Hybrid Spherical- and Planar-Wave Channel Modeling and Estimation for Terahertz Integrated UM-MIMO and IRS Systems

Yuhang Chen, Renwang Li, Student Member, IEEE, Chong Han, Member, IEEE, Shu Sun, Member, IEEE, and Meixia Tao, Fellow, IEEE

Abstract—Integrated ultra-massive multiple-input multiple-output (UM-MIMO) and intelligent reflecting surface (IRS) systems are promising for 6G and beyond Terahertz (0.1-10 THz) communications, to effectively bypass the barriers of limited coverage and line-of-sight blockage. However, excessive dimensions of UM-MIMO and IRS enlarge the near-field region, while strong THz channel sparsity in the far-field is detrimental to spatial multiplexing. Moreover, channel estimation (CE) requires recovering the large-scale channel from severely compressed observations due to limited RF-chains. To tackle these challenges, a hybrid spherical- and planar-wave channel model (HSPM) is introduced for the cascaded channel of the integrated system. The spatial multiplexing gains under near-field and far-field regions are analyzed, which are found to be limited by the segmented channel with a lower rank. Furthermore, a compressive sensing-based CE framework is developed, including a sparse channel representation method, a separate-side estimation (SSE) and a dictionary-shrinkage estimation (DSE) algorithms. Numerical results verify the effectiveness of the HSPM, the capacity of which is only $5 \times 10^{-4}$ bits/s/Hz deviated from that obtained by the ground-truth spherical-wave-model, with 256 elements. While the SSE achieves improved accuracy for CE than benchmark algorithms, the DSE is more attractive in noisy environments, with 1 dB lower normalized-mean-square error than SSE.

Index Terms—Terahertz integrated ultra-massive multiple-input-multiple-output (UM-MIMO) and intelligent reflecting surface (IRS) systems, channel modeling, spatial multiplexing gain, channel estimation.

I. INTRODUCTION

O

WNING abundant bandwidth of multi-GHz up to even Terahertz (THz), the THz spectrum ranging from 0.1 to 10 THz has attracted upsurging attention from academia and industry in recent years. THz wireless communications have the capability to support Terabit-per-second high data rates, which are envisioned as a pillar candidate for 6G wireless networks [2], [3], [4]. However, the THz wave suffers from large free-space attenuation, strong molecular absorption, and high non-line-of-sight (NLoS) propagation losses incurred from reflection, scattering, and diffraction. Therefore, it is challenging to achieve robust wireless transmission in complex occlusion environments, especially when line-of-sight (LoS) is blocked [5]. Moreover, power amplifiers with low efficiency at THz frequencies have constrained output power, which results in the low reception signal-to-noise ratio (SNR), thus constraining the communication distance [6].

To overcome the distance limitation, the ultra-massive multiple-input multiple-output (UM-MIMO) systems are exploited in the THz band [7]. Thanks to the sub-millimeter wavelength, hundreds and even thousands of antennas can be deployed in UM-MIMO, which provides high array gain to compensate for propagation losses. Furthermore, as a key technology to enable intelligent propagation environments in 6G systems, the intelligent reflecting surface (IRS) has been advocated in the literature [8], [9], [10]. The IRS is equipped with a metamaterial surface of the integrated circuit, which can be programmed to enable passive beamforming with high energy efficiency [8]. At lower frequencies, IRS is majorly used to increase the achievable data rates. By contrast, in the THz band, the IRS can effectively bypass the barrier of LoS blockage problem by precisely controlling the reflection of incident THz signals [2], [11]. Therefore, integrated UM-MIMO and IRS systems are promising for THz wireless communications by solving the distance limitation and LoS blockage problems.

Channel modeling, analysis, and channel estimation (CE) arise as three interrelated open challenges of the THz integrated UM-MIMO and IRS systems. First, while most existing work on channel modeling in IRS-assisted systems only
considers the far-field propagation [12], the near-field region is expanded with an enlarged dimension of antenna arrays in UM-MIMO and IRS relative to the sub-millimeter wavelength of the THz wave. The consideration of near-field spherical-wave propagation is imperatively needed [13], [14]. Second, each segmented channel in the integrated IRS and UM-MIMO systems can be in near-field and far-field, whose multiplexing capability concerning the cascaded channel remains unclear. Moreover, due to the large reflection, scattering, and diffraction losses, THz channels are generally sparse and dominated by an LoS and only a few NLoS paths [15]. As a result, THz multi-antenna channels suffer from limited multiplexing capability imposed by the number of multi-paths instead of the number of antennas as in the microwave band. Spatial multiplexing capability needs to be assessed and possibly enhanced in THz integrated UM-MIMO and IRS systems.

Third, hybrid UM-MIMO structures with low hardware cost are commonly deployed in THz systems, which exploit a much smaller number of RF-chains than antennas [16]. This hybrid architecture is helpful in reducing power and hardware costs, which, however, causes a research problem for CE. That is, with the enormous amount of antennas in UM-MIMO and passive reflecting elements lacking signal processing abilities of IRS, CE has to recover a high-dimensional channel relating to the numbers of antennas and passive elements from the severely compressed low-dimensional received signal on RF-chains. Moreover, the consideration of spherical-wave propagation alters the structure of channel models, leading to traditional solutions based on planar-wave propagation to become ineffective. Therefore, new CE methods to address these problems are needed.

A. Related Works

1) Channel Modeling and Analysis: In the literature, mainly two categories of MIMO channel models without IRS are considered, namely, the spherical-wave model (SWM) and the planar-wave model (PWM), which address the near-field and far-field effects, respectively [17], [18]. As an improvement to PWM and SWM, we proposed a hybrid spherical- and planar-wave channel model (HSPM) for THz UM-MIMO systems in [13], which accounted for PWM within the sub-array and SWM among subarrays. Compared to PWM and SWM, HSPM is more effective by deploying a few channel parameters to achieve high accuracy in the near-field. A similar idea that divides the entire antenna array into several subarrays to approximate SWM has also been considered in [19]. In the IRS-assisted communication systems, an alternative physically feasible Rayleigh fading model was proposed in [12] under far-field assumption. By taking both near-field effect and IRS into consideration, authors in [20] considered SWM for THz integrated IRS and UM-MIMO systems. However, SWM suffers from high complexity with the massive number of elements in the UM-MIMO and IRS [13]. To date, an effective model addressing the near-field effect in UM-MIMO and IRS systems is still required.

In IRS and UM-MIMO systems, channel analyses on sum rate, power gain, spectral efficiency (SE), energy efficiency (EE) and spatial degrees-of-freedom (SDoF) have been explored in the literature [21], [22], [23], [24], [25], [26]. In microwave systems, authors in [21] characterized the capacity limit by jointly optimizing the IRS reflection coefficients, and MIMO transmit covariance matrix. The distribution and outage probability of the sum rate were derived in [22] by considering the SWM of the LoS and PWM of the NLoS. A closed-form expression of the power gain was derived in [23], and the near-field and far-field behaviors were analyzed. At higher frequencies, the ergodic capacity under the Saleh-Valenzuela model was derived and optimized in [24]. Furthermore, as a critical metric to assess the spatial-multiplexing capability of the channel, the available SDoF or channel rank for the large intelligent surface systems was studied in [25] by driving the analytical expressions. In THz UM-MIMO systems, to enhance the limited spatial multiplexing, a widely-spaced multi-subarray (WSMS) structure with enlarged subarray spacing was proposed in [26]. The channel rank is enhanced by a factor equal to the number of subarrays, making WSMS a promising structure. However, the rank analysis in the integrated UM-MIMO with WSMS structure and IRS systems is still lacking in the literature.

2) Channel Estimation: CE for IRS-assisted MIMO systems has been explored in the literature [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], and [38], which can be categorized into two main categories, namely, estimation of the segmented channels from user equipment (UE) to IRS and IRS to base station (BS), and estimation of the cascaded channel without separating each segment. On one hand, since the passive IRS lacks signal processing capability, it is hard to separate each channel segment. The segmented CE schemes often require special hardware design, e.g., inserting active IRS elements or using full-duplex equipment, which increases the hardware cost [27], [28], [29], [30]. In [27] and [28], a few IRS elements were activated during pilot reception. The deep-learning tool was then assisted for CE with considerable estimation accuracy. By deploying a full-duplex BS, a two-timescale CE method was proposed in [29]. The segmented CE problem was formulated as a matrix factorization problem and solved in [30], which operates with purely passive IRS. However, this scheme does not address the near-field effect.

On the other hand, since most precoding designs are based on the knowledge of the cascaded channel, the estimation of which has been explored in most existing schemes [31], [32], [33], [34], [35], [36], [37], [38]. In [31], a two-stage atomic norm minimization problem was formulated, by which the super-resolution channel parameter estimation was conducted to obtain the channel state information. Theoretical analysis of the required pilot overhead and a universal CE framework were proposed in [32], which are effective in guiding the design of training and CE. However, all of them are limited to being applicable with fully digital MIMO structures. By exploiting the channel sparsity in the mmWave and THz bands, compressive sensing (CS) based CE methods were explored in [33], [34], [35], [36], [37], and [38]. These schemes deploy the spatial discrete Fourier transform (DFT) based on-grid codebook to sparsely represent the channel, which is beneficial in achieving reduced training overhead. On the downside, the near-field effect was not incorporated in the
DFT codebook, which results in limited estimation accuracy of these schemes in the near-field region. Recently, CE for large-scale MIMO systems by considering near-field channel model was conducted in [39]. However, this scheme suffers from high computational complexity due to the large codebook size. The cascaded channels in the IRS-assisted systems further preclude its direct adoption to IRS-assisted systems. In conclusion, a more effective CE method to assess the near-field effect is demanded in the THz integrated UM-MIMO and IRS systems.

B. Contributions

To fill the aforementioned research gaps, in this work, we first model the cascaded channel and study the spatial multiplexing in THz integrated UM-MIMO and IRS systems by considering both near-field and far-field effects. Based on that, we propose a CS-based CE framework. In particular, we develop a subarray-based on-grid codebook to sparsely represent the channel. Then, a separate side estimation (SSE) and a spatial correlation inspired dictionary shrinkage estimation (DSE) algorithms are proposed to realize low-complexity CE. In our prior and shorter version [1], we introduced the CE algorithms. Furthermore, we perform substantially more extensive analysis of the integrated systems. In this work, we further derive the SSE algorithm to assess the near-field effect is improved based on the widely-spaced architectural design. Moreover, we present that spatial multiplexing can be improved based on the widely-spaced architectural design.

We develop a subarray-based codebook to sparsely represent the channel. Then, a separate side estimation (SSE) and a spatial correlation inspired dictionary shrinkage estimation (DSE) algorithms are proposed to realize low-complexity CE. In our prior and shorter version [1], we introduced the CE algorithms. Furthermore, we perform substantially more extensive analysis of the integrated systems. In this work, we further derive the SSE algorithm to assess the near-field effect is improved based on the widely-spaced architectural design. Moreover, we present that spatial multiplexing can be improved based on the widely-spaced architectural design.

