Physical Beam Sharing Multiuser Systems: Beamforming Designs and OTA Testings

Yan-Yin He, Hsiao-Chien Chen, and Shang-Ho Tsai, *Senior Member, IEEE*

***Abstract*— This paper proposes a novel physical beam sharing system to use few radio-frequency (RF) chains to support more users. The concept of beam sharing is to design a beam pattern to simultaneously benefit all users. In this proposed system, every user can receive a number of data streams equal to the number of RF chains at the base station (BS) and achieve full multiplexing gain. A special design example is that the BS can support multiple users simultaneously even with only one RF chain. As a result, the computational and hardware costs of the BS can be significantly reduced due to the decreased RF chains in deploying a dense communications network. We first derive beamforming designs for this proposed system in a MISO channel with one data stream. Then, the results are further extended to MIMO channels with multiple data streams. Simulation results show that the proposed schemes with the derived solutions outperform conventional systems in several interested parameter settings, and have advantages in terms of fairness and complexity compared with NOMA (non-orthogonal multiple access) systems. Moreover, we also implemented 28 GHz hardware platforms with an 8×8 antenna array to verify the feasibility and advantages of the proposed schemes as well as measure the resulting beam patterns. The over-the-air (OTA) testings validate the concept of the proposed physical beam sharing and show that it outperforms conventional physical beamforming schemes.**

***Index Terms*— 5G, beyond 5G, 6G, beam sharing, physical beamforming, hybrid precoding, MU-MIMO system, millimeter wave, over-the-air, OTA testing, beam pattern measurement.**

I. Introduction

PHYSICAL (analog) beamforming plays important roles to overcome severe path-loss in millimeter wave (mmWave) communications in the 5G and incoming 6G communications systems [1], [2]. In particular, hybrid beamforming/precoding systems that combine analog and digital beamforming have been extensively studied in recent years to reduce cost and power consumption. Unlike the traditional fully-digital MIMO precoding systems [3], in which many expensive RF chains are needed when large-scale antenna array is applied, the hybrid precoding technique can leverage between cost and data rate and has become a promising solution in 5G/6G communications.

The less costly hybrid precoders were proved to reach comparable performance as the fully-digital precoders [4]–[8]. Hybrid precoders for single-user (SU) systems were studied in [4], [9]–[13]. For hybrid precoding in multiuser (MU) systems, research has also been conducted in [5]–[8], [14]–[16]. More specifically, the authors in [5] proposed a heuristic MU hybrid precoding with one data stream that can attain the performance of fully-digital beamforming and then modified it to be more realistic, where only finite resolution phase shifters (PSs) can be used. The study in [6] considered MU-MIMO channels and aimed at optimizing the system with limited channel feedback. However, each user was limited to receive one data stream. The study in [7] designed a hybrid transceiver system based on the combination of the Tomlinson–Harashima precoding and the block-diagonal geometric mean decomposition, which can improve the performance of sum rate and error rate. The authors in [15] provided a weighted-based hybrid precoding scheme with architecture only utilizing switches and inverters. For reducing hardware cost and power consumption in massive MU-MIMO systems, the authors in [16] developed efficient algorithms with finite-alphabet precodings in the low-resolution digital-to-analog converters (DACs) system. Besides, for orthogonal frequency-division multiplexing based (OFDM-based) systems, the authors in [8] designed hybrid precodings for single-user multiple-input multiple-output (SU-MIMO) and multiuser multiple-input single-output (MU-MISO) systems, where heuristic hybrid precoding design was proposed, and the study in [14] proposed a wideband hybrid precoder for downlink multiuser massive MIMO system, where minimum total transmit power was considered.

From the discussion above, in current hybrid precoding MU-MIMO schemes, each user receives data from dedicated data streams, and the analog precoder tends to multiplex users in the spatial domain. In such schemes, there is a constraint that the number of RF chains at the BS should be larger or equal to the product of the user number and stream number [5], [6]. As a result, the hardware complexity of the RF chains becomes high as the number of users or the number of data streams grows. In addition, this constraint also limits the number of data streams of each user and thus it is harmful to spatial multiplexing gain.

From a view point of network deployment, 5G/6G communications demand huge amount of small cells to achieve

Manuscript received 24 May 2021; revised 23 November 2021 and 21 January 2022; accepted 23 January 2022. Date of publication 8 February 2022; date of current version 12 August 2022. This work was supported by the Ministry of Science and Technology (MOST), Taiwan, under Grant MOST 110-2221-E-A49-02 and Grant MOST 109-2221-E-009-106. The associate editor coordinating the review of this article and approving it for publication was H. Pishro-Nik. *(Corresponding author: Shang-Ho Tsai.)*

The authors are with the Department of Electrical Engineering, National Yang Ming Chiao Tung University (previous National Chiao Tung University), Hsinchu 30010, Taiwan (e-mail: shanghot@alumni.usc.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TWC.2022.3148180.

Digital Object Identifier 10.1109/TWC.2022.3148180

dense networks. The costs of the small cells directly relate to the operation expense. Using few RF chains to support multiple users can significantly reduce the hardware and computational costs of BSs and small cells, and thus it benefits the corresponding deployment.

Recently some research has been done to use few RF chains to support multiple users, where the number of RF chains can be smaller than the number of users. This scheme is called the constant envelope precoding (CEP) [17]–[19]. However, to support multiple data streams with only one RF chain, the CEP needs to switch the beamforming coefficients frequently, and the data rate is limited by the switching speed. In [20], the authors proposed a joint constellation design that encodes the data of multiple users into the same symbol, and optimized the precoder to minimize the weighted symbol error rates (SERs). However, these schemes considered only one data stream. Hence designing a system that supports multiple data streams as well as allows the number of RF chains to be smaller than the product of user number and stream number has not been widely researched yet.

To address the issue, we propose a new beam sharing system that uses few RF chains to support more users and multiple data streams. This system utilizes hybrid precoding. In this proposed system, different users use different subcarriers to achieve multiple access, which is one of the scenarios in current 5G standards, *i.e.,* orthogonal frequency division multiple access (OFDMA) scheme. The BS forms a sharing physical beam that benefits all users simultaneously. Each user can exactly receive the same number of data streams with the number of RF chains at the BS and attain full multiplexing gain. In an extreme case that the BS has only one RF chain, the proposed scheme can still support multiple users via the sharing physical beamforming.

It is worth pointing out that some of the hybrid beamforming OFDM systems share analog beam in different subcarriers, *e.g.*, SU-OFDM systems [12], [13] and MU-OFDM systems [21], [22]. However, there are several differences between these systems and the proposed beam sharing system explained as follows: In [12] and [13], the problems and results are for single-user systems and thus are different from the proposed multiuser scheme. The MU-OFDM systems considered in [21] and [22] attain multiple access in spatial domain like that in [6], which may be regarded as an application of MU-MIMO systems. On the other hand, the proposed scheme achieves multiple access in frequency domain, which can be treated as an application of OFDMA systems. When the proposed beam sharing is used, the proposed scheme can attain full multiplexing gain. Therefore, the proposed scheme is different from the MU-MIMO systems. The pros and cons of the proposed and the MU-MIMO hybrid beamforming systems are also provided below:

First, the proposed scheme attains multiple access in frequency domain; while the MU-MIMO scheme attains it in spatial domain. The advantages of the proposed scheme are multiuser interference (MUI)-free and attaining full multiplexing gain. However the disadvantage is that each user can utilize only a partial of the whole bandwidth. On the other hand, although each user in MU-MIMO can utilize the whole bandwidth, it does not achieve full multiplexing gain and there may be multiuser interference depending on how the precoder is designed.

Let us see a toy example. A system has 2 users, 2 subcarriers, and 2 RF chains. When the proposed beam sharing scheme is used, each user only occupies one subcarrier (1/2 bandwidth), but each of them can receive 2 data streams without MUI. On the other hand, for the MU-MIMO scheme, each user can utilize the whole bandwidth, but each of them can receive only one data stream and there may be MUI depending on how the precoder is designed [6], [21], [22].

The same beam is also shared by different users in NOMA systems, *e.g.,* see [23] and [24]. However there are still several differences between the proposed and the NOMA systems described as follows: In most of the literature on NOMA, including [23] and [24], each beam in the NOMA systems is for one specific direction. Hence, user pairing is important in NOMA, see *e.g.,* [23] and [24], to group the users with highly correlated channels into a cluster so that the users in this cluster can share this one-directional beam. To extend, multiple clusters can be grouped for users in different directions or with uncorrelated channels, and users in the same cluster again share a one-directional beam; users in different clusters are served by different beams in NOMA. On the other hand, the proposed beam sharing scheme does not need to pair users in advance, because it has the ability to use only one RF chain to form multiple beams simultaneously towards multiple directions (this is also shown later in the OTA experiments). Hence in the proposed beam sharing scheme, the channels of the users can be uncorrelated and the locations of the users can be far apart; at the same time, the designed sharing beam can cover the users simultaneously. Moreover, the proposed scheme can achieve full multiplexing gain while this is not the case for the NOMA.

In the proposed system, we design the analog precoder to maximize the achievable sum rate. First, we consider MISO channels in both high and low signal-to-noise ratio (SNR) regions. Solutions are obtained for a two-user channel and then extended to multiple-user channels. Inspired by the derived results for the MU-MISO channels, we further investigated MIMO channels and obtained corresponding results. We found that in the MIMO channels, though the required derivation is somewhat more complicated than that in the MISO channels, solutions can still be obtained via similar derivations.

Furthermore, simulation results are provided, where the proposed schemes can outperform conventional MU-MIMO and OFDMA schemes in several interested parameter settings, and prove the correctness of the derived results as well. Moreover, the proposed schemes are also compared to NOMA systems and have advantages in terms of fairness and complexity.

Finally, to verify the feasibility and the advantages of the proposed beam sharing scheme in practical wireless channels, we implemented the hardware platforms and performed OTA experiments. These hardware platforms are at 28 GHz mmWave frequency band, which consist of an 8×8 uniform planar array (UPA). The OTA testings show that the proposed beam sharing scheme indeed works well in practical wireless environments as well as outperforms conventional

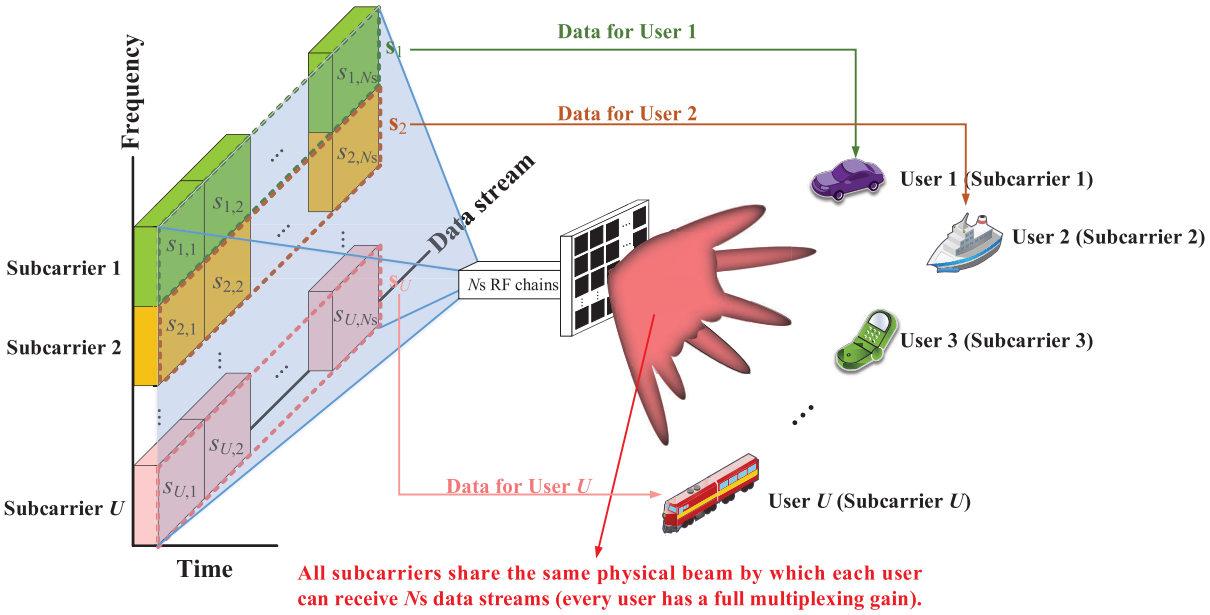


Fig. 1. A system model of the proposed system.

beamforming schemes. We also OTA measured the beam pattern of the proposed sharing beam and showed that the measured beam pattern is close to the expected theoretical one. These nice results in both computer and hardware experiments reveal good potential of using the proposed beam sharing systems and the corresponding precoding solutions in the next generation communications.

The remaining parts are organized as follows: In Sec. II, we present the system model and formulate the problem. In Sec. III, solutions for MU-MISO systems are derived in both high and low SNR regimes. Sec. IV further generalizes the results to MU-MIMO systems to increase multiplexing gain. Sec. V provides the complexity analysis for all the proposed solutions. In Sec. VI, simulation results are provided. The 28 GHz hardware platforms with an 8×8 UPA are implemented and OTA experiments are conducted in Sec. VII. Finally, conclusions and future works are summarized in Sec. VIII.

The following notations and operations are used in this paper: Boldface upper-case letters denote matrices $\mathbf{A}$. Boldface lower-case letters denote vectors $\mathbf{a}$. a^* is the complex conjugate of scalar a, $\Re\{a\}$ is the real part value of a, and $|a|$ is the magnitude of a. $\|\mathbf{A}\|$ is the 2-norm of $\mathbf{A}$ and $\|\mathbf{A}\|_{\mathrm{F}}$ is the Frobenius norm of $\mathbf{A}$. $\mathbf{A}^T$, $\mathbf{A}^H$, and $\mathbf{A}^{\perp}$ are denoted as the transpose, Hermitian, and orthogonal (null) matrices of $\mathbf{A}$, respectively. $\mathbf{A} \perp \mathbf{B}$ means that $\mathbf{A}$ is orthogonal to $\mathbf{B}$. $\mathrm{tr}\,(\mathbf{A})$ is the trace of $\mathbf{A}$. $\det(\mathbf{A})$ is the determinant of $\mathbf{A}$. $\mathbb{C}^{M\times N}$ represents the set of $(M \times N)$-dimensional complex-valued matrices and $\mathbb{E}\{\cdot\}$ is regarded as the expectation.

