MMSE Hybrid Beamforming for Multi-User Millimeter Wave MIMO Systems

Wenqian Ren, Jin Deng, and Xiantao Cheng, *Member, IEEE*

***Abstract*— Millimeter wave (mmWave) massive multiple-input multiple-output (MIMO) system has been recognized as a key technique for 5G communications. In terms of power consumption and hardware cost, it is hardly affordable to assign one radio frequency (RF) chain to each antenna in massive MIMO. Therefore, hybrid analog and digital beamforming is often adopted to reduce the number of RF chains. In this letter, we address the hybrid precoding and combining for multi-user mmWave MIMO systems. The goal is to minimize the bit error rate (BER) in downlink data transmission from the base station (BS) to multiple users. To fulfill this, we choose the sum of the users' mean square errors (sum-MSE) as the optimization objective, and develop an iterative algorithm to design the BS precoding matrices and the combining matrices of each user by gradually minimizing the sum-MSE. Compared with the current hybrid beamforming schemes, the proposed scheme can achieve much better BER performance.**

***Index Terms*— Millimeter wave (mmWave), multiple-input multiple-output (MIMO), hybrid analog and digital beamforming (HBF), minimum mean square error (MMSE).**

I. Introduction

DUE to the large bandwidth available, millimeter wave (mmWave) system manifests a wide application prospect in various wireless communications [1], [2]. To combat the severe propagation loss (including path loss, penetration loss, etc.) at the mmWave frequency band, massive multiple-input multiple-output (MIMO) is generally required [3]. If one dedicated radio frequency (RF) chain is assigned to each antenna in mmWave MIMO, the power consumption and hardware cost will be prohibitively high [4], [5]. As a result, the traditional full-digital beamforming is no longer applicable and the hybrid beamforming is preferred, which can effectively reduce the number of RF chains in mmWave MIMO [6], [7].

The hybrid beamforming (HBF) design can be regarded as an optimization problem. There are various design schemes based on different objective functions and diverse optimization criteria in the literature. So far, most existing schemes are designed to maximize the achievable rate [8], [9], [10], [11], [12], [13], [14], while the schemes for optimizing the bit error rate (BER) performance are rather scare [15], [16]. In [8], a spatially sparse beamforming algorithm via orthogonal matching pursuit (OMP) method is proposed. In [9], the hybrid beamforming problem is solved by resorting to the alternating minimization approach. In [10], a heuristic hybrid beamforming design is proposed. It is further extended to orthogonal frequency division multiplexing (OFDM) transmission in [11]. It is noticed that the schemes in [8], [9], [10], and [11] are designed mainly for mmWave single-user (SU) MIMO systems.[1] In [12], a two-stage zero-forcing (ZF) hybrid beamforming scheme is developed for mmWave multi-user (MU) MIMO systems. In [13], several hybrid beamforming schemes are proposed for mmWave MU MIMO with the aim of maximizing the system sum-rate. In [14], a linear successive allocation (SLA) based method is used to design the hybrid precoder and fully digital combiner for mmWave MU MIMO. Observe that the mean square error (MSE) performance of data symbols is closely related to the BER performance, the authors in [15] and [16] try to improve the BER performance by minimizing the symbol MSE. The hybrid precoding and combiner schemes in [15] and [16] are applicable for SU MIMO and MU MIMO, respectively.

In this letter, we will develop a hybrid beamforming scheme for mmWave MU MIMO systems. The novel contributions include:

(1) Unlike most of the existing hybrid schemes (e.g., [12], [13]), which try to maximize the achievable rate, we aim to minimize the BER. To achieve this, we take the sum of symbol MSEs (sum-MSE) of multiple users as the objective function. It is worthwhile to point out that our sum-MSE objective is not exactly the same as that in [16]. The main difference is that our objective uses different scaling factors for different users (see $\{\beta_k\}$ in (4) below), while the objective in [16] uses the same scaling factor (denoted by β) for all the users. Apparently, the objective in [16] is a special case of ours. With the more tunable parameters $\{\beta_k\}$, our objective function is more flexible than that in [16].

(2) We develop a novel iterative algorithm to design the hybrid precoding matrices of the base station (BS) and the hybrid combining matrices of each user. During each iteration, we update $\{\beta_k\}$, the precoding matrices and the combining matrices in turn, by monotonically decreasing the designed sum-MSE. Therefore, the developed algorithm is guaranteed to converge. Although the generalized eigenvalue decomposition (GEVD) HBF scheme in [16] adopts a similar sum-MSE objective, it obtains the receiver analog/digital combiner by

Manuscript received 31 May 2023; revised 15 August 2023 and 27 September 2023; accepted 22 October 2023. Date of publication 25 October 2023; date of current version 12 December 2023. This work was supported by Natural Science Foundation of Sichuan Province under Grant 2022NSFSC0486. The associate editor coordinating the review of this letter and approving it for publication was A. Li. *(Corresponding author: Xiantao Cheng.)*

The authors are with the National Key Laboratory of Wireless Communications, University of Electronic Science and Technology of China (UESTC), Chengdu 611731, China (e-mail: wenqianren@std.uestc.edu.cn; uestc_st_dj@163.com; xiantaocheng@163.com).

