Hybrid Beamforming Design for Multi-User Multi-Stream Communications with Terahertz Massive MIMO

Ziwei Wan[1], Tong Qin[1], Zhen Gao[1(✉)], Chun Hu[1], Yuezu Lv[1], Tuan Li[1], Chunli Zhu[1], and Jiening Mao[2]

[1] School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, People's Republic of China
{ziweiwan,qintong,gaozhen16,bit_hc,yzlv,tuanli,chunlizhu}@bit.edu.cn
[2] Zhongxing Telecommunication Equipment (ZTE) Corporation, Shenzhen, Guangdong 518000, People's Republic of China
mao.jiening@zte.com.cn

Abstract. The hybrid beamforming (HBF) design for terahertz (THz) massive multiple-input multiple-output (mMIMO) communications is an essential but challenging problem in future wireless communications. In this paper, we propose an efficient HBF design framework for downlink multi-user multi-stream transmission in broadband THz mMIMO systems. First, the analytic solutions for digital precoder are derived to minimize the sum-mean-square error between the transmitted and received symbols. Then, we propose a tractable criterion for the codebook-based fully-connected analog precoder design, and further extend it to the case where the dynamic partially-connected structure is considered. A low-complexity energy-based greedy antenna grouping scheme is proposed for the dynamic hybrid structure. Simulation results demonstrate the effectiveness and the superiority of the proposed scheme in terms of sum rate and bit error rate over its counterparts.

Keywords: Terahertz (THz) · Hybrid beamforming (HBF) · Massive MIMO · Multi-user MIMO

1 Introduction

Given the spectrum crunch caused by ever-increasing connected devices and applications, terahertz (THz) technique is considered as a promising candidate for promoting future wireless communications [1]. Particularly, the cooperation between THz techniques and the state-of-the-art massive multiple-input multiple-output (mMIMO) system has received tremendous interest [1,2]. Equipped with a large number of antennas, mMIMO is able to compensate the severe path loss in the THz band and achieve finer-grain spatial multiplexing. However, it is impractical to deploy the power-aggressive fully-digital THz mMIMO, where one radio frequency chain (RFC) is dedicated for each

J. Zhao (Ed.): WiSATS 2023, LNICST 509, pp. 59–73, 2023.
https://doi.org/10.1007/978-3-031-34851-8_5

antenna. As a remedy, hybrid beamforming (HBF) architecture, which was initially applied to millimeter-wave (mmWave) mMIMO communications [3,4], has been considered for realizing THz mMIMO [5–7]. Similarly to the mmWave cases, it has been reported that the HBF architecture only requires much less RFCs to drive the entire antenna array to strike a balance between the hardware complexity, the power consumption, and the system performance.

Some preliminary THz HBF schemes have been proposed recently [5–7]. In [5], the HBF for THz non-orthogonal multiple access (NOMA) systems was investigated. The inversion of channel matrix with phase quantization was adopted as the analog precoder, while the low-complexity zero-forcing precoding was conducted for the digital counterpart. The authors of [6] proposed a two-stage hybrid beamforming scheme for THz mMIMO systems, which consists of a codebook searching algorithm and a regularized channel inversion method for analog and digital precoding, respectively. However, the prior works [5,6] did not consider the energy-efficient dynamic partially-connected structure in THz hybrid mMIMO, where the antennas are divided into several non-overlapped groups and the connection scheme between the antennas groups and the RFCs can be adjusted based on the real-time channel state information (CSI). A dynamic array-of-subarrays structure and the corresponding HBF design were proposed [7] for THz mMIMO, while the work [7] did not consider multi-user communications, which is vital in future THz cellular systems. The HBF design for multi-user multi-stream communications with THz mMIMO has not been well addressed at this stage.

In this paper, we propose a HBF design for multi-user multi-stream communications with THz mMIMO. The dynamic partially-connected structure is also considered to realize more resilient THz communications. Compared with the prior works [5–7], we consider a more generic communication scenario, namely multi-user multi-stream communications. We derive the analytic solutions for the digital precoder design to minimizes the sum-mean-square error (SMSE) between the transmitted and received symbols. As for the analog precoding in the dynamic partially-connected structure, a low-complexity codebook-based scheme is proposed based on a tractable criterion. Simulations are also conducted to verify the performance of the proposed scheme.

Notations: Lower-case and upper-case boldface letters denote vectors and matrices, respectively. $(\cdot)^T$, $(\cdot)^H$, and $\mathrm{tr}\,(\cdot)$ denote the transpose, conjugate transpose, and trace of an input matrix, respectively. $(\cdot)^{(i)}$ and $(\cdot)_{i,j}$ represent the i-th column vector and the i-th row and j-th column element of an input matrix, respectively. $\mathbf{I}_N$ denotes an $N \times N$ identity matrix, while $\mathbf{0}_{M\times N}$ denotes an all-zero matrix of size $M \times N$. $\mathrm{Re}(a)$ is the real part of a. $\|\cdot\|_2$, $\|\cdot\|_{\mathrm{F}}$, $\mathrm{diag}\,(\cdot)$, and $\mathbb{E}\{\cdot\}$ represent the ℓ_2-norm, Frobenius norm, (block) diagonalization, and statistical expectation, respectively. For a set $\mathcal{I}$, $|\mathcal{I}|_{\mathrm{c}}$ is its cardinality, and $\mathrm{setdiff}\,(\mathcal{I}, \{i\})$ denotes the operation that removes the element i from $\mathcal{I}$.

2 System Model

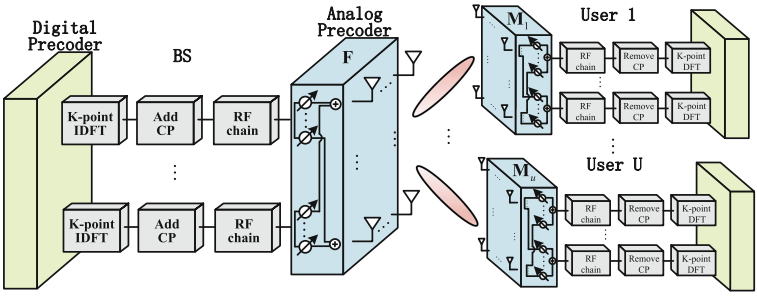

Fig. 1. The considered THz mMIMO system with HBF architecture. The BS supports multiple users with multi-stream transmission for each user.

We consider a multi-user THz mMIMO system with hybrid architecture at both the base station (BS) and users, as shown in Fig. 1. The BS employs N_t antennas and $M_t \ll N_t$ RFCs to simultaneously serve U users, and each user employs N_r antennas and $M_r \ll N_r$ RFCs to support $N_s \leq M_r$ data streams. Without loss of generality, we assume $N_s = M_r$ and $M_t = UM_r$. The analog part of the hybrid architecture consists of phase shifters and its specific structure will be discussed in the sequel. The orthogonal frequency division multiplexing (OFDM) with K sub-carriers and sampling period T_s is adopted in the system.