II. System and Problem Formulation

A. System Model

A system model of the proposed system is shown in Fig. 1. Individual users use different subcarriers and the BS forms a sharing physical beam (analog precoding) to serve all U users simultaneously. The BS has a number N_{RF} of RF chains. Because individual users use different subcarriers, which is one of the resource allocation scenarios in the current 5G standards, *i.e.,* OFDMA mode, there is no MUI and each user can receive a number N_{S} of data streams, where $N_{\mathrm{S}} \leq N_{\mathrm{RF}}$. That is, all the U users can achieve full multiplexing gain simultaneously.

A special advantage of the proposed system is that there is no constraint of $N_{\mathrm{RF}} \geq U \cdot N_{\mathrm{S}}$, where N_{S} is the number of received data streams for each user. Hence even the BS has only one RF chain, it can still support multiple users simultaneously via the sharing physical beam. This advantage becomes more pronounced for dense small cell networks, because reducing RF chains can significantly decrease hardware and computational costs, and this benefits the corresponding deployment. Let the analog precoder be $\mathbf{F}$. Note that the gain and phase of the antenna array are new adjustable architectures (see [25, p. 1]). Thus $\mathbf{F} \in \mathbb{C}^{N_{\mathrm{T}}\times N_{\mathrm{RF}}}$. The digital precoder at the uth subcarrier is represented by $\mathbf{B}_u \in \mathbb{C}^{N_{\mathrm{RF}}\times N_{\mathrm{S}}}$. The data vector for the uth user (at the uth subcarrier) is denoted by $\mathbf{s}_u \in \mathbb{C}^{N_{\mathrm{S}}\times 1}$.

The transmitted data vector $\mathbf{x}_u \in \mathbb{C}^{N_{\mathrm{T}}\times 1}$ from the BS to the uth user (at the uth subcarrier) is given by $\mathbf{x}_u = \mathbf{F}\mathbf{B}_u\mathbf{s}_u$. Because all subcarriers share the same physical beam, only a sharing analog precoder $\mathbf{F}$ is designed to benefit all users. Each user at individual subcarrier can use different digital precoder $\mathbf{B}_u$ to refine the performance. Using the new architecture of antenna array in the proposed system [25], the constant-modulus constraints in the analog precoder are not necessarily needed, where such gain and phase adjustable architectures have been announced by several vendors; for example, observed from the design document of the Anokiwave Inc. [25, p. 1], there are 5 bits used to control the phase and 5 bits used to control the gain. Using the new

architecture of antenna array, the total power constraint is enforced by normalizing the total hybrid precoder such that $\|\mathbf{F}\mathbf{B}_u\|_{\mathrm{F}}^2 = P_{\mathrm{T}}$.

For the channel model, the narrowband and broadband geometric channels with clustering are applied, see [6, (4)] and [13, (3) and (4)] respectively. The channel of the uth user is $\mathbf{H}_u \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{T}}}$ and the noise $\mathbf{n}_u \in \mathbb{C}^{N_{\mathrm{R}} \times 1}$ is an i.i.d. additive white Gaussian noise with $\mathbf{n}_u \sim \mathcal{CN}(0, \sigma_{\mathrm{n}}^2\mathbf{I})$. The received signal $\mathbf{r}_u \in \mathbb{C}^{N_{\mathrm{R}} \times 1}$ of the uth user can be expressed by $\mathbf{r}_u = \mathbf{H}_u\mathbf{F}\mathbf{B}_u\mathbf{s}_u + \mathbf{n}_u$. The RF postcoder $\mathbf{W}_u \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{RF}}}$ and the baseband postcoder $\mathbf{Z}_u \in \mathbb{C}^{N_{\mathrm{RF}} \times N_{\mathrm{S}}}$ are used to process the received signal. The signal after postcoding is then given by

$$\mathbf{y}_u = \mathbf{Z}_u^H\mathbf{W}_u^H\mathbf{H}_u\mathbf{F}\mathbf{B}_u\mathbf{s}_u + \mathbf{Z}_u^H\mathbf{W}_u^H\mathbf{n}_u \tag{1}$$

B. Problem Formulation

We consider to maximize the sum rate for the proposed system. The data rate of the uth user can be expressed by

$$C_u = \log_2 \det\left(\mathbf{I} + \sigma_{\mathrm{s}}^2\mathbf{N}_u^{-1}\mathbf{Z}_u^H\mathbf{W}_u^H\mathbf{H}_u\mathbf{F}\mathbf{B}_u \times \mathbf{B}_u^H\mathbf{F}^H\mathbf{H}_u^H\mathbf{W}_u\mathbf{Z}_u\right) \tag{2}$$

where $\mathbf{N}_u = \sigma_{\mathrm{n}}^2\mathbf{Z}_u^H\mathbf{W}_u^H\mathbf{W}_u\mathbf{Z}_u$ is the noise covariance matrix after postcoding, and the power of the data streams is $\mathbb{E}[\mathbf{s}_u\mathbf{s}_u^H] = \sigma_{\mathrm{s}}^2\mathbf{I}$.

There are four matrix variables $(\mathbf{F}, \mathbf{B}_u, \mathbf{W}_u, \mathbf{Z}_u)$ involved in the original problem, and we seek to design the hybrid precoder $\mathbf{F}\mathbf{B}_u$ that maximizes the sum rate. Research that dealt with similar problems usually designed the precoder and postcoder separately to simplify the optimization problem, and the unconstrained postcoder $\mathbf{W}_u\mathbf{Z}_u$ can be obtained by achieving the minimum mean square error (MMSE) [4], [5], [9] given by

$$\mathbf{W}_u\mathbf{Z}_u = ((\mathbf{B}_u^H\mathbf{F}^H\mathbf{H}_u^H\mathbf{H}_u\mathbf{F}\mathbf{B}_u + \frac{\sigma_{\mathrm{n}}^2}{\sigma_{\mathrm{s}}^2}\mathbf{I})^{-1}\mathbf{B}_u^H\mathbf{F}^H\mathbf{H}_u^H)^H \tag{3}$$

Hence we focus on the design of the hybrid precoder, $\mathbf{F}$ and $\mathbf{B}_u$. The problem is written as

$$\begin{aligned} &\arg\max_{\mathbf{B}_u,\mathbf{F}} \sum_{u=1}^{U} \log_2 \det\left(\mathbf{I} + \frac{\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{n}}^2}\mathbf{H}_u\mathbf{F}\mathbf{B}_u\mathbf{B}_u^H\mathbf{F}^H\mathbf{H}_u^H\right) \\ &\text{s.t. } \|\mathbf{F}\mathbf{B}_u\|_{\mathrm{F}}^2 = P_{\mathrm{T}} \end{aligned} \tag{4}$$

It is worth pointing out that according to (4), there is only a sharing analog precoder $\mathbf{F}$ in the equation; however, it should be designed to make the sum rate among all U users maximum, which implies that the beam, generated by the analog precoder $\mathbf{F}$, is shared by all subcarriers to benefit all users simultaneously.

In the following discussion, we assume that the number of scattering paths is larger or equal to the number of RF chains so that supporting multiple data streams is possible. Also all RF chains are utilized, *i.e.,* $N_{\mathrm{S}} = N_{\mathrm{RF}}$, to attain the best achievable sum rate.

III. Proposed MISO Precoding Schemes

We start with the simplest MISO channels, where every user is equipped with single antenna and receives one data stream. The digital precoder $b_u \in \mathbb{C}^{1\times 1}$ then becomes a complex scalar. Without loss of generality, b_u is set to 1. Hence only the analog precoder $\mathbf{f} \in \mathbb{C}^{N_{\mathrm{T}}\times 1}$ is designed. The channel between the BS and the uth user is $\mathbf{h}_u^H \in \mathbb{C}^{1\times N_{\mathrm{T}}}$. The problem for the MISO channels is reformulated as follows:

$$\begin{aligned} &\arg\max_{\mathbf{f}} \sum_{u=1}^{U} \log_2\left(1 + \frac{\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{n}}^2} \mid \mathbf{h}_u^H\mathbf{f} \mid^2\right) \\ &\text{s.t. } \|\mathbf{f}\|^2 = P_{\mathrm{T}} \end{aligned} \tag{5}$$

From (5), a sharing analog precoder $\mathbf{f}$ should be designed to maximize the sum rate for U users, which is difficult to be solved. To simplify the problem, we first let $\gamma = \frac{\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{n}}^2}$ be the SNR and wrap $\gamma \mid \mathbf{h}_u^H\mathbf{f} \mid^2$ as one variable c_u. The objective function in (5) can be reformulated by $\sum_{u=1}^{U}\log_2(1+c_u)$. When the value of SNR is small, $\log_2(1+c_u) \approx c_u$. On the other hand, when the value of SNR is large, $\log_2(1+c_u) \approx \log_2(c_u)$. According to the different approximations mentioned above, the objective function can then be separated into high and low SNR regimes, and discussed respectively in the following subsections.

A. High SNR Regime in MISO Channels

From (5), the problem in high SNR regime is rewritten as follows:

$$\begin{aligned} &\arg\max_{\mathbf{f}} \sum_{u=1}^{U} \log_2\left(\gamma \mid \mathbf{h}_u^H\mathbf{f} \mid^2\right) \\ &\text{s.t. } \|\mathbf{f}\|^2 = P_{\mathrm{T}} \end{aligned} \tag{6}$$

Performing singular value decomposition (SVD) to every user channel, $\mathbf{h}_u^H = u_u d_u \mathbf{v}_u^H$, the objective function becomes $\sum_{u=1}^{U}\log_2\left(\mathbf{f}^H\mathbf{v}_u\mathbf{v}_u^H\mathbf{f}\right)$. Define $\mathbf{V} \in \mathbb{C}^{N_{\mathrm{T}}\times U}$ as the collection of the right singular vectors $\mathbf{v}_u$: $\mathbf{V} = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_U\end{bmatrix}$. Applying QR decomposition to $\mathbf{V}$ yields

$$\mathbf{V} = \mathbf{Q}\mathbf{R} \tag{7}$$

where $\mathbf{Q} \in \mathbb{C}^{N_{\mathrm{T}}\times U}$ and $\mathbf{R} \in \mathbb{C}^{U\times U}$. One can find the orthonormal basis $\mathbf{q}_u \in \mathbb{C}^{N_{\mathrm{T}}\times 1}$ such as $\mathbf{Q} = \begin{bmatrix}\mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_U\end{bmatrix}$. Let r_{iu} be the ith row and uth column in matrix $\mathbf{R}$, and then $\mathbf{v}_u = \sum_{i=1}^{u} r_{iu}\mathbf{q}_i$, for $u = 1 \ldots U$. Because the column space of $\mathbf{Q}$ and its left null space $\mathbf{Q}^{\perp} \in \mathbf{C}^{N_{\mathrm{T}}\times(N_{\mathrm{T}}-U)}$ form a basis for $\mathbb{C}^{N_{\mathrm{T}}\times 1}$, any solution for precoder $\mathbf{f} \in \mathbb{C}^{N_{\mathrm{T}}\times 1}$ can be described by

$$\mathbf{f} = \mathbf{Q}\mathbf{g} + \mathbf{Q}^{\perp}\mathbf{g}^{\perp} = \sum_{u=1}^{U} g_u\mathbf{q}_u + \mathbf{Q}^{\perp}\mathbf{g}^{\perp} \tag{8}$$

where $\mathbf{g} = \begin{bmatrix}g_1 & g_2 & \cdots & g_U\end{bmatrix}^T$ and $\mathbf{g}^{\perp} \in \mathbb{C}^{(N_{\mathrm{T}}-U)\times 1}$. From the discussion above, (6) can then be reformulated as

$$\begin{aligned} &\arg\max_{g_i,\mathbf{g}^{\perp}} \sum_{u=1}^{U} \log_2\left(\mid \sum_{i=1}^{u} g_i^* r_{iu} \mid^2\right) \\ &\text{s.t. } \sum_{u=1}^{U} \mid g_u \mid^2 + \|\mathbf{g}^{\perp}\|^2 = P_{\mathrm{T}} \end{aligned} \tag{9}$$

Now the variables are transformed into the weight variables g_u and the weight vector $\mathbf{g}^{\perp}$. We discuss the solution from the simplest case by first setting $U = 2$.

1) A Two-User Case: The problem in (9) can be rewritten as follows:

$$\begin{aligned} &\arg\max_{g_1,g_2,\mathbf{g}^{\perp}} \log_2\left(\mid g_1 \mid^2 \mid g_1^* r_{12} + g_2^* r_{22} \mid^2\right) \\ &\text{s.t. } \mid g_1 \mid^2 + \mid g_2 \mid^2 + \|\mathbf{g}^{\perp}\|^2 = P_{\mathrm{T}} \end{aligned} \tag{10}$$

The following Lemma finds the optimal phase settings for g_1 and g_2, and tunes the phase of the objective function to the real axis.

Lemma 1: The optimal phase settings for g_1 and g_2 in problem (10), denoted by ϕ_1 and ϕ_2 respectively, for the two-user case in MISO channels should satisfy the following condition:

$$\phi_1 - \phi_2 = -\left(\theta_{22} - \theta_{12}\right), \quad \textit{for } \forall \mid g_1 \mid, \ \mid g_2 \mid > 0 \tag{11}$$

where θ_{12} is the phase of r_{12} and θ_{22} is the phase of r_{22}.

Proof: Let $g_i = |g_i|e^{j\phi_i}$, $r_{12} = |r_{12}|e^{j\theta_{12}}$, and $r_{22} = |r_{22}|e^{j\theta_{22}}$, where ϕ_i and $\theta_{iu} \in [0, 2\pi]$. The objective function inside the logarithm in (10) can be expressed as

$$\mid g_1 \mid^2 \left(\mid r_{12} \mid^2 \mid g_1 \mid^2 + \mid r_{22} \mid^2 \mid g_2 \mid^2 + 2\mathbb{R}\{g_1 g_2^* r_{12}^* r_{22}\}\right) \tag{12}$$

$\mathbb{R}\{g_1 g_2^* r_{12}^* r_{22}\}$ can be further extended as $|g_1||g_2||r_{12}||r_{22}| \cos\left(\phi_1 - \phi_2 + \theta_{22} - \theta_{12}\right)$; thus, the objective function can be maximized by setting $\cos\left(\phi_1 - \phi_2 + \theta_{22} - \theta_{12}\right) = 1$. ∎

After applying Lemma 1, the problem can be reformulated as

$$\begin{aligned} &\arg\max_{|g_1|,|g_2|,\|\mathbf{g}^{\perp}\|} \log_2\left(\mid r_{12} \mid\mid g_1 \mid^2 + \mid r_{22} \mid\mid g_1 \mid\mid g_2 \mid\right)^2 \\ &\text{s.t. } \mid g_1 \mid^2 + \mid g_2 \mid^2 + \|\mathbf{g}^{\perp}\|^2 = P_{\mathrm{T}} \end{aligned} \tag{13}$$

Since all of the parameters in front of $\mid g_u \mid$ are real and positive, the power should not be allocated to $\|\mathbf{g}^{\perp}\|$ but should be fully dedicated to $\mid g_1 \mid$ and $\mid g_2 \mid$ instead to maximize the sum rate. Thus, the power constraint in problem (13) therefore becomes $\mid g_1 \mid^2 + \mid g_2 \mid^2 = P_{\mathrm{T}}$. Thus, from (8) and setting $\mathbf{g}^{\perp} = 0$, the solution is then introduced in the following proposition.

