



**HAL**  
open science

## Channel Estimation for Stacked Intelligent Metasurface-Assisted Wireless Networks

Xianghao Yao, Jiancheng An, Lu Gan, Marco Di Renzo, Chau Yuen

► **To cite this version:**

Xianghao Yao, Jiancheng An, Lu Gan, Marco Di Renzo, Chau Yuen. Channel Estimation for Stacked Intelligent Metasurface-Assisted Wireless Networks. *IEEE Wireless Communications Letters*, 2024, 13 (5), pp.1349-1353. <10.1109/lwc.2024.3369874>. <hal-04734227>

**HAL Id: hal-04734227**

**<https://hal.science/hal-04734227v1>**

Submitted on 13 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# Channel Estimation for Stacked Intelligent Metasurface-Assisted Wireless Networks

Xianghao Yao, Jiancheng An, *Member, IEEE*, Lu Gan,

Marco Di Renzo, *Fellow, IEEE*, and Chau Yuen, *Fellow, IEEE*

## Abstract

Emerging technologies, such as holographic multiple-input multiple-output (HMIMO) and stacked intelligent metasurface (SIM), are driving the development of wireless communication systems. Specifically, the SIM is physically constructed by stacking multiple layers of metasurfaces and has an architecture similar to an artificial neural network (ANN), which can flexibly manipulate the electromagnetic waves that propagate through it at the speed of light. This architecture enables the SIM to achieve HMIMO precoding and combining in the wave domain, thus significantly reducing the hardware cost and energy consumption. In this letter, we investigate the channel estimation problem in SIM-assisted multi-user HMIMO communication systems. Since the number of antennas at the base station (BS) is much smaller than the number of meta-atoms per layer of the SIM, it is challenging to acquire the channel state information (CSI) in SIM-assisted multi-user systems. To address this issue, we collect multiple copies of the uplink pilot signals that propagate through the SIM. Furthermore, we leverage the array geometry to identify the subspace that spans arbitrary spatial correlation matrices. Based on partial CSI about the channel statistics, a pair of subspace-based channel estimators are proposed. Additionally, we compute the mean square error (MSE) of the proposed channel estimators and optimize the phase shifts of the SIM to minimize the MSE. Numerical results are illustrated to analyze the effectiveness of the proposed channel estimation schemes.

This work is partially supported by Sichuan Science and Technology Program under Grant 2023YFSY0008 and 2023YFG0291, and partially supported by Yibin Science and Technology Program under Grant YBP-002. The work of M. Di Renzo is supported in part by the Horizon Europe projects COVER-101086228, UNITE-101129618, and INSTINCT-101139161, and the ANR projects NF-PERSEUS 22-PEFT-004 and PASSIONATE ANR-23-CHR4-0003-01. The research of Chau Yuen is supported by the Ministry of Education, Singapore, under its MOE Tier 2 (Award number MOE-T2EP50220-0019).

X. Yao and L. Gan are with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan, 611731, China. L. Gan is also with the Yibin Institute of UESTC, Yibin 644000, China (e-mail: xianghao\_yao@163.com; ganlu@uestc.edu.cn). J. An and C. Yuen are with the School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore 639798 (e-mail: jiancheng\_an@163.com; chau.yuen@ntu.edu.sg). M. Di Renzo is with Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des Signaux et Systèmes, 3 Rue Joliot-Curie, 91192 Gif-sur-Yvette, France (e-mail: marco.di-renzo@universite-paris-saclay.fr).

## Index Terms

Holographic MIMO, stacked intelligent metasurface (SIM), channel estimation, spatial correlation.

## I. INTRODUCTION

Next generation wireless networks promise to provide ten-fold transmission rate improvement. Therefore, there is an urgent need for advanced wireless technologies to support ultra-reliability, low latency, and high data rate [1]–[3]. Among the emerging technology candidates, holographic multiple-input multiple-output (HMIMO) technology has drawn increasing interest [4]. By integrating a large number of radiating and sensing elements, the energy can be focused into more precise areas to achieve high energy and spectral efficiencies [4], [5]. However, HMIMO is still in its initial stage of research and there is still a gap between theoretical performance and experimental validation. To address this issue, metasurface technology is a promising candidate for realizing HMIMO due to its low power consumption and reduced number of radio frequency (RF) chains [6]–[9].

To further improve the flexibility of metasurfaces for generating diverse RF waveforms, the authors of [10] proposed a novel stacked intelligent metasurface (SIM) technology, which realizes advanced signal processing capabilities thanks to an architecture that resembles a deep neural network in the wave domain. Leveraging this advanced architecture, the authors of [11], [12] designed an SIM-assisted HMIMO system that implements transmit precoding and receiver combining as the transmitted wave propagates through the SIM, thereby decreasing the number of RF chains. Also, an SIM greatly reduces the power consumption, precoding delay, and hardware cost compared to its digital counterpart [10]. However, an SIM cannot efficiently probe the channels of each passive meta-atom, as the number of RF chains is smaller than the number of meta-atoms. To fully exploit the potential gains brought by SIM-assisted HMIMO systems, the accurate estimation of the channels becomes crucial. Furthermore, to effectively reduce the multi-user interference, the number of meta-atoms is generally larger than the number of antennas at the base station (BS) [10]. This leads to an increase in the dimension of the channel matrix and poses great challenges for channel estimation in SIM-assisted HMIMO systems [13].

Against this background, in this letter, we investigate the channel estimation in SIM-assisted multi-user HMIMO communication systems. Specifically, we propose an estimation protocol by accumulating the received signals over multiple time blocks to tackle the underdetermined channel estimation problem. By transmitting orthogonal pilot sequences to the BS in each block, the interference between multiple users can be mitigated. Based on the proposed model, we then present four distinct channel estimation techniques, including the least squares (LS), the minimum mean square error (MMSE), the reduced-subspace LS (RSLs), and the RSLs relying on isotropic scattering statistics (RSLs-iso). The latter two methods exhibit enhanced estimation accuracy even with limited knowledge of the spatial correlation matrix. Furthermore, the phase shifts of the meta-atoms in the SIM are optimized for improving the performance of the considered channel estimators. Extensive simulation results are illustrated to prove the effectiveness of the proposed channel estimation techniques.