To fill the aforementioned research gaps, in this work, we first model the cascaded channel and study the spatial multiplexing in THz integrated UM-MIMO and IRS systems by considering both near-field and far-field effects. Based on that, we propose a CS-based CE framework. In particular, we develop a subarray-based on-grid codebook to sparsely represent the channel. Then, a separate side estimation (SSE) and a spatial correlation inspired dictionary shrinkage estimation (DSE) algorithms are proposed to realize low-complexity CE. In our prior and shorter version [1], we introduced the CE algorithms. Furthermore, we perform substantially more extensive analysis of the integrated systems. In this work, we further derive the SSE algorithm to assess the near-field effect is improved based on the widely-spaced architectural design. Moreover, we present that spatial multiplexing can be improved based on the widely-spaced architectural design.

We introduce the HSPM for the cascaded channel in the THz integrated UM-MIMO and IRS systems. Based on that, we analyze the spatial multiplexing gain of the cascaded channel. The HSPM accounts for PWM within the subarray and SWM among subarrays, which achieves better accuracy than PWM and lower complexity than SWM. In addition, different from existing studies that model the channel based on subarrays and analyze the performance by evaluating the modeling accuracy [13], [19], in this work, the spatial multiplexing gain of the cascaded channel is analyzed when the segmented channels satisfy the near-field and far-field conditions, respectively. We prove that the rank of the cascaded channel is constrained by the individual channel with a lower rank. Moreover, we present that spatial multiplexing can be improved based on the widely-spaced architectural design.

We develop a subarray-based codebook to sparsely represent the channel. Then, a separate side estimation (SSE) and a spatial correlation inspired dictionary shrinkage estimation (DSE) algorithms are proposed to realize low-complexity CE. In our prior and shorter version [1], we introduced the CE algorithms. Furthermore, we perform substantially more extensive analysis of the integrated systems. In this work, we further derive the SSE algorithm to assess the near-field effect is improved based on the widely-spaced architectural design. Moreover, we present that spatial multiplexing can be improved based on the widely-spaced architectural design.

The remainder of this paper is organized as follows. The system and channel models are introduced in Sec. II. Spatial multiplexing analysis is presented in Sec. III. The subarray-based codebook and the SSE, DSE CE algorithms are proposed in IV. Extensive performance evaluation and numerical analysis are conducted in Sec. V. Finally, the paper is concluded in Sec. VI.

Notation: $a$ is a scalar. $\mathbf{a}$ denotes a vector. $\mathbf{A}$ represents a matrix. $\mathbf{A}(m, n)$ stands for the element at the $m^{th}$ row and $n^{th}$ column in $\mathbf{A}$. $\mathbf{A}(i, :)$ denotes the $i^{th}$ row of $\mathbf{A}$. $\mathbf{A}(:, j)$ refers to the $j^{th}$ column of $\mathbf{A}$. $\mathbf{p}(m : n)$ denotes the $m^{th}$ to $n^{th}$ elements of $\mathbf{p}$. $\mathbf{p}(m)$ refers to the $m^{th}$ element of $\mathbf{p}$. $(\cdot)^{T}$ defines the transpose. $(\cdot)^{H}$ refers to the conjugate transpose. $\parallel \cdot \parallel_{0}$ denotes the pseudo inverse. $\parallel \cdot \parallel_{0}$ defines the $l_{0}$-norm. $\parallel \cdot \parallel_{2}$ stands for the $l_{2}$-norm. $\mathbb{C}^{M \times N}$ depicts the set of $M \times N$-dimensional complex-valued matrices. $\otimes$ refers to Kronecker product. $\circ$ denotes the Khatri-Rao product. $\odot$ depicts a proportional sign.

II. System Overview

A. System Model

As illustrated in Fig. 1, we consider a THz integrated UM-MIMO and IRS communication system. The WSMS THz UM-MIMO with planar-shaped antenna arrays is equipped at both BS and UE. The direct channel between BS and UE is blocked and inaccessible due to the occlusion propagation environment [31], [34]. The communication link is assisted by a planar-shaped IRS with $M$ passive reflecting elements, which is connected to BS via an IRS controller. Moreover, we consider that the IRS can be divided into $K_{m}$ planar-shaped subarrays, with $M = K_{m}N_{am}$, where $N_{am}$ denotes the number of passive reflecting elements on each subarray.

In the WSMS design at BS, $K_{b}$ subarrays are deployed, each of which contains $N_{ab}$ antennas. The total number of antennas is obtained as $N_{b} = K_{b}N_{ab}$. On one hand, within the subarray, the antenna spacing $d = \lambda/2$, where $\lambda$ denotes carrier wavelength. On the other hand, the subarray spacing is multiple times half-wavelengths [26]. Each subarray is connected to one RF-chain. In the THz UM-MIMO systems, a much smaller number of RF-chains antennas is often adopted for lower hardware cost and higher EE [16]. Therefore, we have $K_{b} \ll N_{b}$. Similarly, UE is composed of $N_{u}$ antennas, which can be divided into $K_{u}$ subarrays, each connected to one RF-chain. Each subarray contains $N_{au}$ antennas, satisfying $N_{a} = K_{u}N_{au}$ and $K_{u} \ll N_{u}$.
By considering an uplink transmission, the received signal \( y \in \mathbb{C}^{N_{sb}} \) at BS is denoted as

\[
y = \mathbf{W}^H \mathbf{H}^{\text{cas}} \mathbf{F}_s + \mathbf{W}^H \mathbf{n}, \tag{1}
\]

where \( N_{sb} \) denotes the number of signal streams at BS, \( \mathbf{W} = \mathbf{W}_{\text{RF}} \mathbf{W}_{\text{BB}} \in \mathbb{C}^{N_{sb} \times N_{au}} \) represents the combining matrix, with \( \mathbf{W}_{\text{RF}} \in \mathbb{C}^{N_{sb} \times K_b} \) and \( \mathbf{W}_{\text{BB}} \in \mathbb{C}^{K_b \times N_{sb}} \) denoting the analog and digital combining matrices, respectively. The cascaded channel matrix from the UE to BS is depicted as \( \mathbf{H}^{\text{cas}} \in \mathbb{C}^{N_{su} \times N_{su}} \). The beamforming matrix at the UE is represented as \( \mathbf{F}_s = \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \in \mathbb{C}^{N_{su} \times N_{su}} \), where \( N_{su} \) depicts the transmitted number of signal streams at UE, \( \mathbf{F}_{\text{RF}} \in \mathbb{C}^{N_{su} \times K_u} \) and \( \mathbf{F}_{\text{BB}} \in \mathbb{C}^{K_u \times N_{su}} \) refer to the analog and digital beamforming matrices, respectively. Moreover, \( s \in \mathbb{C}^{N_{su}} \) describes the transmitted signal, while \( n \in \mathbb{C}^{N_{su}} \) represents additive white Gaussian noise (AWGN). The analog beamforming and combining are completed by phase shifters. Therefore, each element of \( \mathbf{W}_{\text{RF}} \) and \( \mathbf{F}_s \) satisfies constant module constraint, which can be expressed as \( \mathbf{W}_{\text{RF}}(i,j) = \frac{1}{\sqrt{N_u}} e^{jw_{i,j}}, \mathbf{F}_s(i,j) = \frac{1}{\sqrt{N_u}} e^{j\pi f_{i,j}} \), where \( w_{i,j}, f_{i,j} \in [0, 2\pi] \) denote the phase shift value. In addition, \( \mathbf{W}_{\text{BB}} \) and \( \mathbf{F}_{\text{BB}} \) are usually set as identity matrices during the training process for CE. In this case, there is \( N_{su} = K_u \) and \( N_{sb} = K_b \).

**B. Channel Model**

The cascaded channel matrix \( \mathbf{H}^{\text{cas}} \) in (1) can be represented as

\[
\mathbf{H}^{\text{cas}} = \mathbf{H}_{\text{IRS--BS}} \mathbf{P} \mathbf{H}_{\text{UE--IRS}}, \tag{2}
\]

where \( \mathbf{H}_{\text{IRS--BS}} \in \mathbb{C}^{N_u \times M} \) stands for the segmented channel from IRS to BS, \( \mathbf{P} = \text{diag}\{p\} \in \mathbb{C}^{M \times M} \) denotes the passive beamforming matrix at IRS, where \( p = [e^{j\phi_1}, \ldots, e^{j\phi_M}]^T \) refers to the phase shift of the \( m \)-th element of IRS, \( m = 1, \ldots, M \). In addition, \( \mathbf{H}_{\text{UE--IRS}} \in \mathbb{C}^{M \times N_u} \) depicts the segmented channel from UE to IRS. The segmented channels can be characterized based on different modeling assumptions. The PWM and SWM are explored by addressing far-field and near-field effects, respectively [13]. Particularly, the receiver (Rx) is in the far-field of the antenna array at the transmitter (Tx) when the communication distance \( D \) is larger than the Rayleigh distance \( \frac{2S^2}{\lambda} \), where \( S \) denotes the array aperture. In this case, the wave is approximated to propagate in a plane, and PWM can be adopted. By contrast, SWM has to be considered when the communication distance is smaller than the Rayleigh distance, where Rx is located in the near-field, and the propagation travels in a sphere.

As an improvement to the PWM and SWM, we proposed the idea of HSPM in [13] in THz UM-MIMO systems, which possesses less complexity than SWM and achieves better modeling accuracy than PWM in the near-field condition. In the following, we first introduce the PWM, SWM and HSPM for the cascaded channel \( \mathbf{H}^{\text{cas}} \). To facilitate the description, during the introduction of different channel models, we use Tx to represent IRS in \( \mathbf{H}_{\text{IRS--BS}} \) and UE in \( \mathbf{H}_{\text{UE--IRS}} \), and use Rx to denote BS in \( \mathbf{H}_{\text{IRS--BS}} \) and IRS in \( \mathbf{H}_{\text{UE--IRS}} \), respectively. Moreover, we consider that Tx is composed of \( N_t \) elements and \( K_t \) subarrays, while Rx employs \( N_r \) antennas and \( K_r \) subarrays, respectively.

1) **PWM**: The PWM suitable for far-field propagation region can be denoted as [18]

\[
\mathbf{H}_P = \sum_{p=1}^{N_p} a_p \mathbf{a}_p \mathbf{a}_p^H, \tag{3}
\]

where \( a_p \) represents the complex gain of the \( p \)-th propagation path, \( p = 1, \ldots, N_p \), with \( N_p \) denoting the total number of paths. The array steering vectors at Rx and Tx are denoted as \( \mathbf{a}_{\text{Rx}} = \mathbf{a}_{N_t} (\psi_{\text{px}}, \psi_{\text{pz}}) \in \mathbb{C}^{N_t} \) and \( \mathbf{a}_{\text{Tx}} = \mathbf{a}_{N_t} (\psi_{\text{tx}}, \psi_{\text{tz}}) \in \mathbb{C}^{N_t} \), respectively. Without loss of generality, by considering an \( N \) element planar-shaped array on the x-z plane with physical angle pair \((\theta, \phi)\), the array steering vector \( \mathbf{a}_{N}\left(\psi_x, \psi_z\right) \in \mathbb{C}^{N} \) can be expressed as

\[
\mathbf{a}_{N}(\psi_x, \psi_z) = \left[1, e^{j\frac{2\pi}{\lambda} \psi_x}, \ldots, e^{j\frac{2\pi}{\lambda} \psi_N}\right]^T, \tag{4}
\]

where \( \psi_n = \frac{d_{nx}}{\lambda} \psi_x + \frac{d_{nz}}{\lambda} \psi_z \). The virtual angles, \( d_{nx} \) and \( d_{nz} \) stand for the distances between the \( n \)-th antenna to the first antenna on x- and z-axis, respectively.

