Channel Estimation for Hybrid RIS Aided MIMO Communications via Atomic Norm Minimization

Rafaela Schroeder*, Jiguang He*, and Markku Juntti*
*Centre for Wireless Communications, FI-90014, University of Oulu, Finland
E-mail:{rafaela.schroeder, jiguang.he, markku.juntti}@oulu.fi

Abstract—Reconfigurable intelligent surfaces (RISs) have been introduced as a remedy for mitigating blockages in millimeter wave (mmWave) and terahertz (THz) communications networks. However, perfect or nearly perfect channel state information (CSI) is fundamental in order to achieve their full potential. Traditionally, an RIS is fully passive without any baseband processing capabilities, which poses great challenges for CSI acquisition. Thus, we focus on the hybrid RIS architecture, where a small portion of RIS elements are active and able to processing the received pilot signals for estimating the corresponding channel. The channel estimation (CE) is done by resorting to off-the-grid compressive sensing technique, i.e., atomic norm minimization, for extracting channel parameters through two stages. Simulation results show that the proposed scheme outperforms the passive RIS CE under the same training overhead.

I. INTRODUCTION

Terahertz (THz) and millimeter-wave (mmWave) multiple-input multiple-output (MIMO) communications are a promising technology to address the ever-increasing demand for high data rate in 5G and beyond [1], [2], [3]. Large antenna arrays are essential at both the transmitter and receiver to compensate for the severe path loss. It is well known that the mmWave/THz exhibit features, such as, frequent blockages, inherent sparsity and large path loss. Specially, THz channels suffer severe path loss due to the higher frequencies [2].

Reconfigurable intelligent surface (RIS) has emerged as a solution to mitigate the blockages in mmWave/THz MIMO systems. The RIS consists of a large number of elements that can be controlled to modify its signal response in order to achieve a certain objective, for example, reflecting the signals towards the receiver [4]. Namely, the RIS enables the control of the wireless propagation environment, and it has been verified as an innovative technology to enhance indoor localization [5], security [6] and improve the performance of mmWave/THz MIMO systems, in terms of spectral efficiency (SE) [7], [8].

The CSI acquisition plays an important role in design of the RIS phase control matrix and active beamforming vectors at the transceivers. Compared with MIMO systems, CSI acquisition becomes more challenging in RIS-aided ones due to increased complexity. In the literature, CSI acquisition has been investigated for both RIS architectures, i.e., *passive RIS* and *hybrid RIS*. The architecture with both passive and *active* elements is called *hybrid* RIS. The channel estimation (CE) methods for passive RIS-aided mmWave/THz MIMO systems have been intensively studied in [4], [9], [10], [11],

[12], [13]. Ardah *et al.* [4] proposed a two-stage procedure for CE of mmWave MIMO systems with a RIS. The CE problem was solved by following the parallel factor (PARAFAC) tensor decomposition in [13]. In our recent work [9], we also considered two-stage CE with the aid of atomic norm minimization (ANM).

In order to ease the CE procedure, Taha *et al.* [14] introduced a few *active* elements at the RIS with baseband capabilities, which allow the CE at the RIS despite the higher power consumption. Inspired by [14], the hybrid RIS architecture has been further studied in [15], [16], [17], [18]. In [16], we made a comparison in terms of CE between the passive and hybrid architectures for MIMO systems, which concludes that passive RIS is superior to the hybrid one when CE is done at the RIS.

In this paper, we develop a new two-stage CE method under the hybrid RIS architecture. Unlike our previous work [16], the proposed CE requires fewer number of active RIS elements and only one-way training (either uplink or downlink). We consider uplink training, where the mobile station (MS) sends pilots to the base station (BS) via the RIS. The pilot signal is also received at the RIS by the active elements, while the remaining elements reflect the signal towards the BS. In the first CE stage, we aim at recovering the parameters of the MS-RIS channel via ANM based on the received signal at the RIS. In the second CE stage, we aim at recovering the parameters of RIS-BS channel from the received signal at the BS. We evaluate the performance in terms of the mean square error (MSE) of the angle of departure (AoD), angle of arrival (AoA), and the products of path gains, and SE. In the evaluation, we also incorporate the path loss model, which has a nonnegligible impact on CE. It is verified that our proposed method can simplify the CSI acquisition for RIS aided mmWave/Thz MIMO systems with a small number of active RIS elements and bring better CE performance compared with that for the passive RIS [9].