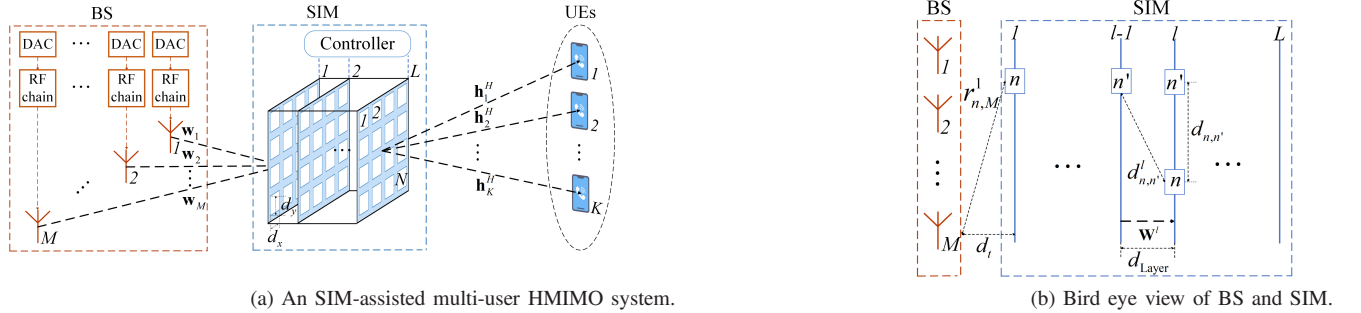


Fig. 1. SIM-assisted HMIMO transmission.

## II. SYSTEM AND CHANNEL MODELS

In this section, we introduce an SIM-assisted multi-user HMIMO communication system. As shown in Fig. 1a, a BS equipped with a uniform linear array (ULA) having  $M$  antennas serves  $K$  single-antenna user equipments (UEs). An SIM consists of  $L$  metasurface layers, each with  $N$  meta-atoms. By adjusting the electromagnetic (EM) properties of all the metasurface layers with an intelligent controller, the SIM can perform wave-based processing [10].

Let  $\mathcal{M} = \{1, 2, \dots, M\}$ ,  $\mathcal{L} = \{1, 2, \dots, L\}$ ,  $\mathcal{N} = \{1, 2, \dots, N\}$ , and  $\mathcal{K} = \{1, 2, \dots, K\}$  denote the indices of the antennas, metasurface layers, meta-atoms on each layer, and UEs, respectively. The distance between the  $m$ -th antenna and the  $n$ -th meta-atom of the first metasurface

layer is denoted by  $r_{n,m}^1$ . The distance between adjacent meta-atoms is denoted by  $r_{\text{atom}}$ , while the length and width of each meta-atom are given by  $d_x$  and  $d_y$ , respectively. The horizontal distance between the antenna array and the first metasurface layer is denoted by  $d_t$ , while the spacing between each metasurface layer is denoted by  $d_{\text{Layer}}$ .

Furthermore, let  $e^{j\theta_n^l}$  be the transmission coefficient of the  $n$ -th meta-atom on the  $l$ -th metasurface layer, with  $\theta_n^l \in [0, 2\pi)$ . The matrix of transmission coefficients for the  $l$ -th metasurface layer is given by  $\Phi^l = \text{diag}(e^{j\theta_1^l}, e^{j\theta_2^l}, \dots, e^{j\theta_N^l}) \in \mathbb{C}^{N \times N}$  [14]. The matrix of propagation channels from the  $(l-1)$ -th to the  $l$ -th metasurface layer is denoted by  $\mathbf{W}^l \in \mathbb{C}^{N \times N}, \forall l \neq 1, l \in \mathcal{L}$ . According to Rayleigh-Sommerfeld's diffraction theory [15], [16], the  $(n, n')$ -th entry of  $\mathbf{W}^l$  can be written as

$$w_{n,n'}^l = \frac{d_x d_y d_{\text{Layer}}}{d_{n,n'}^l{}^2} \left( \frac{1}{2\pi d_{n,n'}^l} - \frac{1}{\lambda} j \right) e^{j2\pi d_{n,n'}^l / \lambda}, \quad (1)$$

where  $\lambda$  is the wavelength,  $d_{n,n'}^l$  is the transmission distance from the  $n'$ -th meta-atom of the  $(l-1)$ -th metasurface layer to the  $n$ -th meta-atom of the  $l$ -th metasurface layer. The vector of propagation channels from the  $m$ -th antenna to the first metasurface layer is denoted by  $\mathbf{w}_m \in \mathbb{C}^{N \times 1}$  and can be obtained by replacing  $d_{n,n'}^l$  in (1) with  $r_{n,m}^1$ , as shown in Fig. 1b.

The SIM-enabled wave-based beamforming matrix can be written as  $\mathbf{G} = \Phi^L \mathbf{W}^L \dots \Phi^2 \mathbf{W}^2 \Phi^1 \in \mathbb{C}^{N \times N}$ . Moreover, we assume a quasi-static flat fading channel model, i.e., the channel parameters within a coherent block are time-invariant [10], [17]. The baseband equivalent channel from the last metasurface layer to the  $k$ -th UE is denoted by  $\mathbf{h}_k^H \in \mathbb{C}^{1 \times N}, \forall k \in \mathcal{K}$ , which is modeled by  $\mathbf{h}_k \sim \mathcal{CN}(\mathbf{0}, \beta_k \mathbf{R})$ , where  $\beta_k$  denotes the distance-dependent path loss between the last metasurface layer and the  $k$ -th UE,  $\mathbf{R} \in \mathbb{C}^{N \times N}$  denotes the covariance matrix characterizing the spatial correlation between each meta-atom on the last metasurface layer. According to [18], we can write the  $(n, n')$ -th entry of  $\mathbf{R}$  as

$$[\mathbf{R}]_{n,n'} = \iint_{-\pi/2}^{\pi/2} f(\varphi, \theta) \times e^{j2\pi(d_{n,n'}^H \sin(\varphi) \cos(\theta) + d_{n,n'}^V \sin(\theta)) / \lambda} d\theta d\varphi, \quad (2)$$

where  $\varphi, \theta, f(\varphi, \theta), d_{n,n'}^H$  and  $d_{n,n'}^V$  denote the azimuth angle, elevation angle, spatial scattering function, the horizontal and vertical distances between the  $n$ -th and  $n'$ -th meta-atoms on the same metasurface layer, respectively. We assume that all metasurface layers have a square structure,

and the metasurface centers are in the same horizontal line. Furthermore, all layers are parallel to each other, and the transmit antenna array is parallel to the SIM.