Digital Object Identifier 10.1109/LCOMM.2023.3327451

[1] In [10] and [11], multiple single-antenna users are also considered.

minimizing the lower/upper bound of its sum-MSE. In a nut shell, our algorithm is quite different from that in [16].

(3) The combination of the artfully designed objective function and the novel iterative algorithm, enables us to significantly lower the sum-MSE, thereby effectively improving the BER performance. This is verified through simulations.

The rest of the letter is organized as follows. The system model is briefly introduced in Section II. In Section III, we elaborately derive the MMSE hybrid beamforming scheme for MU MIMO systems. In Section IV, simulation results are provided to verify the advantages of the proposed scheme over the counterparts. Finally, we conclude the letter in Section V.

Notations: For a matrix $\mathbf{A}$, $\mathrm{Tr}\{\mathbf{A}\}$ and $|\mathbf{A}|$ denote the trace and the determinant of $\mathbf{A}$, respectively. $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and Hermitian transpose, respectively. $E\{\cdot\}$ denotes the expectation operator. $\mathrm{vec}(\mathbf{A})$ outputs a vector by catenating the columns of $\mathbf{A}$. $\mathrm{Re}\{\cdot\}$ denotes the real part of the argument. $\otimes$ denotes the Kronecker product. $\arg\{\cdot\}$ extracts the phase of the argument.

II. System Model

Consider a downlink multi-user MIMO system, in which a BS communicates with K independent users simultaneously. The BS is equipped with N_t transmit antennas and N_t^{RF} RF chains, while each user is equipped with N_r transmit antennas and N_r^{RF} RF chains. In view of the high power consumption and hardware cost of mmWave RF divices, we assume that $N_t^{\mathrm{RF}} \ll N_t$ and $N_r^{\mathrm{RF}} \ll N_r$.

The downlink channel $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$ from the BS to the k-th user can be expressed as

$$\mathbf{H}_k = \sqrt{\frac{N_r N_t}{L}} \sum_{l=1}^{L} \alpha_{l,k} \mathbf{a_r}(\theta_{l,k}^r) \mathbf{a}_t^H(\theta_{l,k}^t), \quad k = 1, 2, \cdots, K, \tag{1}$$

where K is the number of users, the subscript 'k' corresponds to the k-th user, L is the number of multipaths, $\alpha_{l,k} \sim \mathcal{CN}(0,1)$ is the complex gain of the l-th multipath, and $\theta_{l,k}^r$ ($\theta_{l,k}^t$) is the angle of arrival (departure) of the l-th multipath. Assume that uniform linear arrays (ULAs) are used at the BS and the users, then $\mathbf{a_r}(\theta_{l,k}^r) = \frac{1}{\sqrt{N_r}}[1, e^{j\pi \sin\theta_{l,k}^r}, \cdots, e^{j\pi(N_r-1)\sin\theta_{l,k}^r}]^T$ and $\mathbf{a_t}(\theta_{l,k}^t) = \frac{1}{\sqrt{N_t}}[1, e^{j\pi \sin\theta_{l,k}^t}, \cdots, e^{j\pi(N_t-1)\sin\theta_{l,k}^t}]^T$ are the array response vectors of the users and the BS, respectively.

Before data transmission, the BS first applies an $N_t^{\mathrm{RF}} \times KN_s$ baseband digital precoding matrix $\mathbf{V}_{\mathrm{B}}$, where N_s is the number of data streams for each user. We can write $\mathbf{V}_{\mathrm{B}}$ as $\mathbf{V}_{\mathrm{B}} = [\mathbf{V}_{\mathrm{B}_1}, \cdots, \mathbf{V}_{\mathrm{B}_K}]$, where $\mathbf{V}_{\mathrm{B}_k}$ corresponds to the k-th user. Then an $N_t \times N_t^{\mathrm{RF}}$ RF precoding matrix $\mathbf{V}_{\mathrm{RF}}$ is used. Each entry in $\mathbf{V}_{\mathrm{RF}}$ bears constant modulus, i.e., $|\mathbf{V}_{\mathrm{RF}}(i,j)|^2 = 1, \forall i, j$. Therefore, the transmitted signal at the BS can be expressed as

$$\mathbf{x} = \mathbf{V}_{\mathrm{RF}} \mathbf{V}_{\mathrm{B}} \mathbf{s} = \sum_{k=1}^{K} \mathbf{V}_{\mathrm{RF}} \mathbf{V}_{\mathrm{B}_k} \mathbf{s}_k, \tag{2}$$

where $\mathbf{s}_k$ is the $N_s \times 1$ symbol vector for the k-th user, and the total information symbol $\mathbf{s} = [\mathbf{s}_1^T, \cdots \mathbf{s}_K^T]^T$ satisfies $E[\mathbf{s}\mathbf{s}^H] = \mathbf{I}$. To meet the power budget at the BS, the precoding matrices $\mathbf{V}_k = \mathbf{V}_{\mathrm{RF}} \mathbf{V}_{\mathrm{B}_k}$ are constrained by $\sum_{k=1}^K \mathrm{Tr}\{\mathbf{V}_k \mathbf{V}_k^H\} \le P$, where P is the total transmitted power.