Notation: A bold capital letter $\mathbf{A}$ denotes a matrix and a lowercase letter $\mathbf{a}$ denotes a column vector, and $()^{\mathsf{H}}$, $()^*$, and $()^{\mathsf{T}}$ denote the Hermitian transpose, conjugate, and transpose, 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. System Model

We assume a RIS-assisted MIMO system with one multi-antenna BS, one multi-element RIS, and one multi-antenna MS. The number of antennas or reflecting elements at the BS, RIS, and MS are denoted by N_{B}, N_{R} and N_{M}, respectively. We assume that the direct MS-BS channel suffers from blockage, which motivates the use of the RIS to assist the data transmission between the BS and the MS. We adopt an uniform linear array (ULA) for the antennas/elements in this paper, while it is possible to be extended for an uniform planar array (UPA). Regarding CE, we consider uplink training with pilot transmission from the MS.

A. Channel Model

The propagation channel is composed 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 the channel parameters stay constant during the coherence time. By adopting the geometric channel model, the channel between MS and RIS $\mathbf{H}_{\mathrm{M,R}}$ is

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

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 $j = \sqrt{-1}$. The array response matrices $\mathbf{A}(\boldsymbol{\theta}_{\mathrm{M,R}})$ and $\mathbf{A}(\boldsymbol{\phi}_{\mathrm{M,R}})$ are expressed as

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

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

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

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

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. The entire MS-RIS-BS channel based on (1) and (4) is expressed as

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

where $\boldsymbol{\Omega} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{R}}}$ is the phase control matrix at the RIS with unit-modulus constraint on the diagonal. We assume that
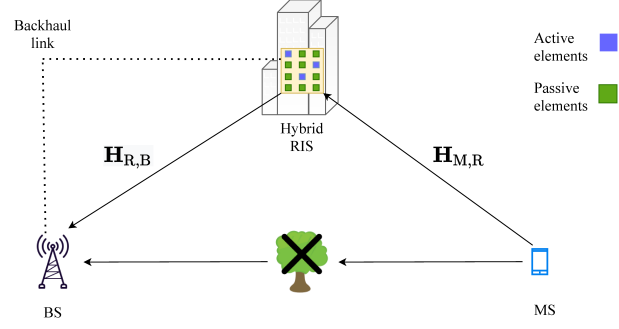

Fig. 1. Hybrid RIS-aided communication system.

the RIS is composed of discrete phase sifters. Thus, the phase control matrix is $[\boldsymbol{\Omega}]_{k,k} = \exp(j\omega_k)$, where $\omega \in [0, 2\pi]$. We also define the effective channel $\mathbf{G} \in \mathbb{C}^{L_{\mathrm{R,B}} \times L_{\mathrm{M,R}}}$ as

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

which depends on the angle differences associated with the RIS, products of path gains, and the RIS phase control matrix.

B. Passive RIS

The passive RIS without any baseband capabilities has been explored in the literature [12], [4], [10]. CE for the passive RIS is more challenging because the individual channels in (1) and (4) can only be estimated at the MS (downlink) or BS (uplink). With the assumption of block-fading channel, we divide one coherence time into two sub-intervals: one for CE and the other for data transmission. Moreover, the CE sub-interval is further divided into K blocks, and within kth block MS sends a series of training sequences i.e., column vectors of $\mathbf{X}_k$, for $k = 1, \cdots, K$ to the BS. The pilot signal is first reflected at the RIS by $\boldsymbol{\Omega}_k$, and further combined at the BS by $\mathbf{W}_k$. The received signal $\mathbf{Y}_{\mathrm{P}k}$ at the BS is summarized as follows:

$$\mathbf{Y}_{\mathrm{P}k} = \beta_2 \mathbf{W}_k^{\mathsf{H}} \mathbf{H}_{\mathrm{R,B}} \boldsymbol{\Omega}_k \mathbf{H}_{\mathrm{M,R}} \mathbf{X}_k + \mathbf{W}_k^{\mathsf{H}} \mathbf{Z}_k, \text{for } k = 1, \cdots, K, \quad (7)$$

where $\mathbf{Z}_k$ is the additive Gaussian noise with each entry distributed as $\mathcal{CN}(0, \sigma^2)$ and $\beta_2 = \sqrt{\frac{1}{\beta(d_1, d_2)}}$, where $\beta(d_1, d_2)$ denotes the overall path loss of the BS-RIS-MS link with d_1 and d_2 being the distance between the MS and RIS and that between the RIS and BS, respectively. We target at the recovery of the channel parameters in (1) and (4) from the received signal in (7).

C. Hybrid RIS

The hybrid RIS with baseband capabilities simplifies the CE procedure for RIS-aided MIMO systems [14]. The hybrid RIS enabled communication system is illustrated in the Fig. 1. Thanks to the availability of baseband capabilities, the two tandem channels can be observed and estimated individually at the RIS.