### III. CHANNEL ESTIMATION FOR SIM-ASSISTED MULTI-USER HMIMO SYSTEMS

In this section, we investigate the channel estimation problem in SIM-assisted multi-user systems. As customary, we assume that the number of meta-atoms  $N$  on each metasurface layer is generally larger than the number of antennas  $M$  at the BS [10]. Additionally, the SIM cannot probe the wireless channels associated with each meta-atom individually, making it difficult to estimate the channels from the UEs to the last layer of the SIM in practice. To overcome this challenge, we consider collecting multiple copies of the pilot signals to probe the uplink channel.

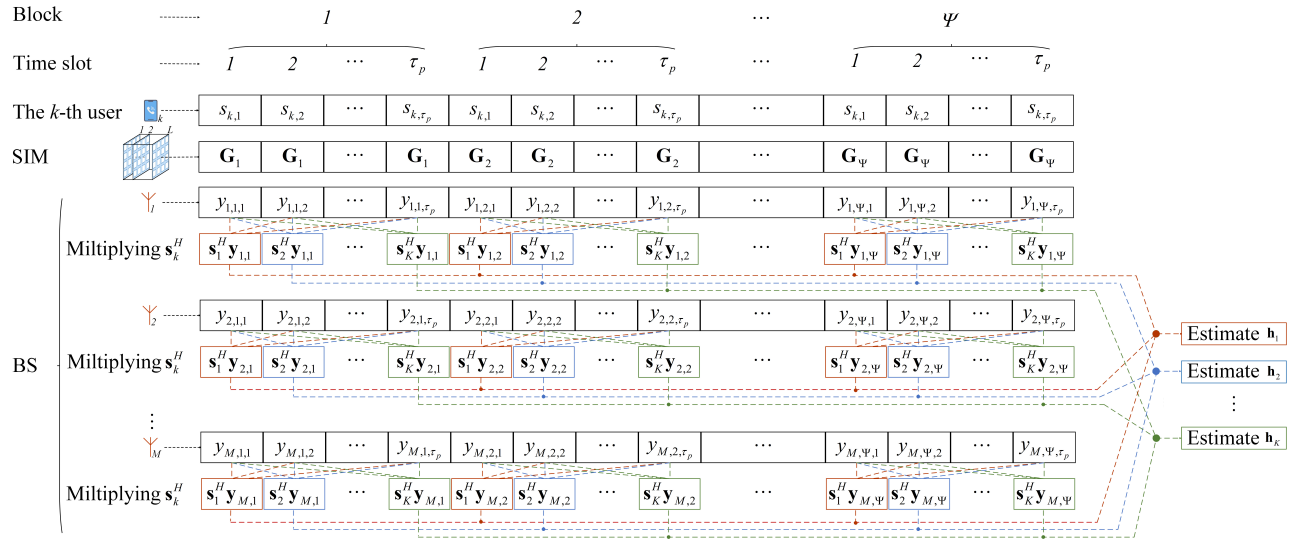


Fig. 2. The proposed channel estimation protocol.

As shown in Fig. 2, we consider  $\Psi$  blocks, each containing  $\tau_p$  time slots. All the UEs send the same pilot sequence of length  $\tau_p$  over different blocks. Let  $s_{k,t}$  represent the pilot signal sent by the  $k$ -th UE in the  $t$ -th time slot of each block. The average power of the pilot signal of the  $k$ -th UE is denoted by  $p_k$ . Therefore, the signal received by the  $m$ -th antenna in the  $t$ -th time slot of the  $\psi$ -th block can be written as

$$y_{m,\psi,t} = \mathbf{w}_m^H \mathbf{G}_\psi^H \sum_{k'=1}^K \mathbf{h}_{k'} \sqrt{p_{k'}} s_{k',t} + n_{m,\psi,t}, \quad (3)$$

where  $\mathbf{G}_\psi \in \mathbb{C}^{N \times N}$  denotes the beamforming matrix of the SIM in the  $\psi$ -th block, and  $n_{m,\psi,t} \sim \mathcal{CN}(0, \sigma_n^2)$  denotes the additive white Gaussian noise at the  $m$ -th antenna in the  $t$ -th time slot

of the  $\psi$ -th block. For simplicity, we assume the same noise power  $\sigma_n^2$  for all the antennas. As shown in Fig. 2, by collecting the signals received at the  $m$ -th antenna in the  $\psi$ -th block, we obtain

$$\underbrace{\begin{bmatrix} y_{m,\psi,1} \\ y_{m,\psi,2} \\ \vdots \\ y_{m,\psi,\tau_p} \end{bmatrix}}_{\mathbf{y}_{m,\psi}} = \sum_{k'=1}^K \sqrt{p_{k'}} \underbrace{\begin{bmatrix} s_{k',1} \\ s_{k',2} \\ \vdots \\ s_{k',\tau_p} \end{bmatrix}}_{\mathbf{s}_{k'}} \mathbf{w}_m^H \mathbf{G}_\psi^H \mathbf{h}_{k'} + \underbrace{\begin{bmatrix} n_{m,\psi,1} \\ n_{m,\psi,2} \\ \vdots \\ n_{m,\psi,\tau_p} \end{bmatrix}}_{\mathbf{n}_{m,\psi}}, \quad (4)$$

