# Transmit Power Minimization in Stacked Intelligent Metasurface-aided Multi-User Systems

Haoxian Niu*, Jiancheng An†¶, Shining Lin*, Lu Gan*, Michail Matthaiou§ and Symeon Chatzinotas‡
*School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China
†School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore 639798, Singapore
§Centre for Wireless Innovation (CWI), Queen's University Belfast, U.K.
‡The Interdisciplinary Centre for Security, Reliability, and Trust (SnT), Luxembourg, L-1855 Luxembourg
Email: haoxian_niu@163.com, jiancheng.an@ntu.edu.sg, 202221011710@std.uestc.edu.cn, ganlu@uestc.edu.cn, m.matthaiou@qub.ac.uk, symeon.chatzinotas@uni.lu
¶Corresponding author: Jiancheng An

***Abstract*—Stacked Intelligent Metasurfaces (SIM), emerging as a revolutionary programmable electromagnetic architecture, have demonstrated unprecedented capabilities in manipulating wireless propagation environments. However, the existing research on SIM-aided downlink communication does not consider the fairness between users. Therefore, this paper studies a SIM-aided digital-analog hybrid system, which aims to fairly guarantee the communication quality of each user while minimizing the transmission power. The hybrid system leverages SIM for channel improvement with digital precoding further eliminating the interference between users. To this end, we formulate a transmit power minimization problem under quality-of-service constraints, solved by an efficient alternating optimization algorithm. Simulation results demonstrate that compared to conventional fully digital MIMO systems, the proposed SIM-aided hybrid system achieves 6.93 dBm lower transmit power under the same signal-to-interference-plus-noise ratio (SINR) constraints for users. This work reveals SIM's powerful wave-based beamforming capability in channel enhancement, providing effective solutions for energy-efficient networks with low hardware costs.**

***Index Terms*—Stacked intelligent metasurfaces (SIM), joint active and passive beamforming, wave-based beamforming, phase shift optimization.**

## I. Introduction

THE sixth-generation (6G) mobile networks hold promise for supporting emerging applications, including holographic video, augmented reality, and so on. To fulfill these visions, extremely large-scale antenna arrays (ELAA) are proposed as one of the key technological drivers for realizing 6G goals, thanks to the significant spatial multiplexing and beamforming gains [1]. Nevertheless, ELAA for 6G implies a great increase in the number of antennas, which accordingly brings high hardware costs and extremely huge energy consumption, as each antenna is typically equipped with a high-precision digital-to-analog converter (DAC) to drive radio frequency (RF) chain. Therefore, it is crucial to develop technologies that simultaneously have low hardware costs, low power consumption, and high spectral efficiency for realizing the next-generation wireless communication networks.

To address these challenges, the stacked intelligent metasurface (SIM) is emerging as a revolutionary new technology [2], [3]. A SIM is composed of multiple layers of programmable metasurfaces, each with a large number of meta-atoms. This structure endows SIM with the powerful capability to perform signal processing directly in the electromagnetic wave domain, which makes it feasible to use low-resolution power-efficient DAC to drive RF chains [2], [4], [5]. Moreover, thanks to the immense degrees of freedom provided by SIM, even without baseband digital precoding, SIMs deployed adjacent to antennas can independently and effectively eliminate inter-user interference, while reducing the number of RF chains [4]. Specifically, An et al. proposed a SIM-aided holographic MIMO transceiver and applied SIM to approximate an end-to-end diagonal channel matrix. And an efficient gradient descent algorithm was proposed to adjust the phase shifts of the SIM. After being optimized, SIM deployed near the transmitter and receiver can perform precoding and combining at the speed of light in the electromagnetic wave domain without digital beamforming at the baseband, achieving significant spatial gains and a substantial increase in channel capacity compared to traditional MIMO systems [2]. In addition, the authors of [4] considered a SIM-aided multi-user MISO communication system and developed an alternating optimization (AO) algorithm to optimize the SIM's phase shifts with the objection of maximizing the sum rate. By alternately optimizing the phase shifts of the SIM and the power allocation at the base station (BS), the sum rate of the SIM-aided system achieves a 200% increase compared to traditional MISO systems.

Nevertheless, the existing works on SIM-aided communication systems typically consider maximizing the sum rate of communication systems as a communication metrix [4], [6]–[12]. Although these studies provide important benchmarks for characterizing the upper limit of communication capacity in SIM-aided communication systems, there may be scenarios where some users achieve high communication rates while others are not served, causing the unfairness of users.

Motivated by the aforementioned observations, in this paper, we propose a SIM-aided hybrid analog-digital system architecture, which can effectively and fairly guarantee the communication rate for all users while minimizing transmit power.