We consider M out of N_{R} RIS elements are active and $N_{\mathrm{RF,R}} \leq M$ RF chains are implemented at the RIS. We assume uplink training, where the MS sends the training matrix $\mathbf{X} \in \mathbb{C}^{N_{\mathrm{M}} \times T}$ to the BS via the RIS with T being the number of
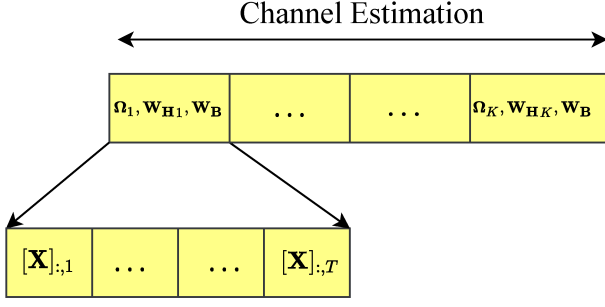
Fig. 2. Uplink training procedure for the hybrid RIS.

training beams. We further divide the CE sub-interval into K blocks, each block with T channel uses.

The pilot signal is received at RIS by the M active RIS elements indexed by set $\mathbb{A}$, i.e., $|\mathbb{A}| = M$, while the passive RIS elements reflect their received signals with a phase control matrix $\boldsymbol{\Omega}_k$. Note that here $\boldsymbol{\Omega}_k$ is different from that in passive RIS since $[\boldsymbol{\Omega}_k]_{i,i} = 0$ for $i \in \mathbb{A}$. After reflection, the signal is further combined at the BS by $\mathbf{W}_B \in \mathbb{C}^{N_B \times N_{C,B}}$, where $N_{C,B}$ is the number of columns of the combining matrix. . In the training procedure, we keep the combining matrix $\mathbf{W}_B$ and the training matrix $\mathbf{X}$ constant over all the blocks. However, the phase control matrix $\boldsymbol{\Omega}_k$ varies from block to block. Fig. 2 sketches the proposed uplink training procedure with a special focus on usage of training matrix, combining matrix, and RIS phase control matrix. The received signal at the hybrid RIS is summarized as

$$\mathbf{Y}_{Hk} = \beta_1 \mathbf{W}_{Hk} \mathbf{H}_{M,R} \mathbf{X} + \mathbf{W}_{Hk} \mathbf{Z}_{1k}, \text{ for } k = 1, \cdots, K, \quad (8)$$

where $\mathbf{W}_{Hk}$ is the row-selection matrix containing M rows of a $N_R \times N_R$ identity matrix. In the case of $N_{RF,R} < M$, further left-multiplying $\mathbf{W}_{Hk}$ by a combining matrix $\mathbf{M}_c \in \mathbb{C}^{N_{RF,R} \times M}$ is required, $\mathbf{Z}_{1k} \in \mathbb{C}^{N_R \times T}$ is the Gaussian noise with each entry distributed as $\mathcal{CN}(0, \sigma^2)$, and $\beta_1 = \sqrt{\frac{1}{\beta(d_1)}}$ with $\beta(d_1)$ being the path loss for the MS-RIS channel. The received signal at BS is expressed as

$$\mathbf{Y}_k = \beta_2 \mathbf{W}_B^H \mathbf{H}_{R,B} \boldsymbol{\Omega}_k \mathbf{H}_{M,R} \mathbf{X} + \mathbf{W}_B^H \mathbf{Z}_{2k}, \quad (9)$$

where $\mathbf{Z}_{2k} \in \mathbb{C}^{N_B \times T}$ is the Gaussian noise at the BS. We collect the received signals across all the K blocks at the RIS, as $[\mathbf{Y}_{H1}^T, \cdots, \mathbf{Y}_{HK}^T]^T \in \mathbb{C}^{MK \times T}$, which can be summarized as

$$\mathbf{Y}_H = \beta_1 \mathbf{W}_H \mathbf{H}_{M,R} \mathbf{X} + \bar{\mathbf{Z}}_1, \quad (10)$$

where $\mathbf{W}_H = [\mathbf{W}_{H1}^T, \cdots, \mathbf{W}_{HK}^T]^T \in \mathbb{C}^{MK \times N_R}$, $\bar{\mathbf{Z}}_1 = [(\mathbf{W}_{H1}\mathbf{Z}_{11})^T, \cdots, (\mathbf{W}_{HK}\mathbf{Z}_{1K})^T]^T \in \mathbb{C}^{MK \times T}$. The collected received signals at the BS $\mathbf{Y} = [\mathbf{Y}_1, \cdots, \mathbf{Y}_K] \in \mathbb{C}^{N_{C,B} \times TK}$ are expressed as

$$\mathbf{Y} = \beta_2 \mathbf{W}_B^H \mathbf{H}_{R,B} \mathbf{U} + \bar{\mathbf{Z}}_2, \quad (11)$$