where  $\mathbf{y}_{m,\psi} \in \mathbb{C}^{\tau_p \times 1}$  and  $\mathbf{n}_{m,\psi} \in \mathbb{C}^{\tau_p \times 1}$  denote the received signal and noise vector by the  $m$ -th antenna, respectively, in the  $\psi$ -th block. The pilot sequence sent by the  $k$ -th UE is  $\mathbf{s}_k \in \mathbb{C}^{\tau_p \times 1}$ . To eliminate the interference between multiple UEs in each block, we adopt the mutually orthogonal pilot sequence strategy, such that  $\sum_{t=1}^{\tau_p} s_{k,t} s_{k',t} = \tau_p$ , for  $k = k'$  and  $\sum_{t=1}^{\tau_p} s_{k,t} s_{k',t} = 0$ , for  $k \neq k'$ . One efficient pilot design is to employ the columns of the  $\tau_p$ -point discrete Fourier transform (DFT) matrix as the pilot sequences for different UEs. Note that the length of  $\mathbf{s}_k$  needs to satisfy  $\tau_p \geq K$  in order to separate the pilot signals sent by  $K$  UEs within a block. In this letter, to minimize the pilot overhead, we consider  $\tau_p = K$  in each block. Additionally, we note that the random phase drift due to practical hardware imperfections can be effectively solved by a compensation technique [19], and is ignored. After left-multiplying  $\mathbf{s}_k^H$  on both sides of (4) and collecting the output signals associated to the  $M$  antennas, we arrive at

$$\underbrace{\begin{bmatrix} \mathbf{s}_k^H \mathbf{y}_{1,\psi} \\ \mathbf{s}_k^H \mathbf{y}_{2,\psi} \\ \vdots \\ \mathbf{s}_k^H \mathbf{y}_{M,\psi} \end{bmatrix}}_{\check{\mathbf{y}}_{k,\psi}} = \tau_p \sqrt{p_k} \underbrace{\begin{bmatrix} \mathbf{w}_1^H \\ \mathbf{w}_2^H \\ \vdots \\ \mathbf{w}_M^H \end{bmatrix}}_{\check{\mathbf{W}}} \mathbf{G}_\psi^H \mathbf{h}_k + \underbrace{\begin{bmatrix} \mathbf{s}_k^H \mathbf{n}_{1,\psi} \\ \mathbf{s}_k^H \mathbf{n}_{2,\psi} \\ \vdots \\ \mathbf{s}_k^H \mathbf{n}_{M,\psi} \end{bmatrix}}_{\check{\mathbf{n}}_{k,\psi}}, \quad (5)$$

where  $\check{\mathbf{y}}_{k,\psi} \in \mathbb{C}^{M \times 1}$  denotes the matched output signals,  $\check{\mathbf{W}} \in \mathbb{C}^{M \times N}$  denotes the channel propagation matrix from the first metasurface layer to the BS, and  $\check{\mathbf{n}}_{k,\psi} \in \mathbb{C}^{M \times 1}$  denotes the noise vector, in the  $\psi$ -th block. Upon collecting the matched output signals over  $\Psi$  blocks, we

obtain

$$\underbrace{\begin{bmatrix} \check{\mathbf{y}}_{k,1} \\ \check{\mathbf{y}}_{k,2} \\ \vdots \\ \check{\mathbf{y}}_{k,\Psi} \end{bmatrix}}_{\tilde{\mathbf{y}}_k} = \tau_p \sqrt{p_k} \underbrace{\begin{bmatrix} \check{\mathbf{W}}\mathbf{G}_1^H \\ \check{\mathbf{W}}\mathbf{G}_2^H \\ \vdots \\ \check{\mathbf{W}}\mathbf{G}_\Psi^H \end{bmatrix}}_{\mathbf{P}} \mathbf{h}_k + \underbrace{\begin{bmatrix} \check{\mathbf{n}}_{k,1} \\ \check{\mathbf{n}}_{k,2} \\ \vdots \\ \check{\mathbf{n}}_{k,\Psi} \end{bmatrix}}_{\tilde{\mathbf{n}}_k}, \quad (6)$$

where  $\tilde{\mathbf{y}}_k \in \mathbb{C}^{\Psi M \times 1}$ ,  $\tilde{\mathbf{n}}_k \in \mathbb{C}^{\Psi M \times 1}$  and  $\mathbf{P} \in \mathbb{C}^{\Psi M \times N}$  denote the received signal, the noise and the transmission matrix from the last metasurface layer to the BS, respectively, over  $\Psi$  blocks. Furthermore, the variance of  $\tilde{\mathbf{n}}_k$  is  $\mathbb{E} \left\{ (\tilde{\mathbf{n}}_k - \mathbb{E} \{ \tilde{\mathbf{n}}_k \}) (\tilde{\mathbf{n}}_k - \mathbb{E} \{ \tilde{\mathbf{n}}_k \})^H \right\} = \tau_p \sigma_n^2 \mathbf{I}_{\Psi M}$ , where  $\mathbb{E} \{ \cdot \}$  denotes the expectation,  $\mathbf{I}_{\Psi M} \in \mathbb{C}^{\Psi M \times \Psi M}$  is the identity matrix. Based on the considered system model, the minimum number of blocks for the estimation of  $\mathbf{h}_k$  is  $\lceil N/M \rceil$ . We consider  $\Psi = \lceil N/M \rceil$  to minimize the pilot overhead. Based on (6), we then focus on the estimation of  $\mathbf{h}_k$ .