We denote the transmitted data symbols for the u-th ($1 \leq u \leq U$) user at the k-th ($1 \leq k \leq K$) sub-carrier as $\mathbf{x}_u[k] \in \mathbb{C}^{N_s \times 1}$. All the U users' data symbols are then concatenated as $\mathbf{x}[k] = \left[\mathbf{x}_1^H[k], ..., \mathbf{x}_U^H[k]\right]^H \in \mathbb{C}^{UN_s \times 1}$. We assume $\mathbb{E}\left\{\mathbf{x}[k]\mathbf{x}^H[k]\right\} = \mathbf{I}_{UN_s}$. Vector $\mathbf{x}[k]$ is precoded by the digital precoder $\mathbf{W}[k] = [\mathbf{W}_1[k], ..., \mathbf{W}_U[k]] \in \mathbb{C}^{M_t \times UN_s}$, where $\mathbf{W}_u[k] \in \mathbb{C}^{M_t \times N_s}$ is the dedicated digital precoders for the u-th user. After K-point inverse discrete Fourier transformation (DFT) and adding a cyclic prefix (CP), the BS forms the transmitted signals by using a frequency-flat analog precoder $\mathbf{F} \in \mathbb{C}^{N_t \times M_t}$. At the u-th user, the received signals are first combined with the frequency-flat analog combiner $\mathbf{M}_u \in \mathbb{C}^{N_r \times M_r}$. Then, after the removal of the CP and performing DFT, the digital combiner $\mathbf{V}_u[k] \in \mathbb{C}^{M_r \times N_s}$ is adopted to combine the received signals at the k-th sub-carrier. Therefore, in the downlink transmission, the received symbols at the k-th sub-carrier of the u-th user can be modeled as

$$\hat{\mathbf{x}}_u[k] = \mathbf{V}_u^H[k]\mathbf{M}_u^H\Big(\mathbf{H}_u[k]\mathbf{FW}[k]\mathbf{x}[k] + \mathbf{n}_u[k]\Big), \tag{1}$$

where $\mathbf{H}_u[k] \in \mathbb{C}^{N_r \times N_t}$ is the effective frequency-domain channel at the k-th sub-carrier between the BS and u-th user. $\mathbf{n}_u[k] \in \mathbb{C}^{N_r \times 1}$ is the additive white Gaussian noise (AWGN) vector at the u-th user with zero mean and power σ^2.

The THz mMIMO channel is modeled as a sum of the contributions of N_c scattering clusters [1,7] and each cluster contains N_p propagation paths. Therefore, the channel impulse response (CIR) of THz mMIMO between the BS and the u-th user can be written as

$$\overline{\mathbf{H}}_u[d] = \sum_{i=1}^{N_c}\sum_{l=1}^{N_p} g_{il}^u(d)\mathbf{a}_r\left(\theta_{il}^{ru},\phi_{il}^{ru}\right)\mathbf{a}_t^H\left(\theta_{il}^{tu},\phi_{il}^{tu}\right), \tag{2}$$

where $g_{il}^u(d) = \sqrt{N_t N_r/(N_c N_p)}\alpha_{il}^u p(dT_s - \tau_{il}^u)$ is the delay-domain channel coefficient, $\alpha_{il}^u \sim \mathcal{CN}(0,1)$ is the complex gain of the i-th path in the l-th cluster, $p(\cdot)$ is the pulse shaping filtering function, τ_{il} is the delay-offset of the i-th path in the l-th cluster, $\theta_{il}^{ru}(\phi_{il}^{ru})$ and $\theta_{il}^{tu}(\phi_{il}^{tu})$ are the azimuth (elevation) angles of arrival (AoAs) and departure (AoDs) of the i-th path in the l-th cluster, respectively, and $\mathbf{a}_r(\theta_{il}^{ru},\phi_{il}^{ru}) \in \mathbb{C}^{N_r\times 1}$ and $\mathbf{a}_t(\theta_{il}^{tu},\phi_{il}^{tu}) \in \mathbb{C}^{N_t\times 1}$ are the normalized receive and transmit steering vectors, respectively. The superscript "u" in (2) corresponds to the u-th user. Throughout the paper, we assume that the half-wavelength spaced uniform planar array (UPA) is deployed at both the BS and users. Taking the BS as an example, the steering vector of the UPA with $N_t = N_y \times N_z$ elements can be written as

$$\mathbf{a}_t(\theta,\phi) = \frac{1}{\sqrt{N_y N_z}}\big[1,...,e^{j\pi(n\sin\theta\cos\phi + m\sin\phi)}, \\ ...,e^{j\pi((N_y-1)\sin\theta\cos\phi+(N_z-1)\sin\phi)}\big]^T, \tag{3}$$

where $1 \le n < N_y$ and $1 \le m < N_z$. The receive steering vector $\mathbf{a}_r(\theta,\phi)$ can be similarly obtained. With the CIR in (2), the effective frequency-domain channel in (1) can be expressed as

$$\mathbf{H}_u[k] = \sum_{d=0}^{D-1}\overline{\mathbf{H}}_u[d]\, e^{-j\frac{2\pi k}{K}d}, \tag{4}$$

where $D > \max_{u,i,l}\{\tau_{il}^u\}/T_s$ is the length of the CP. Note that during the precoding design, the u-th user only knows its own CSI $\mathbf{H}_u[k]$, $\forall k$, while the global CSI $\mathbf{H}_u[k]$, $\forall u,k$, are available at the BS, which is a commonly adopted assumption in literature [4–8] for the investigation of the performance upper bound.

3 Problem Formulation

We propose to minimize the mean square error (MSE) between the transmitted and received symbols for the joint design of the analog and digital precoders. The u-th user's MSE at the k-th sub-carrier can be expressed as

$$\xi_u[k] = \mathbb{E}\left\{\left\|\beta^{-1}[k]\hat{\mathbf{x}}_u[k] - \mathbf{x}_u[k]\right\|_2^2\right\}, \tag{5}$$

where $\beta[k]$ is the normalization factor for the digital precoder $\mathbf{W}[k]$. By substituting (1) into (5), $\xi_u[k]$ can be further expressed as

$$\begin{aligned}\xi_u[k] =& \mathrm{tr}\left(\mathbf{V}_u^H[k]\,\mathbf{H}_{\mathrm{eff}}^u[k]\,\tilde{\mathbf{W}}[k]\,\tilde{\mathbf{W}}^H[k]\,(\mathbf{H}_{\mathrm{eff}}^u[k])^H\mathbf{V}_u[k]\right)\\ &-\mathrm{tr}\left(\tilde{\mathbf{W}}_u^H[k]\,(\mathbf{H}_{\mathrm{eff}}^u[k])^H\mathbf{V}_u[k]\right)-\mathrm{tr}\left(\mathbf{V}_u^H[k]\,\mathbf{H}_{\mathrm{eff}}^u[k]\,\tilde{\mathbf{W}}_u[k]\right)\\ &+\sigma^2\beta^{-2}[k]\,\mathrm{tr}\left(\mathbf{V}_u^H[k]\,\mathbf{M}_u^H\mathbf{M}_u\mathbf{V}_u[k]\right)+N_s,\end{aligned}\tag{6}$$