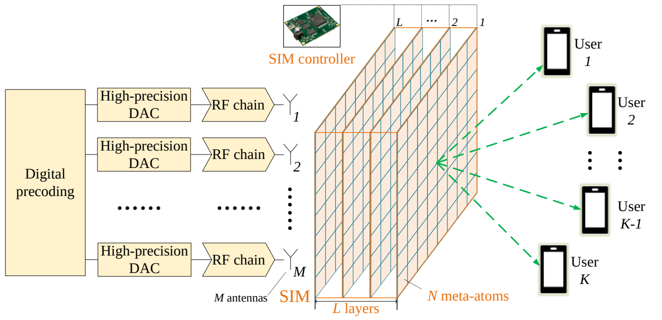


Fig. 1. The proposed SIM-aided analog system.

The system utilizes the wave domain beamforming capability of SIM to improve the channel and digital beamforming to eliminate inter-user interference, which greatly improves the energy efficiency and reduces the hardware cost.

### A. Notations

Bold lowercase and uppercase letters denote vectors and matrices, respectively; $(\cdot)^H$ represents the Hermitian transpose; $\Im(c)$ denotes the imaginary part of a complex number $c$; $\|\cdot\|$ is the Frobenius norm; $\mod(x,n)$ returns the remainder after division of $x$ by $n$; diag$(\mathbf{v})$ produces a diagonal matrix with the elements of $\mathbf{v}$ on its main diagonal; $M \otimes N$ represents the Kronecker product of the matrices $M$ and $N$; vec$(\cdot)$ represents the vectorization operator; $\mathbb{C}^{x\times y}$ represents the space of $x \times y$ complex-valued matrices; The distribution of a circularly symmetric complex Gaussian random vector with a mean vector $\mathbf{v}$ and a covariance matrix $\mathbf{\Sigma}$ is expressed as $\sim \mathcal{CN}(\mathbf{v}, \mathbf{\Sigma})$, where $\sim$ stands for "distributed as".

## II. System Model

### A. Multiuser MISO System Based on SIM

Fig. 1 shows a SIM-aided hybrid analog-digital multiuser MIMO downlink communication system based on SIM, where the SIM is strategically deployed near the BS to facilitate beamforming process. The BS is equipped with a uniform linear array comprising $M$ antennas, serving $K$ users. The SIM, composed of $L$ layers of programmable metasurfaces and $N$ meta-atoms on each layer, is designed to perform analog beamforming directly in the electromagnetic wave domain, thereby enhancing the efficiency and effectiveness of the communication system. We can customize the wireless channel environment by using an intelligent controller, such as a field programmable gate array (FPGA), to manipulate the phase shift of each metasurface layer of the SIM. Let $\mathcal{L} = \{1,2,\ldots,L\}, \mathcal{N} = \{1,2,\ldots,N\}, \mathcal{M} = \{1,2,\ldots,M\}, \mathcal{K} = \{1,2,\ldots,K\}$ represent the sets of metasurface layers, meta-atoms on each layer, and UEs, respectively.

In this paper, we primarily consider downlink communication, where the data streams are first processed by the digital parts and then sent from the antennas in the form of electromagnetic waves to the SIM. The electromagnetic signals are automatically processed as they pass through the configured SIM. And after passing through the SIM, directional beams associated with different users are automatically generated.

We define the phase shift matrix of the $l$-th layer as $\mathbf{\Phi}^l, l \in \mathcal{L}$, which can be written as

$$\mathbf{\Phi}^l = \text{diag}(e^{j\theta_1^l}, e^{j\theta_2^l}, \ldots, e^{j\theta_N^l}) \in \mathbb{C}^{N\times N}, l \in \mathcal{L}, \quad (1)$$

where $\theta_n^l \in [0, 2\pi], l \in \mathcal{L}, n \in \mathcal{N}$ denotes the phase shift of the $n$-th meta-atom of the $l$-th metasurface layer.

And the transmission coefficient matrix from the $(l-1)$-th metasurface layer to the $l$-th layer can be defined as $\mathbf{W}^l \in \mathbb{C}^{N\times N}, \forall l \neq 1, l \in \mathcal{L}$, which is related to the physical distribution of the metasurface.

Specifically, assuming that all metasurface layers have an isomorphic lattice arrangement, each layer of the SIM can be modeled as a uniform planar array. [4], [7] The distance between the $n$-th and $\tilde{n}$-th meta-atoms on the same metasurface layer can be expressed as

$$r_{n,\tilde{n}} = r_e\sqrt{(n_z - \tilde{n}_z)^2 + (n_x - \tilde{n}_x)^2}, \quad (2)$$

where $r_e$ represents the element spacing between adjacent meta-atoms on the same transmit metasurface. Furthermore, $n_x$ and $n_z$ denote the coordinates of the $n$-th meta-atom on that layer, determining the position of the $n$-th meta-atom along the $x$-axis and $z$-axis, respectively, which can be given by the following equation:

$$n_x = \mod(n-1, n_{\text{max}}) + 1, \quad n_z = \left\lceil \frac{n}{n_{\text{max}}} \right\rceil, \quad (3)$$

where $n_{\text{max}}$ represents the number of meta-atoms in each row of the metasurface. Since the SIM is constructed by stacking multiple identical metasurfaces and each metasurface is typically considered as a square metasurface array, generally we have $N = n_{\text{max}}^2$.