Specifically, when  $\mathbf{R}$  is known, the MMSE estimate of  $\mathbf{h}_k$  is given by

$$\hat{\mathbf{h}}_{k,\text{MMSE}} = \sqrt{p_k} \beta_k \mathbf{R} \mathbf{P}^H (\tau_p p_k \beta_k \mathbf{P} \mathbf{R} \mathbf{P}^H + \sigma_n^2 \mathbf{I}_{\Psi M})^{-1} \tilde{\mathbf{y}}_k, \quad (7)$$

for  $k \in \mathcal{K}$ , where  $\{ \cdot \}^{-1}$  denotes the inverse of a matrix. Although the MMSE estimator can suppress the noise, it requires the knowledge of the covariance matrix. If the matrix  $\mathbf{R}$  is not known, on the other hand, the LS estimator can be utilized, yielding

$$\hat{\mathbf{h}}_{k,\text{LS}} = \mathbf{P}^\dagger \tilde{\mathbf{y}}_k / (\tau_p \sqrt{p_k}), k \in \mathcal{K}, \quad (8)$$

where  $\{ \cdot \}^\dagger$  denotes the pseudo-inverse of a matrix. Compared to the MMSE estimator, the LS estimator does not need any statistical information of  $\mathbf{h}_k$ , but it is very sensitive to the noise and its performance degrades in deep fading channels.

To address this issue, we then propose two subspace-based channel estimators, which relax the amount of information about  $\mathbf{R}$  while achieving better performance than the LS estimator. By replacing  $\mathbf{P} \mathbf{R} \mathbf{P}^H$  with its singular value decomposition (SVD), i.e.,  $\mathbf{U}_1 \mathbf{\Lambda}_1 \mathbf{U}_1^H$ , (7) can be

rewritten as

$$\begin{aligned}\hat{\mathbf{h}}_{k,\text{MMSE}} &= \frac{\mathbf{P}^{-1}\mathbf{P}\mathbf{R}\mathbf{P}^H \left( \mathbf{P}\mathbf{R}\mathbf{P}^H + \frac{\sigma_n^2 \mathbf{I}_\eta}{\tau_p p_k \beta_k} \right)^{-1} \tilde{\mathbf{y}}_k}{\tau_p \sqrt{p_k}} \\ &= \frac{\mathbf{P}^{-1}\mathbf{U}_1 \left( \mathbf{\Lambda}_1 \left( \mathbf{\Lambda}_1 + \frac{\sigma_n^2 \mathbf{I}_\eta}{\tau_p p_k \beta_k} \right)^{-1} \right) \mathbf{U}_1^H \tilde{\mathbf{y}}_k}{\tau_p \sqrt{p_k}},\end{aligned}\quad (9)$$

where the diagonal matrix  $\mathbf{\Lambda}_1 \in \mathbb{C}^{\eta \times \eta}$ ,  $\mathbf{U}_1 \in \mathbb{C}^{\Psi M \times \eta}$  and  $\eta$  represent the non-zero eigenvalues, the matrix constituted by the corresponding orthonormal eigenvectors, and the rank of  $\mathbf{P}\mathbf{R}\mathbf{P}^H$ , respectively. The invertibility of the transmission matrix  $\mathbf{P}$  can be readily achieved by adjusting the phase shifts of the SIM. We notice that, as  $p_k \rightarrow \infty$ ,  $\left( \mathbf{\Lambda}_1 \left( \mathbf{\Lambda}_1 + \frac{\sigma_n^2 \mathbf{I}_\eta}{\tau_p p_k \beta_k} \right)^{-1} \right) \rightarrow \mathbf{I}_\eta$ . Leveraging this asymptotic approximation, the proposed RLS estimator is given by

$$\hat{\mathbf{h}}_{k,\text{RSLs}} = \mathbf{P}^{-1}\mathbf{U}_1\mathbf{U}_1^H \tilde{\mathbf{y}}_k / (\tau_p \sqrt{p_k}), k \in \mathcal{K}, \quad (10)$$

which is similar to the MMSE for large values of  $p_k$ . The RLS estimator leverages the rank deficiency of the spatial correlation matrix  $\mathbf{R}$  resulting from the reduced meta-atom spacing in HMIMO systems [18]. Specifically, it projects the  $\Psi M$ -dimensional received signal  $\tilde{\mathbf{y}}_k$  onto a low-dimensional space, effectively eliminating the noise caused by other  $(\Psi M - \eta)$  dimensions. Hence, it achieves improved estimation accuracy compared to the LS estimator. Compared to the MMSE estimator, the RLS estimator requires only prior information about the slowly varying matrix  $\mathbf{U}_1$  which can be more readily obtained compared with  $\mathbf{R}$ .

However,  $\mathbf{U}_1$  in (10) still relies on prior knowledge of the considered scattering environment. When considering an isotropic scattering environment, the  $(n, n')$ -th entry of the spatial correlation matrix  $\mathbf{R}_{\text{iso}}$  is  $[\mathbf{R}_{\text{iso}}]_{n,n'} = \text{sinc} \left( 2\sqrt{d_{n,n'}^H{}^2 + d_{n,n'}^V{}^2} / \lambda \right)$  [4], [20]. The span of the isotropic channel, i.e.,  $f(\varphi, \theta) > 0$ , is across the entire angular domain, including the span of arbitrary practical channels [18]. More specifically, the subspace generated by the columns of  $\mathbf{R}_{\text{iso}}$  includes the subspace generated by the columns of more general matrices  $\mathbf{R}$ . Hence, we can utilize  $f(\varphi, \theta)$  for isotropic channels by employing the information about  $\mathbf{R}_{\text{iso}}$ . Taking this into account, the RLS estimator can be rewritten as

$$\hat{\mathbf{h}}_{k,\text{RSLs-iso}} = \mathbf{P}^{-1}\mathbf{U}_2\mathbf{U}_2^H \tilde{\mathbf{y}}_k / (\tau_p \sqrt{p_k}), k \in \mathcal{K}, \quad (11)$$

where  $\mathbf{U}_2 \in \mathbb{C}^{\Psi M \times \eta_{\text{iso}}}$  is obtained by computing the SVD, i.e.,  $\mathbf{P}\mathbf{R}_{\text{iso}}\mathbf{P}^H = \mathbf{U}_2\mathbf{\Lambda}_2\mathbf{U}_2^H$ , and  $\eta_{\text{iso}}$  represents the rank of  $\mathbf{P}\mathbf{R}_{\text{iso}}\mathbf{P}^H$ . Accordingly, we can use the statistical channel state information (CSI) based on the array arrangement to achieve more accurate estimates of  $\mathbf{h}_k$  than the LS estimator.