2) **SWM**: The SWM is universally applicable to different communication distances, which individually calculates channel responses of all antenna pairs between Tx and Rx to obtain...
the ground-truth channel. Due to high complexity, SWM is usually deployed in the near-field region, where the PPM becomes inaccurate. By denoting the communication distance of the $p^{th}$ propagation path from the $n^t_{th}$ transmitted antenna as $D_{p}^{n_{th}}$, the channel response of each antenna pair can be depicted as [18]

$$ H_S(n_x, n_z) = \sum_{p=1}^{N_p} |\alpha_{p,n_{th}}|^2 e^{-j2\pi D_{p}^{n_{th}}}, \tag{5} $$

where $H_S \in \mathbb{C}^{N_x \times N_z}$ denotes the SWM channel matrix, $\alpha_{p,n_{th}}$ represents the complex path gain.

3) HSPM: The HSPM accounts for PPM within one subarray and PPM among subarrays, which can be denoted as (6), shown at the bottom of the page [13], where $D_{p}^{k_{th}}$ stands for the distance between the $k_{th}$ transmitted and $k_{th}$ received subarray. The array steering vectors of the $p^{th}$ path for the corresponding subarray pairs are denoted as $\mathbf{a}_{tp} = \mathbf{a}_{N_{ar}}(\psi_{tp}, \psi_{tp}),$ and $\mathbf{a}_{tp}^{k_{th}} = \mathbf{a}_{N_{at}}(\psi_{tk_{th}}, \psi_{tk_{th}}),$ respectively, which have similar forms as (4). The virtual angles $\psi_{tp}$ and $\psi_{tp}^{k_{th}}$ are denoted as $\psi_{tp} = \sin(\theta_{tp}^{k_{th}} \cos \phi_{tp}^{k_{th}}),$ $\psi_{tp}^{k_{th}} = \sin(\theta_{tp} \cos \phi_{tp})$ and $\theta_{tp}$ and $\phi_{tp}$ stand for the azimuth and elevation angle pairs at Rx and Tx, respectively. Moreover, $N_{ar}$ and $N_{at}$ depict the number of antennas on the subarrays at Rx and Tx, respectively. We point out that the PPM and SWM are two special cases of HSPM when $K_{t} = K_{r} = 1$ and $K_{t} = N_{t},$ $K_{r} = N_{r}.$ Moreover, by dividing virtual subarrays, HSPM can be directly deployed in other array structures, such as uniform-linear-array and uniform-circular-array.

C. HSPM for THz Integrated UM-MIMO and IRS Systems

In the considered WSSM structures, traditional PPM becomes inaccurate, while SWM suffers from high modeling complexity [13]. By contrast, HSPM achieves higher accuracy than the PPM and lower complexity than the SWM, which is attractive to be adopted to model each segment of the cascaded channel in (2). By replacing the segmented channels $H_{BS}^{n_{th}}$ and $H_{UE-IRS}$ of $H_{BS}^{cas}$ in (2) by the expression in (6), HSPM for the cascaded channel $H_{cas}$ is represented as:

$$ H_{HSPM}^{cas} = \frac{N^U_{P}}{ \prod_{p_{u_i}}(\alpha_{p_{u_i}})} \sum_{p_{u_i}=1}^{N^U_{P}} \left\{ \sum_{k_{th}=1}^{K_{th}} G_{p_{u_i},k_{th}} E_{k_{th},1} \ldots \sum_{k_{th}=1}^{K_{th}} G_{p_{u_i},k_{th}} E_{k_{th},k_{th}} \right\} \left\{ \sum_{k_{th}=1}^{K_{th}} G_{p_{u_i},k_{th}} E_{k_{th},1} \ldots \sum_{k_{th}=1}^{K_{th}} G_{p_{u_i},k_{th}} E_{k_{th},k_{th}} \right\} \left\{ \sum_{k_{th}=1}^{K_{th}} G_{p_{u_i},k_{th}} E_{k_{th},1} \ldots \sum_{k_{th}=1}^{K_{th}} G_{p_{u_i},k_{th}} E_{k_{th},k_{th}} \right\}, \tag{7} $$

where $\alpha_{p_{u_i}}$ denotes the path gain for the $p_{u_i}$th path of $H_{IRS-BS},$ $p_{u_i} = 1, \ldots, N^U_{P},$ $N^U_{P}$ refers to the number of propagation paths in $H_{IRS-BS}$ path. The matrix $G_{p_{u_i},k_{th}} \in \mathbb{C}^{N_{am}} \times N_{am}$ is represented as

$$ G_{p_{u_i},k_{th}} = e^{-j2\pi \psi_{p_{u_i},k_{th}}(a_{tp}, a_{tp}) H_{k_{th}}(\alpha_{p_{u_i}})}, \tag{8} $$

where $D_{k_{th}}(\alpha_{p_{u_i}})$ stands for the communication distance between the $k_{th}$ subarray at BS and $k_{th}$ subarray at the IRS for the $p_{u_i}$th path. Moreover, the received and transmitted array steering vectors are denoted as $\mathbf{a}_{k_{th}} = \mathbf{a}_{N_{am}}(\psi_{k_{th}}, \phi_{k_{th}})$ and $\mathbf{a}_{p_{u_i}} = \mathbf{a}_{N_{am}}(\psi_{p_{u_i},k_{th}}, \phi_{p_{u_i},k_{th}})$ as (4). The virtual angles $\psi_{p_{u_i},k_{th}}$ and $\phi_{p_{u_i},k_{th}}$ are denoted as $\psi_{p_{u_i},k_{th}} = \sin(\theta_{p_{u_i},k_{th}}^{k_{th}} \cos \phi_{p_{u_i},k_{th}}^{k_{th}}),$ $\phi_{p_{u_i},k_{th}} = \sin(\theta_{p_{u_i}} \cos \phi_{p_{u_i}})$ and $\theta_{p_{u_i}}$ and $\phi_{p_{u_i}}$ stand for the azimuth and elevation angle pairs at Rx and Tx, respectively. Moreover, $N_{am}$ and $N_{am}$ depict the number of antennas on the subarrays at Rx and Tx, respectively. Based on (8) and (9), the $(n_{ab}, n_{au})$th element for the production of $G_{p_{u_i},k_{th}} E_{k_{th},k_{th}}(\alpha_{p_{u_i}}) \in \mathbb{C}^{N_{am} \times N_{am}}$ in (7) can be represented as

$$ (G_{p_{u_i},k_{th}} E_{k_{th},k_{th}}(\alpha_{p_{u_i}}))(n_{ab}, n_{au}) = \frac{N^U_{P}}{ \prod_{p_{u_i}}(\alpha_{p_{u_i}})} \sum_{p_{u_i}=1}^{N^U_{P}} \left\{ \sum_{k_{th}=1}^{K_{th}} G_{p_{u_i},k_{th}} E_{k_{th},1} \ldots \sum_{k_{th}=1}^{K_{th}} G_{p_{u_i},k_{th}} E_{k_{th},k_{th}} \right\} \left\{ \sum_{k_{th}=1}^{K_{th}} G_{p_{u_i},k_{th}} E_{k_{th},1} \ldots \sum_{k_{th}=1}^{K_{th}} G_{p_{u_i},k_{th}} E_{k_{th},k_{th}} \right\} \left\{ \sum_{k_{th}=1}^{K_{th}} G_{p_{u_i},k_{th}} E_{k_{th},1} \ldots \sum_{k_{th}=1}^{K_{th}} G_{p_{u_i},k_{th}} E_{k_{th},k_{th}} \right\} \left\{ \sum_{k_{th}=1}^{K_{th}} G_{p_{u_i},k_{th}} E_{k_{th},1} \ldots \sum_{k_{th}=1}^{K_{th}} G_{p_{u_i},k_{th}} E_{k_{th},k_{th}} \right\}, \tag{10} $$

where the aggregated phase $\psi_{p_{u_i},k_{th}}$ can be denoted as

$$ \psi_{p_{u_i},k_{th}} = (n_{ab} - 1)\psi_{p_{u_i},k_{th}} + (n_{au} - 1)\psi_{p_{u_i},k_{th}}, \tag{11} $$

$$ H_{HSPM} = \sum_{p=1}^{N_p} |\alpha_p|^2 \left\{ e^{-j2\pi D_p^{11}(a_{tp}^1, a_{tp}^1)^H} \ldots e^{-j2\pi D_p^{1K_{1}}(a_{tp}^1, a_{tp}^1)^H} \right\} \left\{ e^{-j2\pi D_p^{K_{1}1}(a_{tp}^1, a_{tp}^1)^H} \ldots e^{-j2\pi D_p^{K_{1}K_{1}}(a_{tp}^1, a_{tp}^1)^H} \right\} \left\{ e^{-j2\pi D_p^{K_{1}1}(a_{tp}^1, a_{tp}^1)^H} \ldots e^{-j2\pi D_p^{K_{1}K_{1}}(a_{tp}^1, a_{tp}^1)^H} \right\}, \tag{6} $$

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
where 

\[ H \] 

different channel models. In this section, we analyze the spatial 

and far-field conditions, the segmented channels can adopt 

during the process of analyzing the ranks of PWM, SWM and 

N 

k 

on the 

p 

tion path 

ζ 

HSPM channel in (6) 

distinguishability of propagation paths, the angles of different paths 

ζ 

axis, respectively. Similarly, the aggregated phases 

k 

m 

pitr,nam, 

ζ 

HSPM channel in (6) satisfies the far-field condition 

ζ 

u,i 

K 

m 

i,b 

k 

tp 

r 

n 

am 

n 

amz 

0 

1 

\[ \begin{align*} 
H_{HSPM}(n_{r},:) &= \sum_{p=1}^{N_p} \beta_{tp}^{\lambda_{r}} \begin{bmatrix} a_{tp}^{k_{r}}(n_{ar})(a_{tp}^{k_{r}})^{\dagger} & \ldots & 0 \end{bmatrix}, \\
& \ldots \\
& \sum_{p=1}^{N_p} \beta_{tp}^{K_{r}k_{r}} \begin{bmatrix} a_{tp}^{k_{r}}(n_{ar})(a_{tp}^{K_{r}k_{r}})^{\dagger} \end{bmatrix}. 
\end{align*} \] 

(13) 

Each row of \( H_{HSPM} \) is a linear combination of \( K_{r}K_{r}N_{p} \) 

linearly independent vectors as 

\[ \begin{bmatrix} (a_{tp}^{k_{r}})^{\dagger} \ldots 0 \end{bmatrix}, \] 

(14a) 

\[ \begin{bmatrix} 0 \ldots (a_{tp}^{k_{r}})^{\dagger} \end{bmatrix}, \] 

(14b) 

\[ \ldots \] 

\[ \begin{bmatrix} 0 \ldots 0 \end{bmatrix}, \] 

(14c) 

where \( k_{r} = 1, \ldots, K_{r} \) and 0 is an all-zero vector of dimension 

1 \times N_{at}. 