Proposition 1: In the two-user MISO channels, the precoder design is with the following form:

$$\mathbf{f} = \mid g_1 \mid e^{j\phi_1}\mathbf{q}_1 + \mid g_2 \mid e^{j\phi_2}\mathbf{q}_2 \tag{14}$$

where ϕ_1 and ϕ_2 satisfy the relationship described in Lemma 1; while the gains $\mid g_1 \mid$ and $\mid g_2 \mid$ satisfy $\mid g_1 \mid = \sqrt{P_{\mathrm{T}}}\cos(\theta)$ and $\mid g_2 \mid = \sqrt{P_{\mathrm{T}}}\sin(\theta)$. The parameter θ can be found by grid search in the first quadrant of the following equation.

$$\arg\max_{\theta} \mid r_{12} \mid \cos^2(\theta) + \mid r_{22} \mid \cos(\theta)\sin(\theta), \theta \in [0, \frac{\pi}{2}] \tag{15}$$

Proof: See Appendix A. ∎

2) General Multiuser Cases: For $U > 2$, one can show that similar argument and solution for Proposition 1 do not apply. Fortunately, in the large-scale antenna array architectures, one can have the orthogonality approximation that $\mathbf{v}_i^H \mathbf{v}_j \approx 0$, *e.g.,* see [6] and [26]. As N_{T} increases, the validation of this approximation becomes more pronounced, especially for mmWave systems because the beam width becomes narrow. Therefore, the column vector $\mathbf{v}_u$ from each user in $\mathbf{V} \in \mathbb{C}^{N_{\mathrm{T}} \times U}$ may be approximated as orthonormal basis that spans the column space of $\mathbf{V}$, and $\mathbf{R}$ is thus approximated as a diagonal matrix. In this case, $\mathbf{f}$ in (8) can then be reformulated as

$$\mathbf{f} = \mathbf{V}\mathbf{g} + \mathbf{V}^{\perp}\mathbf{g}^{\perp} = \sum_{u=1}^{U} g_u \mathbf{v}_u + \mathbf{V}^{\perp}\mathbf{g}^{\perp} \tag{16}$$

where $\mathbf{V}^{\perp} \in \mathbb{C}^{N_{\mathrm{T}} \times (N_{\mathrm{T}} - U)}$ represents the left null space of $\mathbf{V} \in \mathbb{C}^{N_{\mathrm{T}} \times U}$. From (7), when every column in $\mathbf{V}$ is approximately orthogonal to one another, it implies that $\mathbf{R}$ is an identity matrix at the same time. Therefore, according to (9), the parameters inside the logarithm for the uth user in (6) can be represented as $\mid \sum_{i=1}^{u} g_i^* r_{iu} \mid^2 = \mid g_u \mid^2$, where $r_{iu} \approx 0, \forall i \neq u$ and $r_{uu} \approx 1$ under the orthogonality approximation.

When N_{T} is sufficiently large, $\mathbf{v}_i^H \mathbf{v}_j \approx 0$, the problem in (6) is rewritten as

$$\begin{aligned} &\arg\max_{g_u,\mathbf{g}^{\perp}} \sum_{u=1}^{U} \log_2\left(\mid g_u \mid^2\right) \\ &\text{s.t. } \sum_{u=1}^{U} \mid g_u \mid^2 + \|\mathbf{g}^{\perp}\|^2 = P_{\mathrm{T}} \end{aligned} \tag{17}$$

To maximize the sum rate, the power should not be allocated to variables that do not appear in the objective function by setting $\|\mathbf{g}^{\perp}\|^2 = 0$. Thus, the power constraint in (17) is set as $\sum_{u=1}^{U} \mid g_u \mid^2 = P_{\mathrm{T}}$. Finally, a solution for MISO channels is derived in the following proposition.

Proposition 2: When N_{T} is sufficiently large, $\mathbf{v}_i^H \mathbf{v}_j \approx 0$, and then (6) can be approximated to (17), $\mid g_u \mid = \sqrt{\frac{P_{\mathrm{T}}}{U}}, \forall u \in 1, \ldots, U$ is an optimal solution to the problem in (17) and is an asymptotically optimal solution to the original problem in (6) as N_{T} approaches ∞.

Proof: We apply the Arithmetic-Geometric Mean Inequality to complete the proof. From this Inequality, we can find an upper bound by applying the constraint for the objective function, such as $\sqrt[U]{\prod_{u=1}^{U} \mid g_u \mid^2} \leq \frac{\sum_{u=1}^{U} |g_u|^2}{U} = \frac{P_{\mathrm{T}}}{U}$. Note that $\mid g_u \mid$ are non-negative numbers. To reach the upper bound and let the equality hold, we let $\mid g_1 \mid = \mid g_2 \mid = \cdots = \mid g_U \mid = \sqrt{\frac{P_{\mathrm{T}}}{U}}$, and the maximum sum rate can be reached. Proposition 2 is then proved. ∎

After reformulating $\mathbf{f}$ as (16) and setting $\mathbf{g}^{\perp} = 0$, based on Proposition 2, we obtain the optimal solution of $\mid g_u \mid$, and the analog precoder $\mathbf{f}$ can be therefore expressed as $\mathbf{f} = \sqrt{\frac{P_{\mathrm{T}}}{U}} \sum_{u=1}^{U} \mathbf{v}_u$.

B. Low SNR Regime in MISO Channels

Since γ and $\mathbf{f}$ are independent of u, they can be pulled out of the summation, and then the optimization problem in low SNR region can be given as follows:

$$\begin{aligned} &\arg\max_{\mathbf{f}} \ \gamma \mathbf{f}^H \left(\sum_{u=1}^{U} \mathbf{h}_u \mathbf{h}_u^H \right) \mathbf{f} \\ &\text{s.t. } \|\mathbf{f}\|^2 = P_{\mathrm{T}} \end{aligned} \tag{18}$$

The solution to (18) is provided in the following proposition.

Proposition 3: Let $\mathbf{H} = \sum_{u=1}^{U} \mathbf{h}_u \mathbf{h}_u^H$. *The analog precoder* $\mathbf{f}$ *in the low SNR region should be the scaled eigenvector corresponding to the largest eigenvalue of the matrix* $\mathbf{H}$ *which satisfies the power constraint.*

Proof: This is a typical quadratic problem and the proof can be found, *e.g.,* in [27]. ∎

IV. Proposed MIMO Precoding Schemes

The problem for the MIMO channels is given by

$$\begin{aligned} &\arg\max_{\mathbf{B}_u, \mathbf{F}} \sum_{u=1}^{U} \log_2 \det \left(\mathbf{I} + \gamma \mathbf{H}_u \mathbf{F} \mathbf{B}_u \mathbf{B}_u^H \mathbf{F}^H \mathbf{H}_u^H \right) \\ &\text{s.t. } \ \|\mathbf{F}\mathbf{B}_u\|_{\mathrm{F}}^2 = P_{\mathrm{T}} \end{aligned} \tag{19}$$

To design the digital precoder $\mathbf{B}_u$, we first assumed that the analog precoder $\mathbf{F}$ is available. Wrapping $\mathbf{H}_u\mathbf{F}$ into the new variable $\tilde{\mathbf{H}}_u \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{S}}}$ and performing SVD to $\tilde{\mathbf{H}}_u$, $\tilde{\mathbf{H}}_u = \tilde{\mathbf{U}}_u \tilde{\mathbf{D}}_u \tilde{\mathbf{V}}_u^H$, the objective function becomes

$$\sum_{u=1}^{U} \log_2 \det \left(\mathbf{I} + \gamma \tilde{\mathbf{D}}_u \tilde{\mathbf{V}}_u^H \mathbf{B}_u \mathbf{B}_u^H \tilde{\mathbf{V}}_u \tilde{\mathbf{D}}_u^H \right). \tag{20}$$

Since the dimension of $\mathbf{B}_u$ is exactly the same as the dimension of $\tilde{\mathbf{V}}_u \in \mathbb{C}^{N_{\mathrm{S}} \times N_{\mathrm{S}}}$, the optimal solution is obtained by $\mathbf{B}_u = \tilde{\mathbf{V}}_u$ following the same design as the fully-digital precoder, *e.g.,* see [4] and [5]. In addition, $\mathbf{B}_u$ is a unitary matrix and it satisfies $\mathbf{B}_u \mathbf{B}_u^H = \mathbf{B}_u^H \mathbf{B}_u = \mathbf{I}_{N_{\mathrm{S}}}$. After designing the digital precoder $\mathbf{B}_u$, one can focus on the design of the analog precoder $\mathbf{F}$. The problem is then rewritten as follows:

$$\begin{aligned} &\arg\max_{\mathbf{F}} \sum_{u=1}^{U} \ \log_2 \det \left(\mathbf{I} + \gamma \mathbf{F}^H \mathbf{H}_u^H \mathbf{H}_u \mathbf{F} \right) \\ &\text{s.t. } \|\mathbf{F}\|_{\mathrm{F}}^2 = P_{\mathrm{T}} \end{aligned} \tag{21}$$

where the sharing analog precoder $\mathbf{F}$ again should be designed to optimally match U different user channels for maximizing the sum rate, which is difficult to be solved. Thus following the same argument in the MISO channels, the objective function in (21) can also be separated into high SNR and low SNR regions, and discussed respectively in the following subsections.

A. High SNR Regime in MIMO Channels

From (21), the identity matrix $\mathbf{I}$ can be ignored in the high SNR regime and the objective function for the MIMO channels is as follows:

$$\sum_{u=1}^{U} \log_2 \det \left(\gamma \mathbf{F}^H \mathbf{H}_u^H \mathbf{H}_u \mathbf{F} \right) \tag{22}$$

where the desired sharing analog precoder $\mathbf{F} \in \mathbb{C}^{N_{\mathrm{T}} \times N_{\mathrm{S}}}$ is a matrix now. We follow similar derivation in subsection III-A by taking SVD to each user channel $\mathbf{H}_u$, such that $\mathbf{H}_u = \mathbf{U}_u \mathbf{D}_u \mathbf{V}_u^H$. The rank of channel $\mathbf{H}_u$ is denoted as $K = \mathrm{rank}(\mathbf{H}_{\mathrm{u}})$. Using the property that $\mathbf{U}_u^H \mathbf{U}_u = \mathbf{I}$ and taking non-zero singular values and their corresponding singular vectors, such that $\mathbf{D}_u \in \mathbb{C}^{K \times K}$ and $\mathbf{V}_u \in \mathbb{C}^{N_{\mathrm{T}} \times K}$, the objective function in (22) can be reformulated as

$$\sum_{u=1}^{U} \log_2 \det \left(\gamma \mathbf{F}^H \mathbf{V}_u \mathbf{\Sigma}_u \mathbf{V}_u^H \mathbf{F} \right) \tag{23}$$

where $\mathbf{\Sigma}_u = \mathbf{D}_u^H \mathbf{D}_u$ and $\mathbf{\Sigma}_u \in \mathbb{C}^{K \times K}$.

Following similar argument as in the MISO channels and defining $\mathbf{V} \in \mathbb{C}^{N_{\mathrm{T}} \times KU}$ as the collection of the right singular matrices $\mathbf{V}_u$: $\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_2 & \cdots & \mathbf{V}_U \end{bmatrix}$, $\mathbf{V}$ also yields similar relationship described in (7), but the dimensions change as $\mathbf{Q} \in \mathbb{C}^{N_{\mathrm{T}} \times KU}$ and $\mathbf{R} \in \mathbb{C}^{KU \times KU}$. One can find the collection of the orthonormal basis $\mathbf{Q}_u \in \mathbb{C}^{N_{\mathrm{T}} \times K}$ such as $\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 & \cdots & \mathbf{Q}_U \end{bmatrix}$ and express $\mathbf{V}_u$ using the orthonormal bases in $\mathbf{Q}_i$ and the corresponding weights in $\mathbf{R}_{iu}$ as $\mathbf{V}_u = \sum_{i=1}^{u} \mathbf{Q}_i \mathbf{R}_{iu}$. The analog precoder $\mathbf{F} \in \mathbb{C}^{N_{\mathrm{T}} \times N_{\mathrm{S}}}$ can be seen column-wise such as $\mathbf{F} = \begin{bmatrix} \mathbf{f}_1 & \mathbf{f}_2 & \cdots & \mathbf{f}_{N_{\mathrm{S}}} \end{bmatrix}$, and the column space of $\mathbf{Q}$ and its left null space $\mathbf{Q}^{\perp} \in \mathbb{C}^{N_{\mathrm{T}} \times (N_{\mathrm{T}} - KU)}$ form a basis for $\mathbb{C}^{N_{\mathrm{T}} \times 1}$; thus, any solution for precoder $\mathbf{F} \in \mathbb{C}^{N_{\mathrm{T}} \times N_{\mathrm{S}}}$ can then be described by

$$\mathbf{F} = \mathbf{Q}\mathbf{G} + \mathbf{Q}^{\perp}\mathbf{G}^{\perp} = \sum_{u=1}^{U} \mathbf{Q}_u \mathbf{G}_u + \mathbf{Q}^{\perp}\mathbf{G}^{\perp} \tag{24}$$

where $\mathbf{G} \in \mathbb{C}^{KU \times N_{\mathrm{S}}}$ such as $\mathbf{G} = \begin{bmatrix} \mathbf{G}_1 & \mathbf{G}_2 & \cdots & \mathbf{G}_U \end{bmatrix}^T$, and $\mathbf{G}^{\perp} \in \mathbb{C}^{(N_{\mathrm{T}} - KU) \times N_{\mathrm{S}}}$. From (23) and the discussion above, the problem can be reformulated as

$$\begin{aligned} &\arg\max_{\mathbf{G}_i, \mathbf{G}^{\perp}} \sum_{u=1}^{U} \ \log_2 \det \left(\gamma \left(\sum_{i=1}^{u} \mathbf{G}_i^H \mathbf{R}_{iu} \right) \mathbf{\Sigma}_u \left(\sum_{i=1}^{u} \mathbf{R}_{iu}^H \mathbf{G}_i \right) \right) \\ &\text{s.t. } \sum_{u=1}^{U} \|\mathbf{G}_u\|_{\mathrm{F}}^2 + \|\mathbf{G}^{\perp}\|_{\mathrm{F}}^2 = P_{\mathrm{T}} \end{aligned} \tag{25}$$

The result is similar to (9) in the MISO channels but the desired weight variables $\mathbf{G}_u$ and $\mathbf{G}^{\perp}$ are both matrices in the MIMO channels. Let us discuss from the simplest case by setting $U = 2$ and $L = 1$.