where $\mathbf{H}_{\mathrm{eff}}^u[k] \triangleq \mathbf{M}_u^H\mathbf{H}_u[k]\mathbf{F} \in \mathbb{C}^{M_r\times M_t}$ represents the effective baseband channel, $\tilde{\mathbf{W}}[k] = \beta^{-1}[k]\,\mathbf{W}[k]$ is an effective digital precoder, and $\tilde{\mathbf{W}}_u[k] = \beta^{-1}[k]\,\mathbf{W}_u[k]$. Note that we leverage the independence among the transmitted symbols $\mathbf{x}[k]$ and the AWGN vector $\mathbf{n}_u[k]$ for obtaining (6). We assume that the analog part at the user is of fully-connected structure, while that at the BS is of dynamic partially-connected structure. The HBF design can be formulated as the following optimization problem:

$$\underset{\{\mathbf{W}[k]\}_{k=1}^K,\{\mathbf{V}[k]\}_{k=1}^K,\mathbf{F},\mathbf{M}}{\text{minimize}} \sum_{k=1}^{K}\sum_{u=1}^{U}\xi_u[k] \tag{7a}$$

$$\text{s.t.}\sum_{k=1}^{K}\mathrm{tr}\left(\mathbf{F}\mathbf{W}[k]\mathbf{W}^H[k]\mathbf{F}^H\right)=P, \tag{7b}$$

$$|\mathbf{F}_{i,j}|\in\{0,\frac{1}{\sqrt{N_t}}\},\left|(\mathbf{M}_u)_{i,j}\right|=\frac{1}{\sqrt{N_r}},\forall i,j,u, \tag{7c}$$

where $\mathbf{V}[k] = \mathrm{diag}(\mathbf{V}_1[k],...,\mathbf{V}_U[k])$, $\mathbf{M} = \mathrm{diag}(\mathbf{M}_1,...,\mathbf{M}_U)$, and P is the total transmit power. Note that $|\mathbf{F}_{i,j}| = 0$ means that there is no connection between the i-th RF chain and the j-th transmit antenna.

4 Proposed Hybrid Beamforming Scheme

In this section, we first derive the digital precoding solution based on the min-SMSE criterion. Then, we further propose a codebook-based scheme for analog precoding design under dynamic partially-connected structure.

4.1 Digital Precoding Design

We first focus on the digital precoder/combiner design while fixing the analog precoder/combiner. Based on (7), the digital precoder/combiner design problem can be rewritten as

$$\underset{\{\mathbf{W}[k]\}_{k=1}^K,\{\mathbf{V}[k]\}_{k=1}^K}{\text{minimize}} \sum_{k=1}^{K}\sum_{u=1}^{U}\xi_u[k] \tag{8a}$$

$$\text{s.t.}\sum_{k=1}^{K}\mathrm{tr}\left(\mathbf{F}\mathbf{W}[k]\mathbf{W}^H[k]\mathbf{F}^H\right)=P. \tag{8b}$$

Given the limited CSI at the user side and the independence among the sub-carriers, we decompose the problem (8) into the subproblems that minimizing $\xi_u[k]$ for each u and k when designing the digital combiners. For the unconstrained $\mathbf{V}_u^H[k]$, we let the partial derivative $\frac{\partial \xi_u[k]}{\partial \mathbf{V}_u^H[k]}$ ($\xi_u[k]$ is shown in (6)) equal to zero and thus obtain

$$\mathbf{V}_u^H[k] = \tilde{\mathbf{W}}_u^H[k] \left(\mathbf{H}_{\text{eff}}^u[k]\right)^H \mathbf{A}^{-1}, \tag{9}$$

where

$$\mathbf{A} \triangleq \mathbf{H}_{\text{eff}}^u[k] \tilde{\mathbf{W}}[k] \tilde{\mathbf{W}}^H[k] \left(\mathbf{H}_{\text{eff}}^u[k]\right)^H + \sigma^2 \beta^{-2}[k] \mathbf{M}_u^H \mathbf{M}_u.$$

On the other hand, we design the digital precoder $\mathbf{W}[k]$ at the BS in order to minimize the SMSE of all users, namely $\xi[k] = \sum_{u=1}^{U} \xi_u[k]$, which can be written as

$$\begin{aligned} \xi[k] =& \beta^{-2}[k] \operatorname{tr}(\mathbf{B}) - \beta^{-1}[k] \operatorname{tr}\left(\mathbf{W}^H[k] \mathbf{H}_{\text{eff}}^H[k] \mathbf{V}[k]\right) \\ & - \beta^{-1}[k] \operatorname{tr}\left(\mathbf{V}^H[k] \mathbf{H}_{\text{eff}}[k] \mathbf{W}[k]\right) \\ & + \sigma^2 \beta^{-2}[k] \operatorname{tr}\left(\mathbf{V}^H[k] \mathbf{M}^H \mathbf{M} \mathbf{V}[k]\right) + U N_s, \end{aligned} \tag{10}$$

where $\mathbf{B} \triangleq \mathbf{V}^H[k] \mathbf{H}_{\text{eff}}[k] \mathbf{W}[k] \mathbf{W}^H[k] \mathbf{H}_{\text{eff}}^H[k] \mathbf{V}[k]$ and $\mathbf{H}_{\text{eff}}[k] = \left[\left(\mathbf{H}_{\text{eff}}^1[k]\right)^H, \ldots, \left(\mathbf{H}_{\text{eff}}^U[k]\right)^H\right]^H \in \mathbb{C}^{UM_r \times M_t}$. To seek the optimal digital precoder with the transmit power constraint, we introduce a Lagrangian multiplier $\mu[k]$ and construct the Lagrangian function as

$$L = \xi[k] + \mu[k] \left(\operatorname{tr}\left(\mathbf{F}\mathbf{W}[k] \mathbf{W}^H[k] \mathbf{F}^H\right) - \tilde{P}\right), \tag{11}$$

where $\tilde{P} = P/K$ denotes the transmit power for each sub-carrier after uniform power allocation among sub-carrier. Letting the partial derivative $\frac{\partial L}{\partial \mathbf{W}[k]}$ equal to zero yields

$$\mathbf{W}[k] = \beta[k] \mathbf{C}^{-1} \mathbf{H}_{\text{eff}}^H[k] \mathbf{V}[k], \tag{12}$$

where $\mathbf{C} \triangleq \mathbf{H}_{\text{eff}}^H[k] \mathbf{V}[k] \mathbf{V}^H[k] \mathbf{H}_{\text{eff}}[k] + \mu[k] \beta^2[k] \mathbf{F}^H \mathbf{F}$. Similarly, let the partial derivative $\frac{\partial L}{\partial \beta[k]}$ equal to zero and we arrive at

$$\begin{aligned} & \beta[k] \operatorname{Re}\left\{\operatorname{tr}\left(\mathbf{V}^H[k] \mathbf{H}_{\text{eff}}[k] \mathbf{W}[k]\right)\right\} \\ & \quad - \operatorname{tr}(\mathbf{B}) - \sigma^2 \operatorname{tr}\left(\mathbf{V}^H[k] \mathbf{M}^H \mathbf{M} \mathbf{V}[k]\right) = 0. \end{aligned} \tag{13}$$

We introduce the following Lemma 1 to further simplify (13).