Similarly at the k-th user, it uses a $N_r \times N_r^{\mathrm{RF}}$ RF combining matrix $\mathbf{W}_{\mathrm{RF}_k}$, followed by a $N_r^{\mathrm{RF}} \times N_s$ baseband digital combining matrix $\mathbf{W}_{\mathrm{B}_k}$. The entries in $\mathbf{W}_{\mathrm{RF}_k}$ are also of constant modulus, i.e., $|\mathbf{W}_{\mathrm{RF}_k}(m,l)|^2 = 1, \forall m, l$. Let $\mathbf{W}_k = \mathbf{W}_{\mathrm{RF}_k} \mathbf{W}_{\mathrm{B}_k}$, the received signal vector at the k-th user can be expressed as

$$\begin{aligned}\mathbf{y}_k &= \mathbf{W}_k^H(\mathbf{H}_k \mathbf{x} + \mathbf{n}_k) \\ &= \mathbf{W}_k^H \mathbf{H}_k \mathbf{V}_k \mathbf{s}_k + \mathbf{W}_k^H \mathbf{H}_k \sum_{i \neq k}^{K} \mathbf{V}_i \mathbf{s}_i + \tilde{\mathbf{n}}_k,\end{aligned} \tag{3}$$

where the $N_s \times 1$ vector $\mathbf{n}_k \sim \mathcal{CN}(0, \sigma_k^2 \mathbf{I}_{N_t})$ denotes the additive complex Gaussian noise vector, and $\tilde{\mathbf{n}}_k = \mathbf{W}_k^H \mathbf{n}_k$. Since $\mathbf{n}_k$ and $\mathbf{s_k}$ are independent, we have $E[\mathbf{s}_k \mathbf{n}_k^H] = \mathbf{0}$.

The MSE between the received signal and the transmitted signal at k-th user can be defined as

$$\begin{aligned}\mathrm{MSE}_k &\triangleq E\left\{\left\|\mathbf{s}_k - \beta_k^{-1}\mathbf{y}_k\right\|^2\right\} \\ &= \mathrm{Tr}\left\{(\mathbf{s}_k - \beta_k^{-1}\mathbf{y}_k)(\mathbf{s}_k - \beta_k^{-1}\mathbf{y}_k)^H\right\} \\ &= \mathrm{Tr}\Big\{\beta_k^{-2}\mathbf{W}_k^H \mathbf{H}_k \mathbf{V}_k \mathbf{V}_k^H \mathbf{H}_k^H \mathbf{W}_k \\ &\quad + \beta_k^{-2}\mathbf{W}_k^H \mathbf{H}_k \sum_{i \neq k}^{K} (\mathbf{V}_i \mathbf{V}_i^H) \mathbf{H}_k^H \mathbf{W}_k \\ &\quad - \beta_k^{-1}\mathbf{W}_k^H \mathbf{H}_k \mathbf{V}_k - \beta_k^{-1}\mathbf{V}_k^H \mathbf{H}_k^H \mathbf{W}_k \\ &\quad + \beta_k^{-2}\sigma_k^2 \mathbf{W}_k^H \mathbf{W}_k + \mathbf{I}_{N_s}\Big\},\end{aligned} \tag{4}$$

where β_k is a scaling factor. Then the sum-MSE for K users can be written as $\sum_{k=1}^K \mathrm{MSE}_k$.

As a result, the optimization problem for hybrid beamforming design can be formulated as

$$\begin{aligned}\underset{\{\mathbf{V}_{\mathrm{B}_k}\}, \mathbf{V}_{\mathrm{RF}}, \{\mathbf{W}_{\mathrm{B}_k}\}, \{\mathbf{W}_{\mathrm{RF}_k}\}, \{\beta_k\}}{\text{minimize}} \quad & \sum_{k=1}^{K} \mathrm{MSE}_k \\ \text{subject to} \quad & \sum_{k=1}^{K} \mathrm{Tr}\left\{\mathbf{V}_k \mathbf{V}_k^H\right\} \le P \\ & |\mathbf{V}_{\mathrm{RF}}(i,j)|^2 = 1, \quad \forall i, j \\ & |\mathbf{W}_{\mathrm{RF}_k}(m,l)|^2 = 1, \quad \forall m, l.\end{aligned} \tag{5}$$

In this letter, we assume that $\{\mathbf{H}_k\}$ are known. This is reasonable since accurate channel estimation is available for mmWave hybrid MIMO systems (see e.g., [17]). Besides, although we focus on flat-fading channels, the proposed scheme can be easily extended to the OFDM system over frequency-selective channels.

III. MMSE Hybrid Beamforming Design

A. Hybrid Precoder Design

In this subsection, we temporarily assume that the combining matrices $\mathbf{W}_{\mathrm{B}_k}$ and $\mathbf{W}_{\mathrm{RF}_k}$ are known and fixed, and thus focus on the hybrid precoding design. The optimization problem can be reduced to

$$\begin{aligned}\underset{\mathbf{V}_{\mathrm{B}_k},\mathbf{V}_{\mathrm{RF}},\beta_k}{\text{minimize}} \quad & \sum_{k=1}^{K}\mathrm{MSE}_k \\ \text{subject to} \quad & \mathrm{Tr}\left\{\mathbf{V}_k\mathbf{V}_k^H\right\}=P_k,\forall k, \\ & \sum_{k=1}^{K}P_k\le P, \\ & \left|\mathbf{V}_{\mathrm{RF}}(i,j)\right|^2=1,\quad \forall i,j. \end{aligned} \tag{6}$$