1) A Two-User Case in LOS MIMO Channels: By setting the number of scattering paths $L = 1$, the rank of the channels therefore becomes $K = 1$, and the number of data streams is also limited to $N_{\mathrm{S}} = 1$. The problem for the two-user in LOS environments can be written as

$$\begin{aligned} &\arg\max_{g_i, \mathbf{g}^{\perp}} \sum_{u=1}^{U} \log_2 \det \left(\gamma \left(\sum_{i=1}^{u} g_i^* r_{iu} \right) \mid d_u \mid^2 \left(\sum_{i=1}^{u} r_{iu}^* g_i \right) \right) \\ &\text{s.t. } \sum_{u=1}^{U} \mid g_u \mid^2 + \|\mathbf{g}^{\perp}\|^2 = P_{\mathrm{T}} \end{aligned} \tag{26}$$

where $\mid d_u \mid^2$ is the only one non-zero element in $\mathbf{\Sigma}_u$. The following proposition is therefore derived based on the result in the subsection III-A.1.

Proposition 4: In MIMO channels with the number of scattering paths $L = 1$ and the number of users $U = 2$, the precoder solution is exactly the same as the two-user case in the MISO channels, i.e., Proposition 1.

Proof: By extending (26) and removing uncorrelated constants γ and $\mid d_u \mid^2$, the problem is exactly the same optimization problem defined in (9). ∎

2) General Multiuser Cases: For $U > 2$, $L > 1$ and $N_{\mathrm{S}} > 1$, one can show that similar argument and solution for Proposition 4 do not satisfy. Assume that N_{T} is sufficiently large, and the orthogonality approximation that $\mathbf{V}_i \perp \mathbf{V}_j \approx \mathbf{0}$ is adopted here as well. Then, the column vectors in $\mathbf{V}_u$ from each user in $\mathbf{V} \in \mathbb{C}^{N_{\mathrm{T}} \times KU}$ may be approximated as orthonormal bases that span the column space of $\mathbf{V}$, and $\mathbf{R} \in \mathbb{C}^{KU \times KU}$ is approximated as a diagonal matrix. In this case, $\mathbf{F}$ in (24) can then be expressed as

$$\mathbf{F} = \mathbf{V}\mathbf{G} + \mathbf{V}^{\perp}\mathbf{G}^{\perp} = \sum_{u=1}^{U} \mathbf{V}_u \mathbf{G}_u + \mathbf{V}^{\perp}\mathbf{G}^{\perp} \tag{27}$$

where $\mathbf{V}^{\perp} \in \mathbb{C}^{N_{\mathrm{T}} \times (N_{\mathrm{T}} - KU)}$ represents the left null space of $\mathbf{V} \in \mathbb{C}^{N_{\mathrm{T}} \times KU}$. When N_{T} is sufficiently large, $\mathbf{V}_i^H \mathbf{V}_j \approx \mathbf{0}$. Substituting (27) into the objective function in (23), the problem can then be reformulated as

$$\begin{aligned} &\arg\max_{\mathbf{G}_u, \mathbf{G}^{\perp}} \sum_{u=1}^{U} \log_2 \det\left(\gamma \mathbf{G}_u^H \mathbf{\Sigma}_u \mathbf{G}_u\right) \\ &\text{s.t.} \sum_{u=1}^{U} \|\mathbf{G}_u\|_{\mathrm{F}}^2 + \|\mathbf{G}^{\perp}\|_{\mathrm{F}}^2 = P_{\mathrm{T}} \end{aligned} \tag{28}$$

Since γ and $\mathbf{\Sigma}_u$ in the objective function of the problem in (28) are all positive constants, to maximize the sum rate, the solution to the weights for the left null space $\mathbf{G}^{\perp}$ is similar to that in the MISO channels, *i.e.*, $\mathbf{G}^{\perp}$ should be set to zero. Thus, the power constraint in (28) is set as $\sum_{u=1}^{U} \|\mathbf{G}_u\|_{\mathrm{F}}^2 = P_{\mathrm{T}}$.

Although we follow almost the same argument as that in the MISO channels, the problem for MIMO channels is still somewhat different owing to the diagonal matrix $\mathbf{\Sigma}_u \in \mathbb{C}^{K \times K}$. An extra optimization step to pull out the diagonal matrix is introduced in the following Lemma.

Lemma 2: Let $\tilde{\mathbf{G}}_u \in \mathbb{C}^{N_{\mathrm{S}} \times N_{\mathrm{S}}}$ be the first N_{S} rows in $\mathbf{G}_u$, which corresponds to the N_{S} largest elements in $\mathbf{\Sigma}_u$. $\mathbf{G}_u \in \mathbb{C}^{K \times N_{\mathrm{S}}}$ should be shrunk as the square matrix $\tilde{\mathbf{G}}_u \in \mathbb{C}^{N_{\mathrm{S}} \times N_{\mathrm{S}}}$ by setting all variables other than $\tilde{\mathbf{G}}_u$ to zeros and ignoring them. The problem in (28) can then be reformulated by $\tilde{\mathbf{G}}_u$ as

$$\begin{aligned} &\underset{\tilde{\mathbf{G}}_u}{arg\,max} \sum_{u=1}^{U} \log_2 \det\left(\gamma \tilde{\mathbf{G}}_u^H \tilde{\mathbf{\Sigma}}_u \tilde{\mathbf{G}}_u\right) \\ &s.t. \sum_{u=1}^{U} \|\tilde{\mathbf{G}}_u\|_{\mathrm{F}}^2 = P_{\mathrm{T}} \end{aligned} \tag{29}$$

where $\tilde{\mathbf{\Sigma}}_u \in \mathbb{C}^{N_{\mathrm{S}} \times N_{\mathrm{S}}}$ contains the N_{S} largest elements in $\mathbf{\Sigma}_u$.

Proof: See Appendix B. ∎

Since all matrices inside the determinant are square matrices, the objective function in (29) can be further simplified as follows:

$$\sum_{u=1}^{U} \log_2 \det\left(\gamma \tilde{\mathbf{\Sigma}}_u\right) + \sum_{u=1}^{U} \log_2 \det\left(\tilde{\mathbf{G}}_u^H \tilde{\mathbf{G}}_u\right) \tag{30}$$

where the first term is independent of $\tilde{\mathbf{G}}_u$. Defining $\tilde{\mathbf{P}}_u = \tilde{\mathbf{G}}_u^H \tilde{\mathbf{G}}_u$, the optimization problem can then be transformed into

$$\begin{aligned} &\arg\max_{\tilde{\mathbf{P}}_u} \sum_{u=1}^{U} \log_2 \det\left(\tilde{\mathbf{P}}_u\right) \\ &\text{s.t.} \sum_{u=1}^{U} \mathrm{tr}(\tilde{\mathbf{P}}_u) = P_{\mathrm{T}} \end{aligned} \tag{31}$$

The solution of the general cases for the MIMO channels is then introduced in the following proposition.

Proposition 5: The optimization solution for the problem in (31) is shown to be

$$\tilde{\mathbf{P}}_u = \frac{P_{\mathrm{T}}}{UN_{\mathrm{S}}} \mathbf{I}_{N_{\mathrm{S}}} \tag{32}$$

Proof: Two upper bounds can then be obtained by adopting Hadamard's Inequality and Arithmetic-Geometric Mean Inequality:

$$\begin{aligned} &\textstyle\sum_{u=1}^{U} \log_2 \det\left(\tilde{\mathbf{P}}_u\right) \\ &\leq \log_2 \left(\textstyle\prod_{i=1}^{N_{\mathrm{S}}} \tilde{p}_{1,ii} \prod_{i=1}^{N_{\mathrm{S}}} \tilde{p}_{2,ii} \cdots \prod_{i=1}^{N_{\mathrm{S}}} \tilde{p}_{U,ii}\right) \\ &\leq \log_2 \left(\frac{\sum_{i=1}^{N_{\mathrm{S}}} \tilde{p}_{1,ii} + \cdots + \sum_{i=1}^{N_{\mathrm{S}}} \tilde{p}_{U,ii}}{UN_{\mathrm{S}}}\right)^{UN_{\mathrm{S}}} \\ &= \log_2 \left(\frac{P_{\mathrm{T}}}{UN_{\mathrm{S}}}\right)^{UN_{\mathrm{S}}}, \end{aligned}$$

where $\tilde{p}_{u,ii}$ is the ith diagonal term in $\tilde{\mathbf{P}}_u$ and the sum of $\tilde{p}_{u,ii}$ is the trace of $\tilde{\mathbf{P}}_u$ such that $\sum_{u=1}^{U} \mathrm{tr}\left(\tilde{\mathbf{P}}_u\right) = \sum_{u=1}^{U} \sum_{i=1}^{N_{\mathrm{S}}} \tilde{p}_{u,ii} = P_{\mathrm{T}}$, which is related to the power constraint. Based on these theorems, $\tilde{\mathbf{P}}_u = \frac{P_{\mathrm{T}}}{UN_{\mathrm{S}}} \mathbf{I}_{N_{\mathrm{S}}}$ can reach the upper bound of $\sum_{u=1}^{U} \log_2 \det\left(\tilde{\mathbf{P}}_u\right)$. Proposition 5 is then proved. ∎

Under the relationship of $\tilde{\mathbf{P}}_u = \tilde{\mathbf{G}}_u^H \tilde{\mathbf{G}}_u$, since all the elements in matrix $\tilde{\mathbf{G}}_u$ are variables that we can adjust, there are multiple ways to satisfy the optimal solution in (32). The most straightforward way to design $\tilde{\mathbf{G}}_u$ is setting it as a diagonal matrix and an optimal solution for $\tilde{\mathbf{G}}_u$ is

$$\tilde{\mathbf{G}}_u = \sqrt{\frac{P_{\mathrm{T}}}{UN_{\mathrm{S}}}} \mathbf{I}_{N_{\mathrm{S}}} \tag{33}$$

Apart from the design in (33), any design of $\tilde{\mathbf{G}}_u$ that satisfies only one element in the column to be non-zero with non-repeated row can also fulfill the requirement in (32). For example, another possible form of $\tilde{\mathbf{G}}_u$ can be $\sqrt{\frac{P_{\mathrm{T}}}{UN_{\mathrm{S}}}}$ multiplies any $N_{\mathrm{S}} \times N_{\mathrm{S}}$ permutation matrix given by

$$\tilde{\mathbf{G}}_u = \sqrt{\frac{P_{\mathrm{T}}}{UN_{\mathrm{S}}}} \mathbf{P}_{permut} \tag{34}$$

Finding the optimal solution of $\tilde{\mathbf{G}}_u$ is equivalent to finding the solution of the analog precoder $\mathbf{F}$. Since there are numerous solutions of $\tilde{\mathbf{G}}_u$, there are also multiple possible combinations of the analog precoder $\mathbf{F}$ corresponding to the different designs of $\tilde{\mathbf{G}}_u$ based on (27). For example, if $\tilde{\mathbf{G}}_u$ is designed as in (33), the ith column $\mathbf{f}_i$ in $\mathbf{F}$ should be $\mathbf{f}_i = \sqrt{\frac{P_\mathrm{T}}{UN_\mathrm{S}}} \sum_{u=1}^{U} \mathbf{v}_{u,i}$ where $\mathbf{v}_{u,i}$ is the ith column in the matrix $\mathbf{V}_u$.

If one decides the non-zero element in each column in matrix $\tilde{\mathbf{G}}_u$ randomly such as in (34), each column in $\mathbf{F}$ can be decided by combining U singular vectors with equal weight $\sqrt{\frac{P_\mathrm{T}}{UN_\mathrm{S}}}$, where each vector is randomly selected from the N_S dominant singular vectors in $\mathbf{V}_u \in \mathbb{C}^{N_\mathrm{T} \times K}$ and each singular vector in $\mathbf{V} \in \mathbb{C}^{N_\mathrm{T} \times KU}$ can only be selected once.

B. Low SNR Regime in MIMO Channels

In low SNR region, the problem is expressed as

$$\begin{aligned} &\arg\max_{\mathbf{F}} \det\left(\mathbf{I} + \gamma \mathbf{F}^H \left(\sum_{u=1}^{U} \mathbf{H}_u^H \mathbf{H}_u\right) \mathbf{F}\right) \\ &\text{s.t.} \quad \|\mathbf{F}\|_\mathrm{F}^2 = P_\mathrm{T} \end{aligned} \tag{35}$$

where $\mathbf{F} \in \mathbb{C}^{N_\mathrm{T} \times N_\mathrm{S}}$ and $\mathbf{H}_u \in \mathbb{C}^{N_\mathrm{R} \times N_\mathrm{T}}$ with rank K. Let $\mathbf{H} = \sum_{u=1}^{U} \mathbf{H}_u^H \mathbf{H}_u$. Proposition 6 presents the precoder design for this problem.

Proposition 6: Let $\tilde{\mathbf{U}} \in \mathbb{C}^{N_\mathrm{T} \times N_\mathrm{S}}$ be the set of eigenvectors corresponding to the N_S largest eigenvalues of $\mathbf{H}$ and $\mathbf{D} \in \mathbb{C}^{N_\mathrm{S} \times N_\mathrm{S}}$ represent allocated power to each data stream obtained by water-filling algorithm. The precoder $\mathbf{F}$ in the low SNR region is given by $\mathbf{F} = \tilde{\mathbf{U}}\mathbf{D}$.