Lemma 1. *Our proposed precoding design in the sequel will render*

$$\operatorname{Re}\left\{\operatorname{tr}\left(\mathbf{V}^H[k] \mathbf{H}_{\text{eff}}[k] \tilde{\mathbf{W}}[k]\right)\right\} = \operatorname{tr}\left(\mathbf{V}^H[k] \mathbf{H}_{\text{eff}}[k] \tilde{\mathbf{W}}[k]\right), \tag{14}$$

i.e., all the diagonal elements of $\mathbf{V}^H[k] \mathbf{H}_{\text{eff}}[k] \tilde{\mathbf{W}}[k]$ *are real numbers.*

Proof. Note that

$$\mathrm{tr}\left(\mathbf{V}^H[k]\,\mathbf{H}_{\mathrm{eff}}[k]\,\tilde{\mathbf{W}}[k]\right)=\sum_{u=1}^{U}\mathrm{tr}\left(\mathbf{V}_u^H[k]\,\mathbf{H}_{\mathrm{eff}}^u[k]\,\tilde{\mathbf{W}}_u[k]\right), \tag{15}$$

and thus **Lemma** 1 can be proven by noticing from (27) that $\mathbf{V}_u^H[k]\,\mathbf{H}_{\mathrm{eff}}^u[k]\,\tilde{\mathbf{W}}_u[k]$, $\forall u$, are positive semi-definite matrices, whose diagonal elements are obviously real numbers. □

According to Lemma 1, we can safely remove the symbol $\mathrm{Re}\{\cdot\}$ in (13). Then, based on (13), we have

$$\mu[k]\,\beta^2[k]=\sigma^2\mathrm{tr}\left(\mathbf{V}^H[k]\,\mathbf{M}^H\mathbf{M}\mathbf{V}[k]\right)/\tilde{P}. \tag{16}$$

Apparently, if we substitute (16) into (12), $\tilde{\mathbf{W}}[k]$ can be solved accordingly, even though the specific values of $\mu[k]$ and $\beta[k]$ are unknown. The actual digital precoder can be obtained as $\mathbf{W}[k]=\beta[k]\,\tilde{\mathbf{W}}[k]$ with

$$\beta[k]=\sqrt{\tilde{P}/\mathrm{tr}\left(\mathbf{F}\tilde{\mathbf{W}}[k]\,\tilde{\mathbf{W}}^H[k]\,\mathbf{F}^H\right)}. \tag{17}$$

With the closed-form solutions (9) and (12), it is natural to conduct an alternating optimization, i.e., to optimize $\mathbf{W}[k]$ or $\mathbf{V}_u[k]$ by fixing another as constant and repeat until convergence, at the cost of prohibitive computational complexity. Fortunately, it has been verified in [9] that if an appropriate initial value $\mathbf{V}_{\mathrm{ini}}$ is chosen and assigned to $\mathbf{V}[k]$, a sufficiently good performance can be guaranteed even without the alternating optimization. Specifically, we set

$$\mathbf{V}_{\mathrm{ini}}=\mathrm{diag}\left(\mathbf{V}_{\mathrm{ini}}^1,...,\mathbf{V}_{\mathrm{ini}}^U\right), \tag{18}$$

where $\{\mathbf{V}_{\mathrm{ini}}^u\}_{u=1}^U$ are arbitrary unitary matrices. We propose the following procedure to design the digital precoder/combiner. Firstly, $\mathbf{V}_{\mathrm{ini}}$ in (18) is used as the initial solution of the digital combiners of all users and is substituted into (12) to obtain the digital precoder at the BS. Then, the obtained digital precoder is substituted into (9) to update the digital combiners at the users. It has been reported in [9] that the above procedure without alternating optimization guarantees $\mathbf{V}^H[k]\,\mathbf{H}_{\mathrm{eff}}[k]\,\tilde{\mathbf{W}}[k]$ to be a (nearly) diagonal matrix, eliminating the inter-user and inter-stream interferences. Interested readers are referred to [9] for more details.

4.2 Analog Precoding with Fully-Connected Structure

We first consider the analog precoding design for the fully-connected architecture at the BS with the obtained digital part, and the proposed scheme will be extended to the dynamic partially-connected architecture in the next subsection. The analog precoding design problem with the known digital precoder/combiner can be written as

$$\underset{\{\mathbf{M}_u\}_{u=1}^{U},\mathbf{F}}{\text{minimize}} \sum_{k=1}^{K} \xi[k], \tag{19a}$$

$$\text{s.t.} \quad \left|(\mathbf{M}_u)_{i,j}\right| = \frac{1}{\sqrt{N_r}}, |\mathbf{F}_{i,j}| = \frac{1}{\sqrt{N_t}}, \forall i, j, u. \tag{19b}$$

Based on (13) and (10), we can obtain that when the proposed digital precoder/combiner are applied, the following equation holds

$$\begin{aligned}\xi[k] &= UN_s - \beta^{-1}[k]\,\mathrm{tr}\left(\mathbf{V}^H[k]\,\mathbf{H}_{\text{eff}}[k]\,\mathbf{W}[k]\right) \\ &= UN_s - \mathrm{tr}\left(\mathbf{V}^H[k]\,\mathbf{H}_{\text{eff}}[k]\,\tilde{\mathbf{W}}[k]\right),\end{aligned} \tag{20}$$

which indicates that we can improve $\xi[k]$ by maximizing $\mathrm{tr}\big(\mathbf{V}^H[k]\,\mathbf{H}_{\text{eff}}[k]\,\tilde{\mathbf{W}}[k]\big)$.

To further characterize the impact of $\mathbf{H}_{\text{eff}}$ on $\xi[k]$, in the rest of this section, we assume that $\mathbf{F}^H\mathbf{F} = \mathbf{I}_{M_t}$ and $\mathbf{M}^H\mathbf{M} = \mathbf{I}_{UM_r}$, which can be perfectly satisfied via the proposed codebook-based analog precoding design, as detailed hereinafter. Recall that $\{\mathbf{V}_{\text{ini}}^u\}_{u=1}^{U}$ are unitary matrices, so we have $\mu[k]\,\beta^2[k] = \sigma^2\mathrm{tr}\left(\mathbf{M}^H\mathbf{M}\right)/\tilde{P} \approx \sigma^2 UM_r/\tilde{P}$ based on (16). Therefore, according to (12), the obtained digital precoder (without power constraint) can be written as $\tilde{\mathbf{W}}[k] = \left(\mathbf{H}_{\text{eff}}^H[k]\,\mathbf{H}_{\text{eff}}[k] + \varepsilon\mathbf{I}_{M_t}\right)^{-1}\mathbf{H}_{\text{eff}}^H[k]\,\mathbf{V}_{\text{ini}}$, that is