However, the angles of different paths to different received 

subarrays might be the same. In this case, vectors in (14) 

can be linearly dependent, which reduces the rank of the 

HSPM channel. Thus, there is 

\[ \text{Rank}(H_{HSPM}) \leq \min\{K_{r}N_{p}, K_{r}N_{r}, N_{r}\}. \] 

To prove the left-hand side inequality, we consider an extreme case. For a fixed propagation 

path, the angles among different subarray pairs between Tx and Rx are the same. In this case, the HSPM equals to the 

channel model in [26], whose rank has been proved to be 

equal to \( \min\{K_{r}N_{p}, K_{r}N_{r}, N_{r}\} \), which lower bounds the 

rank of the HSPM. Till here, we have completed the proof for 

Lemma 1.

Similar conclusions can be drawn from the above HSPM channel rank analysis based on linearly independent vectors 

and that in [25]. That is, the spatial multiplexing capability 

can be improved from 1 to \( \min\{K_{r}N_{p}, K_{r}N_{r}, N_{r}\} \) by 

adopting the WSMS in the THz band. Moreover, our analysis 
based on WSMS and HSPM is more practical and effective 

for the THz UM-MIMO systems [13], [26]. 

B. Cascaded Channel Rank Analysis 

To analyze the rank of the cascaded channel, we first introduce the following lemma. 

Lemma 2: For matrices \( A \in \mathbb{C}^{M \times N} \), \( B \in \mathbb{C}^{N \times N} \) and \( C \in \mathbb{C}^{N \times K} \), where \( B \) is a diagonal matrix, and \( rank(A) = R_{a} \), 

\[ \text{rank}(ABC) \leq \min\{R_{a}, R_{c}\}, \] 

(15) 

where the equality holds when \( A \) and \( C \) are full-rank matrices. 

Lemma 2 can be easily deduced from matrix theory. Therefore, 

we omit the proof here. Next, we analyze the rank of the 
cascaded channel. We adopt the PWM in the far-field region, 

while the SWM and HSPM are deployed for the near-field 

region, respectively.

1) Both Segmented Channels Satisfy the Far-Field Condition: In this case, both \( H_{IRS-BS} \) and \( H_{UE-IRS} \) in (2) adopt 

PWM, whose ranks equal to \( N_{p}^{\text{PB}} \) and \( N_{p}^{\text{UI}} \), respectively. 

By denoting \( H_{cas}^{\text{PWM}} \) in (2) as \( H_{cas}^{\text{PWM}} \), from Lemma 2, we can state that 

\[ \text{rank}(H_{cas}^{\text{PWM}}) \leq \min\{N_{p}^{\text{PB}}, N_{p}^{\text{UI}}\}. \] 

Moreover, when 

\[ N_{p}^{\text{UI}} = N_{p}^{\text{PB}} = N_{p}, \text{rank}(H_{cas}^{\text{PWM}}) \leq N_{p}. \] 

2) One of the Segmented Channels Satisfies the Far-Field Condition, While the Other Satisfies the Near-Field Condition: We first consider \( H_{IRS-BS} \) satisfies the far-field condition and adopts the PWM, while \( H_{UE-IRS} \) meets the near-field condition, which deploys SWM or HSPM. Since the number of 

propagation paths in the THz channel is much smaller.
than the number of elements in the UM-MIMO and IRS, 
\( \text{rank}(H_{\text{IRS-BS}}) = N^\text{IB} \prec \text{rank}(H_{\text{UE-IRS}}) \). From Lemma 2, we can obtain that \( \text{rank}(H^{\text{cas}}) \leq N^\text{IB} \). A similar deduction can be drawn when \( H_{\text{IRS-BS}} \) meets the near-field condition while \( H_{\text{UE-IRS}} \) satisfies the far-field condition. Thus, when \( N^\text{IB}_{\text{p}} = N^\text{IB}_{\text{p}} = N_{\text{p}} \), there is \( \text{rank}(H^{\text{cas}}) \leq N_{\text{p}} \).

3) Both Segmented Channels Satisfy the Near-Field Condition: We denote \( H^{\text{cas}} \) as \( H_{\text{SMW}}^{\text{cas}} \) when both \( H_{\text{IRS-BS}} \) and \( H_{\text{UE-IRS}} \) meet near-field condition and adopt SWM. In this case, \( H_{\text{IRS-BS}}, \overline{P} \) and \( H_{\text{UE-IRS}} \) are full-rank matrices. Based on Lemma 2, we know that \( \text{rank}(H_{\text{SMW}}^{\text{cas}}) = \min\{M, N_{\text{u}}, N_{\text{b}}\} \). When \( N_{\text{u}} = N_{\text{b}} = M = N \), \( \text{rank}(H_{\text{SMW}}^{\text{cas}}) = N \). Similarly, we denote \( H^{\text{cas}} \) as \( H_{\text{HSPM}}^{\text{cas}} \) when both \( H_{\text{IRS-BS}} \) and \( H_{\text{UE-IRS}} \) adopt HSPM. By considering \( K_{\text{m}} = K_{\text{u}} = K_{\text{b}} = K \) and \( N^\text{IB}_{\text{p}} = N^\text{IB}_{\text{p}} = N_{\text{p}} \), we have \( \text{rank}(H_{\text{HSPM}}^{\text{cas}}) \leq K^2 N_{\text{p}} \).

From the above analysis, we can state that in THz integrated UM-MIMO and IRS systems, the total rank of the cascaded channel is limited by the segmented channel with a smaller rank. This suggests that the rank of the cascaded channel is increased only when both segmented channels meet the near-field condition, which inspires us to enlarge the array size and obtain a larger near-field region. It is worth noting that the above discussions are not dependent on the IRS beamingforming matrix \( \overline{P} \). Therefore, we claim that given fixed segmented channels, the channel rank can not be improved by the IRS.

We will show in Sec. V that the capacity of the THz integrated UM-MIMO and IRS system based on HSPM is close to that based on the ground truth SWM, which reveals the accuracy of the HSPM. In addition, the HSPM possesses lower complexity compared to the SWM [13]. Therefore, we adopt the HSPM for both segmented channels for CE.

### IV. Channel Estimation

In this section, we present the CS-based CE framework for THz integrated UM-MIMO and IRS communication systems, which is composed of three steps, namely, on-grid sparse channel representation, signal observation and sparse recovery algorithm. Specifically, the sparse channel representation is based on an on-grid codebook, by which the channel matrix is expressed as the production of the codebook and a sparse matrix. We first introduce the traditional DFT codebook, which is shown to be ineffective in the considered integrated systems. Inspired by this, we propose a subarray-based codebook by considering the characteristic of the HSPM channel, which possesses higher sparsity and accuracy than the DFT codebook. Second, we introduce the training procedure to obtain the channel observation and formulate the CE problem as a sparse recovery problem. Third, to obtain the CE result, we develop the low-complexity SSE algorithm with high accuracy. The spatial correlation inspired DSE algorithm is further developed, which possesses lower complexity compared to the SSE at the cost of slightly degraded accuracy.

**A. On-Grid Sparse Channel Representation**

1) Traditional DFT-Based Sparse Channel Representation: In the literature, the spatial DFT-based on-grid codebook is widely deployed [33], [34], [35], [36], [37], [38]. This codebook treats the entire antenna array as a unit, and considers that the virtual spatial angles \( \psi_x = \sin \theta \cos \phi \) and \( \psi_z = \sin \phi \) are taken form a uniform grid composed of \( N_x \) and \( N_z \) points, respectively. \( \theta \) and \( \phi \) denote the azimuth and elevation angles, while \( N_x \) and \( N_z \) refer to the number of antennas on the \( x \)- and \( z \)-axis, respectively. In this way, the channel matrix is sparsely represented as

\[
H = A_{\text{Dr}} A_{\text{Dz}} A_{\text{Dt}}^H
\]

where \( A_{\text{Dr}} \in \mathbb{C}^{N_r \times N_r} \) and \( A_{\text{Dz}} \in \mathbb{C}^{N_z \times N_z} \) refer to the two-dimensional DFT on-grid codebooks at Rx and Tx, respectively, which hold a similar form, and can be represented as

\[
A_{\text{D}} = [a_N(-1, -1) \ldots a_N \left( 2(n_x-1) - 1 \right) - 1, 2(n_x-1) - 1, \ldots a_N \left( 2(n_x-1) - 1 \right) - 1].
\]

The sparse on-grid channel with complex gains on the quantized spatial angles is depicted by \( A_{\text{D}} \in \mathbb{C}^{N_x \times N_z} \).

To assess the performance of the DFT codebook, we first evaluate the sparsity of the on-grid channel \( A_{\text{D}} \) in (16) in different systems. Moreover, since in practice, there does not exist a grid whose amplitude is strictly equal to 0, we consider that the sparsity of the on-grid channel equals the number of grids whose amplitude is greater than a small value, e.g., 0.01. First, as illustrated in Fig. 2(a), the amplitude of \( A_{\text{D}} \) is shown by deploying a compact array without enlarging the subarray spacing. The on-grid channel is sparse, the number
of grids with an amplitude larger than 0.01 is only 397, which is much smaller than the preset total number of grids, i.e., 262144. By contrast, in Fig. 2(b), the amplitude of the on-grid channel in the WSMS is plotted. The on-grid channel contains 2755 grids with amplitude larger than 0.01. Therefore, the DFT codebook lacks sparsity in representing the HSPM.

2) Proposed Subarray-Based Sparse Channel Representation: We observe that the HSPM in (6) views each subarray as a unit, each block of which is the product of the array steering vectors for the subarrays at Rx and Tx, respectively. Inspired by this, we consider a subarray-based on-grid codebook. At Rx, the virtual spatial angles for each subarray are considered to be taken from fixed $N_{arx} = N_{array}N_{array}$ grids, where $N_{arx}$ and $N_{array}$ refer to the number of elements on $x$- and $z$-axis of the subarray at Rx, respectively. The corresponding DFT codebook is expressed as $U_{Dr} = [a_{N_{arx}}(-1,-1), \ldots , a_{N_{arx}}(2(N_{arx} - 1)N_{array} - 1, 2(N_{arx} - 1) - 1), \ldots , a_{N_{arx}}(2N_{array} - 1, 2N_{array} - 1)], n_{arx} = 1, \ldots , N_{arx}, n_{array} = 1, \ldots , N_{array}$. We define $\bar{\mathbf{A}} \in \mathbb{C}^{N_r \times N_t}$ as the subarray-based codebook at Rx, which deploys $K_r$ U-rays on its diagonal as

$$\bar{\mathbf{A}}_r = \text{blkdiag} [U_{Dr1}, \ldots , U_{DrT}] .$$

(17)

The on-grid codebook matrix at Tx $\bar{\mathbf{A}}_t \in \mathbb{C}^{N_t \times N_t}$ is constructed similarly. It is observed that the columns in $\bar{\mathbf{A}}_r$ are orthogonal with each other. Therefore, the on-grid representation of the HSPM in (6) based on the subarray-based codebook can be denoted as

$$H_{HSPM} \approx \bar{\mathbf{A}} \bar{\mathbf{A}}^H$$

(18)

where $\bar{\mathbf{A}} \in \mathbb{C}^{N_r \times N_t}$ is a sparse matrix. If all spatial angles were taken from the grids and not equal to each other, $\bar{\mathbf{A}}$ would contain $K_rK_tN_p$ non-zero elements.