Proof: This is a typical water-filling optimization problem and the proof can be found, *e.g.,* in [28]. ■

V. Complexity Analysis

In this section, we use Big O, denoted as $\mathcal{O}(\cdot)$, to calculate the complexity orders for different propositions [29]. For simplifying the procedure of analysis, let us first give the complexity for some operations. Suppose that a is a scalar, $\mathbf{X}$ and $\mathbf{Y}$ are $p \times q$ matrices, $\mathbf{Z}$ is a $q \times r$ matrix, and $\mathbf{D}$ is a $r \times r$ diagonal matrix. Then, when $p \ll q$ or $p = q$, the complexity order for applying SVD operation to $\mathbf{X}$ is of $\mathcal{O}(q^3)$. However, if $p = 1$ and $\mathbf{X}$ becomes a vector $\mathbf{x}$, the complexity order for applying the SVD is of $\mathcal{O}(q)$. When $q \gg r$, the complexity order for applying QR decomposition to $\mathbf{Z}$ is of $\mathcal{O}(qr^2)$. The complexity orders for scalar-matrix multiplication $a\mathbf{X}$, matrix addition $\mathbf{X} + \mathbf{Y}$, and matrix multiplication $\mathbf{YZ}$ are of $\mathcal{O}(pq)$, $\mathcal{O}(pq)$, and $\mathcal{O}(pqr)$, respectively. The complexity order for matrix multiplication $\mathbf{ZD}$ is of $\mathcal{O}(qr)$. Moreover, the complexity order for water-filling over n eigenmodes is of $\mathcal{O}(n^2)$. According to above descriptions, the complexity orders for different propositions are described as follows:

MISO two-user case in high SNR regime (Proposition 1): In the first step, we apply the SVD to the channels of the two users, such as $\mathbf{h}_1^H \in \mathbb{C}^{1 \times N_\mathrm{T}}$ and $\mathbf{h}_2^H \in \mathbb{C}^{1 \times N_\mathrm{T}}$, where since the channels are vectors, the complexity of applying the SVD to them is $\mathcal{O}(2N_\mathrm{T})$. Then, we collect their right singular vectors as $\mathbf{V} = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{bmatrix} \in \mathbb{C}^{N_\mathrm{T} \times 2}$ and apply QR decomposition to $\mathbf{V}$, where the complexity is $\mathcal{O}(4N_\mathrm{T})$. Next, we use (15) to apply N points grid search to find optimal θ, and thus the dominant complexity is $\mathcal{O}(N)$. Finally, (14) is used to design the analog precoder, where the complexity is dominated by two scalar-vector multiplications and one vector addition and thus the complexity is $\mathcal{O}(3N_\mathrm{T})$. Therefore, the complexity for Proposition 1 is $\mathcal{O}(9N_\mathrm{T} + N)$.

MISO multiuser case in high SNR regime (Proposition 2): In the first step, we apply the SVD to the vector channels of U users, where the complexity is $\mathcal{O}(UN_\mathrm{T})$. Then, we use $\mathbf{f} = \sqrt{\frac{P_\mathrm{T}}{U}} \sum_{u=1}^{U} \mathbf{v}_u$ to design the analog precoder, where the complexity is dominated by one scalar-vector multiplication and $U-1$ vector additions and thus the complexity is $\mathcal{O}(UN_\mathrm{T})$. Therefore, the complexity for Proposition 2 is $\mathcal{O}(2UN_\mathrm{T})$.

MISO multiuser case in low SNR regime (Proposition 3): According to Proposition 3, $\mathbf{H} = \sum_{u=1}^{U} \mathbf{h}_u \mathbf{h}_u^H$ should be first calculated, whose complexity is dominated by U vector multiplications $\mathbf{h}_u \mathbf{h}_u^H$ and $U-1$ matrix additions. The complexity for both vector multiplication and matrix addition is $\mathcal{O}(N_\mathrm{T}^2)$. Thus, the complexity for calculating $\sum_{u=1}^{U} \mathbf{h}_u \mathbf{h}_u^H$ is $\mathcal{O}((2U-1)N_\mathrm{T}^2)$. Then, we apply the SVD to $\mathbf{H}$ to obtain the eigenvectors and the complexity is $\mathcal{O}(N_\mathrm{T}^3)$. Finally, a selected eigenvector is scaled to be the analog precoder which satisfies the power constraint, where the complexity is dominated by one scalar-vector multiplication and thus the complexity is $\mathcal{O}(N_\mathrm{T})$. Therefore, the complexity for Proposition 3 is $\mathcal{O}(N_\mathrm{T}^3 + (2U-1)N_\mathrm{T}^2 + N_\mathrm{T})$.

MIMO two-user case with LOS channel in high SNR regime (Proposition 4): From Proposition 4, we know that the solution is the same as that for MISO two-user case in high SNR regime (Proposition 1). Therefore, the complexity for Proposition 4 is $\mathcal{O}(9N_\mathrm{T} + N)$.

MIMO multiuser case in high SNR regime (Proposition 5): In the first step, we apply the SVD to the matrix channels of U users $\mathbf{H}_u$, where the complexity is $\mathcal{O}(UN_\mathrm{T}^3)$. Then, the ith column of analog precoder can be designed by $\mathbf{f}_i = \sqrt{\frac{P_\mathrm{T}}{UN_\mathrm{S}}} \sum_{u=1}^{U} \mathbf{v}_{u,i}$, where the complexity of finding $\mathbf{f}_i$ is dominated by one scalar-vector multiplication and $U-1$ vector additions. Thus the complexity is $\mathcal{O}(UN_\mathrm{T})$. However, since the analog precoder has N_S columns, the total complexity of finding $\mathbf{F}$ is $\mathcal{O}(N_\mathrm{S}UN_\mathrm{T})$. Therefore, the complexity for Proposition 5 is $\mathcal{O}(UN_\mathrm{T}^3 + N_\mathrm{S}UN_\mathrm{T})$.

MIMO multiuser case in low SNR regime (Proposition 6): According to Proposition 6 and (35), $\mathbf{H} = \sum_{u=1}^{U} \mathbf{H}_u^H \mathbf{H}_u$ should be first calculated, whose complexity is dominated by U matrix multiplications $\mathbf{H}_u^H \mathbf{H}_u$ and $U-1$ matrix additions. The complexity for each matrix multiplication is $\mathcal{O}(N_\mathrm{R}N_\mathrm{T}^2)$ and for each matrix addition is $\mathcal{O}(N_\mathrm{T}^2)$. Thus, the complexity for calculating $\mathbf{H} = \sum_{u=1}^{U} \mathbf{H}_u^H \mathbf{H}_u$ is $\mathcal{O}((UN_\mathrm{R} + U - 1)N_\mathrm{T}^2)$. Then, we apply the SVD to $\mathbf{H}$ to obtain the $\tilde{\mathbf{U}} \in \mathbb{C}^{N_\mathrm{T} \times N_\mathrm{S}}$ and the complexity is $\mathcal{O}(N_\mathrm{T}^3)$. Moreover, the complexity for using water-filling over N_S eigenmodes to find the diagonal $\mathbf{D}$ is $\mathcal{O}(N_\mathrm{S}^2)$. Finally, analog precoder is designed by $\mathbf{F} = \tilde{\mathbf{U}}\mathbf{D}$, where the complexity is $\mathcal{O}(N_\mathrm{S}N_\mathrm{T})$. Therefore, the complexity for Proposition 6 is $\mathcal{O}(N_\mathrm{T}^3 + (UN_\mathrm{R} + U - 1)N_\mathrm{T}^2 + N_\mathrm{S}N_\mathrm{T} + N_\mathrm{S}^2)$.

TABLE I
COMPLEXITY SUMMARY FOR THE PROPOSED PROPOSITIONS

Complexity analysis for MISO channels			
Method	Proposition 1	Proposition 2	Proposition 3
Complexity order	$\mathcal{O}(9N_{\mathrm{T}}+N)$	$\mathcal{O}(2UN_{\mathrm{T}})$	$\mathcal{O}(N_{\mathrm{T}}^3+(2U-1)N_{\mathrm{T}}^2+N_{\mathrm{T}})$
Complexity analysis for MIMO channels			
Method	Proposition 4	Proposition 5	Proposition 6
Complexity order	$\mathcal{O}(9N_{\mathrm{T}}+N)$	$\mathcal{O}(UN_{\mathrm{T}}^3+N_{\mathrm{S}}UN_{\mathrm{T}})$	$\mathcal{O}(N_{\mathrm{T}}^3+(UN_{\mathrm{R}}+U-1)N_{\mathrm{T}}^2+N_{\mathrm{S}}N_{\mathrm{T}}+N_{\mathrm{S}}^2)$

The complexity for all the proposed propositions is summarized in Table I, where one can see that the complexity is mainly affected by the number N_{T} of the transmit antennas. Also Propositions 1, 2, and 4 have complexity in linear order of N_{T}, because obtaining singular vectors for vectors is much simpler than for matrices.

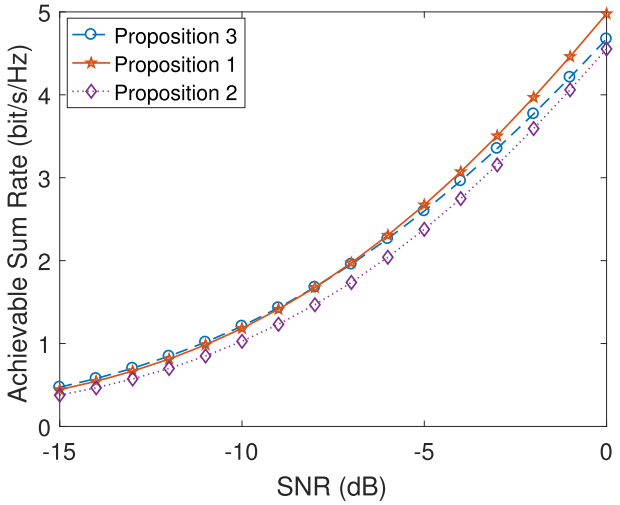
VI. SIMULATION RESULTS

In this section, simulation results are demonstrated to show the performance of the proposed system. The signal-to-noise ratio is $\gamma = \frac{\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{n}}^2}$. If it is not specifically mentioned, the narrowband geometric channel model is utilized [6, (4)], the number of transmit antennas is $N_{\mathrm{T}} = 8$, and the number of receive antennas is $N_{\mathrm{R}} = 1$ in MISO channels; while the number of transmit antennas is $N_{\mathrm{T}} = 64$, and the number of receive antennas is $N_{\mathrm{R}} = 4$ in MIMO channels. The number of scattering paths is $L = 4$, and the number of transmitted data streams is $N_{\mathrm{S}} = 4$. The number of channel realizations in each experiment is 10^4.

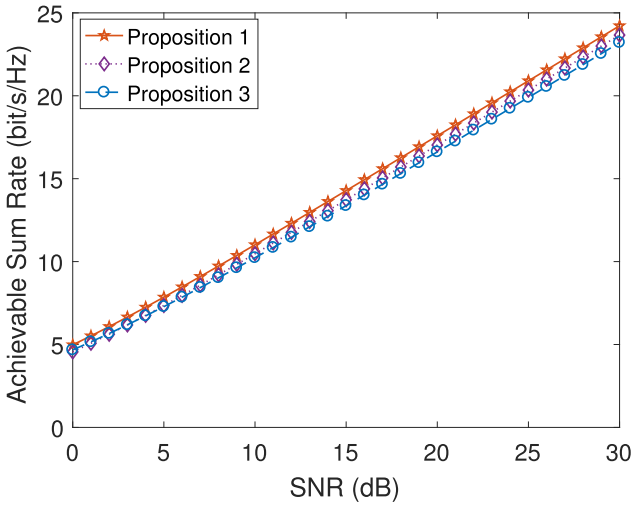
A. Experiment 1: Proposed Systems in MISO Channels

In the first experiment, we demonstrate the performance for MISO channels. Let us see the two-user case in MISO channels, which is shown in Figs. 2(a) and 2(b) for low SNR and high SNR regimes respectively. In Fig. 2(a), Proposition 3 is the best among other propositions when the value of SNR is smaller than -8 dB and gradually enters the low SNR regime, since it is the asymptotically optimal solution of the two-user case for the low SNR regime but others are not. However, when the value of SNR is higher than -7 dB and gradually enters the high SNR regime, Proposition 1 outperforms other two schemes observed from Figs. 2(a) and 2(b), since it is the asymptotically optimal solution of the two-user case for the high SNR regime. In addition, since Proposition 2 is derived for the high SNR regime based on applying additional orthogonality approximation, its performance is inferior to Proposition 1 as expected. Thus for the low SNR regime shown in Fig. 2(a), Proposition 2 is then worse than Proposition 1 and has the worst performance. However, for the high SNR regime shown in Fig. 2(b), Proposition 2 asymptotically achieves the performance of Proposition 1 and does not sacrifice too much performance. Moreover, it outperforms Proposition 3 because Proposition 3 is the solution for the low SNR regime.

Next, let us see the cases with more than two users in MISO channels. Let the numbers of users be three and five, and the number of antennas be $N_{\mathrm{T}} = 16$. Figs. 3(a) and 3(b) respectively show the performance in low SNR and high



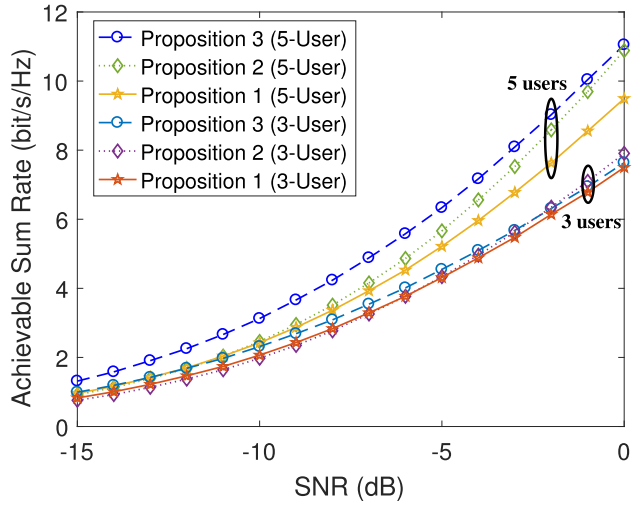
(a) low SNR region



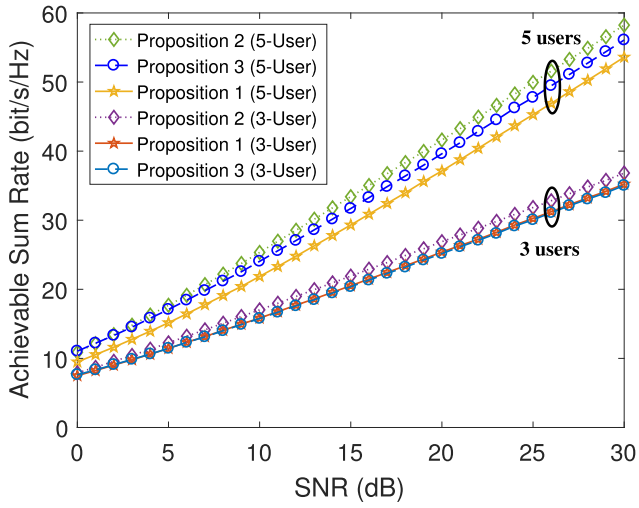
(b) high SNR region

Fig. 2. Performance comparison: MISO channels and two-user case.