$$\tilde{\mathbf{W}}_u[k] = \left(\mathbf{H}_{\text{eff}}^H[k]\,\mathbf{H}_{\text{eff}}[k] + \varepsilon\mathbf{I}_{M_t}\right)^{-1}(\mathbf{H}_{\text{eff}}^u[k])^H\mathbf{V}_{\text{ini}}^u, \tag{21}$$

where $\varepsilon \triangleq \sigma^2 UM_r/\tilde{P}$ and $\mathbf{V}_{\text{ini}}$ is the initial value for the design of digital part, as shown in (18). Moreover, by substituting $\mathbf{M}^H\mathbf{M} \approx \mathbf{I}_{UM_r}$ into (9), we have

$$\begin{aligned}\mathbf{V}_u^H[k] \approx &\tilde{\mathbf{W}}_u^H[k]\,(\mathbf{H}_{\text{eff}}^u[k])^H \\ &\left(\mathbf{H}_{\text{eff}}^u[k]\,\tilde{\mathbf{W}}[k]\,\tilde{\mathbf{W}}^H[k]\,(\mathbf{H}_{\text{eff}}^u[k])^H + \sigma^2\beta^{-2}[k]\,\mathbf{I}_{M_r}\right)^{-1}.\end{aligned} \tag{22}$$

To simplify the representation in (22), we consider the singular value decomposition (SVD) of $\mathbf{H}_{\text{eff}}[k]$, that is

$$\mathbf{H}_{\text{eff}}[k] = \mathbf{L}[k]\,\boldsymbol{\Sigma}[k]\,\mathbf{R}^H[k], \tag{23}$$

where $\mathbf{L}[k] \in \mathbb{C}^{UM_r \times UM_r}$ and $\mathbf{R}[k] \in \mathbb{C}^{M_t \times M_t}$ are unitary matrices, and $\boldsymbol{\Sigma}[k] = \mathrm{diag}\left(b_1[k], ..., b_{M_t}[k]\right)$ is a diagonal matrix with the non-zero singular values $b_n[k]$, $1 \le n \le M_t$, on its diagonal, since we have assumed that $M_t = UM_r$ and the channel matrices are rank-sufficient for multi-stream communications. By denoting $\mathbf{L}[k] = \left[\mathbf{L}_1^H[k], ..., \mathbf{L}_U^H[k]\right]^H$ with each $\mathbf{L}_u[k] \in \mathbb{C}^{M_r \times UM_r}$, we obtain

$$\mathbf{H}_{\text{eff}}^u[k] = \mathbf{L}_u[k]\,\boldsymbol{\Sigma}[k]\,\mathbf{R}^H[k]. \tag{24}$$

By exploiting (23) and (24), we simplify (22) as

$$\mathbf{V}_u^H[k] = \mathbf{D}\left(\mathbf{D}^H\mathbf{D} + \sigma^2\beta^{-2}[k]\,\mathbf{I}_{M_r}\right)^{-1}, \tag{25}$$

where the auxiliary matrix $\mathbf{D}$ can be expressed as

$$\mathbf{D} = \left(\mathbf{V}_{\text{ini}}^{u}\right)^{H} \mathbf{S}_u [k] \left(\mathbf{I}_{UM_r} + \varepsilon \boldsymbol{\Sigma}^{-2} [k]\right)^{-1} \mathbf{S}_u^{H} [k]. \tag{26}$$

With (25)–(26), the effective channel for the u-th user $\mathbf{V}_u^H [k] \mathbf{H}_{\text{eff}}^u [k] \tilde{\mathbf{W}}_u [k]$ can be written as

$$\mathbf{V}_u^H [k] \mathbf{H}_{\text{eff}}^u [k] \tilde{\mathbf{W}}_u [k] = \mathbf{D}\left(\mathbf{D}^H \mathbf{D} + \sigma^2 \beta^{-2} [k] \mathbf{I}_{M_r}\right)^{-1} \mathbf{D}^H. \tag{27}$$

For ease of further analysis, we focus on a large signal-to-noise-ratio (SNR) limit such that $\varepsilon = \sigma^2 U M_r / \tilde{P} \to 0$. This assumption has been adopted in the previous work on mmWave HBF [9], and it is also reasonable in THz mMIMO communications, since higher antenna gain and array gain can be achieved by THz antenna elements [1] with the known CSI during precoding stage. In such cases, $\mathbf{D}$ in (26) approximates to $\left(\mathbf{V}_{\text{ini}}^{u}\right)^{H}$, and thus we have

$$\begin{aligned} \text{tr}\left(\mathbf{V}_u^H [k] \mathbf{H}_{\text{eff}}^u [k] \tilde{\mathbf{W}}_u [k]\right) &= \text{tr}\left(\mathbf{D}^H \mathbf{D}\left(\mathbf{D}^H \mathbf{D} + \sigma^2 \beta^{-2} [k] \mathbf{I}_{M_r}\right)^{-1}\right) \\ &\overset{\varepsilon \to 0}{=} \frac{M_r}{1 + \sigma^2 \beta^{-2} [k]}, \end{aligned} \tag{28}$$

where

$$\beta^2 [k] = \frac{\tilde{P}}{\text{tr}\left(\left(\boldsymbol{\Sigma}^2 [k] + \varepsilon \mathbf{I}_{M_t}\right)^{-1} \boldsymbol{\Sigma}^2 [k] \left(\boldsymbol{\Sigma}^2 [k] + \varepsilon \mathbf{I}_{M_t}\right)^{-1}\right)} \overset{\varepsilon \to 0}{=} \frac{\tilde{P}}{\sum\limits_{i=1}^{M_t} \frac{1}{b_i^2 [k]}}, \tag{29}$$

according to (17).

Substituting (27) and (29) into (20), we can rewrite $\xi [k]$ as

$$\xi [k] = U N_s - \sum_{u=1}^{U} \text{tr}\left(\mathbf{V}_u^H [k] \mathbf{H}_{\text{eff}}^u [k] \tilde{\mathbf{W}}_u [k]\right) = \frac{U N_s \sigma^2}{\frac{\tilde{P}}{\sum\limits_{i=1}^{M_t} \frac{1}{b_i^2 [k]}} + \sigma^2}. \tag{30}$$