The amplitude of the on-grid channel $\bar{\mathbf{A}}$ in (18) using the proposed codebook is plotted in Fig. 2(e), by considering the same channel as in Fig. 2(b). The number of grids with an amplitude larger than 0.01 is 1609, which is 1164 smaller than that by using the DFT codebook in Fig. 2(b). In addition, to reveal the accuracy of the on-grid channel, we calculate the difference between the practical channel $H_{HSPM}$ and the reconstructed channels approximated by the on-grid channel and the codebooks in (16) and (18) as $||\mathbf{A}_{Dr} \mathbf{A}_{Dr}^H - H_{HSPM}||^2_2$ and $||\bar{\mathbf{A}} \bar{\mathbf{A}}^H - H_{HSPM}||^2_2$, respectively. The approximation error based on the proposed codebook is around 4 dB lower than that based on the DFT codebook. To this end, we state that by possessing higher sparsity and lower approximation error, the proposed codebook is more efficient than the traditional DFT codebook.

B. Training Process and Problem Formulation

1) Training Process for Channel Observation: In this work, we are focusing on estimating the multiplied channel matrix defined as $H_{\text{mul}} = H_{\text{UE}}^T \otimes H_{\text{IRS-BS}} \in \mathbb{C}^{N_t, N_t \times M}$, since obtaining which is enough for conducting precodings [31], [32], [33], [34], [35], [36], [37], [38]. To effectively estimate $H_{\text{mul}}$, observations at BS, UE and IRS are needed. By contrast, if only two-level observation is obtained, e.g., at BS and UE levels, the information at the IRS is not sufficient to support the recovery of the IRS component in $H_{\text{mul}}$ [31], [34]. However, the hybrid UM-MIMO structures are commonly deployed in the THz band, which exploits a much smaller number of RF-chains than the number of antennas [16]. Therefore, BS can only obtain an RF-dimensional observation at each uplink training slot that transmits one signal vector. It is hard to recover the high-dimensional channel matrix from this severely compressed low-dimensional signal [40]. To reconstruct a high-dimensional channel observation on BS, UE and IRS, we consider an uplink pilot training procedure. As illustrated in Fig. 3, the training is conducted on three levels, namely, BS training, UE training, and IRS training levels, respectively. During the training process, BS transmits known pilot signals to BS via IRS in $T = T_1T_2T_3$ training slots for uplink CE, where $T_i$ denote the number of training slots for BS, IRS and UE, respectively. At the $(b, u, i)^{th}$ slot, BS deploys the training beamformer $\mathbf{F}_u \in \mathbb{C}^{N_x \times N_u}$ and transmits pilot signal $s_{b,u,i} \in \mathbb{C}^{N_u}$, $b = 1, \ldots , T_b$, $u = 1, \ldots , T_u$, $i = 1, \ldots , T_i$. In the meantime, the training phase shift vector $\mathbf{p}_i \in \mathbb{C}^{M}$ and combiner $\mathbf{W}_b \in \mathbb{C}^{N_t \times N_u}$ are deployed at IRS and BS, respectively, to obtain the received signal $y_{b,u,i} \in \mathbb{C}^{N_h}$ at BS as

$$y_{b,u,i} = \mathbf{W}_b^H H_{\text{IRS-BS}} \mathbf{diag}\{\mathbf{p}_i\} H_{\text{UE-IRS}} \mathbf{F}_u s_{b,u,i} + n_{b,u,i} ,$$

(19)

where $n_{b,u,i} = \mathbf{W}_b^H n_{b,u,i} \in \mathbb{C}^{N_h}$, and $n_{b,u,i} \in \mathbb{C}^{N_h}$ refers to the received AWGN.

The BS training is first conducted, in which totally $T_b$ different training combiners are used to obtain the received signal as (19). By collecting $y_{b,u,i}, b = 1, \ldots , T_b$ as $y_{u,i} =$
The received signal after BS training can be expressed as
\[
y_{u,i} = (f_{i}^T \otimes W^H) H^{\text{mul}} p_i + n_{u,i},
\]
where \(f_i = F_a s_{a,i} \in \mathbb{C}^{N_a} \) stands for the equivalent training beamformer, \(W = [W_1, \ldots, W_{T_{a}}] \in \mathbb{C}^{N_a \times N_{T_{a}}} \) denotes the training combiner, and \(n_{u,i} = [n_{1,1,i}, \ldots, n_{T_{a},1,i}]^T \in \mathbb{C}^{N_a T_{a}} \) represents the collected noise.

After one round of BS training, UE changes its beamformer \(\tilde{F}_u\) to complete the UE training. Particularly, \(T_{a}\) beamformers are used to obtain the received signal as \(20\). By collecting \(y_{u,i}\) for \(u = 1, \ldots, T_{u}\) as \(y_i = [y_{T_{a},1,i}, \ldots, y_{1,1,i}]^T \in \mathbb{C}^{N_a T_{a} T_{u}}\), we can obtain
\[
y_i = (F^T \otimes W^H) H^{\text{mul}} p_i + n_i,
\]
where \(F = [f_1, \ldots, f_{T_{u}}] \in \mathbb{C}^{N_a \times T_{u}}\) denotes UE training beamforming matrix. Moreover, \(n_i = [n_{1,1,i}, \ldots, n_{T_{a},1,i}]^T \in \mathbb{C}^{N_a T_{a} T_{u}}\) represents noise. Finally, the phase shift vector at IRS \(p_i\) is changed to control the IRS training. After obtaining each \(y_i\) as \(21\), \(i = 1, \ldots, T_i\), we stack \(y_i\) as \(Y = [y_{1,1}, \ldots, y_{T_r}] \in \mathbb{C}^{N_a T_{a} T_{u} T_{i}}\), which can be represented as
\[
Y = (F^T \otimes W^H) H^{\text{mul}} P + N,
\]
where \(P = [p_1, \ldots, p_{T_r}] \in \mathbb{C}^{M \times T_{r}}\) refers to the training phase shift matrix. In addition, \(N = [n_1, \ldots, n_{T_r}] \in \mathbb{C}^{N_a T_{a} T_{u} T_{i}}\) represents the stacked noise.

In this work, CE refers to estimating the multiplied channel matrix \(H^{\text{mul}}\) in \(22\) before link establishment, which is useful in configuring the initial beamforming at BS, IRS, and UE \([35, 37]\). Considering the dynamic channel between UE and IRS, the IRS-UE channel needs to be estimated more frequently as \([29]\) after link establishment. Based on the proposed codebook in \((18)\), \(H^{\text{mul}}\) can be represented as
\[
H^{\text{mul}} \approx \left( \overline{A}^{\text{UE-IRS}} T_{a}^T \otimes \overline{A}^{\text{IRS-UE}} \right) \mathbf{P} + \mathbf{A} \Lambda A^T,
\]
where \(\overline{A}^{\text{UE-IRS}} \in \mathbb{C}^{N_a \times N_a}\) and \(\overline{A}^{\text{IRS-UE}} \in \mathbb{C}^{M \times M}\) denote the codebook matrices for the UE-IRS channel at UE and IRS, respectively, \(\overline{A}^{\text{IRS-BS}} \in \mathbb{C}^{N_s \times N_s}\) denotes the sparse on-grid channel. The codebook matrices for the IRS-BS channel at BS and IRS are denoted as \(\overline{A}^{\text{IRS-BS}} \in \mathbb{C}^{N_a \times N_a}\) and \(\overline{A}^{\text{IRS-BS}} \in \mathbb{C}^{M \times M}\), respectively. \(\overline{A}^{\text{IRS-BS}} \in \mathbb{C}^{N_s \times M}\) stands for the corresponding sparse matrix. Moreover, \(\mathbf{A} = (\overline{A}^{\text{UE-IRS}} \otimes \overline{A}^{\text{IRS-BS}}) \in \mathbb{C}^{N_s N_a \times N_s N_a}\) represents the combined codebook matrix at the left-hand side, \(\hat{\Lambda} = (\overline{A}^{\text{UE-IRS}} \otimes \overline{A}^{\text{IRS-BS}}) \in \mathbb{C}^{N_s N_a \times M^2}\) depicts the multiplied sparse matrix, \(\hat{A}_t = \overline{A}^{T_{a}} (\overline{A}^{\text{UE-IRS}} \otimes \overline{A}^{\text{IRS-BS}}) \in \mathbb{C}^{M^2 \times M}\) represents the combined transmit codebook matrix.

It is worth noticing that \(\overline{A}^{\text{UE-IRS}} = \overline{A}^{\text{IRS-BS}} = \overline{A}^{\text{IRS}}\), where \(\overline{A}^{\text{IRS}}\) denotes the on-grid codebook matrix at the IRS. Therefore, the multiplied channel matrix is transformed as
\[
H^{\text{mul}} \approx \mathbf{A} \Lambda A^T,
\]
where \(\Lambda \in \mathbb{C}^{N_s N_s \times M}\) denotes the sparse on-grid channel matrix, which is a function of \(A_s\), \(A_t\) and \(\hat{\Lambda} A_s = \Lambda_{\text{IRS}} \in \mathbb{C}^{M \times M}\) denotes the codebook matrix on the right-hand side. In addition, we point out that the rows of the non-zero elements in \(\Lambda\) correspond to the grid points in \(A_s\), while the columns of non-zero elements in \(\Lambda\) indicate the grid points in \(A_t\).

2) Problem Formulation: By combining the on-grid channel representation in \(24\) with the channel observation in \(22\), we can obtain
\[
Y = (F^T \otimes W^H) A_s \Lambda A_t P + N.
\]
The CE problem can be formulated as a sparse signal recovery problem as
\[
\begin{align}
\min & \quad \| \Lambda \|_0, \\
\text{s.t.} & \quad \| Y - (F^T \otimes W^H) A_s \Lambda A_t P \|_0 \leq \epsilon,
\end{align}
\]
where \(\epsilon\) is a constant to measure the estimation error. In addition, the \(l_0\) norm in problem \(26\) is usually transformed into the \(l_1\) norm due to its non-convexity \([40]\).

To solve the problem in \(26\), the received signal \(Y\) can be vectorized as \(y_{\text{vec}} = \text{vec}(Y) \in \mathbb{C}^{N_a T_{a} T_{u} T_{i}}\) to obtain \(y_{\text{vec}} = \Phi \Psi \mathbf{h} + n_{\text{vec}}\), where \(\Phi = (F^T \otimes F^T) \in \mathbb{C}^{N_a T_{a} T_{u} T_{i} \times N_a T_{a} T_{u} T_{i}}\) defines the measurement matrix, the overall codebook matrix is \(\Psi = (A_s^T \otimes A_t) \in \mathbb{C}^{N_a T_{a} T_{u} T_{i} \times N_s N_a M}\) and \(n_{\text{vec}} = \text{vec}(N) \in \mathbb{C}^{N_s T_{a} T_{u} T_{i}}\) represents the vectorized noise. Various of greedy algorithms such as orthogonal matching pursuit (OMP) \([35]\) and compressive sampling matching pursuit (CoSaMP) \([41]\) can be used to recover \(h\) from \(y_{\text{vec}}\).

However, the dimension of \(\Psi\) is proportional to the number of antennas at BS \(N_a\), UE \(N_u\) and the number of passive reflecting elements at IRS \(M\). In our considered UM-MIMO and IRS systems, the dimension becomes unacceptably large, the computational complexity of existing algorithms surges. Inspired by this, we propose low-complexity SSE and DSE CE algorithms.