Next, we consider the distance between meta-atoms between adjacent metasurface layers. For simplicity, let $d_{\text{layer}}$ denotes the interlayer spacing of the SIM, assuming that the spacing between each layer of the SIM is uniform and that each layer of the SIM metasurfaces is parallel to each other. Then, the distance from the $n$-th meta-atom of the $(l-1)$-th layer to the $\tilde{n}$-th meta-atom of the $l$-th layer can be expressed as

$$d_{n,\tilde{n}}^l = \sqrt{r_{n,\tilde{n}}^2 + d_{\text{layer}}^2}, \quad l \in \mathcal{L}/\{1\}. \quad (4)$$

Consider the interlayer transmission process of the SIM. After each layer of the SIM's meta-atoms receives the electromagnetic wave signal from the previous layer or the BS, each meta-atom becomes a new point wave source, continuing to propagate the electromagnetic signal to the next layer. According to the Rayleigh-Sommerfeld diffraction theory, the $(n,\tilde{n})$-th coefficient of $\mathbf{W}^l$ can be obtained by [13]

$$w_{n,\tilde{n}}^l = \frac{A_t \cos\psi_{n,\tilde{n}}^l}{d_{n,\tilde{n}}^l}\left(\frac{1}{2\pi d_{n,\tilde{n}}^l} - j\frac{1}{\lambda}\right)e^{j2\pi d_{n,\tilde{n}}^l/\lambda}, \quad (5)$$

where $A_t$ is the area of each meta-atom on the metasurface, $\lambda$ is the wavelength, while $\psi_{n,\tilde{n}}^l$ represents the angle between the propagation direction and the normal direction of the $(l-1)$-st transmit metasurface layer.

Thus, the transmission matrix of SIM could be expressed as follow [2]

$$\mathbf{G} = \mathbf{\Phi}^L \mathbf{W}^L \mathbf{\Phi}^{L-1} \mathbf{W}^{L-1} \dots \mathbf{\Phi}^2 \mathbf{W}^2 \mathbf{\Phi}^1 \in \mathbb{C}^{N \times N}. \quad (6)$$

Considering the transmission process from the BS to the SIM. We assume that the uniform linear array (ULA) at the BS is parallel to the SIM's metasurface layers, and the center of the array is aligned with the center of the SIM. And the spacing between antennas is $\lambda/2$. The relative distance between the $m$-th antenna of the ULA and the $n$-th meta-atom of the first metasurface layer on $z$ -axis and the $x$ -axis can be expressed as $d_{m,n}^z = \left(n_z - \frac{n_{\max}+1}{2}\right) r_e - \left(m - \frac{M+1}{2}\right) \frac{\lambda}{2}$ and $d_{m,n}^x = (n_x - \frac{n_{\max}+1}{2}) r_e$, respectively. Then, the distance from the $m$-th antenna at the BS to the $n$-th meta-atom of the first layer of the SIM can be represented by

$$d_{m,n}^l = \sqrt{{d_{m,n}^z}^2 + {d_{m,n}^x}^2 + d_{\text{layer}}^2}. \quad (7)$$

Moreover, the transmission matrix from the BS to the first layer of the SIM is calculated similarly to the interlayer propagation matrix. Let $\mathbf{W}^1 = [\mathbf{w}_1^1, \mathbf{w}_2^1, \dots, \mathbf{w}_M^1] \in \mathbb{C}^{N \times M}$ denotes the transmission channel from antennas of BS to the input layer of the SIM, where $\mathbf{w}_m^1, m \in \mathcal{M}$ represents the channel from the $m$-th antenna to the input layer of the SIM.

### *B. Channel Model*

We model the channel as a millimeter wave channel [14]. Specifically, due to the typically limited scattering in millimeter wave channels, we use a cluster channel model with $Q$ scattering clusters. We assume that each cluster contains a single propagation path from the SIM to the user. Thus, the transmission channel between the output layer of the SIM and the users can be written as

$$\mathbf{h}_k = \sqrt{\frac{N}{Q}} \sum_{q=1}^{Q} \beta_{q,k} \mathbf{a}(\xi_{q,k}, \phi_{q,k}), \forall k, \quad (8)$$

where $\beta_{q,k}$ represents the complex gain of the $q$-th path for the $k$-th user, $\xi_{q,k} \in [0, 2\pi]$ and $\phi_{q,k} \in [0, 2\pi]$ represent the corresponding azimuth and elevation angles of arrival, respectively. Additionally, $\mathbf{a}(\xi_{q,k}, \phi_{q,k})$ denotes the steering vector with respect to the last layer of the SIM. For the sake of convenience, we first define the electrical angles in the x and z directions as

$$\vartheta_{q,k} = \kappa d_x \sin \xi_{q,k} \cos \phi_{q,k}, \forall k, \quad (9)$$

$$\gamma_{q,k} = \kappa d_z \sin \xi_{q,k} \sin \phi_{q,k}, \forall k, \quad (10)$$