SNR regions. From Fig. 3(a), since Proposition 3 is the asymptotically optimal solution for the low SNR regime, it outperforms other schemes as expected. From Fig. 3(b), Proposition 2 has the best performance instead of Proposition 1, which is because Proposition 1 is the solution for the two-user case. When the number of the users is more than two, Proposition 1 is no longer the asymptotically optimal solution for the high SNR regime. However, Proposition 2 is the asymptotically optimal solution of multiuser cases for the

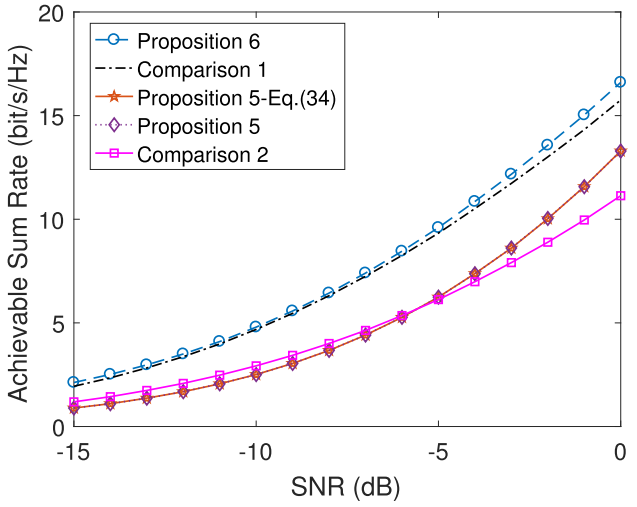


(a) low SNR region

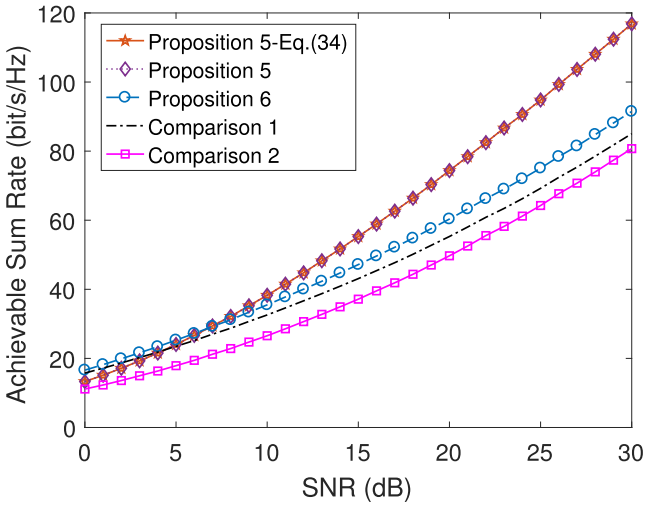


(b) high SNR region

Fig. 3. Performance comparison: MISO channels and more than two users.



(a) low SNR region



(b) high SNR region

Fig. 4. Performance comparison: MIMO channels and $U = 4$.

high SNR regime and thus becomes the best solution instead of Proposition 1.

In addition, as Figs. 3(a) and 3(b) show, when the number of users increases, the performance behavior of each proposition can be more pronounced. Thus, their performance gaps become more obvious when the number of users increases.

B. Experiment 2: Proposed Systems in MIMO Channels and Comparison to Other Precoder Designs

In the experiment, we demonstrate the performance of the proposed solutions in MIMO channels (Propositions 5 and 6) and compare them to other precoder designs. Two conventional precoding schemes used for comparison are described as follows:

Comparison 1: Combine all right singular vectors of all user channels and arrange them according to their eigenvalues to form a matrix $\mathbf{V}$. Then, the analog precoder $\mathbf{F}$ is designed as the first N_{S} right singular vectors in $\mathbf{V}$.

Comparison 2: Applying SVD to individual user channels: $\mathbf{H}_u = \mathbf{U}_u\mathbf{D}_u\mathbf{V}_u^H$, where $\mathbf{D}_u$ is a diagonal matrix with elements $s_{u,i}$ in descending order and the corresponding columns in right singular matrix $\mathbf{V}_u$ are represented as $\mathbf{v}_{u,i}$. Since there are U users, for the same ith column of $\mathbf{V}_u$, there are U vectors $\mathbf{v}_{u,i}$. The one corresponding to the largest singular value $s_{u,i}$ is chosen as the ith column in matrix $\mathbf{F}$. Let us take the two-user case as an example, the ith column of $\mathbf{F}$ is determined by $\mathbf{f}_i = \begin{cases} \mathbf{v}_{1,i}, \text{ if } \mid s_{1,i} \mid \geq \mid s_{2,i} \mid \\ \mathbf{v}_{2,i}, \text{ if } \mid s_{1,i} \mid < \mid s_{2,i} \mid \end{cases}$.

The reason to use these two heuristic schemes is because in the literature, many precoder designs are based on the singular vectors, *e.g.,* see [8], [9], [11]. Thus we attempt to provide comparisons with some reasonable and heuristic schemes, and show that although these schemes indeed perform not bad, better performance can be achieved by properly designing the precoder using the proposed algorithms. Let the number U of users be 4. Figs. 4(a) and 4(b) show the performance comparisons in low and high SNR regions respectively.

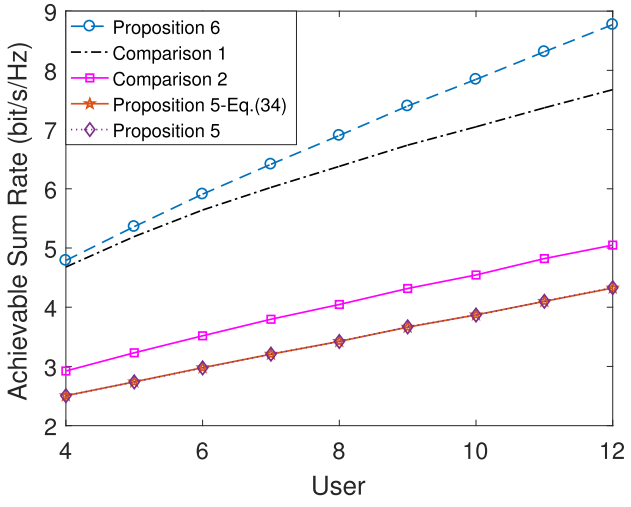
Several observations are summarized. First, the proposed solutions outperform the two conventional schemes in both low and high SNR regions (see Proposition 6 in Fig. 4(a) for

low SNR region and Proposition 5 in Fig. 4(b) for high SNR region).

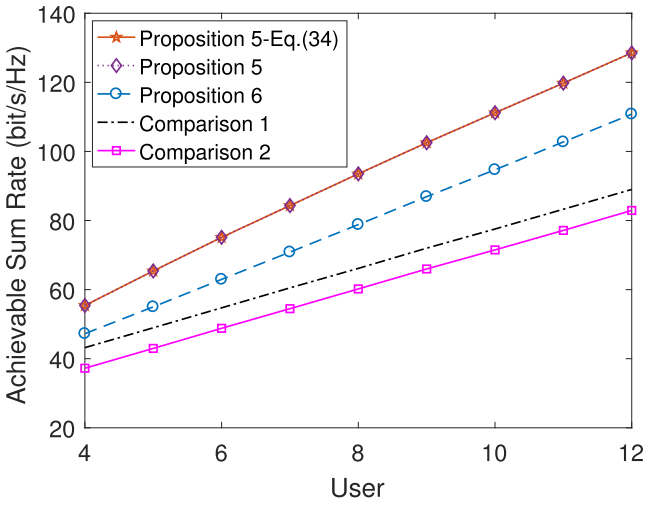
Second, these two figures also reveal that different solutions should be properly used for different SNR regions to achieve the best performance; for low SNR region, the proposed low SNR solution (Proposition 6) outperforms the proposed high SNR solution (Proposition 5) and other two comparing schemes, observed from Fig. 4(a). From the same figure, without using proper solution according to SNR, in low SNR region the conventional scheme Comparison 1 even outperforms Proposition 5 (high SNR solution). Similar situation occurs for high SNR region in Fig. 4(b).

Third, recalling in Proposition 5, we proved that multiple solutions exist for high SNR region. That is, (33) and (34) are all possible solutions. Here we use (33) as the solution for Proposition 5 and compare it to (34) with random permutation matrix in every realization. Figs. 4(a) and 4(b) show that these multiple solutions all lead to the same performance (see the overlapped curves with star and diamond marks). This also shows the correctness of the proposed analytical result.

Figs. 5(a) and 5(b) respectively show the performance comparisons for various numbers of users in the low and high SNR regimes, where the values of SNR are fixed as -10 dB and 15 dB, respectively. From the two figures, the performance of each scheme increases when the number of users grows. Moreover, the performance gaps between the proposed and other schemes become larger, which shows that the advantages of the proposed scheme become more pronounced as the number of users increases.



(a) SNR = -10 dB



(b) SNR = 15 dB

Fig. 5. Performance comparison: MIMO channels with various numbers of users.

C. Experiment 3: Comparison of Different User Distances in Proposed Systems With MIMO Channels

In this experiment, we compare the performance of the proposed systems with various user closeness. Let the number U of users be 4, and for the channel model, we set the number of paths $L = 16$. In addition, we adjust the standard deviation σ of the Laplacian distribution of the AOA/AOD scattering paths to reflect how close the users are in position. That is, if the users are close and in similar environments, σ should be small; otherwise if the users are far apart, σ is large.

Fig. 6 shows the performance comparison of the proposed systems with various standard deviations of AoA/AoD $\sigma = 2$, 4, and ∞. Observed that the proposed scheme with small value of σ outperforms that with large value of σ. The result is reasonable because the proposed scheme benefits from user closeness. When the users are close, the sharing physical beam can be more concentrated in a narrow direction. When the users are far apart, the sharing physical beam should point several and wide directions to take care of all users. As a result, the energy dissipates and the performance degrades.

D. Experiment 4: Comparison of Proposed System and Spatial Multiplexing Precoded OFDMA System

In this experiment, we show the performance of the proposed and the spatial multiplexing precoded OFDMA systems. Let $U = 2$ and the two users are in two different subcarriers. Also, the number of RF chains is two. For the proposed system, the two users share the two data streams from a sharing physical beam, and hence each user can receive two data streams. For the spatial multiplexing precoded OFDMA system, each user can receive only one dedicated data stream. The broadband geometric channel model in [13, (3) and (4)] and the number of scattering paths $L = 16$ are applied in this simulation, where the carrier frequency is 28 GHz, and the bandwidth is 61.44 MHz. The delay spread is 32, the number of the subcarriers is 512, and the subcarrier spacing is 120 KHz. The transmit power and the noise power are determined according to the value of SNR.

Fig. 7 shows the performance comparison of the proposed and the spatial multiplexing precoded OFDMA systems. One can observe that the proposed system outperforms the spatial multiplexing precoded OFDMA system as the value of SNR exceeds 10 dB. This result is reasonable because in the spatial multiplexing precoded OFDMA system, both spatial and frequency domains are used for multiple access. Thus, the

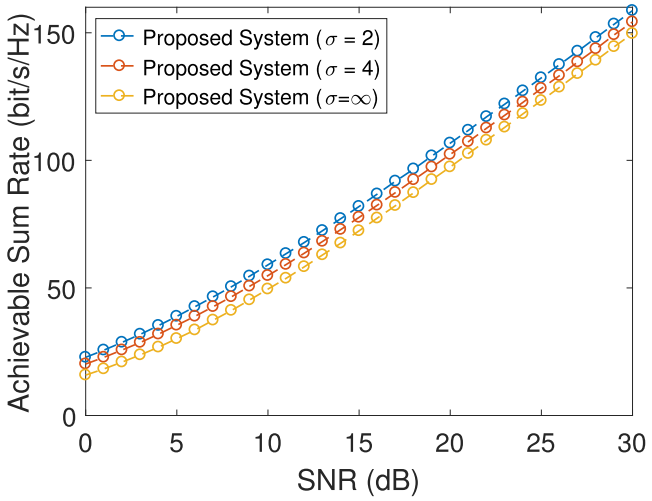


Fig. 6. Performance comparison of different user closeness in proposed systems: $U = 4$.

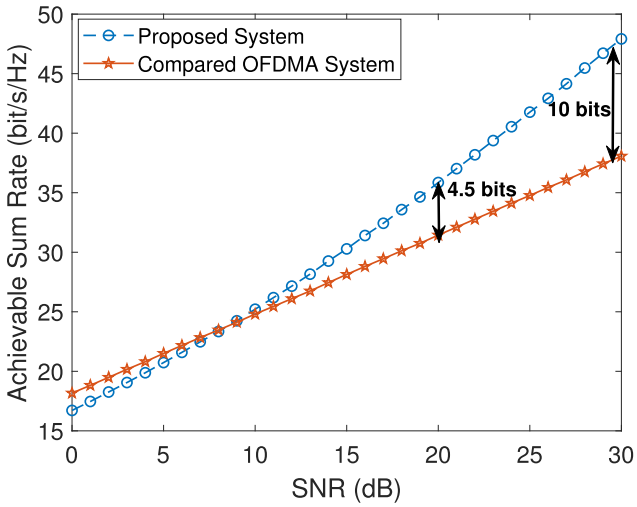


Fig. 7. Performance comparison between proposed and spatial multiplexing precoded systems in a two-user case.

number of data streams is only one for each user. On the other hand, although the proposed system forms a sharing physical beam with two directions and dissipates the received precoding gain of individual users, it only uses frequency domain for multiple access, and the number of data streams is two for each user (full multiplexing gain).

E. Experiment 5: Comparison of the Proposed and NOMA Systems

In this experiment, we add the NOMA system [24] as a benchmark and compare it to the proposed scheme. There are one BS and two users ($U = 2$) in the system, where the two users are in a cluster. The BS is equipped with one RF chain and $N_{\mathrm{T}} = 64$ transmit antennas. Each user is equipped with $N_{\mathrm{R}} = 1$ antenna. The narrowband geometric channel model with clustering is applied, and the number of scattering paths is $L = 16$. We assume that the channel strengths of the users are different by setting different complex gains and changing

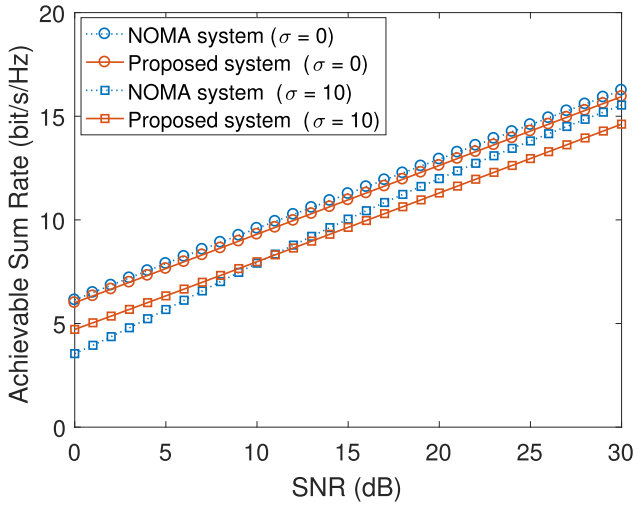


Fig. 8. Sum rate performance between the proposed and the NOMA systems.