where P_k is the transmit power for the k-th user. For brevity, let $\bar{\mathbf{H}}_k=\mathbf{W}_{\mathrm{B}_k}^H\mathbf{W}_{\mathrm{RF}_k}^H\mathbf{H}_k$ be the equivalent channel, and $\mathbf{V}_{\mathrm{D}_k}=\beta_k^{-1}\mathbf{V}_{\mathrm{B}_k}$ be the unnormalized baseband precoder, hence the MSE_k for the k-th user in (4) can be rewritten as

$$\begin{aligned}\mathrm{MSE}_k=\mathrm{Tr}\Bigg\{&\bar{\mathbf{H}}_k\mathbf{V}_{\mathrm{RF}}\sum_{i=1}^{K}(\mathbf{V}_{\mathrm{D}_i}\mathbf{V}_{\mathrm{D}_i}^H)\mathbf{V}_{\mathrm{RF}}^H\bar{\mathbf{H}}_k^H \\ &-\bar{\mathbf{H}}_k\mathbf{V}_{\mathrm{RF}}\mathbf{V}_{\mathrm{D}_k}-\mathbf{V}_{\mathrm{D}_k}^H\mathbf{V}_{\mathrm{RF}}^H\bar{\mathbf{H}}_k^H \\ &+\beta_k^{-2}\sigma_k^2\mathbf{W}_k^H\mathbf{W}_k+\mathbf{I}_{N_s}\Bigg\}. \end{aligned}\tag{7}$$

Based on the first constraint in (6), we have

$$\beta_k^{-1}=\sqrt{\frac{\mathrm{Tr}\left\{\mathbf{V}_{\mathrm{RF}}\mathbf{V}_{\mathrm{D}_k}\mathbf{V}_{\mathrm{D}_k}^H\mathbf{V}_{\mathrm{RF}}^H\right\}}{P_k}},\tag{8}$$

By substituting $\{\beta_k^{-1}\}_{k=1}^K$ into (7) and (6), and using the Cauchy inequality, we can readily obtain the optimal $\{P_k\}_{k=1}^K$: $P_k=\sigma_k a_k P/(\sum_{k=1}^K\sigma_k a_k)$, where $a_k=\sqrt{\mathrm{Tr}\left\{\mathbf{V}_{\mathrm{RF}}\mathbf{V}_{\mathrm{D}_k}\mathbf{V}_{\mathrm{D}_k}^H\mathbf{V}_{\mathrm{RF}}^H\right\}\mathrm{Tr}\left\{\mathbf{W}_k^H\mathbf{W}_k\right\}}$.

Substituting (8) into (7), we can use the Karush-Kuhn-Tucker (KKT) conditions to obtain the optimal $\mathbf{V}_{\mathrm{D}_k}$ in terms of $\mathbf{V}_{\mathrm{RF}}$, as shown below

$$\mathbf{V}_{\mathrm{D}_k}=\left(\mathbf{V}_{\mathrm{RF}}^H\left(\sum_{i=1}^{K}\bar{\mathbf{H}}_i^H\bar{\mathbf{H}}_i\right)\mathbf{V}_{\mathrm{RF}}+\sigma_k^2 w_k\mathbf{V}_{\mathrm{RF}}^H\mathbf{V}_{\mathrm{RF}}\right)^{-1}\mathbf{V}_{\mathrm{RF}}^H\bar{\mathbf{H}}_k^H,\tag{9}$$

where $w_k\triangleq\frac{\mathrm{Tr}\{\mathbf{W}_k^H\mathbf{W}_k\}}{P_k}$.

With $\{\mathbf{V}_{\mathrm{D}_k}\}$ and $\{\beta_k\}$, let us consider $\mathbf{V}_{\mathrm{RF}}$. Due to the non-convex unit-modulus constraints, it is hard to obtain $\mathbf{V}_{\mathrm{RF}}$ by directly solving (6). Using the identity that $\mathrm{Tr}(\mathbf{ABCD})=\mathrm{vec}(\mathbf{A}^T)^T(\mathbf{D}^T\otimes\mathbf{B})\mathrm{vec}(\mathbf{C})$, we can rewrite the objective function in (6) as

$$f(\mathbf{v})=\sum_{k=1}^{K}\mathrm{MSE}_k(\mathbf{V}_{\mathrm{RF}})=\mathbf{v}^H\mathbf{A}\mathbf{v}-2\mathrm{Re}(\mathbf{v}^H\mathbf{B}),\tag{10}$$

where $\mathbf{v}=\mathrm{vec}(\mathbf{V}_{\mathrm{RF}})$, $\mathbf{B}=\sum_{k=1}^K\bar{\mathbf{H}}_k^H\mathbf{V}_{\mathrm{D}_k}^H$, $\mathbf{A}=(\sum_{k=1}^K\mathbf{V}_{\mathrm{D}_k}\mathbf{V}_{\mathrm{D}_k}^H)^T\otimes(\sum_{k=1}^K\bar{\mathbf{H}}_k^H\bar{\mathbf{H}}_k)$, and the constants independent of $\mathbf{V}_{\mathrm{RF}}$ are ignored.