### C. Sparse Recovery Algorithms

1) SSE Algorithm: The SSE algorithm separately estimates positions of non-zero grids on each side of the multiplied channel \(H^{\text{mul}}\) in \(24\). It is necessary to estimate multiple non-zero grids. Despite the LoS dominant property \([42]\), multi-paths still exist according to the channel measurement result in the THz band \([43]\). Moreover, with the massive number of antennas in the UM-MIMO, it can be concluded from \textit{Lemma 1} that the channel rank is around \(K_r K_t N_p\). In addition, as the pre-fixed spatial grids are discrete while the actual channel grids are continuously distributed, the required number of grids to accurately estimate the channel is usually larger than the channel rank, i.e., \(K_r K_t N_p\). In the SSE algorithm, since the non-zero grids on left- and right-hand side codebook matrices \(A_s\) and \(A_t\) relate to the non-zero rows and columns of \(\Lambda\), respectively, we consider to estimate them separately. The procedures of the SSE algorithm are summarized in \textbf{Algorithm 1} and explained as follows.
Algorithm 1 SSE Algorithm

Input: Received signal $Y$ in (25), combined training matrices at UE, IRS and BS, $F$, $P$ and BS $W$, the codebook matrices $A_r$ and $A_t$

Initialization: $\Pi_r = \emptyset$, $\Pi_t = \emptyset$, $B_r = (F^T \otimes W^H)A_r$, $B_t = A_t P$

Stage 1: Estimate non-zero grid points in $A_r$

1. $y_{sumr} = \sum_{i=1}^{M} (YB_r^H)(;i)$
2. $y = y_{sumr}$, $B = B_r$, $f \propto K_r N_p U_p^H N_{PB}$

Use Algorithm 2 to obtain the estimated grid point $\Pi_r$

Stage 2: Estimate non-zero grid points in $A_t$

3. $y_{sumt} = \left( \sum_{i=1}^{N_p} (B^H Y) (\Pi_r(i,:)) \right)^T$
4. $y = y_{sumt}$, $B = B_t$, $f \propto K_t N_p U_p^H N_{PB}$

Use Algorithm 2 to obtain the estimated grid point $\Pi_t$

Stage 3: Recover the channel matrix

5. $A_r = B_r(;,\Pi_r)$, $A_t = B_t(;,\Pi_t)$

6. $\hat{A}(\Pi_r, \Pi_t) = \hat{A}(Y; A_t)^H$

Output: Estimated channel $H = A_r \hat{A}(A_t)^H$

At Stage 1, the non-zero grid points $\Pi_r$ in $A_r$ is estimated. Specifically, by adding the columns of $Y$ in Step 2, $y_{sumr} \in C^{N_r T_r T_t}$ can be expressed as $y_{sumr} = \left( (F^T \otimes W^H) A_r s_{sumr} + n_{sumr} \right) \in C^{N_r T_r T_t}$, where $s_{sumr} = \left( \sum_{i=1}^{T_r} (A_r P P^H A_r^H)(;i) \right) \in C^{N_r T_r}$ denotes the equivalent transmit signal, and $n_{sumr} = \left( \sum_{i=1}^{T_r} (N P^H A_r^H)(;i) \right) \in C^{N_r T_r T_t}$ refers to the equivalent noise. Due to the sparsity of $A_r$, $s_{sumr}$ is a sparse vector; the non-zero positions in $s_{sumr}$ relate to the non-zero rows of $A_r$. Therefore, the positions of non-zero rows of $A_r$ can be determined by estimating the non-zero positions of $s_{sumr}$, which is completed in Step 4.

Similarly, at Stage 2, the non-zero grid points $\Pi_r$ in $A_t$ is estimated. Since the positions of the non-zero rows of $A_t$ have been determined in the previous stage, using these rows of $Y$ to compose $y_{sumt}$ is enough to determine the non-zero columns of $A_t$, which is shown in Step 6. Moreover, $\Pi_t$ is also determined by Algorithm 2 in Step 8. Followed by that, at Stage 3 of Algorithm 1, the estimated $A_r$ and $A_t$ is first obtained in Step 10. The sparse on-grid channel matrix is then estimated in Step 11. Based on these estimated matrices, the channel matrix is finally recovered as illustrated in Step 11, which completes Algorithm 1.

To estimate the positions of non-zero grids with received signal $y$ and measurement matrix $B$, Algorithm 2 first calculates the correlation between $B$ and the residual vector $r$ in Step 2. The most correlative column index is expressed as $n$, which is regarded as the newly founded grid index and added to the grid set $\Pi$. The estimated signal $s$ on the grids specified by $\Pi$ is calculated in Step 4. Then, the residual vector is updated in Step 5 by removing the effect of the non-zero grid point that has been estimated in the previous step. By repeating these procedures, $\Pi$ indexes are selected as the estimated non-zero grid points.

2) DSE Algorithm: The computational complexity of the SSE algorithm majorly comes from the production in Step 2 of Algorithm 2 in Step 4 and Step 8 of Algorithm 1, which are around $O(N_{\text{sub}} T_u T_b N_u N_b)$ and $O(T MH)$ in each iteration, respectively. These values become large with the increased number of antennas in UM-MIMO and elements in IRS. The DSE algorithm addresses this problem by exploiting the spatial correlation among subarrays. Specifically, in the HSPM channel (6), for the entire array on the left-hand side, the spatial angles from subarrays on the right-hand side are close. Therefore, if we separately consider the codebooks between each subarray on the right-hand side and the entire array on the left-hand side, the positions of non-zero grids would be close. Inspired by this, DSE first calculates positions of non-zero grids for the codebook between the first subarray on the right-hand side and the entire array on the left-hand side, which are saved as benchmark grids. For the remaining subarrays on the right-hand side, the grid searching space is shrunk by limiting the potential grids in the neighbor of the benchmark grids for reduced complexity.

The grid shrinkage of the DSE algorithm operates at Stage 1 and Stage 2 of Algorithm 1, which are detailed in Algorithm 3. The input to DSE is the summarized channel observation $y_{sum}$, the sensing matrix $\Phi$, the codebook relating to the subarray at Tx and the entire subarray at Rx $A_{\text{sub}}$, number of iterations $I$ and number of subarrays at right-hand side $K$. At Stage 1 of Algorithm 1, these parameters are obtained as $y_{sum} = y_{sumr}$, $\Phi = F^T \otimes W^H$, $A_{\text{sub}} = U_m^T \otimes \Phi_{\text{IRS}}$, $I \propto K_r N_p U_p^H N_{PB}$ and $K = K_m$, where $U_m$ denotes the spatial DFT matrix for the subarray at UE. At Stage 2 of Algorithm 1, these parameters are calculated as $y_{sum} = y_{sumt}$, $\Phi = P^T$, $A_{\text{sub}} = U_m^T \otimes \Phi_{\text{IRS}}$, $I \propto K_t N_p U_p^H N_{PB}$ and $K = K_m$, where $U_m$ refers to the spatial DFT matrix for the subarray at the IRS.

For the $k$th subarray on the right-hand side, the DSE algorithm first obtains the sensing matrix $Q$ and measurement matrix $B$ at Step 2 and Step 3 of Algorithm 3, respectively. For the first subarray, non-zero grids relating to $A_{\text{sub}}$ are directly estimated and recorded in $\Pi_1$ as benchmark grids. The neighboring $q$ elements for each grid in $\Pi_1$ are then selected as the potential searching grids for the remaining subarrays, which are saved as $\tilde{\Pi}$, as shown in Step 7 to 10 in Algorithm 3. For the remaining subarray pairs, only the grids in $\tilde{\Pi}$ will be searched, as illustrated in Step 4 to 6. Finally, in Step 11, the determined grid positions for subarrays are transformed to positions for the entire array and saved in $\Pi$. 

Algorithm 2 Grid Position Estimation

Input: Received signal $y$, measurement matrix $B$, iterations $I$

Initialization: $\Pi = \emptyset$, $r = y$, $\tilde{s} = 0_{\text{size}(B, 2)}$

1. for $i = 1, \ldots, I$
2. $n = \text{argmax} \|B_\Pi r\|_2^2$
3. $\Pi = \Pi \cup n_r$
4. $\tilde{s}(\Pi) = B^T(:, \Pi)y$
5. $r = y - B\tilde{s}$
6. end for

Output: Estimated grid position $\Pi$
Algorithm 3 DSE Algorithm for Grid Position Estimation

| Input: | Received signal $y_{\text{sum}}$, sensing matrix $\Phi$, codebook matrix for subarray $A_{\text{sub}}$, number of iterations $T$, number of subarrays $K$ |
|--------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|        | $\Pi$ = $\emptyset$, $\Pi_{1:k} = \emptyset$, $N_a = \text{size}(A_{\text{sub}}, 1)$ |
|        | 1. for $k = 1 : K$
|        | 2. $Q = \Phi(:, (k - 1) * N_a + 1 : kN_a)$
|        | 3. $B = QA_{\text{sub}}$
|        | 4. if $k > 1$
|        | 5. $B = QA_{\text{sub}}(:, \tilde{F})$
|        | 6. end if
|        | 7. Use Algorithm 2 to obtain the estimated grid point $\hat{\Pi}_k$
|        | 8. if $k = 1$
|        | 9. Construct $\tilde{F}$ by selecting $q$ neighboring grids of points in $\Pi_{1}$
|        | 10. end if
|        | 11. Transform positions in $\Pi_{k}$ to positions in $\Pi = \Pi_{1:k}$
|        | 12. end for

| Output: | Estimated grid position $\Pi$ |

3) Computational Complexity: The total computational complexity of the SSE algorithm can be approximated as $O(\{N_{s}T_uT_\text{b}N_{u}N_{b} + T_\text{i}M\})$. The computational complexity of the DSE algorithm also mainly comes from Step 4 and Step 8 of Algorithm 1, the total computational complexity of the DSE algorithm can be approximated as $O(\{T_\text{i} \left( N_{s}T_uT_\text{b}N_{u}N_{b}N_{k} + T_\text{i}M \right)\})$.

V. PERFORMANCE EVALUATION

In this section, we first numerically assess the system capacities by deploying different channel models for the THz integrated UM-MIMO and IRS systems. Then, the performances of the proposed SSE and DSE CE algorithms are extensively evaluated.

A. Simulation Setup

The simulation parameters and important notations used in this paper are summarized in TABLE I. We employ the system in Fig. 1, where the complex gain of the THz channel is generated based on the channel model in [15]. To evaluate capacity, the IRS beamforming matrix $\hat{P}$ in (1) is randomly generated, while the phase of each element of $\hat{P}$ follows a uniform distribution over $[0, 2\pi]$. In the CE process, we adopt the HSPM channel model in (6) for both segmented channels $H_{\text{UE-IRS}}$ and $H_{\text{IRS-BS}}$. The spatial angles of azimuth and elevation in the HSPM are randomly generated, following uniform distributions over $[0, \pi]$. The training process from (19) to (22) is deployed. The phase of each element of $W_{b}, P_{i}$ and $\tilde{F}_{u}$ are randomly generated, following uniform distribution over $[0, 2\pi]$. The resulting measurement matrices $F^T \otimes W^H$ and $\hat{P}$ in (25) can be regarded as random matrices. It has been demonstrated in Sec. III in [44] that the restricted isometry property (RIP) condition can be achieved with a high probability (as derived in Lemma 5.1 of [44]) if the measurement matrix in CS is randomly distributed. This setup has also been verified to be effective in recent empirical studies [31], [45].