respectively, where $\kappa = 2\pi/\lambda$ is the wavenumber, and $d_x = d_z = \lambda/2$ represent the spacing between meta-atoms in the x- and z-directions, respectively. Therefore, the steering vector of (8) can be expressed as [14]–[16]

$$\begin{aligned} \mathbf{a}(\xi_{q,k}, \phi_{q,k}) &= \mathbf{a}(\vartheta_{q,k}, \gamma_{q,k}) \\ &= \frac{1}{\sqrt{N}} \mathbf{a}_{\mathbf{x}}(\vartheta_{q,k}) \otimes \mathbf{a}_{\mathbf{z}}(\gamma_{q,k}), \forall k, \end{aligned} \quad (11)$$

where we have $N = N_x N_z$, with $N_x$ and $N_z$ representing the number of meta-atoms in the x- and z-directions, respectively. Thus, $a_x(\vartheta_{q,k}) \in \mathbb{C}^{N_x \times 1}$ and $a_z(\gamma_{q,k}) \in \mathbb{C}^{N_z \times 1}$ can be written as

$$\mathbf{a}_{\mathbf{x}}(\vartheta_{q,k}) = [1, e^{j\vartheta_{q,k}}, e^{j2\vartheta_{q,k}}, \dots, e^{j(N_x-1)\vartheta_{q,k}}]^T, \quad (12)$$

$$\mathbf{a}_{\mathbf{z}}(\gamma_{q,k}) = [1, e^{j\gamma_{q,k}}, e^{j2\gamma_{q,k}}, \dots, e^{j(N_z-1)\gamma_{q,k}}]^T, \quad (13)$$

respectively. In this paper, we assume that the channel state information $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K]^H \in \mathbb{C}^{K \times N}$ has been perfectly known by the BS.

## III. Problem formulation

When combining SIM with the digital beamforming, we can further improve the degree of freedom for beamforming and enhance the performance of SIM-aided communication systems. Thus, SIM-aided hybrid systems can achieve better performance with fewer RF chains than traditional fully digital systems. In the SIM-aided hybrid system shown in Fig. 1, transmit signals are first processed by the digital precoding and then the combined data streams are sent to the SIM.

We denote the normalized signal sent to the $k$-th user as $s_k$. It is assumed that $s_k, k \in \mathcal{K}$ are independent random variables with zero mean and unit variance. In addition, we define $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_K] \in \mathbb{C}^{M \times K}$ as the digital beamforming matrix, where $\mathbf{b}_k, k \in \mathcal{K}$ is the beamforming vector for the $k$-th user. Thus, the received signal of the $k$-th user can be expressed as

$$y_k = \mathbf{h}_k^H \mathbf{G} \mathbf{W}^1 \sum\nolimits_{i=1}^{K} \mathbf{b}_i s_i + n_k, \forall k, \quad (14)$$

where $n_k \sim \mathcal{CN}(0, \sigma_k^2)$ represents additive white Gaussian noise at user $k$ and $\sigma_k^2$ is the average noise power. The corresponding signal-to-interference-plus-noise ratio (SINR) of user $k$ is

$$\Gamma_k = \frac{|\mathbf{h}_k^H \mathbf{G} \mathbf{W}^1 \mathbf{b}_k|^2}{\sum_{i \neq k}^{K} |\mathbf{h}_k^H \mathbf{G} \mathbf{W}^1 \mathbf{b}_i|^2 + \sigma_k^2}, \forall k. \quad (15)$$

In this paper, we aim to minimize the transmit power while ensuring the communication quality of each user, in order to achieve low-power communication and enhance energy efficiency.

This is achieved by jointly optimizing the digital beamforming matrix and the phase shifts of the SIM, subject to the SINR constraints for each user. Specifically, the optimization problem is formulated as

$$(\text{P1}): \min_{\mathbf{\Phi}^l, \mathbf{B}} \quad P = \text{tr}(\mathbf{B}\mathbf{B}^H) \quad (16a)$$

$$\text{s.t.} \quad \Gamma_k \geq \Gamma_k^{QoS}, \forall k, \quad (16b)$$

where $\Gamma_k^{QoS}$ represents the SINR requirement for the $k$-th user, $P$ denotes the total transmit power at the BS.

## IV. The proposed alternating optimization algorithm

To solve this problem, we propose an alternating optimization algorithm to alternately optimize the digital precoding matrix $\mathbf{B}$ and the phase shifts of the SIM $\mathbf{\Phi}^l$ to solve the problem.

At first, we randomly initialize the phases of each layer of the SIM, and then alternate the following two steps until the total transmit power converges.