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

We propose a two-stage CE to recover the channel parameters from the received signals in (10) and (11). To be specific, in the first stage, we extract the channel parameters in the

$\mathbf{H}_{M,R}$ from the received signals at the RIS (10). After that, the RIS sends the estimates of the channel parameters to the BS via the error-free backhaul link, and the BS reconstructs $\mathbf{H}_{M,R}$ based on the estimates with the reconstructed one denoted as $\hat{\mathbf{H}}_{M,R}$. In the second stage, we recover the channel parameters in $\mathbf{H}_{R,B}$ from the received signals at the BS (11) by assuming $\mathbf{H}_{M,R} = \hat{\mathbf{H}}_{M,R}$.

III. Hybrid RIS Channel Estimation

The channel estimation for MIMO systems is formulated as a sparse signal recovery problem, which can be addressed by compressive sensing (CS) methods. By using CS, we can explore the sparsity of the channels and reduce the training overhead [4]. CS methods can be classified as on-the-grid or off-the-grid CS methods, such as OMP and ANM [19], respectively. ANM is a advanced technique that can avoid the performance degradation due to basis mismatch, which is a well-known challenge in on-the-grid methods. Motivated by the recent works on off-the-grid methods [9], [20], we also apply ANM for hybrid RIS CE. Again, our target is to recover the channel parameters from the received signal (10) and (11). We divided the CE in two stages, as we described as follows.

A. Preliminaries

We apply the ANM for the estimation of the angles, or equivalent spatial frequencies. By replacing the angular parameters with spatial frequencies, we can express (1) and (4), respectively, as

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

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

where $\mathbf{g}_1 = \sin(\boldsymbol{\phi}_{M,R})$, $\mathbf{f}_1 = \sin(\boldsymbol{\theta}_{M,R})$, $\mathbf{g}_2 = \sin(\boldsymbol{\phi}_{R,B})$, and $\mathbf{f}_2 = \sin(\boldsymbol{\theta}_{R,B})$. The spatial frequencies are within $[0, 1)$, and the resultant atomic set of $\mathbf{H}_{M,R}$ is denoted by $\mathcal{A}_M$, as

$$\mathcal{A}_M = \left\{ \mathbf{Q}_1(\mathbf{f}_1, \mathbf{g}_1) : \mathbf{f}_1 \in \left[0, 1\right), \mathbf{g}_1 \in \left[0, 1\right) \right\}, \quad (14)$$

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

$$\mathcal{A}_N = \left\{ \mathbf{Q}_2(\mathbf{f}_2, \mathbf{g}_2) : \mathbf{f}_2 \in \left[0, 1\right), \mathbf{g}_2 \in \left[0, 1\right) \right\}, \quad (15)$$

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

B. Stage 1 CE

To recover the matrix $\mathbf{H}_{M,R}$, we formulate the atomic norm with respect to the atomic set $\mathcal{A}_M$ as

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

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

$$\|\mathbf{H}_{M,R}\|_{\mathcal{A}_M} = \inf_{\{\mathbf{C},\mathbf{V}\}} \left\{ \frac{1}{2N_M} \text{Tr}(\mathbb{T}(\mathbf{C})) + \frac{1}{2N_R} \text{Tr}(\mathbb{T}(\mathbf{V})) \right\}$$

$$\text{s.t } \begin{bmatrix} \mathbb{T}(\mathbf{C}) & \mathbf{H}_{\text{M,R}} \\ \mathbf{H}_{\text{M,R}}^{\mathsf{H}} & \mathbb{T}(\mathbf{V}) \end{bmatrix} \succeq 0. \tag{17}$$

where $\mathbb{T}(\mathbf{C})$ and $\mathbb{T}(\mathbf{V})$ are 2-level Toeplitz matrices. We can recover the angles $\boldsymbol{\theta}_{\text{M,R}}$ and $\boldsymbol{\phi}_{\text{M,R}}$ by addressing the following convex problem

$$\hat{\mathbf{H}}_{\text{M,R}} = \arg\min_{\mathbf{H}_{\text{M,R}}} \tau \|\mathbf{H}_{\text{M,R}}\|_{\mathcal{A}_{\text{M}}} + \frac{1}{2}\|\beta_1 \mathbf{W}_{\text{H}} \mathbf{H}_{\text{M,R}} \mathbf{X} - \mathbf{Y}_{\text{H}}\|_{\text{F}}^2 \tag{18}$$