We adopt the majorization-minimization (MM) algorithm [18] to obtain $\mathbf{V}_{\mathrm{RF}}$ by iteratively minimizing $f(\mathbf{v})$. At the t-th iteration of the MM algorithm, we need to construct a surrogate function $g(\mathbf{v},\mathbf{v}_{t-1})$, which satisfies that $g(\mathbf{v},\mathbf{v}_{t-1})\ge f(\mathbf{v}),\forall\mathbf{v}$ and $g(\mathbf{v}_{t-1},\mathbf{v}_{t-1})=f(\mathbf{v}_{t-1})$ with $\mathbf{v}_{t-1}$ representing the $\mathbf{v}$ obtained in the $(t-1)$-th MM iteration. By using the results in [18], we can construct the surrogate function as

$$\begin{aligned}g(\mathbf{v},\mathbf{v}_{t-1})=&\lambda_{max}(\mathbf{A})\mathbf{v}^H\mathbf{v}+2\mathrm{Re}(\mathbf{v}^H(\mathbf{A}-\lambda_{max}(\mathbf{A})\mathbf{I})\mathbf{v}_{t-1}) \\ &+\mathbf{v}_{t-1}^H(\lambda_{max}(\mathbf{A})\mathbf{I}-\mathbf{A})\mathbf{v}_{t-1}-2\mathrm{Re}(\mathbf{v}^H\mathbf{B}),\end{aligned}\tag{11}$$

where $\lambda_{max}(\mathbf{A})$ represents the largest eigenvalue of $\mathbf{A}$.

At the t-th MM iteration, $\mathbf{v}_t$ can be obtained by minimizing $g(\mathbf{v},\mathbf{v}_{t-1})$ under the unit-modulus constrains. This is equivalent to solve the following:

$$\begin{aligned}\underset{\mathbf{v}}{\text{minimize}}\quad & \tilde{g}(\mathbf{v},\mathbf{v}_{t-1})=2\mathrm{Re}(\mathbf{v}^H\mathbf{q}_{t-1}) \\ \text{subject to}\quad & |\mathbf{v}_m|^2=1,\quad\forall m=1\cdots N_t\times N_t^{\mathrm{RF}},\end{aligned}\tag{12}$$

where $\mathbf{q}_{t-1}=(\mathbf{A}-\lambda_{max}(\mathbf{A})\mathbf{I}_{N_t\times N_t^{\mathrm{RF}}})\mathbf{v}_{t-1}-\mathbf{B}$, we obtain $\tilde{g}(\mathbf{v},\mathbf{v}_{t-1})$ from $g(\mathbf{v},\mathbf{v}_{t-1})$ by using the fact that the entries in $\mathbf{v}$ bear unit modulus and by keeping only the terms revelent to $\mathbf{v}$. The optimal solution to (12) can be easily obtained as

$$\mathbf{v}_t=-e^{j\arg(\mathbf{q}_{t-1})}.\tag{13}$$

Obviously, we have $f(\mathbf{v}_t)\le g(\mathbf{v}_t,\mathbf{v}_{t-1})\le g(\mathbf{v}_{t-1},\mathbf{v}_{t-1})=f(\mathbf{v}_{t-1})$. That is, each MM iteration will monotonically decrease the sum-MSE. After sufficient MM iterations, we will obtain the new $\mathbf{V}_{\mathrm{RF}}$.

B. Hybrid Combiner Design

With the obtained $\{\mathbf{V}_{\mathrm{B}_k}\}$, $\mathbf{V}_{\mathrm{RF}}$ and $\{\beta_k\}$ in the above subsection, we move to derive the optimal combiner at each user. That is, in this subsection we assume that $\{\mathbf{V}_{\mathrm{B}_k}\}$, $\mathbf{V}_{\mathrm{RF}}$ and $\{\beta_k\}$ are fixed. It is worthwhile to point out that the combining matrices $\mathbf{W}_{\mathrm{B}_k}$ and $\mathbf{W}_{\mathrm{RF}_k}$ are only included in MSE_k. Hence the optimization of $\mathbf{W}_{\mathrm{B}_k}$ and $\mathbf{W}_{\mathrm{RF}_k}$ is equivalent to the minimization of MSE_k. As a result, we consider the following minimization problem:

$$\begin{aligned}\underset{\mathbf{W}_{\mathrm{B}_k},\mathbf{W}_{\mathrm{RF}_k}}{\text{minimize}}\quad & \mathrm{MSE}_k \\ \text{subject to}\quad & |\mathbf{W}_{\mathrm{RF}_k}(m,l)|^2=1,\quad\forall m,l.\end{aligned}\tag{14}$$