#### IV. MSE OF THE PROPOSED CHANNEL ESTIMATORS

In this section, we compute the mean square error (MSE) of the proposed LS, MMSE, RLS, and RLS-iso estimators. According to [21], the MSE of the MMSE and LS channel estimators are given by

$$\text{MSE}_{\text{MMSE}} = \sigma_n^2 \text{tr} \left\{ \beta_k \mathbf{R} (\tau_p p_k \beta_k \mathbf{P}^H \mathbf{P} \mathbf{R} + \sigma_n^2 \mathbf{I}_N)^{-1} \right\}, \quad (12)$$

$$\text{MSE}_{\text{LS}} = \sigma_n^2 \text{tr} \left\{ (\tau_p p_k \mathbf{P}^H \mathbf{P})^{-1} \right\}, \quad (13)$$

where  $\text{tr} \{ \cdot \}$  denotes the trace. Furthermore, the MSE of the RLS estimator is given by

$$\begin{aligned} \text{MSE}_{\text{RLS}} &= \mathbb{E} \left\{ \left\| \mathbf{h}_k - \hat{\mathbf{h}}_{k,\text{RLS}} \right\|^2 \right\} \\ &= \mathbb{E} \left\{ (\mathbf{h}_k - \hat{\mathbf{h}}_{k,\text{RLS}})^H (\mathbf{h}_k - \hat{\mathbf{h}}_{k,\text{RLS}}) \right\} \\ &= \mathbb{E} \left\{ \mathbf{h}_k^H (\mathbf{I}_N - \mathbf{X} - \mathbf{X}^H + \mathbf{X}^H \mathbf{X}) \mathbf{h}_k \right\} \\ &\quad + \mathbb{E} \left\{ (\mathbf{P}^{-1} \mathbf{U}_1 \mathbf{U}_1^H \tilde{\mathbf{n}}_k)^H \mathbf{P}^{-1} \mathbf{U}_1 \mathbf{U}_1^H \tilde{\mathbf{n}}_k \right\} / (\tau_p^2 p_k), \end{aligned} \quad (14)$$

where  $\mathbf{X} = \mathbf{P}^{-1} \mathbf{U}_1 \mathbf{U}_1^H \mathbf{P}$ . Specifically, the first entry in (14) can be rewritten as

$$\begin{aligned} &\mathbb{E} \left\{ \mathbf{h}_k^H (\mathbf{I}_N - \mathbf{X} - \mathbf{X}^H + \mathbf{X}^H \mathbf{X}) \mathbf{h}_k \right\} \\ &= \beta_k \text{tr} \left\{ \mathbf{R} (\mathbf{I}_N - \mathbf{X} - \mathbf{X}^H + \mathbf{X}^H \mathbf{X}) \right\} \\ &= \beta_k \text{tr} \left\{ (\mathbf{I}_N - \mathbf{X}) \mathbf{R} (\mathbf{I}_N - \mathbf{X}^H) \right\} \\ &= 0. \end{aligned} \quad (15)$$

Therefore, (14) can be simplified as

$$\text{MSE}_{\text{RLS}} = \frac{\sigma_n^2}{\tau_p p_k} \text{tr} \left\{ \mathbf{P}^{-1} \mathbf{U}_1 (\mathbf{P}^{-1} \mathbf{U}_1)^H \right\}. \quad (16)$$

Similarly, the MSE of the RLS-iso estimator is given by

$$\text{MSE}_{\text{RLS-iso}} = \frac{\sigma_n^2}{\tau_p p_k} \text{tr} \left\{ \mathbf{P}^{-1} \mathbf{U}_2 (\mathbf{P}^{-1} \mathbf{U}_2)^H \right\}. \quad (17)$$

Moreover, we note that the MSE of these four estimators are highly dependent on the transmission matrix  $\mathbf{P}$ . Hence, the MSE of the considered estimators can be reduced by optimizing the phase shifts of the SIM during the training stage. In mathematical terms, the optimization of the SIM can be formulated as follows:

$$\begin{aligned}
& \min_{\theta_{n,\psi}^l} \text{MSE} \\
& s.t. \quad \mathbf{P} = \left[ \mathbf{G}_1 \check{\mathbf{W}}^H, \mathbf{G}_2 \check{\mathbf{W}}^H, \dots, \mathbf{G}_\Psi \check{\mathbf{W}}^H \right]^H, \\
& \quad \mathbf{G}_\psi = \Phi_\psi^L \mathbf{W}^L \dots \Phi_\psi^2 \mathbf{W}^2 \Phi_\psi^1, \\
& \quad \Phi_\psi^l = \text{diag} \left( e^{j\theta_{1,\psi}^l}, e^{j\theta_{2,\psi}^l}, \dots, e^{j\theta_{N,\psi}^l} \right), l \in \mathcal{L}, \\
& \quad \theta_{n,\psi}^l \in [0, 2\pi), l \in \mathcal{L}, n \in \{1, 2, \dots, N\}.
\end{aligned} \tag{18}$$

Since the optimal configuration of the SIM involves a non-convex optimization problem [22], in this paper, we consider a codebook-based scheme to heuristically produce a suboptimal solution of the phase shifts [23]. Specifically, for each estimator, we randomly generate multiple sets of phase shift vectors as the codebook and select the one that minimizes the corresponding MSE to configure the SIM during channel estimation.