where τ is the regularization parameter set as $\tau \propto \sigma\sqrt{N_{\text{R}}N_{\text{M}}\log(N_{\text{R}}N_{\text{M}})}$. Using the SDP formulation, the problem can be expressed as

$$\hat{\mathbf{H}}_{\text{M,R}} = \arg\min_{\mathbf{H}_{\text{M,R}},\mathbf{C},\mathbf{V}} \frac{\tau}{2N_{\text{M}}} \text{Tr}(\mathbb{T}(\mathbf{C})) + \frac{\tau}{2N_{\text{R}}} \text{Tr}(\mathbb{T}(\mathbf{V}))$$
$$+ \frac{1}{2}\|\beta_1 \mathbf{W}_{\text{H}} \mathbf{H}_{\text{M,R}} \mathbf{X} - \mathbf{Y}_{\text{H}}\|_{\text{F}}^2$$
$$\text{s.t } \begin{bmatrix} \mathbb{T}(\mathbf{C}) & \mathbf{H}_{\text{M,R}} \\ \mathbf{H}_{\text{M,R}}^{\mathsf{H}} & \mathbb{T}(\mathbf{V}) \end{bmatrix} \succeq 0. \tag{19}$$

The solution of the Toeplitz matrices $\mathbb{T}(\mathbf{C})$ and $\mathbb{T}(\mathbf{V})$ leads us to the recovery of the angles $\boldsymbol{\theta}_{\text{M,R}}$ and $\boldsymbol{\phi}_{\text{M,R}}$, respectively, by applying the root MUSIC algorithm [21]. We consider the order information of the angles and the number of paths as a priori information. We estimate the path gain vector $\boldsymbol{\rho}_{\text{M,R}}$ by applying the least squares (LS) method, which results in

$$\hat{\boldsymbol{\rho}}_{\text{M,R}} = \left[\beta_1(\mathbf{X}^{\mathsf{T}} \otimes \mathbf{W}_{\text{H}})\big((\mathbf{A}^*(\hat{\boldsymbol{\theta}}_{\text{M,R}}) \odot \mathbf{A}(\hat{\boldsymbol{\phi}}_{\text{M,R}})\big)\right]^{\dagger} \mathbf{y}_{\text{H}}, \tag{20}$$

where $\mathbf{y}_{\text{H}} = \text{vec}(\mathbf{Y}_{\text{H}})$.

C. Stage 2 CE

In the second stage, we target at the recovery of the remaining channel parameters from the received signal at BS. In order to facilitate the CE, we use the estimate of $\hat{\mathbf{H}}_{\text{M,R}}$ from the first stage, which results in

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

We formulate the atomic norm for the matrix $\mathbf{H}_{\text{R,B}}$ as

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

By following the SDP formulation, we can write

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

where $\mathbb{T}(\mathbf{S})$ and $\mathbb{T}(\mathbf{O})$ are 2-level Toeplitz matrices. In order to recover the angles, we formulate the problem as

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

where $\nu \propto \sigma\sqrt{N_{\text{B}}N_{\text{R}}\log(N_{\text{B}}N_{\text{R}})}$ By following the SDP formulation, we expressed the problem as

$$\hat{\mathbf{H}}_{\text{R,B}} = \arg\min_{\mathbf{H}_{\text{R,B}},\mathbf{S},\mathbf{O}} \frac{\nu}{2N_{\text{R}}} \text{Tr}(\mathbb{T}(\mathbf{S})) + \frac{\nu}{2N_{\text{B}}} \text{Tr}(\mathbb{T}(\mathbf{O}))$$

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

The angles $\boldsymbol{\theta}_{\text{R,B}}$ and $\boldsymbol{\phi}_{\text{R,B}}$ can be recovered based on the solution of $\mathbb{T}(\mathbf{S})$ and $\mathbb{T}(\mathbf{O})$, respectively, via root MUSIC algorithm [21]. Similarly, the estimate of the path gain vector is addressed by LS, as

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

where $\mathbf{y} = \text{vec}(\mathbf{Y})$. Based on the estimation of the channel parameters, we calculate the angle differences and the products of path gains associated with the RIS for the purpose of comparison with the passive RIS benchmark [9]. The angle differences are functions of the estimates of $\boldsymbol{\phi}_{\text{M,R}}$ and $\boldsymbol{\theta}_{\text{R,B}}$, expressed as

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

Moreover, we also define the vectorization of $\hat{\boldsymbol{\Delta}}$ as $\hat{\boldsymbol{\delta}} = \text{vec}(\hat{\boldsymbol{\Delta}})$. The estimated products of path gains $\hat{\boldsymbol{\rho}} \in \mathbb{C}^{L_{\text{R,B}}L_{\text{M,R}} \times 1}$ are defined as follows:

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

The training overhead for hybrid RIS CE is

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

where $N_{\text{RF,B}}$ is the number of RF chains at BS.