***Step 1: Optimization of*** $\mathbf{B}$ ***for given*** $\mathbf{\Phi}^l$

To facilitate the discussion, we first define the equivalent channel from the antennas to the $k$-th users as $\mathbf{g}_k^H = \mathbf{h}_k^H \mathbf{G} \mathbf{W}^1, k \in \mathcal{K}$. When the phase shift matrix of each layer of the SIM is fixed, the equivalent channel $\mathbf{g}_k^H$ remains constant. Thus, (P1) degenerates into a conventional power minimization problem in the multi-user MISO downlink broadcast channel, which can be reformulated as

$$\text{(P2)}: \min_{\mathbf{B}} \quad ||\text{vec}(\mathbf{B})|| \tag{17a}$$

$$\text{s.t.} \quad \frac{|\mathbf{g}_k^H \mathbf{b}_k|^2}{\sum_{i \neq k}^K |\mathbf{g}_k^H \mathbf{b}_i|^2 + \sigma_k^2} \geq \Gamma_k^{QoS}, \forall k. \tag{17b}$$

This problem can be efficiently solved using Second-Order Cone Programming (SOCP) [17] or SDP [18]. Besides, the SINR constraints (17b) hold with equality at the optimal solution of (P2) [19].

***Step 2: Optimization of*** $\mathbf{\Phi}^l$ ***for given*** $\mathbf{B}$

Before discussing how to optimize $\mathbf{\Phi}^l$, we consider a single-user scenario to get some insights for conventional minimum power beamforming design. Specifically, (P2) is reduced to

$$\min_{\mathbf{b}_1} \quad ||\mathbf{b}_1||^2 \tag{18a}$$

$$\text{s.t.} \quad |\mathbf{h}_1^H \mathbf{b}_1|^2 \geq \Gamma^{QoS} \sigma^2. \tag{18b}$$

Apparently, the optimal beamforming vector of the problem (18), $\mathbf{b}_1 = \sqrt{P} \frac{\mathbf{h}_1}{||\mathbf{h}_1||}$, is collinear with the channel vector $\mathbf{h}_1$, which is the well-known maximum ratio transmission (MRT) [20], [21]. Here, we substitute the result of MRT into the problem (18), and can obtain the following problem:

$$\min_{P} \quad P \tag{19a}$$

$$\text{s.t.} \quad P||\mathbf{h}_1^H||^2 \geq \Gamma^{QoS} \sigma^2, \forall k. \tag{19b}$$

It is not difficult to find that the optimal solution of problem (19) is $P = (\Gamma^{QoS} \sigma^2)/||\mathbf{h}_1^H||^2$. Therefore, minimizing the transmit power is equivalent to maximizing the corresponding channel gain.

Although in the multiuser scenario, MRT is not the optimal solution, it can be observed that the less the interference and the greater the channel gain, the lower the transmit power can be. In addition, in the SIM-aided MISO systems, there are generally a large number of the meta-atoms on the output layer. When the number of meta-atoms $N$ tends to infinity (i.e., $N \to \infty$), the user channels exhibit stronger orthogonality, thereby reducing interference, and the MRT performs relatively well [20].

Motivated by these observations, we further employ an alternating optimization algorithm to alternately optimize the phase shifts of each layer to enhance the weighted channel gain for all users, holding the other layers constant while optimizing the phase shifts of one layer. For simplicity, we define the following:

$$\mathbf{h}_{k,l}^H = \begin{cases} \mathbf{h}_k^H \mathbf{\Phi}^L \mathbf{W}^L \mathbf{\Phi}^{L-1} \ldots \mathbf{\Phi}^{l+1} \mathbf{W}^{l+1}, & l \neq L, \\ \mathbf{h}_k^H, & l = L, \end{cases} \tag{20}$$

$$\mathbf{R}_l = \begin{cases} \mathbf{W}^l \mathbf{\Phi}^{l-1} \mathbf{W}^{l-1} \ldots \mathbf{W}^2 \mathbf{\Phi}^1 \mathbf{W}^1, & l \neq 1, \\ \mathbf{W}^1, & l = 1, \end{cases} \tag{21}$$

where $\mathbf{h}_{k,l}^H, k \in \mathcal{K}, l \in \mathcal{L}$ represents the channel from the $l$-th metasurface layer to the $k$-th user, $\mathbf{R}_l, l \in \mathcal{L}$ denotes the transmit channel from the BS to the $l$-th metasurface layer, and $\mathbf{a}_{k,j,l} = \text{diag}(\mathbf{h}_{k,l}^H) \mathbf{R}_l \mathbf{b}_j, k \in \mathcal{K}, l \in \mathcal{L}, j \in \mathcal{K}$ is the coefficient vector of the communication symbols from the $j$-th user arriving at the $k$-th user corresponding to each meta-atom on the $l$-th layer. In addition, we define $\mathbf{v}_l = (e^{-j\theta_1^l}, e^{-j\theta_2^l}, \ldots, e^{-j\theta_N^l})^T$ as the conjugate vector of the phase shifts vector at the $l$-th layer. The combined channel gain can be expressed as

$$\sum_{k=1}^K t_k ||\mathbf{h}_k^H \mathbf{G} \mathbf{W}^1 \mathbf{b}_k||^2 = \sum_{k=1}^K t_k ||\mathbf{v}_l^H \mathbf{a}_{k,k,l}||^2 = \sum_{k=1}^K t_k \text{tr}(\mathbf{A}_{k,k,l} \mathbf{V}_l), \tag{22}$$