It is obvious that $\xi [k]$ decreases with $b_n^2 [k]$ increasing, $\forall n$, which sheds light on how to design the analog precoder/combiner to improve the SMSE. As a heuristic method, we can maximize $\sum\limits_{k=1}^{K} \sum\limits_{i=1}^{M_t} b_i^2 [k] = \text{tr}\left(\mathbf{H}_{\text{eff}} [k] \mathbf{H}_{\text{eff}}^H [k]\right)$. Furthermore, given that $\mathbf{H}_{\text{eff}}$ should be a (nearly) diagonal matrix to eliminate interferences such that $\text{tr}\left(\mathbf{H}_{\text{eff}} [k] \mathbf{H}_{\text{eff}}^H [k]\right) \approx \sum\limits_{i=1}^{M_t} \left|\left(\mathbf{H}_{\text{eff}} [k]\right)_{i,i}\right|^2$, we can maximize the sum of the squares of the modulus of diagonal entries in $\mathbf{H}_{\text{eff}} [k]$, $\sum\limits_{i=1}^{M_t} \left|\left(\mathbf{H}_{\text{eff}} [k]\right)_{i,i}\right|^2$. On the other hand, to realize low-complexity computations and affordable feedback overhead for THz communications, we consider that the analog precoders (i.e., $\mathbf{M}_u^{(m)}$ and $\mathbf{F}^{(m)}$) can only be constructed from the two-dimensional DFT

codebooks $\mathcal{D}_{r,u}$ and $\mathcal{D}_t$, respectively. Using the obtained optimization objective (i.e., $\sum_{i=1}^{M_t} \left|\left(\mathbf{H}_{\text{eff}}[k]\right)_{i,i}\right|^2$) and codebooks, the analog precoding problem turns into

$$\underset{\{\mathbf{M}_u\}_{u=1}^{U},\mathbf{F}}{\text{maximize}} \sum_{k=1}^{K}\sum_{u=1}^{U}\sum_{m=1}^{N_s}\left|\left(\mathbf{M}_u^{(m)}\right)^H \mathbf{H}_u[k]\,\mathbf{F}_u^{(m)}\right|^2 \tag{31a}$$

$$\text{s.t.}\mathbf{M}_u^{(m)} \in \mathcal{D}_{r,u}, \mathbf{F}_u^{(m)} \in \mathcal{D}_t, \forall m, u, \tag{31b}$$

We propose an iterative greedy algorithm to solve (31), as described in **Algorithm 1**. Specifically, in each iteration, we select the optimal codeword-pair from the corresponding codebooks as the columns of analog precoder and combiner based on the criterion in (31). Note that we adopt (two-dimensional) DFT codewords in $\mathcal{D}_t$ and $\mathcal{D}_{t,u}$ so that the orthogonality assumptions $\mathbf{F}^H\mathbf{F} = \mathbf{I}_{M_t}$ and $\mathbf{M}^H\mathbf{M} = \mathbf{I}_{UM_r}$ in previous analysis can be perfectly satisfied.

Algorithm 1. Proposed Analog Precoding Design

Input: The codebooks $\mathcal{D}_t$ and $\mathcal{D}_{r,u}$, and $\mathbf{H}_u[k]$, $\forall u, k$.

Initialization: $\mathcal{U} = \{1, ..., U\}$, $\mathbf{F}_u = \mathbf{M}_u =$ empty matrix, $\forall u$;
while $|\mathcal{U}|_{\text{c}} > 0$ **do**
  $\{u^{\text{opt}}, \mathbf{a}_r^{\text{opt}}, \mathbf{a}_t^{\text{opt}}\} = \underset{u\in\mathcal{U},\mathbf{a}_r\in\mathcal{D}_{r,u},\mathbf{a}_t\in\mathcal{D}_t}{\arg\max} \sum_{k=1}^{K} \left\|\mathbf{a}_r^H\mathbf{H}_u[k]\,\mathbf{a}_t\right\|_2^2$;
  $\mathbf{M}_{u^{\text{opt}}} = [\mathbf{M}_{u^*}|\mathbf{a}_r^{\text{opt}}]$, $\mathbf{F}_{u^{\text{opt}}} = [\mathbf{F}_{u^{\text{opt}}}|\mathbf{a}_t^{\text{opt}}]$;
  $\mathcal{D}_{r,u^{\text{opt}}} = \text{setdiff}\left(\mathcal{D}_{r,u^{\text{opt}}}, \left\{\mathbf{a}_r^{\text{opt}}\right\}\right)$;
  $\mathcal{D}_t = \text{setdiff}\left(\mathcal{D}_t, \left\{\mathbf{a}_t^{\text{opt}}\right\}\right)$;
  if the column number of $\mathbf{M}_{u^{\text{opt}}}$ is M_r, **then**
    $\mathcal{U} = \text{setdiff}\left(\mathcal{U}, \left\{u^{\text{opt}}\right\}\right)$;
  end if
end while

Output: Analog precoders $\mathbf{F}_u$ and combiners $\mathbf{M}_u$, $\forall u$.

4.3 Analog Precoding with Dynamic Partially-Connected Structure

In this subsection, we further consider the analog precoding design when the BS adopts the dynamic partially-connected structure. As shown in Fig. 2, in the considered dynamic structure, both antennas and RFCs are divided into U non-overlapped groups. Each antenna group has N_t/U (which is assume to be an integer) antennas, while each RFC group has N_s RFCs. There is a one-to-one mapping between U antenna groups and U RFC groups, and the fully-connected architecture is applied within each group-pair. Note that the antenna indices in each antenna group can be dynamically adjusted based on CSI, which provides more DoFs for THz communications compared to the fixed partially-connected architecture [5,6]. Let $\mathcal{A}_n \subseteq \{1, ..., N_t\}$ denotes the antenna indices set of the

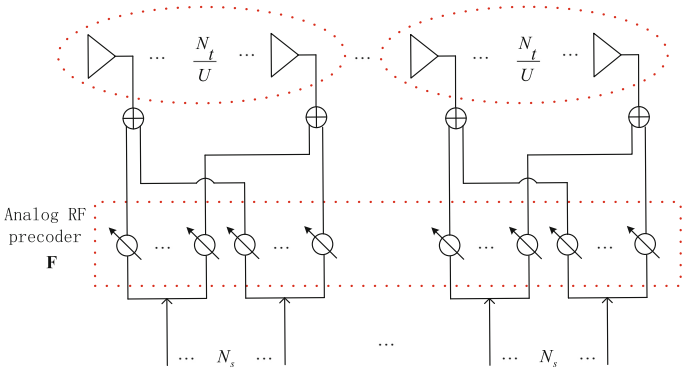

Fig. 2. A dynamic partially-connected hybrid architecture at the BS.

n-th antenna group, $1 \leq n \leq U$, with $|\mathcal{A}_u|_{\mathrm{c}} = N_t/U$ and $\mathcal{A}_u \cap \mathcal{A}_{u'} = \emptyset$ for $u \neq u'$. Accordingly, the problem (31) can be re-formulated under the dynamic architecture as the following optimization problem

$$\underset{\{\mathcal{A}_n\}_{n=1}^{U},\{\mathbf{M}_u\}_{u=1}^{U},\mathbf{F}}{\text{maximize}} \sum_{u=1}^{U}\sum_{m=1}^{N_s}\sum_{k=1}^{K}\left|\left(\mathbf{M}_u^{(m)}\right)^H \mathbf{H}_u\left[k\right]\mathbf{\Lambda}_u\mathbf{F}_u^{(m)}\right|^2 \tag{32a}$$

$$\text{s.t.}\mathbf{M}_u^{(m)} \in \mathcal{D}_{r,u}, \mathbf{F}_u^{(m)} \in \mathcal{D}_t, \forall m, u, \tag{32b}$$

$$(\mathbf{\Lambda}_u)_{i,j} = \begin{cases} 1, & \text{if } i = j \text{ and } i \in \mathcal{A}_u \\ 0, & \text{Otherwise} \end{cases}. \tag{32c}$$