D. Design of Phase Control Matrix and Beamforming Vectors

We briefly present the design of the phase control matrix and the beamforming vectors. We clarify that we follow the same procedure for both the hybrid and passive RIS. We formulate the design of the phase control matrix in order to maximize the power (a.k.a. squared Frobenius norm) of $\mathbf{G}$, defined in (6). Moreover, the design of the beamforming vectors is based on the reconstruction of the entire channel $\hat{\mathbf{H}} = \hat{\mathbf{H}}_{\text{R,B}} \hat{\boldsymbol{\Omega}} \hat{\mathbf{H}}_{\text{M,R}}$. We conduct singular value decomposition (SVD) on $\hat{\mathbf{H}}$ and use the left and right singular vectors associated with the largest singular value as BS and MS beamforming vectors, respectively.

E. Passive RIS CE via ANM

In the first block of CE, i.e., $k = 1$, we target at the extraction of the AoAs at BS and AoDs at MS, with $\mathbf{X}_1 \in \mathbb{C}^{N_{\text{M}} \times N_0}$, $\mathbf{W}_1 \in \mathbb{C}^{N_{\text{B}} \times M_0}$, and $\boldsymbol{\Omega}_1$. Based on the estimates $\{\hat{\boldsymbol{\theta}}_{\text{M,R}}, \hat{\boldsymbol{\phi}}_{\text{R,B}}\}$ in the first stage, we design the sequential beam matrices and the combining matrices for the second stage, resulting in $\mathbf{X}_k = \mathbf{A}(\hat{\boldsymbol{\theta}}_{\text{M,R}}) \in \mathbb{C}^{N_{\text{M}} \times L_{\text{M,R}}}$ and $\mathbf{W}_k = \mathbf{A}(\hat{\boldsymbol{\phi}}_{\text{R,B}}) \in \mathbb{C}^{N_{\text{B}} \times L_{\text{R,B}}}$, for $k = 2, \cdots, K$. In the second stage, we target at the recovery of the angle differences and the products of path gains. Note that during this stage we keep the training matrix and the combining matrix fixed, while the RIS phase control matrix varies from block to block, i.e., $\boldsymbol{\Omega}_2 \neq \boldsymbol{\Omega}_3 \neq \cdots \neq \boldsymbol{\Omega}_K$. More details can be referred to [9].

TABLE I
PARAMETER SETUP FOR THE HYBRID RIS ARCHITECTURE.

	N_R	M	$N_{RF,R}$	K	T	$N_{C,B}$	$N_{RF,B}$	T_H
Setup 1	32	8	8	5	8	8	8	40
Setup 2	32	4	4	5	8	8	8	40
Setup 3	32	2	2	5	8	8	8	40

Fig. 3. System setup.

Fig. 4. MSE of the angular parameters.

IV. PERFORMANCE EVALUATION

In this section, we describe the parameter setup and the performance metrics. The propagation path gains are distributed as $\mathcal{CN}(0,1)$. We set $N_B = 16$, $N_R = 32$, $N_M = 16$, $L_{R,B} = L_{M,R} = 2$. We perform 1000 trials to average out the results. For simplicity, we assume 2-Dimensional Cartesian coordinate system. Fig. 3 shows the locations of the BS, RIS, and MS. We set the location of the BS, RIS and MS as $(0,0)$, (x,y), $(x_T, 0)$, respectively, where $x = d_T - d_x$, $y = d_y$ and $x_T = d_T$. The distance between MS and RIS is $d_1 = \sqrt{d_x^2 + d_y^2}$, while the distance between BS and RIS is $d_2 = \sqrt{(d_T - d_x)^2 + d_y^2}$. The path loss as a function of distance is given by [8], $\beta(d_1) = \beta_0 \left(\frac{d_0}{d_1}\right)^\gamma$ and $\beta(d_1, d_2) = \beta_0 \left(\frac{d_0}{d_1 d_2}\right)^\gamma$, where β_0 is the path loss with the reference distance d_0, defined as $\beta_0 = (\frac{\lambda}{4\pi d_0})^2$, λ denotes the wavelength, c is the speed of the light in meters per second ($c = 3 \times 10^8$), and f_c the carrier frequency in Hz. We set d_0 as 1 [m], the path loss exponent $\gamma = 3$, $f_c = 28$ [MHz], the noise power density as -173 [dBm/Hz], the bandwidth as 100 MHz, and the transmit power as $P_T = 0 : 5 : 20$ [dBm].