where $\mathbf{A}_{k,k,l} = \mathbf{a}_{k,k,l} \mathbf{a}_{k,k,l}^H$ and $\mathbf{V}_l = \mathbf{v}_l \mathbf{v}_l^H$. And we set the weight as $t_k = 1/(\gamma_k \sigma_k^2), k \in \mathcal{K}$. Therefore, the phase shifts of the $l$-th layer can be obtained by solving the following optimization problem:

$$\max_{\mathbf{V}_l} \quad \sum_{k=1}^K t_k \text{tr}(\mathbf{A}_{k,k,l} \mathbf{V}_l) \tag{23a}$$

$$\text{s.t.} \quad |\{\mathbf{v}_l\}_n| = 1, l \in \mathcal{L}, n \in \mathcal{N}, \tag{23b}$$

$$\mathbf{V}_l = \mathbf{v}_l \mathbf{v}_l^H, l \in \mathcal{L}, \tag{23c}$$

The problem (23) is a non-convex problem with a non-convex constraint rank$(\mathbf{V}_l) = 1$. Using the SDR technique to relax this constraint, the problem (23) is reduced to

$$\text{(P3)}: \max_{\mathbf{V}_l} \quad \sum_{k=1}^K t_k \text{tr}(\mathbf{A}_{k,k,l} \mathbf{V}_l) \tag{24a}$$

$$\text{s.t.} \quad \{\mathbf{V}_l\}_{i,i} = 1, i \in \mathcal{N}, l \in \mathcal{L}, \tag{24b}$$

$$\mathbf{V}_l \succeq 0, l \in \mathcal{L}, \tag{24c}$$

Problem (P3) is a convex SDP problem, which can be optimized and solved using CVX. Note that the solution of the problem (24) may not satisfy the rank-one constraint. Thus, after obtaining the high-rank solution $\mathbf{V}_l$, it is necessary to perform multiple Gaussian randomizations to obtain a feasible solution $\mathbf{v}_l$. Existing research has shown that such an SDR method, combined with a sufficiently large number of randomizations, guarantees at least a $\frac{\pi}{4}$-approximation of the optimal objective value for problem (23) [19].

Therefore, by solving (P3), we can obtain the SIM phase shifts of $l$-th layer. In each iteration, we optimize layer by layer from the first layer until we update the phase shift matrix for all layers of the SIM.

For clarity, we summarize the specific steps of the alternating optimization algorithm for SIM-aided hybrid system in Algorithm 1.

**Algorithm 1** The proposed alternating optimization algorithm for SIM-aided hybrid system

1: **Input:** $\mathbf{H}, \mathbf{W}^l, \mathbf{\Phi}^l, \forall l \in \mathcal{L}, \mathbf{B}$;
2: Initialize $\mathbf{\Phi}_l, \forall l \in \mathcal{L}$ randomly and then initialize $\mathbf{B}$ by solving (P2);
3: **Repeat**
4: Update the phase shifts of the $l$-th layer of the SIM in turn by solving the problem (P3);
5: Update the digital beamforming $\mathbf{B}$ by solving the problem (P2);
6: **Until** The total power converges;
7: **Output:** $\mathbf{B}, \mathbf{\Phi}_l, \forall l \in \mathcal{L}$.

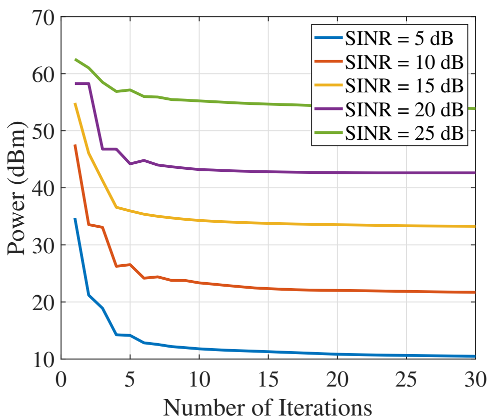


Fig. 2. Transmit power of the SIM-aided hybrid system versus the number of iterations under different SINR constraints for $K$ users, $N = 255, L = 3, K = 5$.

## V. Simulation Results

### A. Simulation Settings

In this section, numerical simulation results demonstrate the effectiveness of the proposed alternating optimization algorithm in the SIM-aided hybrid system. And we outline the parameters of our simulation environment. The simulation is configured with the following settings.

The BS serves $K = 5$ users, each located at a stochastically determined distance ranging from 100 meters to 300 meters relative to the BS. The BS is equipped with $M = 25$ transmit antennas, and the SIM is designed with $N = 225$ meta-atoms and $L = 3$ metasurface layers. The carrier frequency is set at 30 GHz, with the SIM's layer spaced at the wavelength $\lambda$, and the meta-atoms are arrayed with a spacing of $\lambda/2$. The system operates within a bandwidth of 200 MHz. The noise power, denoted as $\sigma_k^2$, is calculated using the formula $1.380649 \times 10^{-23} \times \text{bandwidth} \times 290$, reflecting the thermal noise floor at room temperature.