Let $\hat{\mathbf{H}}_{k,k}=\beta_k^{-1}\mathbf{H}_k\mathbf{V}_{\mathrm{RF}}\mathbf{V}_{\mathrm{B}_k}$ be the equivalent channel, and $\hat{\mathbf{H}}_{k,i}=\beta_k^{-1}\mathbf{H}_k\mathbf{V}_{\mathrm{RF}}\mathbf{V}_{\mathrm{B}_i}$ be the interference information, therefore the MSE_k for the k-th user can be rewritten as

$$\begin{aligned}\mathrm{MSE}_k=\mathrm{Tr}\Bigg\{&\mathbf{W}_{\mathrm{B}_k}^H\mathbf{W}_{\mathrm{RF}_k}^H\hat{\mathbf{H}}_{k,k}\hat{\mathbf{H}}_{k,k}^H\mathbf{W}_{\mathrm{RF}_k}\mathbf{W}_{\mathrm{B}_k} \\ &+\mathbf{W}_{\mathrm{B}_k}^H\mathbf{W}_{\mathrm{RF}_k}^H\sum_{i\ne k}^{K}(\hat{\mathbf{H}}_{k,i}\hat{\mathbf{H}}_{k,i}^H)\mathbf{W}_{\mathrm{RF}_k}\mathbf{W}_{\mathrm{B}_k} \\ &-\mathbf{W}_{\mathrm{B}_k}^H\mathbf{W}_{\mathrm{RF}_k}^H\hat{\mathbf{H}}_{k,k}-\hat{\mathbf{H}}_{k,k}^H\mathbf{W}_{\mathrm{RF}_k}\mathbf{W}_{\mathrm{B}_k} \\ &+\beta_k^{-2}\sigma_k^2\mathbf{W}_{\mathrm{B}_k}^H\mathbf{W}_{\mathrm{RF}_k}^H\mathbf{W}_{\mathrm{RF}_k}\mathbf{W}_{\mathrm{B}_k}+\mathbf{I}_{N_s}\Bigg\}.\end{aligned}\tag{15}$$

According to the KKT conditions, we can obtain the optimal $\mathbf{W}_{\mathrm{B}_k}$ as shown below:

$$\begin{aligned}\mathbf{W}_{\mathrm{B}_k}=&\left(\mathbf{W}_{\mathrm{RF}_k}^H\left(\sum_{i=1}^{K}\hat{\mathbf{H}}_{k,i}\hat{\mathbf{H}}_{k,i}^H\right)\mathbf{W}_{\mathrm{RF}_k}+\beta_k^{-2}\sigma_k^2\mathbf{W}_{\mathrm{RF}_k}^H\mathbf{W}_{\mathrm{RF}_k}\right)^{-1} \\ &\times\mathbf{W}_{\mathrm{RF}_k}^H\hat{\mathbf{H}}_{k,k}.\end{aligned}\tag{16}$$

TABLE I
PROPOSED HYBRID BEAMFORMING DESIGN

Input: $\{\mathbf{H}_k\}$, the number of iterations I
Initialization: $\mathbf{V}_{\rm RF}^0$, $\mathbf{V}_{{\rm B}_k}^0$, $\mathbf{W}_{{\rm B}_k}^0$, $\mathbf{W}_{{\rm RF}_k}^0$.
for $1 \leq i \leq I$**:**
1: Update β_k^{-1} using (8);
2: Compute $\mathbf{V}_{{\rm D}_k}$ using (9), and $\mathbf{V}_{{\rm B}_k} = \beta_k \mathbf{V}_{{\rm D}_k}$;
3: Compute $\mathbf{V}_{\rm RF}$ using the MM algorithm:
  1)During each MM iteration, update $\mathbf{v}$ using (13)
  2) After MM iterations terminate, use the finally obtained $\mathbf{v}$ to update $\mathbf{V}_{\rm RF}$;
4: Compute $\mathbf{W}_{{\rm B}_k}$ using (16);
5: Compute $\mathbf{W}_{{\rm RF}_k}$ using the MM algorithm:
  1) During each MM iteration, update $\mathbf{w}$ using (18)
  2) After MM iterations terminate, use the finally obtained $\mathbf{w}$ to update $\mathbf{W}_{{\rm RF}_k}$;
end for
Output: $\{\beta_k\}$,$\{\mathbf{V}_{{\rm B}_k}\}$, $\mathbf{V}_{\rm RF}$, $\{\mathbf{W}_{{\rm B}_k}\}$, $\{\mathbf{W}_{{\rm RF}_k}\}$

Similar to the updating of $\mathbf{V}_{\rm RF}$, we can use the MM algorithm to obtain $\mathbf{W}_{{\rm RF}_k}$. Since the derivations are quite similar to $\mathbf{V}_{\rm RF}$, we ignore the details and directly provide the results. Specifically, at the s-th MM iteration, we need to solve the following:

$$\begin{aligned} &\underset{\mathbf{w}}{\text{minimize}} \quad 2\text{Re}(\mathbf{w}^H \mathbf{p}_{s-1}) \\ &\text{subject to } |\mathbf{w}_m|^2 = 1, \quad \forall m = 1 \cdots N_r \times N_r^{\rm RF}, \end{aligned} \tag{17}$$

where $\mathbf{p}_{s-1} = (\mathbf{C} - \lambda_{max}(\mathbf{C})\mathbf{I}_{N_r \times N_r^{\rm RF}})\mathbf{w}_{s-1} - \mathbf{D}$, $\mathbf{w} = \text{vec}(\mathbf{W}_{{\rm RF}_k})$, $\mathbf{D} = \hat{\mathbf{H}}_{k,k}\mathbf{W}_{{\rm B}_k}^H$, and $\mathbf{C} = (\mathbf{W}_{{\rm B}_k}\mathbf{W}_{{\rm B}_k}^H)^T \otimes (\beta_k^{-2}\sigma_k^2\mathbf{I} + \sum_{i=1}^K \hat{\mathbf{H}}_{k,i}\hat{\mathbf{H}}_{k,i}^H)$. The optimal solution is given by

$$\mathbf{w}_s = -e^{j\arg(\mathbf{p}_{s-1})} \tag{18}$$

According to the MM algorithm, the sum-MSE will decrease after each iteration. After sufficient MM iterations, we can obtain the new $\mathbf{W}_{{\rm RF}_k}$.