To make the performance comparison fair, we keep the training overhead the same for both the hybrid RIS and the passive RIS. We evaluate the performance considering different number of active elements at RIS. Table I summarizes the parameter setup for the hybrid RIS architecture. In order to guarantee the super-resolution estimation, we assume the spatial frequencies are separated by at least $\left(\frac{4}{N_B}\right)$, $\left(\frac{4}{N_R}\right)$, $\left(\frac{4}{N_M}\right)$.

A. Performance Metrics

We evaluate the performance in terms of the MSE[1] of the estimates of AoDs, AoAs, angle differences and the products

[1]The MSE can be formulated by considering the sine of the angles, i.e., spatial frequencies. We clarify that the results will be consistent with these based on estimated angular parameters.
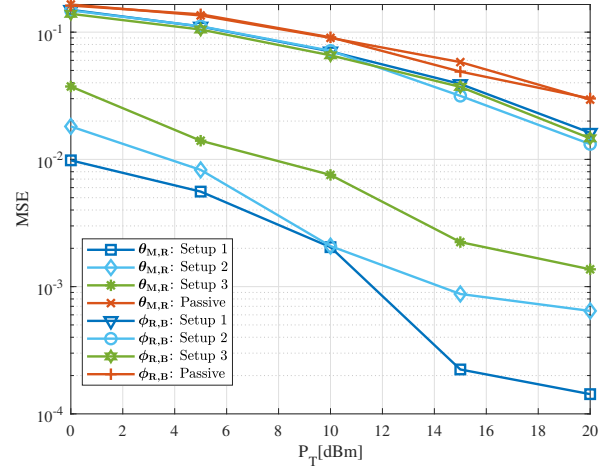
of path gains. Also, we evaluate the performance in terms of average SE in (bits/s/Hz).

B. Simulation Results

The simulation results for the estimation of channel parameters $\boldsymbol{\theta}_{M,R}$ and $\boldsymbol{\phi}_{R,B}$ are shown in Fig. 4, where we consider $d_T = 25$ [m], $d_x = 10$ [m], $d_y = 2$ [m]. We compare our results with the benchmark scheme, i.e., the passive RIS detailed in [9]. For the estimation of $\boldsymbol{\theta}_{M,R}$, the setups 1–3 of the hybrid RIS outperform the passive RIS. The performance of the setup 1 brings the best performance among the hybrid setups, while the setup 3 has the lowest power consumption due to the least number of active elements and RF chains at the RIS. The results can be explained by the effect of path loss on the RIS CE. In our proposed method, we can perform CE at the RIS, where the path loss is proportional to the distance d_1. On the other hand, when the CE is performed at BS, the path loss is proportional to $d_1 d_2$, which degrades the performance.

Regarding the MSE of $\boldsymbol{\delta}$ and $\boldsymbol{\rho}$, the hybrid RIS also has better performance than the passive RIS. Also, we can see in Fig. 5 that the performance of hybrid setups are quite similar. The estimates of the coupled parameters $\boldsymbol{\delta}$ and $\boldsymbol{\rho}$ depend on the estimates of $\boldsymbol{\phi}_{M,R}$ and $\boldsymbol{\rho}_{M,R}$ in the first stage and these of $\boldsymbol{\theta}_{R,B}$ and $\boldsymbol{\rho}_{R,B}$ in the second stage. In the first CE stage, we can obtain a better estimate of the channel parameters due to the lower path loss. However, in the second CE stage, the path loss is more severe, which results in a poor estimation of $\boldsymbol{\theta}_{R,B}$ and $\boldsymbol{\rho}_{R,B}$. The poor estimation in the second stage will dominate the resultant performance in Fig. 5. Fig. 6 shows the average SE of our proposed method and the passive RIS [9]. The setups 1–3 of our proposed method can achieve better performance than the passive RIS for low transmit power, for example $P_T = 5$ [dBm]. However, for high transmit power, $P_T = 20$ [dBm], the hybrid and the passive RIS can achieve similar results. Also, we can obtain similar performances even with reduced number of active elements. Since the active elements are proposed only for collecting received signals, their position does not affect the CE. Moreover, the position of the RIS can
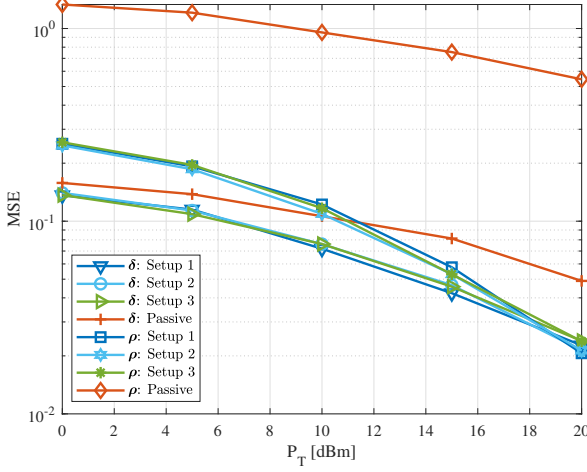
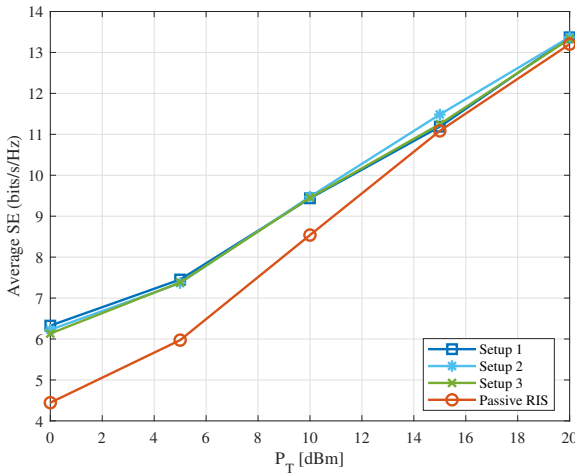Fig. 5. MSE of the products of path gains and angle differences.