### B. Numerical Results

As shown in Fig. 2, we present the convergence performance of the proposed Algorithm 1 for the SIM-aided hybrid beamforming system and compare the convergence

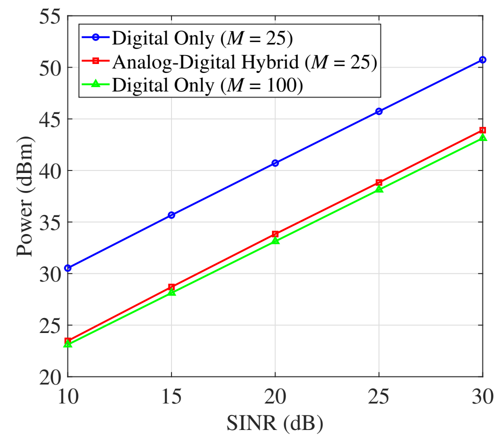


Fig. 3. Comparison for the minimum transmit power of three types of system.

under different settings. The figure indicates that the proposed algorithm converges within approximately 25 iterations under various SINR constraints for different users. Additionally, as the user SINR requirements increase, the final converged transmit power increases, necessitating higher transmit power to satisfy the users' SINR constraints.

Fig. 3 contrasts our proposed SIM-aided hybrid system with traditional digital systems. "Analog-digital Hybrid" denotes the SIM-aided hybrid system, and "Digital Only" indicates the digital system. As the required receive SINR increases, the minimum transmit power for both systems increases to meet the SINR requirements. It is observable that, thanks to the synergistic gain of SIM and digital beamforming, the minimum transmit power of the SIM-aided analog system is approximately 6.93 dBm lower than that of the fully digital system with 25 antennas. This shows that the introduction of SIM can greatly reduce the energy consumption of the system and improve the energy efficiency of the system. In addition, it is noted that the performance of our proposed scheme reaches a purely digital system performance of close to 100 antennas when only 25 antennas are used. Therefore, for achieving the same quality of service, the introduction of SIM can greatly reduce the number of RF chains required and reduce the hardware cost of the system.

Fig. 4 illustrates the variation of the minimum transmit power versus the number of layers under different meta-atom configurations in the SIM-aided hybrid system. As shown in Fig. 4, the system's transmit power gradually decreases as the number of SIM's layers increases. This is because the increased metasurface layer brings additional degrees of freedom and strengthens the capability of wave-domain beamforming. When the number of layers reaches three, the rate of power attenuation tends to stabilize. As the number of metasurface layers exceeds three, the transmit power decreases gradually and reaches its minimum value when the number of layers is around seven. This is partly due to the power attenuation caused by the increased number of layers, and partly because the computational capability provided by the excessive number

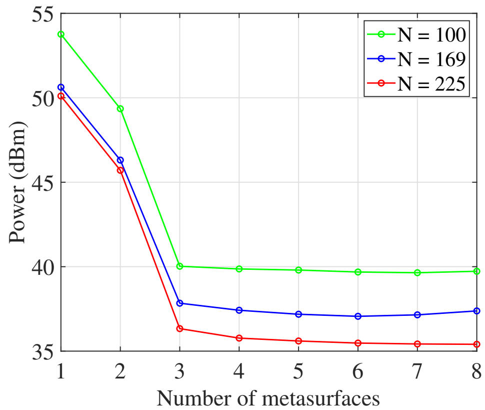


Fig. 4. Transmit power of the SIM-aided hybrid system versus the number of the SIM's metasurface layers.

of layers far exceeds what is needed to solve the optimization problem, leading to a diminishing marginal effect. On the other hand, as the number of meta-atoms increases, the minimum transmit power of the system decreases significantly. Besides, it is noteworthy that the transmitted power of the SIM-aided system with fewer total meta-atoms (3 layers of 100 meta-atoms) is significantly lower than that of the system with more total meta-atoms (2 layers of 225 meta-atoms). This observation indicates that the gain achieved by simply increasing the number of meta-atoms is far less than the gain obtained through the synergistic effect of a multi-layer structure.

## VI. Conclusion

In this study, we thoroughly investigated a multi-user MISO communication system aided by SIM and proposed an effective algorithm to minimize transmit power while ensuring communication quality. By leveraging the channel-enhancing capabilities of SIM in the wave domain and the interference-cancellation capabilities of digital beamforming, the proposed hybrid system demonstrates significant advantages in reducing transmit power and improving energy efficiency. The results show that compared to conventional fully digital system, the proposed SIM-aided hybrid system achieves a transmit power reduction of approximately 6.93 dBm under the same SINR constraints for users. Additionally, increasing the number of SIM's layers and meta-atoms further reduces transmit power, with the synergistic effect of a multi-layer structure proving more effective than simply increasing the number of meta-atoms. Future research can focus on further optimizing SIM configurations and exploring new applications in diverse wireless communication scenarios to fully utilize SIM's potential in enhancing energy efficiency and interference elimination.