## V. NUMERICAL RESULTS

In this section, we illustrate simulation results to evaluate the performance of the proposed channel estimators. The actual channel coefficients are obtained as  $\mathbf{h}_k = \sqrt{\beta_k} \mathbf{R}_k^{\frac{1}{2}} \mathbf{v}$ , with  $\mathbf{v} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ , and  $\beta_k = C_0 d_k^{-\alpha}$ , where  $C_0$  is the path loss at a reference distance of 1 m,  $d_k$  is the distance from the  $k$ -th UE to the center of the last metasurface layer, and  $\alpha$  is the path loss exponent [24]. We assume that  $K$  UEs are randomly located within a disk parallel to the metasurface layer. The horizontal distance between the last metasurface layer and the disk is 25 m, and the radius of the disk is 10 m. Moreover, the carrier frequency is 28 GHz. As for the large-scale path loss, we consider  $C_0 = -30$  dB,  $\alpha = 2.8$ , and the noise power is  $\sigma_n^2 = -104$  dBm. For the SIM, we consider  $d_t = d_{\text{Layer}} = 5\lambda/L$ ,  $d_x = d_y = \lambda/8$  and  $r_{\text{atom}} = \lambda/4$ . The number of metasurface layers and the number of antennas are  $L = 6$  and  $M = 5$ , respectively.

First, we verify the analytical results by assuming that the spatial correlation matrix  $\mathbf{R}_{\text{iso}}$  is known. As shown in Fig. 3a, the derived theoretical NMSEs accurately characterize the performance of all the proposed estimators. Furthermore, the LS estimator is highly sensitive

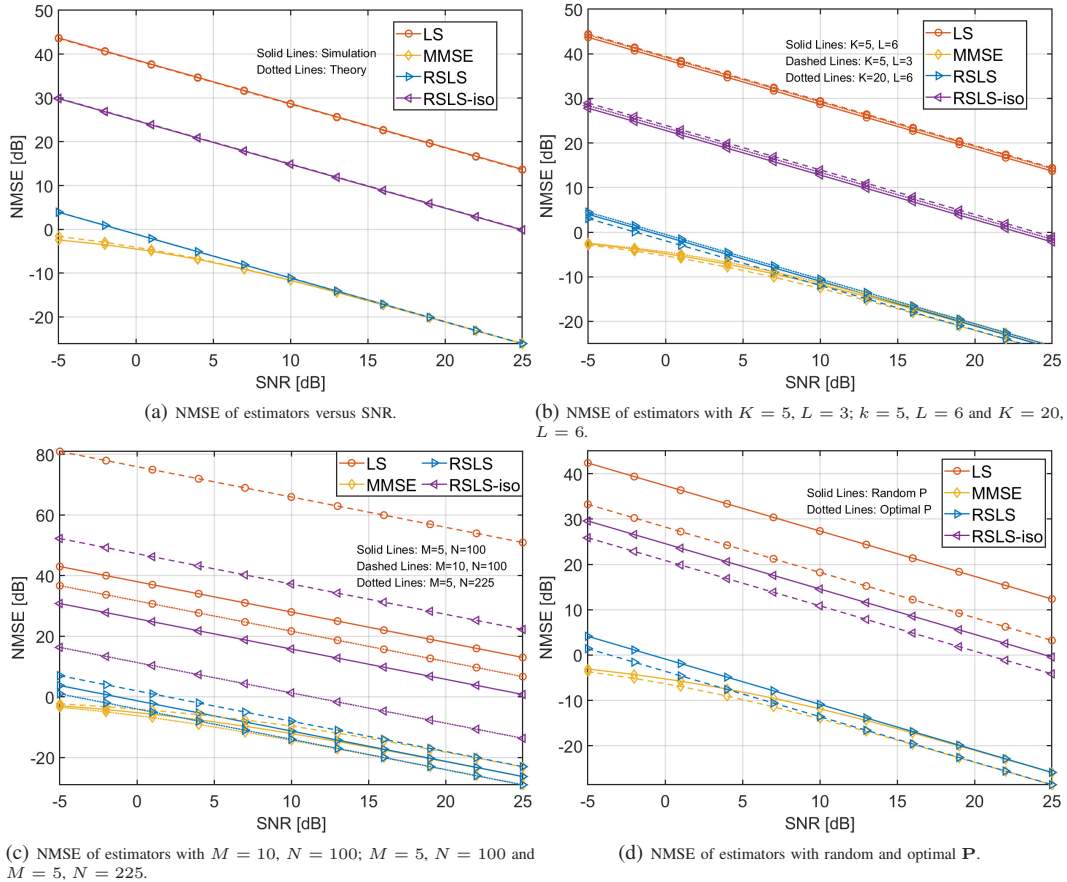


Fig. 3. NMSE of the proposed channel estimators.

to the noise, while the MMSE estimator can effectively suppress the noise and provide more than 35 dB performance gain. The RLS estimator relying on statistical CSI offers the same performance as the MMSE estimator in the high SNR region. Moreover, although the RLS-iso estimator underperforms the MMSE and RLS estimators, it still provides a performance gain of 14 dB over the LS estimator by exploiting the rank deficiency of  $\mathbf{R}_{\text{iso}}$ .

In Fig. 3b, we evaluate the NMSE of the proposed channel estimators considering a different number of UEs and SIM layers. Note that the proposed channel estimation protocol can mitigate the interference among multiple UEs and is slightly affected by  $L$ . Moreover, Fig. 3c evaluates the impact of the number of antennas and meta-atoms on the estimation performance, which demonstrates that the NMSE of the proposed estimators is improved by increasing the number of meta-atoms  $N$ . This is because the rank deficiency of the spatial correlation matrix is more significant as the number of meta-atoms increases [18]. Also, we increase the number of antennas  $M$  from 5 to 10, and observe that the performance of all channel estimators degrades due to the

fact that the properties of the matrix  $\mathbf{P}$  are limited by the reduced number of observation blocks  $\Psi = \lceil N/M \rceil$ . Under all setups, the MMSE estimator and the RLS estimator offer a stronger robustness than the LS estimator and the RLS-iso estimator. Finally, we select the optimal  $\mathbf{P}$  for each estimator from a random codebook of size 100 that minimizes the MSE of the four estimators. Compared to a random SIM configuration, Fig. 3d indicates that a well-optimized SIM is capable of improving the NMSE for all the estimators.