Fig. 6. Average SE.

bring benefits for the hybrid RIS. Since we can perform CE at the RIS, we can obtain better performance when the RIS is closer to the MS.

V. Conclusions

In this paper, we have studied the channel estimation in hybrid RIS aided MIMO systems and proposed a two-stage CE scheme via ANM. The results have shown that our proposed CE method could bring better performance compared to the passive RIS. Nevertheless, it is worth mentioning that due to the presence of the active elements, the power consumption is higher. For this reason, the evaluation of our proposed method in terms of EE is an interesting research direction. Also, the assumption of *a priori* information on the order of the paths should be avoided to simplify the implementation of the method, which is left as our future work. We will further investigate the effect of the position of the RIS on CE.

VI. Acknowledgements

References

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

[2] R. W. Heath, N. Gonzalez-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, 2016.

[3] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[4] 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.

[5] T. Ma, Y. Xiao, X. Lei, W. Xiong, and Y. Ding, "Indoor localization with reconfigurable intelligent surface," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 161–165, 2020.

[6] L. Yang, J. Yang, W. Xie, M. O. Hasna, T. Tsiftsis, and M. Di Renzo, "Secrecy performance analysis of RIS-aided wireless communication systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 12 296–12 300, 2020.

[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, 2019.

[8] 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, 2020.

[9] J. He, H. Wymeersch, and M. Juntti, "Channel estimation for RIS-aided mmWave MIMO systems via atomic norm minimization," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2021.

[10] 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, 2019.

[11] J. Mirza and B. Ali, "Channel estimation method and phase shift design for reconfigurable intelligent surface assisted MIMO networks," *IEEE Trans. on Cogn. Commun. Netw.*, pp. 1–1, 2021.

[12] J. He, M. Leinonen, H. Wymeersch, and M. Juntti, "Channel estimation for RIS-aided mmWave MIMO channels," in *proc. IEEE Globecom*, 2020, pp. 1–6.

[13] G. T. de Araújo and A. L. de Almeida, "PARAFAC-Based channel estimation for intelligent reflective surface assisted MIMO system," in *proc.IEEE SAM*, 2020, pp. 1–5.

[14] A. Taha, M. Alrabeiah, and A. Alkhateeb, "Enabling large intelligent surfaces with compressive sensing and deep learning," *arXiv*, 2019. [Online]. Available: https://arxiv.org/abs/1904.10136.

[15] 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*, 2020, pp. 9175–9179.

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

[17] N. T. Nguyen, Q.-D. Vu, K. Lee, and M. Juntti, "Hybrid relay-reflecting intelligent surface-assisted wireless communication," *arXiv*, 2021. [Online]. Available: https://arxiv.org/abs/2103.03900.

[18] Y. Lin, S. Jin, M. Matthaiou, and X. You, "Tensor-Based algebraic channel estimation for hybrid IRS-Assisted MIMO-OFDM," *IEEE Trans. Commun.*, pp. 1–1, 2021.

[19] 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, 2018.

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

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