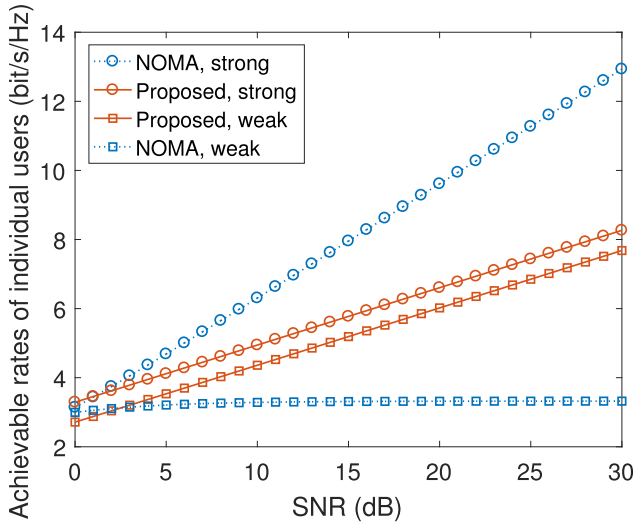
the standard deviation σ of the Laplacian distribution of the AOA/AOD scattering paths like that used in Experiment 3. As mentioned in the introduction, each user in the proposed system only utilizes $\frac{1}{U}$ bandwidth, and this is considered in the simulation.

The sum rate performance of the NOMA and the proposed schemes is shown in Fig. 8. We observe that the NOMA system slightly outperforms the proposed system when the channels of users are highly correlated ($\sigma = 0$). However, when σ increases and channels of the users tend to be uncorrelated, the proposed system can outperform the NOMA system when the SNR is lower than 10 dB. The achievable rates of individual users in both the proposed and the NOMA systems are shown in Figs. 9(a) and 9(b) for $\sigma = 0$ and $\sigma = 10$, respectively. From these figures, we observe that in the proposed system, the two users are fairly taken care of, no matter the user channels are correlated or uncorrelated, *i.e.*, $\sigma = 0$ or $\sigma = 10$, thanks to the MUI-free property and the multi-directional sharing beam. On the other hand, for the NOMA system, since the weak user is significantly affected by the MUI and only one-direction beam is applied, the weak user has much worse performance than the strong user, which is widely seen in the literature on NOMA systems.

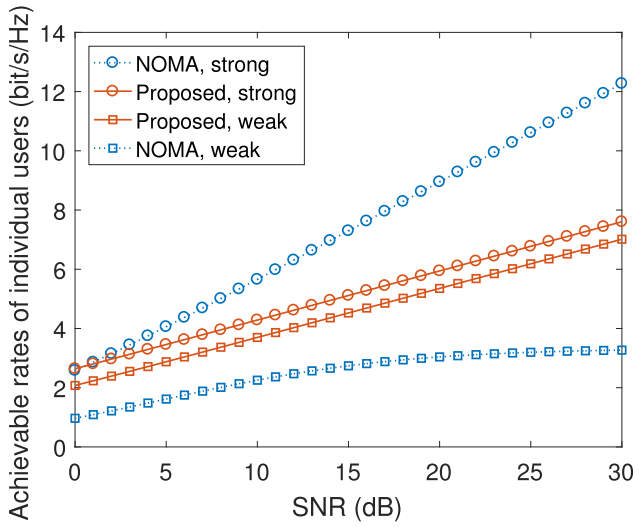
To conclude, the NOMA scheme generally achieves higher sum rate performance than the proposed scheme. However, the proposed system has advantage in terms of fairness as well as low complexity at the user side because there is no need to apply SIC.

VII. Hardware Implementation and OTA Measurement

To verify the feasibility of the proposed ideas in practical situations, the 28 GHz mmWave hardware systems were implemented, where OTA performance and beam pattern of the proposed beam sharing scheme were practically measured. These two OTA measurements are described in the following subsections.



(a) $\sigma = 0$



(b) $\sigma = 10$

Fig. 9. Achievable rates of individual users for the proposed and the NOMA systems.

A. OTA Performance Evaluation for the Proposed Beam Sharing Scheme

The block diagram of this hardware platform is shown in Fig. 10, and the practical hardware platform is shown in Fig. 11. The transmitter uses one mmWave 28 GHz 8×8 antenna array, equipped with one RF chain. Two horn antennas are used to represent as receivers (User 1 and User 2). The gain and phase of this antenna array are both adjustable (see [25]). This setting implements a hardware system discussed in subsection III-A.1.

At the transmitter side, the baseband (BB) signals are modulated as 16 quadrature amplitude modulation (QAM) signals and allocated to two different subcarriers corresponding to individual users with subcarrier spacing of 120 KHz and total bandwidth of 61.44 MHz. After modulated to 28 GHz through the transmitter software defined radio (SDR) and upconverter, the signals are transmitted by the 28 GHz 8×8 antenna array. At the receiver side, the 28 GHz signals received by two horn antennas are processed by downconverter and receiver SDR to obtain BB signals and then calculate their error vector magnitudes (EVMs).

Two precoding cases were considered in this hardware experiment. For Case 1, dedicated beam for User 2 is considered. On the other hand, for Case 2, a sharing beam designed according to Proposition 1 is considered.

Fig. 12 shows the simulated beam patterns of these two cases. In Case 1, the beam is directed to 0 degrees and dedicated to User 2, and therefore User 1 is abandoned. Case 1 using dedicated beam is not uncommonly seen in a beamforming system with only one RF chain. For the proposed sharing beam system in Case 2, the sharing beam takes care of both Users 1 and 2. Hence, both users can simultaneously achieve satisfactory performance.

Fig. 13 shows the demodulated constellations and corresponding measured EVMs for the two testing cases. For Case 1, shown in Fig. 13(a), the beam is dedicated to User 2 and thus the constellation of User 2 is concentrated (EVM $= -28.9$ dB). However, extremely divergent constellation is obtained by User 1 (EVM $= -2.7$ dB). From Fig. 13(a), we observe that only one user can benefit from the dedicated beam, which largely limits the development especially when the RF chains are limited. For Case 2, the proposed sharing beam is applied and the EVM result is shown in Fig. 13(b). We observe that both User 1 and User 2 can simultaneously attain concentrated constellations and satisfactory EVMs (EVM of User 1 $= -24$ dB and EVM of User 2 $= -28.5$ dB).

It is worth pointing out that in this practical hardware experiment, there are several non-ideal effects including inter carrier interference (ICI) caused by carrier frequency offset (CFO), phase noise, jitters, timing mismatch, and Doppler effect, etc. Although these effects can mostly be mitigated using conventional signal processing at the receiver side by the SDR, the residual effects still exist. These residual effects can be observed from the EVM plots in Fig. 13, where the constellation points near the four corners tend to have larger noise than those near zero.

We observe that the proposed beam sharing scheme is feasible in practical systems; meanwhile it can well leverage the performance of individual users compared to the conventional schemes. As a result, more users can be supported using limited RF chains, and this benefits the development of communications networks.

B. OTA Beam Pattern Measurement for the Proposed Beam Sharing Scheme

In subsection VII-A, we showed that two users with different AoAs (angles of arrival) can be covered by the sharing beam generated by only one RF chain, and their EVM performance implicitly proved this concept. To prove this concept in an explicit way, we conduct another hardware OTA experiment to practically measure the resulting beam pattern generated by the 28 GHz 8×8 UPA platform. For this experiment, another measurement platform was implemented, which consists of an automatic three-axis stabilizer to sequentially shift its

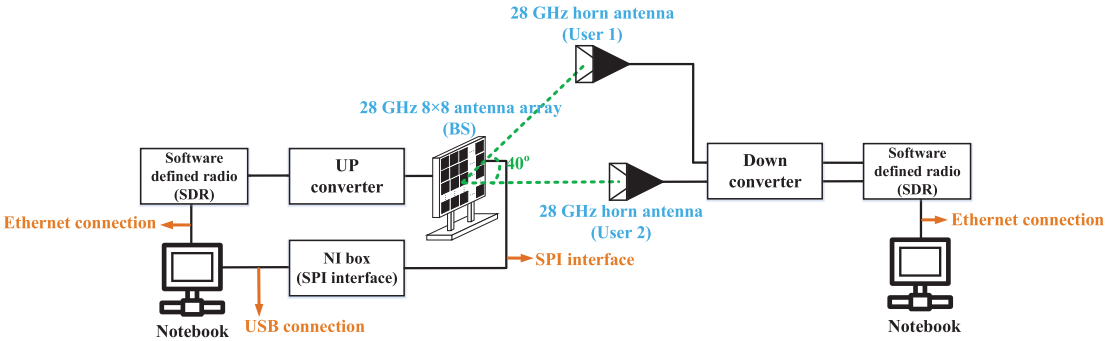


Fig. 10. A block diagram of the hardware platform.

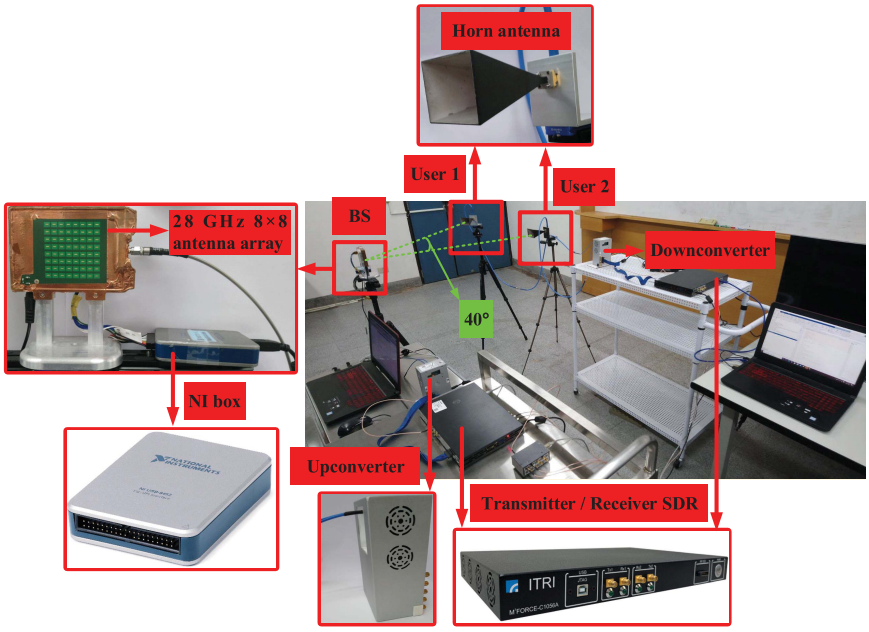


Fig. 11. A photo of the implemented mmWave 28 GHz hardware platform.

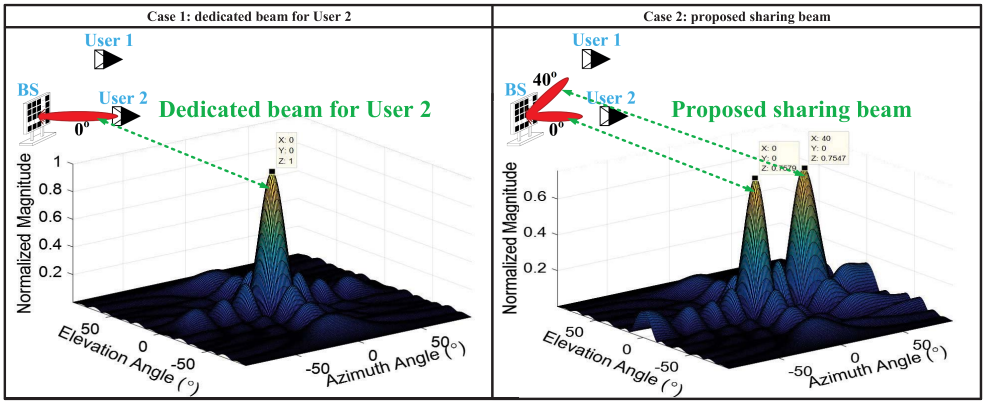
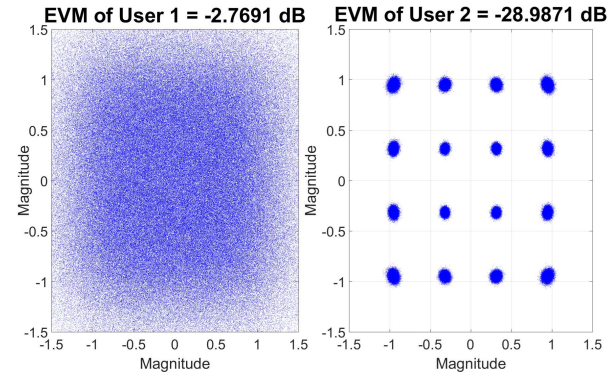


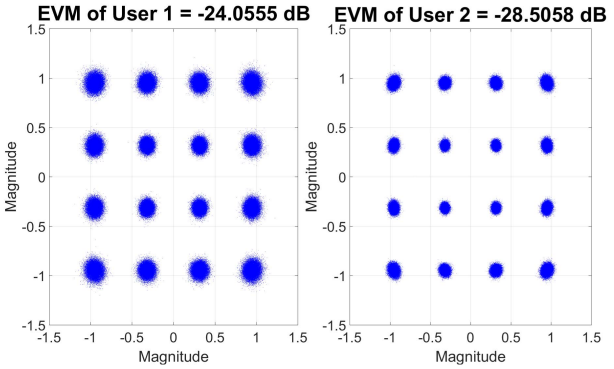
Fig. 12. The illustration figures and simulated beam patterns of different cases.

direction horizontally and vertically, to capture the power of the resulting beam pattern. The implemented measurement platform is shown in Fig. 14.

We let the desired beam be pointed simultaneously to two angles. One is −20 degrees and the other is 20 degrees. The measurement platform rotates the direction every two degrees



(a) Case 1: dedicated beam for User 2



(b) Case 2: proposed sharing beam

Fig. 13. The demodulated constellations and corresponding measured EVMs for different cases and different users.