The training time slot for BS, UE and IRS $T_\text{b}, T_u$ and $T_\text{i}$ satisfy $T_\text{b} \leq N_\text{b}N_\text{k}$, $T_u \leq N_uN_\text{k}$ and $T_\text{i} \leq M$ during our evaluation, to guarantee reduced training overhead. The estimation accuracy is evaluated by the normalized-mean-square-error (NMSE), which is obtained as

$$\text{NMSE} = \frac{\mathbb{E}\left\{\|H - H_{\text{mul}}\|_2^2\right\}}{\mathbb{E}\left\{\|H_{\text{mul}}\|_2^2\right\}},$$

where $H$ denotes the estimated channel, and $H_{\text{mul}}$ refers to the practical channel in (20) by using the HSPM. All the results are obtained by averaging 5000 trials of Monte Carlo simulations.

B. System Capacity Based on Different Channel Models

In order to show the advantages of HSPM over SWM and PWM, we first evaluate the system capacities by using PWM, SWM, and HSPM for the segmented channels in different communication distances and subarray spacing. We consider both cascaded channels $H_{\text{UE-IRS}}$ and $H_{\text{IRS-BS}}$ have one LoS path, i.e., $N_\text{p} = 1$, which is simplified yet practical for the THz systems due to the LoS domination property [42]. In this case, the $H_{\text{PWM}}$ has rank 1 with no spatial multiplexing capability. Moreover, transmit power at the BS is fixed at 20 dBm.

The capacity results over different communication distances from BS to IRS and IRS to UE are illustrated in Fig. 4. It is observed that the capacity of $H_{\text{cas}}^{\text{HSPM}}$ is very close to $H_{\text{cas}}^{\text{SWM}}$, which is much higher than $H_{\text{cas}}^{\text{PWM}}$. In particular, as shown in Fig. 4(a), when the communication distance is 40 m, the capacity of $H_{\text{cas}}^{\text{HSPM}}$ is only $5 \times 10^{-4}$ bits/s/Hz lower than that of $H_{\text{cas}}^{\text{SWM}}$. The capacities of $H_{\text{cas}}^{\text{HSPM}}$ and $H_{\text{cas}}^{\text{SWM}}$ are 37.0 bits/s/Hz higher than the capacity of $H_{\text{cas}}^{\text{PWM}}$. This is explained that in near-field transmission, the PWM loses its effectiveness in characterizing the channel. By contrast, the HSPM possesses spatial multiplexing capability even with only one propagation path. In addition, we have shown in [13] that the HSPM deploys a small number of channel parameters compared to the SWM to achieve high accuracy. Therefore, the HSPM is effective to achieve near-optimal performance with low complexity.

By contrast, the capacities based on PWM, HSPM, and SWM converge when the communication distance far exceeds the Rayleigh distance, which equals 46.3 m. When the communication distance is over three times the Rayleigh distance, i.e., 160 m, the capacities based on different channel models finally approach to be close, where the difference between the PWM and the other channels reduces to 2.2 bits/s/Hz. Therefore, the Rayleigh distance overestimates the accuracy of PWM approximation from SWM. Equivalently, the misuse of PWM could cause severe deterioration of capacity even when the communication distance is equal to or larger than the Rayleigh distance, i.e., the so-called far-field region. As a take-away lesson from our analysis, the HSPM is effective and generally applicable when the communication distance is

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
TABLE I
SIMULATION PARAMETERS AND NOTATIONS

| Notation | Meaning | Value in simulation |
|----------|---------|--------------------|
| $f$      | Carrier frequency | 0.3 THz |
| $B$      | Bandwidth | 5 GHz |
| $\lambda$ | Carrier wavelength | 0.001 m |
| $K_b, K_m, K_u$ | Number of subarrays at the BS, IRS, and UE | Selected in 1, 4 |
| $N_{ab}, N_{am}$ | Number of antennas on a subarray at the BS and IRS | Selected in 64, 256 |
| $N_{bu}$ | Number of antennas on a subarray at the UE | Selected in 4, 16 |
| $N_b, M, N_u$ | Number of antennas at the BS, IRS and UE | |
| $N_{p}^{\text{UT}}, N_{p}^{\text{IR}}$ | Number of paths in $H_{\text{UB--IRS}}$ and $H_{\text{IRS--BS}}$ | $N_p$ |
| $q$ | | 5 |
| $T_u, T_b, T_i$ | Training time slots of UE, BS and IRS | |
| $N_{R}, N_{t}$ | Number of antennas at Rx (IRS or BS) and Tx (UE or IRS) | |
| $K_R, K_t$ | Number of subarrays at Rx (IRS or BS) and Tx (UE or IRS) | |
| $A_L, A_R$ | Codebook matrices at left and right side of $H^{\text{null}}$, respectively | |
| $F_{u}, W_{b}, P_{t}$ | Training beamforming, combining and IRS reflection matrices | |
| $F_{l}, W_{l}$ | Combined training beamforming, combining, IRS reflection matrices | |
| $H_{\text{UB--IRS}}, H_{\text{IRS--BS}}$ | UE beamforming, IRS beamforming and BS combining matrices | |
| $H_{\text{PWM}}, H_{\text{SWM}}, H_{\text{HSPM}}$ | Segmented channels form UE to IRS and IRS to BS, respectively | |
| $H^{\text{null}}_{\text{PWM}}, H^{\text{null}}_{\text{SWM}}, H^{\text{null}}_{\text{HSPM}}$ | Cascaded channels based on the PWM, SWM and HSPM channel matrices | |
| $\mathbf{Y}$ | The multiplied channel matrix to be estimated in (22) | |
| $\Lambda$ | Observation matrix used for CE after training in (22) | |
| }
we fix the number of paths as N of which deploy the traditional DFT codebook. In addition, comparison, including the OMP [35] and CoSaMP [41], both systems. Two classical CS-based algorithms are selected for cation distances and in WSMS and traditional compact array estimation methods, we compare the NMSE performance of C. Performance of SSE and DSE Channel Estimation spectral efficiency [26].

condition number, which can be optimized to achieve high usually smaller. Moreover, the subarray spacing decides the in good condition, and the preferred number of subarrays is achieved by focusing the transmit energy to the sub-channel 10 dB. By contrast, with low SNR, high spectral efficiency is multiplexing at good link conditions, e.g., SNR larger than 0 dB. This is explained that at low SNR with a strong spatial multiplexing gain [13]. By contrast, the capacity based on the PWM remains around 39.9 bits/s/Hz. In addition, as the subarray distance further increases beyond 144λ, the capacity begins fluctuating due to the variation of the eigenvalues of the channel matrix [46]. In this study, we consider the reasonable widely-spaced subarrays, e.g., the subarray spacing is smaller than 144λ = 0.144 m. Therefore, the spatial multiplexing of the THz integrated UM-MIMO and IRS systems can be improved based on the widely-spaced architecture design.

It is worth noticing that in the WSMS architecture, the number of subarrays decides the maximum spatial degree of freedom. More subarrays are preferred for better spatial multiplexing at good link conditions, e.g., SNR larger than 10 dB. By contrast, with low SNR, high spectral efficiency is achieved by focusing the transmit energy to the sub-channel in good condition, and the preferred number of subarrays is usually smaller. Moreover, the subarray spacing decides the condition number, which can be optimized to achieve high spectral efficiency [26].

**C. Performance of SSE and DSE Channel Estimation**

To demonstrate the effectiveness of the proposed channel estimation methods, we compare the NMSE performance of the proposed SSE and DSE algorithms at different communication distances and in WSMS and traditional compact array systems. Two classical CS-based algorithms are selected for comparison, including the OMP [35] and CoSaMP [41], both of which deploy the traditional DFT codebook. In addition, we fix the number of paths as Np = 2 for each channel segment as the LoS and reflected paths are usually the strongest paths in the THz channel [43]. The subarray spacing at BS is selected as 64λ. As illustrated in Fig. 6(a), the estimation NMSE against SNR with HSPM channel is evaluated. The estimation NMSE of SSE becomes lower than that of DSE as the SNR exceeds 0 dB. This gap decreases with the increment of SNR. The NMSE of SSE becomes lower than that of DSE as the SNR exceeds 0 dB. This is explained that at low SNR with a strong LoS path, the potential grid error for the LoS path can be avoided by determining potential searching grids from the first subarray, as operated in the DSE algorithm. In the high SNR region, i.e., SNR>0 dB, the best grids for the entire array are hard to be mapped to the first subarray due to the existence of multi-paths. Therefore, the performance of the DSE becomes worse than the SSE as the SNR increases. By contrast, the SSE algorithm is useful in searching all grids related to the LoS and NLoS paths, despite the higher probability of grid error in the low SNR region. To this end, we can state that the DSE algorithm is more attractive in the low SNR region, i.e., SNR<0 dB.

Furthermore, by extending the communication distance to 80 m that exceeds the Rayleigh distance, e.g., 46.3 m, the estimation NMSEs of different algorithms are evaluated in Fig. 6(b). In this case, communication is conducted in the far-field region. The result is consistent with that in Fig. 6(a), which further reinforces the effectiveness of the proposed HSPM. Specifically, the estimation accuracy of the SSE outperforms the other algorithms at higher SNR larger than 0 dB, while the DSE algorithm achieves the lowest NMSE among the evaluated algorithms when SNR<0 dB. In conclusion, the
Fig. 7. Estimation NMSE against the number of training slots. SSE and DSE algorithms are effective and generally applicable as the communication distance ranges from near- to far-field.

To study the performance of the proposed SSE and DSE algorithms even in the traditional compact array systems without enlarging the subarray spacing, we evaluate their performances in Fig. 6(c) in contrast to the OMP and CoSaMP algorithms. We observe that the estimation NMSEs of the OMP, SSE, and DSE algorithms are close. This is explained that in the traditional compact array systems, the number of subarrays at the BS, UE, and IRS is equal to 1. The subarray-based codebook degenerates into the DFT codebook, and the operations in the DSE and SSE algorithms become the same. This result further demonstrates the effectiveness of the subarray-based codebook in the WSMS systems. Instead of being restricted by the performance of the sparse recovery algorithms, the performances of the OMP and CoSaMP methods are limited by the accuracy of the DFT codebook. Moreover, the NMSE performance of the CoSaMP slightly outperforms the remaining algorithms. This is owing to the benefits of the grid selection mechanism in the CoSaMP [41].

In Fig. 7, the NMSE performances of the DSE and SSE schemes versus the length of training slots in different SNRs are analyzed. The NMSE decreases with the increased length of training slots allocated to both the IRS and BS. In particular, as illustrated in Fig. 7(a), pertaining to the SSE algorithm, when SNR is 10 dB, the NMSE drops sharply first, in which the decrement is around 3.9 dB as $T_i$ increases from 32 to 608. However, as $T_i$ further increases from 608 to 1024, the decline tends smooth, i.e., NMSE decreases by 0.18 dB. This is due to the fact that the NMSE results are determined by the dimensions of the channel and channel observation. The latter one is enlarged with the increment of the pilot length. Thus, the channel is more accurately estimated with longer training slots, especially when the dimensions of the observation and the channel are comparable. In addition, as shown in Fig. 7(a) and Fig. 7(b), to obtain a good CE performance, time slots of lengths 600 and 100 are enough for the IRS and BS training in the considered configuration, respectively, after which the NMSE degradation is very limited by adding the length of the training slots. By deploying the sparsity of the THz channels, the proposed DSE and SSE reduce the training overhead by 31.6% and 60.9% at the IRS and BS, respectively, compared to the traditional least-square (LS) and minimum-mean-square (MMSE) CE methods [34], in which 1024 and 256 slots are required for the IRS and BS training, respectively. Therefore, the proposed CE framework can estimate the channel with reduced training overhead.