It is clear that once $\mathcal{A}_u$, or equivalently $\mathbf{\Lambda}_u$, $\forall u$, are determined, problem (32) will be the same as problem (31) by treating $\mathbf{H}_u\left[k\right]\mathbf{\Lambda}_u$, $\forall u$, as new effective channels. Inspired by this observation, we focus on the antenna grouping design. Basically, the antenna grouping problem is a combinatorial optimization problem and it requires extremely complicated exhaustive search to obtain the optimal solution. As a remedy, we propose a low-complexity alternative in the sequel. Note that

$$\left|\left(\mathbf{M}_u^{(m)}\right)^H \mathbf{H}_u\left[k\right]\mathbf{\Lambda}_u\mathbf{F}_u^{(m)}\right| \leq \sigma_{\max}(\mathbf{H}_u\left[k\right]\mathbf{\Lambda}_u) \tag{33}$$

holds for any normalized vectors $\mathbf{M}_u^{(m)}$ and $\mathbf{F}_u^{(m)}$, where $\sigma_{\max}\left(\mathbf{H}_u\left[k\right]\mathbf{\Lambda}_u\right)$ is the maximal singular value of $\mathbf{H}_u\left[k\right]\mathbf{\Lambda}_u$. (33) can be proven by considering the compatibility between the vector norm and the matrix norm [10]. For ease of analysis, we assume that $\mathbf{M}_u^{(m)}$ and $\mathbf{F}_u^{(m)}$ have been optimized when designing antenna grouping such that $\left|\left(\mathbf{M}_u^{(m)}\right)^H \mathbf{H}_u\left[k\right]\mathbf{\Lambda}_u\mathbf{F}_u^{(m)}\right| = \sigma_{\max}(\mathbf{H}_u\left[k\right]\mathbf{\Lambda}_u)$. In this way, we can focus only on how to design $\mathbf{\Lambda}_u$ to maximize $\sigma_{\max}(\mathbf{H}_u\left[k\right]\mathbf{\Lambda}_u)$. We introduce a notation $\mathbf{H}_u^{\langle i\rangle}\left[k\right]$ representing a matrix obtained by reserving

only the i-th column of $\mathbf{H}_u[k]$ and setting other elements of $\mathbf{H}_u[k]$ to zero. We have [10]

$$\sigma_{\max}(\mathbf{H}_u[k]) \leq \sigma_{\max}\left(\mathbf{H}_u^{\langle i\rangle}[k]\right) + \sigma_{\max}\left(\mathbf{H}_u[k] - \mathbf{H}_u^{\langle i\rangle}[k]\right). \tag{34}$$

Therefore, $\sigma_{\max}(\mathbf{H}_u[k]\,\mathbf{\Lambda}_u)$ can be further expressed as

$$\begin{aligned}\sigma_{\max}(\mathbf{H}_u[k]\,\mathbf{\Lambda}_u) &\overset{(a)}{\geq} \sigma_{\max}(\mathbf{H}_u[k]) - \sum_{i\notin\mathcal{A}_u}\sigma_{\max}\left(\mathbf{H}_u^{\langle i\rangle}[k]\right)\\ &\overset{(b)}{=} \sigma_{\max}(\mathbf{H}_u[k]) - \sum_{i\notin\mathcal{A}_u}\left\|\mathbf{H}_u^{(i)}[k]\right\|_2,\end{aligned} \tag{35}$$

where (a) is obtained by applying (34) for each $i \notin \mathcal{A}_u$, and (b) is because for the rank-1 matrix $\mathbf{H}_u^{\langle i\rangle}[k]$, its maximal singular value equals to its Frobenius norm, and futher equals to the ℓ_2-norm of its i-th column vector. It is observed from (35) that $\sigma_{\max}(\mathbf{H}_u[k]\,\mathbf{\Lambda}_u)$ increases as the sum ℓ_2-norm of the unselected columns in $\mathbf{H}_u[k]$ decreasing. In other words, given $\mathbf{H}_u[k]$, we can reserve the N_t/U columns with the most significant ℓ_2-norms and allocate the corresponding antennas to the u-th user to maximize $\sigma_{\max}(\mathbf{H}_u[k]\,\mathbf{\Lambda}_u)$, and further to reduce SMSE. Based on the analysis above, the proposed low-complexity energy-based greedy antenna grouping algorithm are summarized in Algorithm 2. Note that once the antenna grouping is completed, the analog precoding design in Algorithm 1 can be directly applied to solve (32).

Algorithm 2. Energy-based Greedy Antenna Grouping Design

Input: $\mathbf{H}_u[k]$, $\forall u, k$.

1: Initialization: $\mathcal{U} = \{1, \cdots, U\}$, $\mathcal{A} = \{1, \cdots, N_t\}$, $\mathcal{A}_u = \emptyset$, $\forall u$;
2: $\tilde{\mathbf{H}}_u = \left[\mathbf{H}_u^T[1], \mathbf{H}_u^T[2], \cdots, \mathbf{H}_u^T[K]\right]^T$;
3: **while** $|\mathcal{A}|_c > 0$ **do**
4: $\quad\{u^{\text{opt}}, n^{\text{opt}}\} = \arg\max_{u\in\mathcal{U}, n\in\mathcal{A}} \left\|\tilde{\mathbf{H}}_u^{(n)}\right\|_2$;
5: $\quad\mathcal{A}_{u^{\text{opt}}} = \mathcal{A}_{u^{\text{opt}}} \cup \{n^{\text{opt}}\}$;
6: $\quad\mathcal{A} = \text{setdiff}(\mathcal{A}, \{n^{\text{opt}}\})$
7: $\quad$**if** $|\mathcal{A}_u|_c = N_t/U$, **then**
8: $\quad\quad\mathcal{U} = \text{setdiff}(\mathcal{U}, \{u^{\text{opt}}\})$;
9: $\quad$**end if**
10: **end while**

Output: Antenna grouping sets $\mathcal{A}_u$, $\forall u$.

5 Simulation Results

In this section, we will investigate the performance of the proposed scheme by numerical simulation. We introduce two benchmarks: (i) the simultaneous greedy

hybrid precoding (S-GHP) scheme in [8], and (ii) the design of digital part in [9]. We set $K = 64$ and $D = 16$. For the THz channel model, the number of clusters is assumed to be $N_c = 8$, and each cluster has $N_p = 10$ propagation paths with the angle spreads 7.5°. The azimuth/elevation AoAs and AoDs follow the uniform distribution within $[-\pi/2, \pi/2)$, and each delay-offset is uniformly distributed within $[0, DT_s)$. The "sinc" function is adopted as the pulse shaping filter function $p(\cdot)$. We consider quadrature phase shift keying (QPSK) modulation when conducting BER simulation. The SNR of the system is defined as $10\log_{10}(P/\sigma^2)$ [dB]. The numbers of antennas, RFCs (data streams), and users are provided below each figure.