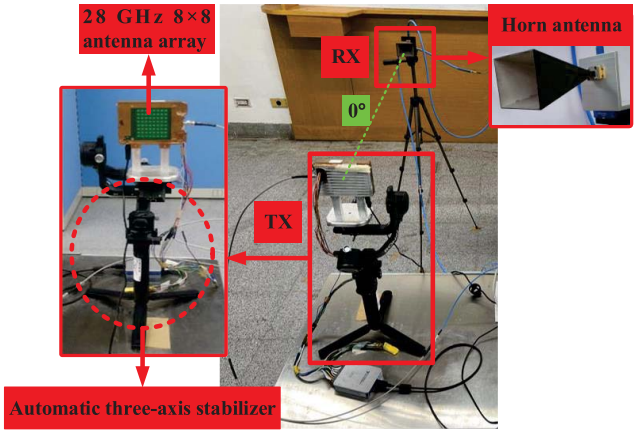
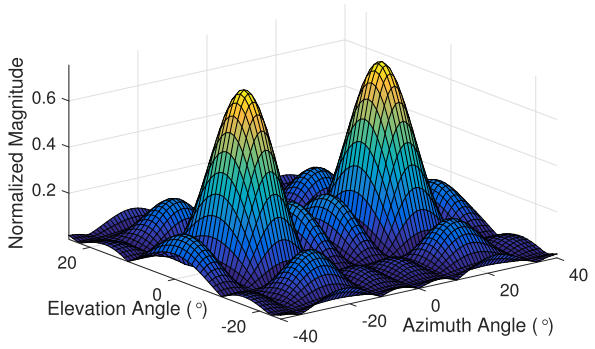


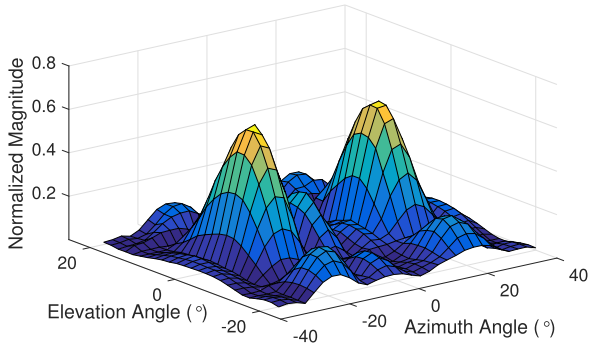
Fig. 14. The implemented platform to measure the beam pattern.

horizontally and vertically to capture the power of the transmit beamforming, and the corresponding theoretical and measured results are shown in Fig. 15.

From this figure, we observe that the measured beam pattern indeed forms beams simultaneously towards two directions, and the trend is close to the theoretical beam pattern. From this hardware experiment, we prove in a more explicit way that the proposed sharing beam is practically feasible in wireless channels.



(a) Theoretical beam pattern



(b) Measured beam pattern

Fig. 15. The theoretical and measured beam patterns of the proposed beam sharing scheme.

VIII. Conclusion and Future Works

We have proposed a novel physical beamforming multiuser system, where a sharing beam is designed to benefit all users simultaneously, and every user can attain full multiplexing gain. Precoding solutions for this proposed system have also been developed for MU-MISO and MU-MIMO channels. Computer simulations have shown that the proposed system outperforms the conventional schemes and has advantages in terms of fairness and complexity compared with NOMA systems. We have also implemented the 28 GHz hardware platforms with an 8×8 UPA to prove the feasibility and advantages of the beam sharing concept in practical wireless channels. The nice properties verified by both the computer simulations and the OTA testings have made the proposed system deserve further investigations for the next generation communications. For the future works, the concept of the proposed beam sharing may be combined with other schemes to inspire potential new techniques, *e.g.,* using the proposed beam sharing concept in the NOMA schemes. Moreover, scheduling based on the new proposed beam sharing scheme can also be investigated.

Appendix A
Proof of Proposition 1

Proof: Since the absolute value $|\, g_u \,|$ guarantees to be a positive value, only the first quadrant needs to be considered.

Two desired variables $\mid g_1 \mid$ and $\mid g_2 \mid$ can be found by single parameter θ by setting $\mid g_1 \mid = \sqrt{P_{\mathrm{T}}}\cos(\theta)$ and $\mid g_2 \mid = \sqrt{P_{\mathrm{T}}}\sin(\theta)$. The optimization problem in (13) can be represented as

$$\begin{aligned} &\arg\max_{\theta} \mid r_{12} \mid \cos^2(\theta) + \mid r_{22} \mid \cos(\theta)\sin(\theta) \\ &\text{s.t. } \cos^2(\theta) + \sin^2(\theta) = 1, \quad \text{for } \theta \in [0, \frac{\pi}{2}] \end{aligned} \tag{36}$$

Since we have already limited the range of θ to the first quadrant, it is fast to use the grid search here based on (36) to find the optimal θ, and the optimal values of $\mid g_1 \mid$ and $\mid g_2 \mid$ can therefore be found. ■

APPENDIX B
PROOF OF LEMMA 2

Proof: We define $\mathbf{P}_u = \mathbf{G}_u^H \mathbf{\Sigma}_u \mathbf{G}_u$ and denote the off-diagonal terms with notation $*$, *i.e.*,

$$\begin{aligned} \mathbf{P}_u &= \mathbf{G}_u^H \mathbf{\Sigma}_u \mathbf{G}_u \\ &= \begin{bmatrix} \sum_{k=1}^{K} \lambda_{u,k} \mid g_{u,k1} \mid^2 & \cdots & * \\ \vdots & \ddots & \vdots \\ * & \cdots & \sum_{k=1}^{K} \lambda_{u,k} \mid g_{u,kN_{\mathrm{S}}} \mid^2 \end{bmatrix} \end{aligned} \tag{37}$$

where $\lambda_{u,k}$ is the non-negative scalar in $\mathbf{\Sigma}_u$ and represents the square of the kth singular value, and $g_{u,ki}$ is the variable in the kth row and ith column of $\mathbf{G}_u$. According to the Hadamard's Inequality, the upper bound of $\det(\mathbf{P}_u)$ is given as $\det(\mathbf{P}_u) \leq \prod_{i=1}^{N_{\mathrm{S}}} p_{u,ii}$, where $p_{u,ii}$ is the ith diagonal term in $\mathbf{P}_u$. To maximize the upper bound while remain the rank N_{S}, referring to (37), the power should be allocated to the variables that multiply with $\lambda_{u,1}$ to $\lambda_{u,N_{\mathrm{S}}}$.

Therefore, any variable $g_{u,ki}$, where $k > N_{\mathrm{S}}$, is set to zero and then ignored. This leads to $\tilde{\mathbf{G}}_u \in \mathbb{C}^{N_{\mathrm{S}} \times N_{\mathrm{S}}}$, and the effective $\mathbf{\Sigma}_u$ also becomes $\tilde{\mathbf{\Sigma}}_u \in \mathbb{C}^{N_{\mathrm{S}} \times N_{\mathrm{S}}}$. $\mathbf{P}_u$ can then be represented by $\mathbf{P}_u = \tilde{\mathbf{G}}_u^H \tilde{\mathbf{\Sigma}}_u \tilde{\mathbf{G}}_u$, and the ith diagonal element can be expressed as $\sum_{k=1}^{N_{\mathrm{S}}} \lambda_{u,k} \mid g_{u,ki} \mid^2$. ■

ACKNOWLEDGMENT

The authors would like to thank all the anonymous reviewers for their constructive suggestions, which have significantly improved the quality of this work. They would also like to acknowledge Roger S. H. Huang and Gene C. H. Chuang in Wistron NeWeb Corporation (WNC) for their technical supports in antenna arrays and mmWave measurements, Wen-Wei Tsai in NYCU for assisting the hardware installation and measurements, and Industrial Technology Research Institute (ITRI) for technical supports in the software defined radio (SDR).

REFERENCES

[1] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Feb. 2014.

[2] S. A. Busari, K. M. S. Huq, S. Mumtaz, L. Dai, and J. Rodriguez, "Millimeter-wave massive MIMO communication for future wireless systems: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 836–869, 2nd Quart., 2018.

[3] J. Litva and T. K. Lo, *Digital Beamforming in Wireless Communications*. Norwood, MA, USA: Artech House, 1996.

[4] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale MIMO systems," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2015, pp. 2929–2933.

[5] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, Apr. 2016.

[6] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6481–6494, Nov. 2015.

[7] T.-Y. Chang and C.-E. Chen, "A hybrid Tomlinson–Harashima transceiver design for multiuser mmWave MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 118–121, Feb. 2018.

[8] F. Sohrabi and W. Yu, "Hybrid analog and digital beamforming for mmWave OFDM large-scale antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1432–1443, Jul. 2017.

[9] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, Jr., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

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

[11] H. Seleem, A. I. Sulyman, and A. Alsanie, "Hybrid precoding-beamforming design with Hadamard RF codebook for mmWave large-scale MIMO systems," *IEEE Access*, vol. 5, pp. 6813–6823, 2017.

[12] Y.-P. Lin, "Hybrid MIMO-OFDM beamforming for wideband mmWave channels without instantaneous feedback," *IEEE Trans. Signal Process.*, vol. 66, no. 19, pp. 5142–5151, Oct. 2018.

[13] A. Alkhateeb and R. W. Heath, "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1801–1818, May 2016.

[14] L. Kong, S. Han, and C. Yang, "Hybrid precoding with rate and coverage constraints for wideband massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4634–4647, Jul. 2018.

[15] Y. Lu, C. Cheng, J. Yang, and G. Gui, "Improved hybrid precoding scheme for mmWave large-scale MIMO systems," *IEEE Access*, vol. 7, pp. 12027–12034, 2019.

[16] C.-J. Wang, C.-K. Wen, S. Jin, and S.-H. Tsai, "Finite-alphabet precoding for massive MU-MIMO with low-resolution DACs," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4706–4720, Jul. 2018.

[17] S. K. Mohammed and E. G. Larsson, "Per-antenna constant envelope precoding for large multi-user MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 1059–1071, Mar. 2013.

[18] M. Kazemi, H. Aghaeinia, and T. M. Duman, "Discrete-phase constant envelope precoding for massive MIMO systems," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2011–2021, May 2017.

[19] J.-C. Chen, C.-K. Wen, and K.-K. Wong, "Improved constant envelope multiuser precoding for massive MIMO systems," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1311–1314, Aug. 2014.

[20] K.-C. Ho and S.-H. Tsai, "A novel multiuser beamforming system with reduced complexity and beam optimizations," *IEEE Trans. Wireless Commun.*, vol. 18, no. 9, pp. 4544–4557, Sep. 2019.

[21] X. Gao, L. Dai, S. Zhou, A. M. Sayeed, and L. Hanzo, "Wideband beamspace channel estimation for millimeter-wave MIMO systems relying on lens antenna arrays," *IEEE Trans. Signal Process.*, vol. 67, no. 18, pp. 4809–4824, Sep. 2019.

[22] C. Pradhan, A. Li, L. Zhuo, Y. Li, and B. Vucetic, "Hybrid-precoding for mmWave multi-user communications in the presence of beam-misalignment," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 6083–6099, Sep. 2020.

[23] A. A. Badrudeen, C. Y. Leow, and S. Won, "Sub-connected structure hybrid precoding for millimeter-wave NOMA communications," *IEEE Wireless Commun. Lett.*, vol. 10, no. 6, pp. 1334–1338, Jun. 2021.

[24] A. A. Badrudeen, C. Y. Leow, and S. Won, "Performance analysis of hybrid beamforming precoders for multiuser millimeter wave NOMA systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 8739–8752, Aug. 2020.

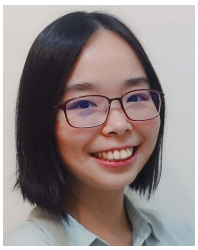
[25] Anokiwave. *AWMF-0108 Developer Kit Product Overview*. Accessed: Sep. 4, 2018. [Online]. Available: http://www.anokiwave.com/specifications/AWMF-0108-DK.pdf

[26] O. El Ayach, R. W. Heath, S. Abu-surra, S. Rajagopal, and Z. Pi, "The capacity optimality of beam steering in large millimeter wave MIMO systems," in *Proc. IEEE 13rd Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2012, pp. 100–104.

[27] K. P. Edwin Chong and H. Stanislaw Zak, *An Introduction to Optimization*, 4th ed. Hoboken, NJ, USA: Wiley, 2013.
[28] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 145–152, Jan. 2004.
[29] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MA, USA: Johns Hopkins Univ. Press, 1996.

Yan-Yin He received the M.S. degree from the Institute of Electrical and Control Engineering, National Chiao Tung University (now National Yang Ming Chiao Tung University, NYCU), Hsinchu, Taiwan, in 2019, where he is currently pursuing the Ph.D. degree.

His research interests include signal processing and transceiver beamforming design for the mmWave communications. He was selected as an Honorary Member of the Phi Tau Phi Scholastic Honor Society in 2018. He was awarded the Scholarship from the Institute of Electrical and Control Engineering, NYCU, in 2019, and the scholarships of Pan Wen-Yuan Foundation in both 2020 and 2021.



Hsiao-Chien (Angie) Chen received the B.S. degree in electrical and computer engineering and the M.S. degree in electrical and control engineering from the National Chiao Tung University (now National Yang Ming Chiao Tung University, NYCU) in 2016 and 2020, respectively, and the M.S. degree in electrical and computer engineering from the KU Leuven, Belgium, in 2020, by one-year overseas study under the Joint-Dual Degree Program.

From October 2016 to December 2017, she dedicated herself to the research of the signal processing in the wireless communication and worked closely with Wistron NeWeb Corporation. In 2018, she received the WeTech Qualcomm Global Scholarship and was recognized as one of the Potential Young Female Talents in the STEM field. She was awarded a Scholarship from NYCU and TOEFL for the overseas study in 2015.

Shang-Ho (Lawrence) Tsai (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Southern California (USC), USA, in August 2005.

From June 1999 to July 2002, he was with Silicon Integrated Systems Corporation (SiS), where he participated the VLSI design for DMT-ADSL systems. From September 2005 to January 2007, he was with MediaTek Inc. (MTK) participating in the VLSI design for MIMO-OFDM systems and standard specifications for IEEE 802.11n. Since February 2007, he has been with the Department of Electrical Engineering, National Chiao Tung University (now National Yang Ming Chiao Tung University), where he is currently a Professor. From June 2013 to December 2013, he was a Visiting Fellow with the Department of Electrical Engineering, Princeton University. His research interests are in the areas of signal processing for communications, statistical signal processing, and machine learning. He was awarded a Government Scholarship for Overseas Study from the Ministry of Education, Taiwan, from 2002 to 2005. He was awarded the Micron Teacher Award at 2018. He was in the Editorial Board of IEEE SigPort from 2018 to 2020. Since 2022, he has been an Editor of *APSIPA Transactions on Signal and Information Processing*.