VI. CONCLUSION

As a promising technology for THz communications, the integrated UM-MIMO and IRS systems can effectively solve the LoS problem in complex occlusion environments. Three challenges arise. First, the huge dimensional antenna array in UM-MIMO and IRS, in contrast with the sub-millimeter wavelength, enlarges the near-field region. Second, the spatial multiplexing and capacity are limited by the THz channel sparsity. Third, the adoption of hybrid beamforming systems in UM-MIMO results in limited RF-chains, with the lack of signal processing capability of the IRS, CE has to recover high-dimensional channel from severely compressed observations.

In this work, we have introduced the HSPM to accurately model the cascaded channel of the THz integrated UM-MIMO and IRS system. We have analyzed the spatial multiplexing under near- and far-field cases and proved that the spatial multiplexing of the cascaded channel is limited by the segmented channel with a lower rank. Additionally, we have developed a subarray-based codebook and the SSE and DSE algorithms to address the CE problem. Extensive simulations are conducted, and results demonstrate the accuracy of the HSPM channel. The capacity based on HSPM is only $5 \times 10^{-4}$ bits/s/Hz lower than that based on the ground-truth SWM. Moreover, the spatial multiplexing gain is improved based on the widely-spaced architecture design. Based on the proposed codebook, the SSE and DSE achieve better estimation accuracy than traditional algorithms. While the SSE possesses the highest accuracy at SNR over 0 dB, the DSE is more attractive at low SNR, whose estimation NMSE is 1 dB lower than the SSE when $\text{SNR} = -10 \text{ dB}$.

REFERENCES

[1] Y. Chen, R. Li, C. Han, and M. Tao, "Hybrid spherical- and planar-wave channel modeling and spatial multiplexing analysis for terahertz integrated UM-MIMO and IRS systems," in Proc. IEEE Int. Conf. Commun., May 2022, pp. 3490-3495.
[2] I. F. Akyildiz, C. Han, Z. Hu, S. Nie, and J. M. Jornet, “Terahertz band communication: An old problem revisited and research directions for the next decade,” *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 4250–4285, Jun. 2022.

[3] T. S. Rappaport et al., “Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond,” *IEEE Access*, vol. 7, pp. 78729–78757, 2019.

[4] Z. Chen et al., “Terahertz wireless communications for 2030 and beyond: A cutting-edge frontier,” *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 66–72, Nov. 2021.

[5] Z. Chen, B. Ning, C. Han, Z. Tian, and S. Li, “Intelligent reflecting surface assisted terahertz communications toward 6G,” *IEEE Wireless Commun.*, vol. 28, no. 6, pp. 110–117, Dec. 2021.

[6] I. F. Akyildiz, C. Han, and S. Nie, “Combating the distance problem in the millimeter wave and terahertz frequency bands,” *IEEE Commun. Mag.*, vol. 56, no. 6, pp. 102–108, Jun. 2018.

[7] I. F. Akyildiz and J. M. Jornet, “Realizing ultra-massive MIMO (1024–1024) communication in the (0.06–10) terahertz band,” *Nano Commun. Netw.*, vol. 8, pp. 46–54, Jun. 2016.

[8] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, “Intelligent reflecting surface-aided wireless communications: A tutorial,” *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, May 2021.

[9] M. Di Renzo et al., “Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead,” *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.

[10] J. Li, B. Guo, M. Tao, Y.-F. Liu, and W. Yu, “Joint design of hybrid beamforming and reflection coefficients in RIS-aided mmWave MIMO systems,” *IEEE Trans. Commun.*, vol. 70, no. 4, pp. 2404–2416, Apr. 2022.

[11] E. Björnson, Ö. Özdogan, and E. G. Larsson, “Reconfigurable intelligent surfaces: Three myths and two critical questions,” *IEEE Commun. Mag.*, vol. 58, no. 12, pp. 90–96, Dec. 2020.

[12] E. Björnson and L. Sanguinetti, “Rayleigh fading modeling and channel hardening for reconfigurable intelligent surfaces,” *IEEE Commun. Lett.*, vol. 10, no. 4, pp. 830–834, Apr. 2021.

[13] Y. Chen, L. Yan, and C. Han, “Hybrid spherical- and planar-wave modeling and DCNN-powered estimation of terahertz ultra-massive MIMO channels,” *IEEE Trans. Commun.*, vol. 69, no. 10, pp. 7063–7076, Oct. 2021.

[14] 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, Jan. 2023.

[15] C. Han, A. O. Bicen, and I. F. Akyildiz, “Multi-ray channel modeling and wideband characterization for wireless communications in the terahertz band,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2402–2421, May 2015.

[16] G. Han, L. Yan, and J. Yuan, “Hybrid beamforming for terahertz wireless communications: Challenges, architectures, and open problems,” *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 198–204, Aug. 2021.

[17] P. Zhang, J. Chen, X. Yang, N. Ma, and Z. Zhang, “Recent research on massive MIMO propagation channels: A survey,” *IEEE Commun. Mag.*, vol. 56, no. 12, pp. 22–29, Dec. 2018.

[18] F. Bohagen, P. Orten, and G. E. Oien, “On spherical vs. plane wave modeling of line-of-sight MIMO channels,” *IEEE Trans. Commun.*, vol. 57, no. 3, pp. 841–849, Mar. 2009.

[19] M. Cui, L. Dai, R. Schober, and L. Hanzo, “Near-field wideband beamforming for extremely large antenna arrays,” 2021, arXiv:2109.10054.

[20] K. Dovelos, S. D. Assimonis, H. Q. Ngo, B. Bellalta, and M. Matthaiou, “Intelligent reflecting surfaces at terahertz bands: Channel modeling and analysis,” in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2021, pp. 1–6.

[21] S. Zhang and R. Zhang, “Capacity characterization for intelligent reflecting surface-assisted MIMO communication,” *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1823–1838, Aug. 2020.

[22] M. Jung, W. Saad, Y. Jang, G. Gong, and S. Choi, “Reliability analysis of large intelligent surfaces (LISs): Rate distribution and outage probability,” *IEEE Wireless Commun. Lett.*, vol. 8, no. 6, pp. 1662–1666, Dec. 2019.

[23] E. Björnson and L. Sanguinetti, “Power scaling laws and near-field behaviors of massive MIMO and intelligent reflecting surfaces,” *IEEE Open J. Commun. Soc.*, vol. 1, pp. 1306–1324, 2020.
Yuhang Chen received the B.E. degree in communication engineering from the Harbin Institute of Technology, China, in 2019. She is currently pursuing the Ph.D. degree with the Terahertz Wireless Communications (TWC) Laboratory, University of Michigan-Shanghai Jiao Tong University (UM-SJTU) Joint Institute, Shanghai Jiao Tong University, China. Her current research interests include terahertz beam management and channel estimation.

Renwang Li (Student Member, IEEE) received the B.S. degree in communication engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2019. He is currently pursuing the Ph.D. degree in information and communication engineering with the Department of Electronic Engineering, Shanghai Jiao Tong University (SJTU), Shanghai, China. His current research interests include convex optimization, reconfigurable intelligent surface, and signal processing.

Chong Han (Member, IEEE) received the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, USA, in 2016. He is currently a John Wu & Jane Sun Endowed Associate Professor with the University of Michigan-Shanghai Jiao Tong University (UM-SJTU) Joint Institute, Shanghai Jiao Tong University, China, and the Director of the Terahertz Wireless Communications (TWC) Laboratory. Since 2021, he has been affiliated with the Department of Electronic Engineering and Cooperative Mediation Innovation Center (CMIC), Shanghai Jiao Tong University. His research interests include terahertz and millimeter-wave communications. He is a member of the ACM. He was a recipient of the 2018 Elsevier NanoComNet (Nano Communication Network) Young Investigator Award, 2017 Shanghai Sailing Program, and 2018 Shanghai ChenGuang Program. He is the TPC Chair to organize multiple IEEE and ACM conferences and workshops, including GC-2023 SAC THz Communications. He has been the Co-Founder and the Vice-Chair of IEEE ComSoc Special Interest Group (SIG) on Terahertz Communications since 2021. He is the Guest Editor of IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, IEEE TRANSACTIONS ON NANOTECHNOLOGY, and IEEE OPEN JOURNAL OF VEHICULAR TECHNOLOGY.

Shu Sun (Member, IEEE) received the B.S. degree in applied physics from Shanghai Jiao Tong University (SJTU), China, in 2012, and the Ph.D. degree in electrical engineering from New York University (NYU), USA, in 2018. She held summer internship positions with Nokia Bell Laboratories in 2014 and 2015. From July 2018 to August 2021, she was a Systems Engineer with the Next Generation and Standards Group, Intel Corporation. She is currently a tenure-track Associate Professor with the Department of Electronic Engineering, SJTU. As a delegate of NYU, she has coauthored a white paper on 5G channel modeling in collaboration with 3GPP members in 2016, and her proposal on path loss models has been adopted by 3GPP since Release 14. Additionally, she has published many highly-cited papers in the field of wireless communications and has received a number of awards, including the 2015 IEEE Donald G. Fink Award, the 2016 IEEE VTC-Spring Best Paper Award, the 2017 IEEE Neil Shepherd Memorial Best Propagation Paper Award, the 2017 Marconi Society Young Scholar Award, the 2018 NYU Dante Youla Award, the 2021 Intel Distinguished Invention Award, and the 2022 Xiaomi Young Scholar of SJTU. Her research interests include millimeter-wave communications, integrated sensing and communication, massive MIMO and beyond, and reconfigurable intelligent surfaces. She serves as a Guest Editor for IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, an Associate Editor for IEEE OPEN JOURNAL OF ANTENNAS AND PROPAGATION, and a Guest Editor for Sensors.

Meixia Tao (Fellow, IEEE) received the B.S. degree in electronic engineering from Fudan University, Shanghai, China, in 1999, and the Ph.D. degree in electrical and electronic engineering from The Hong Kong University of Science and Technology in 2003. She is currently a Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University, China. Her current research interests include wireless edge learning, coded caching, reconfigurable intelligence surfaces, and physical layer multicasting.

Dr. Tao received the 2019 IEEE Marconi Prize Paper Award, the 2013 IEEE Heinrich Hertz Award for Best Communications Letters, the IEEE/CIC International Conference on Communications in China (ICCC) 2015 Best Paper Award, and the International Conference on Wireless Communications and Signal Processing (WCSP) 2012 Best Paper Award. She also received the 2009 IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award. She served as a member of the Executive Editorial Committee of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2015 to 2019. She was also on the Editorial Board of several other journals as an Editor or the Guest Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS and IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. She is also serving as the TPC Co-Chair for IEEE ICC 2023. She is an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY and an Editor-at-Large of the IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY.