, Mar. 1986.

[38] Z. Cheng, M. Tao, and P.-Y. Kam, "Channel path identification in mmWave systems with large-scale antenna arrays," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5549–5562, Sep. 2020.

[39] Z. Zhang, Y. Wang, and Z. Tian, "Efficient two-dimensional line spectrum estimation based on decoupled atomic norm minimization," *Signal Process.*, vol. 163, pp. 95–106, Oct. 2019.

[40] J. Zhang, Y. Zhang, C. Zhong, and Z. Zhang, "Robust design for intelligent reflecting surfaces assisted MISO systems," *IEEE Commun. Lett.*, vol. 24, no. 10, pp. 2353–2357, Oct. 2020.

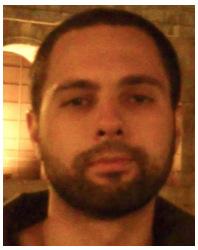
[41] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

[42] L. S. Pillutla and R. Annavajjala, "Bayesian CRLB for joint AoA, AoD and multipath gain estimation in millimeter wave wireless networks," in *Proc. IEEE GLOBECOM*, Dec. 2017, pp. 1–6.

[43] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2012.
[44] K. Venugopal, A. Alkhateeb, N. González-Prelcic, and R. W. Heath, Jr., "Channel estimation for hybrid architecture-based wideband millimeter wave systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1996–2009, Sep. 2017.

Rafaela Schroeder (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from the Federal University of Technology—Paraná (UTFPR), Curitiba, Brazil, in 2017 and 2019, respectively. She is currently pursuing the doctoral degree with the Centre for Wireless Communications (CWC), University of Oulu. Her research interests include millimeter wave MIMO communications, reconfigurable intelligent surfaces, channel estimation, and localization.

Jiguang He (Member, IEEE) received the Ph.D. degree in communications engineering from the University of Oulu, Finland, in 2018. From September 2013 to March 2015, he was with the Key Laboratory of Millimeter Waves, City University of Hong Kong, working on beam tracking over millimeter wave MIMO systems. From June 2015 to August 2021, he was with the Centre for Wireless Communications (CWC), University of Oulu, first as a Doctoral Researcher and then a Post-Doctoral Researcher. He was an Assistant Professor with the Macau University of Science and Technology from August 2021 to March 2022. He is currently a Senior Researcher with the Technology Innovation Institute, Abu Dhabi, United Arab Emirates; and also a Docent (Adjunct Professor) with the University of Oulu. He has participated in many international and national projects, e.g., EU FP7 RESCUE, EU H2020 ARIADNE, and 6G Flagship. His research interests span millimeter wave MIMO communications, reconfigurable intelligent surfaces for simultaneous localization and communication (SLAC), and advanced signal processing techniques.



Glauber Brante (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the Federal University of Technology—Paraná (UTFPR), Curitiba, Brazil, in 2013. In 2012, he was a Visiting Researcher with the Institute of Information and Communication Technologies, Electronics, and Applied Mathematics, Catholic University of Louvain, Belgium. He is currently an Assistant Professor with UTFPR. His research interests include wireless communications, machine-type communications, energy efficiency, and energy harvesting. He was a recipient of the Best Ph.D. Thesis Award in electrical engineering in Brazil in 2014. He was a co-recipient of the 2016 Research Award from the Cuban Academy of Sciences. He was the Co-Editor-in-Chief of the *Journal of Communication and Information Systems* from 2018 to 2020. Since 2018, he has been serving as an Associate Editor for the IEEE COMMUNICATIONS LETTERS and has been serving as an Associate Editor for the IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY since 2019.

Markku Juntti (Fellow, IEEE) received the M.Sc.E.E. and Dr.Sc.E.E. degrees from the University of Oulu, Oulu, Finland, in 1993 and 1997, respectively.

He was with the University of Oulu from 1992 to 1998. From 1994 to 1995, he was a Visiting Scholar with Rice University, Houston, TX, USA. From 1999 to 2000, he was a Senior Specialist with Nokia Networks, Oulu. He has been a Professor of communications engineering with the Centre for Wireless Communications (CWC), University of Oulu, since 2000, where he leads the Communications Signal Processing (CSP) Research Group. He also serves as the Head for the CWC—Radio Technologies (RT) Research Unit. He is also an Adjunct Professor with the Department of Electrical and Computer Engineering, Rice University. He is an author or coauthor in almost 500 papers published in international journals and conference records as well as in books *Wideband CDMA for UMTS* from 2000 to 2010, *Handbook of Signal Processing Systems* in 2013 and 2018 and *5G Wireless Technologies* in 2017. His research interests include signal processing for wireless networks as well as communication and information theory. He was a Secretary of IEEE Communication Society Finland Chapter from 1996 to 1997 and the Chairman from 2000 to 2001. He has been a Secretary of the Technical Program Committee (TPC) of the 2001 IEEE International Conference on Communications (ICC) and the Chair or Co-Chair of the Technical Program Committee of several conferences including 2006 and 2021 IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), the Signal Processing for Communications Symposium of IEEE Globecom 2014, Symposium on Transceivers and Signal Processing for 5G Wireless and mm-Wave Systems of IEEE GlobalSIP 2016, ACM NanoCom 2018, and 2019 International Symposium on Wireless Communication Systems (ISWCS). He has also served as the General Chair for 2011 IEEE Communication Theory Workshop (CTW 2011) and 2022 IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC). He was an Editor of IEEE TRANSACTIONS ON COMMUNICATIONS and an Associate Editor of IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.