## References

[1] M. Cui, Z. Wu, Y. Lu, X. Wei, and L. Dai, "Near-field MIMO communications for 6G: fundamentals, challenges, potentials, and future directions," *IEEE Commun. Mag.*, vol. 61, no. 1, pp. 40–46, Sep. 2023.

[2] J. An, C. Xu, D. W. K. Ng, G. C. Alexandropoulos, C. Huang, C. Yuen, and L. Hanzo, "Stacked intelligent metasurfaces for efficient holographic MIMO communications in 6G," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 8, pp. 2380–2396, Aug. 2023.

[3] J. An, M. Debbah, T. J. Cui, Z. N. Chen, and C. Yuen, "Emerging technologies in intelligent metasurfaces: shaping the future of wireless communications," *arXiv preprint arXiv:2411.19754*, Nov. 2024.

[4] J. An, M. Di Renzo, M. Debbah, and C. Yuen, "Stacked intelligent metasurfaces for multiuser downlink beamforming in the wave domain," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Rome, Italy, May 2023, pp. 1–6.

[5] J. An, C. Yuen, C. Xu, and et al., "Stacked intelligent metasurface-aided MIMO transceiver design," *IEEE Wirel. Commun.*, vol. 31, no. 4, pp. 123–131, Apr. 2024.

[6] D. Darsena, F. Verde, I. Iudice, and V. Galdi, "Design of stacked intelligent metasurfaces with reconfigurable amplitude and phase for multiuser downlink beamforming," *IEEE Open J. Commun. Soc.*, vol. 6, pp. 531–550, Jan. 2025.

[7] A. Papazafeiropoulos, J. An, P. Kourtessis, T. Ratnarajah, and S. Chatzinotas, "Achievable rate optimization for stacked intelligent metasurface-assisted holographic MIMO communications," *IEEE Trans. Wireless. Commun.*, Oct. 2024.

[8] S. Lin, J. An, L. Gan, M. Debbah, and C. Yuen, "Stacked intelligent metasurface enabled LEO satellite communications relying on statistical CSI," *IEEE Wireless. Commun. Lett.*, vol. 13, no. 5, pp. 1295–1299, May 2024.

[9] A. Papazafeiropoulos, P. Kourtessis, S. Chatzinotas, D. I. Kaklamani, and I. S. Venieris, "Performance of double-stacked intelligent metasurface-assisted multiuser massive MIMO communications in the wave domain," *IEEE Trans. Wireless Commun.*, Jan. 2025.

[10] H. Liu, J. An, G. C. Alexandropoulos, D. W. K. Ng, C. Yuen, and L. Gan, "Multi-user MISO with stacked intelligent metasurfaces: a DRL-based sum-rate optimization approach," *IEEE Trans. Cognit. Commun. Networking*, pp. 1–1, Apr. 2025.

[11] H. Niu, J. An, A. Papazafeiropoulos, L. Gan, S. Chatzinotas, and M. Debbah, "Stacked intelligent metasurfaces for integrated sensing and communications," *IEEE Wireless Commun. Lett.*, vol. 13, no. 10, pp. 2807–2811, Oct. 2024.

[12] E. Shi, J. Zhang, J. An, G. Zhang, Z. Liu, C. Yuen, and B. Ai, "Joint AP-UE association and precoding for SIM-aided cell-free massive MIMO systems," *IEEE Trans. Wireless Commun.*, pp. 1–1, Mar. 2025.

[13] X. Lin, Y. Rivenson, N. T. Yardimci, M. Veli, Y. Luo, M. Jarrahi, and A. Ozcan, "All-optical machine learning using diffractive deep neural networks," *Science*, vol. 361, no. 6406, pp. 1004–1008, Jul. 2018.

[14] W. Ma, C. Qi, Z. Zhang, and J. Cheng, "Sparse channel estimation and hybrid precoding using deep learning for millimeter wave massive MIMO," *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 2838–2849, Feb. 2020.

[15] J. An, C. Yuen, Y. L. Guan, M. D. Renzo, M. Debbah, H. V. Poor, and L. Hanzo, "Two-dimensional direction-of-arrival estimation using stacked intelligent metasurfaces," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 10, pp. 2786–2802, Jun. 2024.

[16] J. An, C. Yuen, M. D. Renzo, M. Debbah, H. V. Poor, and L. Hanzo, "Flexible intelligent metasurfaces for downlink multiuser MISO communications," *IEEE Trans. Wireless Commun.*, vol. 24, no. 4, pp. 2940–2955, Jan. 2025.

[17] A. Wiesel, Y. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, 2006.

[18] M. Bengtsson and B. Ottersten, "Optimum and suboptimum transmit beamforming," in *Handbook of Antennas in Wireless Communications*. CRC Press, 2018, pp. 18–1.

[19] 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, Aug. 2019.

[20] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: a difficult problem with a simple solution structure [lecture notes]," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 142–148, Jun. 2014.

[21] T. D. Hua, M. Mohammadi, H. Q. Ngo, and M. Matthaiou, "Swipt in cell-free massive mimo using stacked intelligent metasurfaces," *arXiv preprint arXiv:2503.14032*, Mar. 2025.