C. Summary

The proposed hybrid beamforming scheme is summarized in Table I. It runs in an iterative manner. For the initialization, we set $\mathbf{W}_{{\rm B}_k}^0 = \mathbf{I}_{N_r^{\rm RF} \times N_s}$, which is composed of the first N_s columns of the identity matrix $\mathbf{I}_{N_r^{\rm RF}}$, and set $\mathbf{V}_{{\rm B}_k}^0 = \mathbf{I}_{N_t^{\rm RF} \times N_s}$. Denote $\mathbf{A}_k^r = [\mathbf{a_r}(\theta_{1,k}^r), \cdots, \mathbf{a_r}(\theta_{L,k}^r)]$. When $N_r^{\rm RF} > L$, we set $\mathbf{W}_{{\rm RF}_k}^0 = [\mathbf{A}_k^r, \mathbf{0}_{N_r \times (N_r^{\rm RF} - L)}]$ with $\mathbf{0}_{m \times n}$ being an all-zero matrix of size $m \times n$. Otherwise, we choose $N_r^{\rm RF}$ columns in $\mathbf{A}_k^r$ to construct $\mathbf{W}_{{\rm RF}_k}^0$. Besides, we set $\mathbf{V}_{\rm RF}^0$ in a similar way to $\mathbf{W}_{{\rm RF}_k}^0$.

During each iteration, the proposed scheme updates $\{\beta_k\}$,$\{\mathbf{V}_{{\rm B}_k}\}$, $\mathbf{V}_{\rm RF}$, $\{\mathbf{W}_{{\rm B}_k}\}$, $\{\mathbf{W}_{{\rm RF}_k}\}$ in turn. Each updating operation (see (8), (9), (13), (16) and (18)) will monotonically decrease the sum-MSE objective. Therefore, after sufficient iterations, the proposed scheme is guaranteed to converge.

Suppose that I iterations are used, it is easy to know that the computational complexity of the proposed scheme is $\mathcal{O}(IKT_M N_t^2 (N_t^{\rm RF})^2)$, where T_M represents the number of MM iterations to compute $\mathbf{V}_{\rm RF}$ and $\mathbf{W}_{{\rm RF}_k}$. This complexity is close to those of the schemes in [12], [13], and [16].

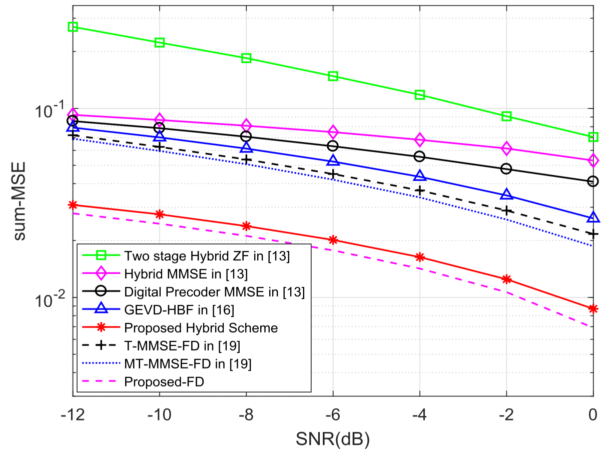

Fig. 1. $K = 2$, Sum-MSE Performance comparison for different schemes over varying SNR, QPSK.

IV. SIMULATION AND DISCUSSION

During simulations, the multipath channels are randomly generated with $L = 4$, $\alpha_{l,k} \sim \mathcal{CN}(0,1)$, $\theta_{l,k}^r(\theta_{l,k}^t) \sim \mathcal{U}[0, 2\pi]$, where $\mathcal{U}[a,b]$ denotes the uniform distribution over the range $[a,b]$. Besides, we use $N_t = 64$, $N_t^{\rm RF} = 4$, $N_r = 16$, $N_r^{\rm RF} = 2$ and $N_s = 2$. If not specified elsewhere, we consider $K = 2$ users. The total transmitted power $P = 1$ and thus the signal-to-noise ratio (SNR) is defined as $1/\sigma^2$. To generate the BER figures below, we use QPSK and 16QAM.