## VI. CONCLUSIONS

In this letter, we proposed channel estimators for SIM-assisted multi-user HMIMO communication systems. Channel estimation is, in fact, a crucial task for subsequent beamforming and signal processing in SIM-assisted communication systems. To this end, we designed an effective channel estimation protocol to solve the underdetermined problem where the number of antennas at the BS is smaller than the number of meta-atoms on the last metasurface layer of the SIM. Four efficient channel estimation techniques were proposed. Additionally, we derived the MSE of the proposed channel estimators and minimized them by appropriately adjusting the phase shifts of the SIM.

## REFERENCES

- [1] J. An, C. Yuen, L. Dai, M. Di Renzo, M. Debbah, and L. Hanzo, "Toward beamfocusing-aided near-field communications: Research advances, potential, and challenges," *arXiv preprint arXiv:2309.09242*, 2023.
- [2] C.-X. Wang, X. You, X. Gao *et al.*, "On the road to 6G: Visions, requirements, key technologies and testbeds," *IEEE Commun. Surv. Tutor.*, vol. 25, no. 2, pp. 905–974, Feb. 2023.
- [3] F. Tang, X. Chen, M. Zhao, and N. Kato, "The roadmap of communication and networking in 6G for the metaverse," *IEEE Wireless Commun.*, vol. 30, no. 4, pp. 72–81, Aug. 2023.
- [4] J. An, C. Yuen, C. Huang *et al.*, "A tutorial on holographic MIMO communications—part I: Channel modeling and channel estimation," *IEEE Commun. Lett.*, vol. 27, no. 7, pp. 1664–1668, Jul. 2023.
- [5] A. Pizzo, L. Sanguinetti, and T. L. Marzetta, "Fourier plane-wave series expansion for holographic MIMO communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 6890–6905, Sep. 2022.
- [6] J. An, C. Yuen, C. Xu, H. Li, D. W. K. Ng, M. Di Renzo, M. Debbah, and L. Hanzo, "Stacked intelligent metasurface-aided mimo transceiver design," *arXiv preprint arXiv:2311.09814*, 2023.
- [7] C. Liu, Q. Ma, Z. J. Luo *et al.*, "A programmable diffractive deep neural network based on a digital-coding metasurface array," *Nat. Electron.*, vol. 5, no. 2, pp. 113–122, Feb. 2022.

- [8] R. Deng, Y. Zhang, H. Zhang *et al.*, “Reconfigurable holographic surface: A new paradigm to implement holographic radio,” *IEEE Veh. Technol. Mag.*, vol. 18, no. 1, pp. 20–28, Mar. 2023.
- [9] C. Huang, S. Hu, G. C. Alexandropoulos *et al.*, “Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges, and trends,” *IEEE Wireless Commun.*, vol. 27, no. 5, pp. 118–125, Oct. 2020.
- [10] J. An, M. Di Renzo, M. Debbah, and C. Yuen, “Stacked intelligent metasurfaces for multiuser beamforming in the wave domain,” in *Proc. IEEE Int. Conf. Commun. (ICC), Rome, Italy*, May. 2023, pp. 1–6.
- [11] J. An, C. Xu, D. W. K. Ng *et al.*, “Stacked intelligent metasurfaces for efficient holographic MIMO communications in 6G,” *IEEE J. Sel. Areas Commun.*, vol. 41, no. 8, pp. 2380–2396, Aug. 2023.
- [12] J. An, M. Di Renzo, M. Debbah, H. V. Poor, and C. Yuen, “Stacked intelligent metasurfaces for multiuser downlink beamforming in the wave domain,” *arXiv preprint arXiv:2309.02687*, 2023.
- [13] X. Peng, L. Zhao, Y. Jiang *et al.*, “Channel estimation for extremely large-scale massive MIMO systems in hybrid-field channel,” in *2023 IEEE/CIC International Conference on Communications in China (ICCC)*. IEEE, Aug. 2023, pp. 1–6.
- [14] M. Di Renzo, A. Zappone, M. Debbah *et al.*, “Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead,” *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.
- [15] X. Lin, Y. Rivenson, N. T. Yardimci *et al.*, “All-optical machine learning using diffractive deep neural networks,” *Sci.*, vol. 361, no. 6406, pp. 1004–1008, Jul. 2018.
- [16] J. An, C. Yuen, Y. L. Guan, M. Di Renzo, M. Debbah, H. V. Poor, and L. Hanzo, “Two-dimensional direction-of-arrival estimation using stacked intelligent metasurfaces,” *arXiv preprint arXiv:2402.08224*, 2024.
- [17] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge university press, 2005.
- [18] Ö. T. Demir, E. Björnson, and L. Sanguinetti, “Channel modeling and channel estimation for holographic massive MIMO with planar arrays,” *IEEE Wireless Commun. Lett.*, vol. 11, no. 5, pp. 997–1001, Feb. 2022.
- [19] C. Wei, H. Jiang, J. Dang *et al.*, “Accurate channel estimation for mmwave massive MIMO with partially coherent phase offsets,” *IEEE Commun. Lett.*, vol. 26, no. 9, pp. 2170–2174, Sep. 2022.
- [20] E. Björnson and L. Sanguinetti, “Rayleigh fading modeling and channel hardening for reconfigurable intelligent surfaces,” *IEEE Wireless Commun. Lett.*, vol. 10, no. 4, pp. 830–834, Dec. 2020.
- [21] M. Biguesh and A. B. Gershman, “Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals,” *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884–893, Mar. 2006.
- [22] Q.-U.-A. Nadeem, J. An, and A. Chaaban, “Hybrid digital-wave domain channel estimator for stacked intelligent metasurface enabled multi-user MISO systems,” *arXiv preprint arXiv:2309.16204*, Sep. 2023.
- [23] J. An, C. Xu, Q. Wu, D. W. K. Ng, M. D. Renzo, C. Yuen, and L. Hanzo, “Codebook-based solutions for reconfigurable intelligent surfaces and their open challenges,” *IEEE Wireless Communications*, pp. 1–8, 2022, Early Access.
- [24] J. An, C. Xu, L. Gan, and L. Hanzo, “Low-complexity channel estimation and passive beamforming for RIS-assisted MIMO systems relying on discrete phase shifts,” *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 1245–1260, Feb. 2022.