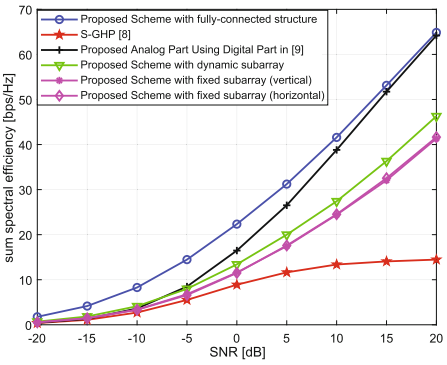

Fig. 3. Sum spectral efficiency achieved by different schemes in an 8-user MIMO system, where $N_t = 16 \times 16$, $N_r = 2 \times 2$, $M_t = 8$, $M_r = N_s = 1$.

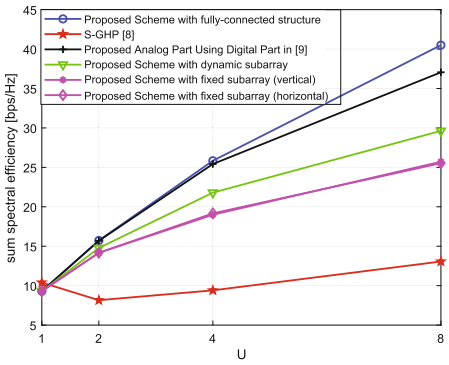

Fig. 4. Sum spectral efficiency achieved by different schemes, where $N_t = 16 \times 16$, $N_r = 1$, $M_r = N_s = 1$, $M_t = UN_s$, SNR = 15 dB.

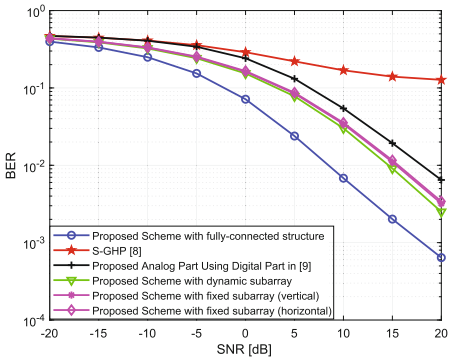

Fig. 5. BER performance achieved by different schemes in a 4-user MIMO system, where $N_t = 16 \times 16$, $N_r = 2 \times 2$, $M_t = 8$, $M_r = N_s = 2$.

Figure 3 illustrates the SSE of proposed schemes against different SNRs. It can be observed that the proposed scheme with fully-connected structure provides a upper-bound for all other schemes, as it requires much more hardware complexity. Also we observe that the proposed scheme and the proposed analog part using digital part in [9] have the same SSE performance in the high SNR. This is because when σ^2 is closed to 0, the proposed digital scheme is equivalent to the digital scheme in [9]. Besides, the proposed scheme with dynamic partially-connected structure outperforms S-GHP scheme [8], the proposed analog part using digital part in [9], and the proposed schemes with different fixed subarray types (vertical and horizontal, cf. [11, Fig. 3]), which indicates that the proposed antenna grouping design improves the precoding performance with dynamic partially-connected structure. Figure 4 illustrates the SSE of the proposed scheme against the number of users. As the number of users increases, the proposed schemes achieve the better performance, and increasing number of users is also better for the proposed scheme with dynamic partially-connected structure than that with different fixed subarray types.

Figure 5 investigates the bit error rate (BER) performance achieved by different schemes. Similarly, We observe that the proposed scheme with fully-connected structure provides a upper-bound for all other schemes. Meanwhile, the proposed scheme with dynamic partially-connected structure also has the better BER performance than the other schemes, benefiting from the proposed HBF scheme which directly minimizes the SMSE between the transmitted and received symbols.

6 Conclusions

This paper proposed an efficient HBF design for multi-user multi-stream communications with THz massive MIMO. By minimizing the SMSE between the transmitted and received symbols, we proposed to design the digital and analog precoding separately. Specifically, first, we derived the analytic solution for digital precoding design based on Lagrangian multiplier method. Then, with the designed digital precoding, we further proposed a tractable criterion for analog precoding design, based on which the analog precoder/combiner are drawn from the pre-defined codebooks to reduce computational complexity and feedback overhead. Moreover, the proposed scheme was extended to dynamic partially-connected architecture. An energy-based greedy antenna grouping algorithm was proposed to optimize the connections between the antennas and radio frequency chains. Simulation results showed that the proposed scheme outperforms its counterparts in term of both the sum spectral efficiency and bit error rate.

References

1. Sarieddeen, H., Alouini, M.S., Al-Naffouri, T.Y.: An overview of signal processing techniques for terahertz communications. Proc. IEEE **109**(10), 1628–1665 (2021)
2. Wan, Z., Gao, Z., Gao, F., Di Renzo, M., Alouini, M.S.: Terahertz massive MIMO with holographic reconfigurable intelligent surfaces. IEEE Trans. Commun. **69**(7), 4732–4750 (2021)
3. Heath, R.W., Gonzalez-Prelcic, N., Rangan, S., Roh, W., Sayeed, A.M.: An overview of signal processing techniques for millimeter wave MIMO systems. IEEE J. Sel. Top. Sig. Process. **10**(3), 436–453 (2016)
4. El Ayach, O., Rajagopal, S., Abu-Surra, S., Pi, Z., Heath, R.W.: Spatially sparse precoding in millimeter wave MIMO systems. IEEE Trans. Wireless Commun. **13**(3), 1499–1513 (2014)
5. Zhang, H., Zhang, H., Dong, J., Leung, V.C., et al.: Energy efficient user clustering and hybrid precoding for terahertz MIMO-NOMA systems. In: ICC 2020–2020 IEEE International Conference on Communications (ICC), pp. 1–5. IEEE (2020)
6. Yuan, H., Yang, N., Yang, K., Han, C., An, J.: Hybrid beamforming for terahertz multi-carrier systems over frequency selective fading. IEEE Trans. Commun. **68**(10), 6186–6199 (2020)
7. Yan, L., Han, C., Yuan, J.: A dynamic array-of-subarrays architecture and hybrid precoding algorithms for terahertz wireless communications. IEEE J. Sel. Areas Commun. **38**(9), 2041–2056 (2020)
8. Rodriguez-Fernández, J., Gonzálcz-Prelcic, N.: Low-complexity multiuser hybrid precoding and combining for frequency selective millimeter wave systems. In: 2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), pp. 1–5. IEEE (2018)
9. Mao, J., Gao, Z., Wu, Y., Alouini, M.S.: Over-sampling codebook-based hybrid minimum sum-mean-square-error precoding for millimeter-wave 3D-MIMO. IEEE Wireless Commun. Lett. **7**(6), 938–941 (2018)
10. Zhang, X.D.: Matrix Analysis and Applications. Cambridge University Press, Cambridge (2017)
11. Sun, Y., et al.: Principal component analysis-based broadband hybrid precoding for millimeter-wave massive MIMO systems. IEEE Trans. Wirel. Commun. **19**(10), 6331–6346 (2020)