Fig.1 presents the sum-MSE performance of the proposed hybrid beamforming scheme. Here the sum-MSE is uniformly defined as: $\sum_{k=1}^K E\{\|\mathbf{s}_k - \tilde{\mathbf{s}}_k\|^2\}$, where $\tilde{\mathbf{s}}_k$ is the output of the receiver combiner and $\tilde{\mathbf{s}}_k = \beta_k^{-1}\mathbf{y}_k$ for the proposed scheme. For comparison, we consider the counterpart schemes in [12], [13], and [16]. Note that the scheme in [12] is the same as the 'two stage hybrid ZF' scheme in [13], we thus choose not to mention [12] in the figure. Moreover, in [13], the 'digital precoder MMSE' scheme adopts fully digital precoder at the BS and hybrid combiner at each user, while 'hybrid MMSE' scheme adopts hybrid precoder and hybrid combiner. Among the hybrid schemes, the proposed scheme performs the best sum-MSE performance. This is reasonable since the schemes in [13] are designed to maximize the sum-rate rather than to minimize the sum-MSE. Although the scheme in [16] considers a similar sum-MSE metric, it obtains the transmitter analog precoder based on the assumption that $\mathbf{V}_{\rm RF}^H\mathbf{V}_{\rm RF} \approx N_t\mathbf{I}_{N_t^{\rm RF}}$, obtains the receiver analog/digital combiner by minimizing the lower/upper bound of its sum-MSE. As mentioned in the introduction section, our sum-MSE is more flexible than that in [16]. Moreover, the proposed scheme is derived by directly minimizing the sum-MSE. Therefore, the proposed scheme remarkably outperforms the scheme in [16] in terms of sum-MSE performance.

In Fig.1, three fully digital (FD) schemes are used for comparison. The 'Proposed-FD' scheme is obtained from the proposed hybrid scheme by setting $N_t^{\rm RF} = N_t$, $N_r^{\rm RF} = N_r$, $\mathbf{V}_{\rm RF} = \mathbf{I}_{N_t}$ and $\mathbf{W}_{\rm RF} = \mathbf{I}_{N_r}$. The total-MMSE (T-MMSE) FD scheme and the modified T-MMSE (MT-MMSE) FD scheme try to minimize the objectives: $\sum_{k=1}^K E\{\|\mathbf{s}_k - \mathbf{y}_k\|^2\}$ and $\sum_{k=1}^K E\{\|\mathbf{s}_k - \beta^{-1}\mathbf{y}_k\|^2\}$, respectively [16], [19]. Due to

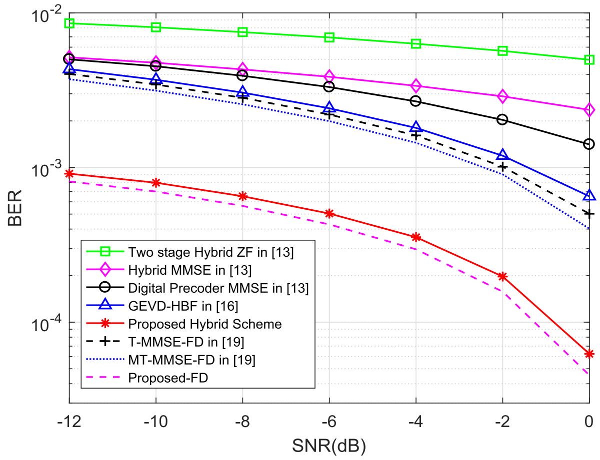

Fig. 2. $K = 2$, BER Performance comparison for different schemes over varying SNR, QPSK.

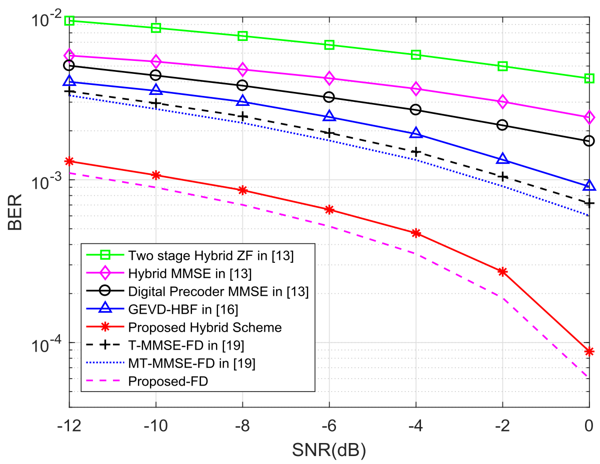

Fig. 3. $K = 2$, BER Performance comparison for different schemes over varying SNR, 16QAM.

our more flexible objective (see (4)), the proposed FD scheme can achieve lower sum-MSE than the other two FD schemes. Besides, the sum-MSE performance of the proposed hybrid scheme is close to that of the proposed FD scheme.

We compare the BER performance for the proposed scheme and the counterpart schemes in Fig.2 and Fig.3, where QPSK and 16QAM are considered, respectively. The BER curves are obtained by averaging the BER results over the two users. As mentioned above, the BER performance is closely related to the sum-MSE performance [15], [16]. Specifically, the lower sum-MSE, the better BER performance. Consistent with the observations in Fig.1, the proposed scheme bears the best BER performance among the hybrid scehmes.

V. Conclusion

In this letter, we proposed a hybrid beamforming scheme for multi-user mmWave MIMO systems. To optimize the BER performance, we chose the symbol sum-MSE of multiple users as the objective function. Then we developed an iterative method to obtain the analog/digital precoder at the BS and the analgo/digital combiner at each user, by monotonically minimizing the sum-MSE objective function. Compared with the hybrid counterparts, the proposed scheme can provide much better BER performance while bearing a similar computational complexity.

References

[1] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, Jun. 2011.

[2] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino, "Millimeter-wave massive MIMO: The next wireless revolution?" *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 56–62, Sep. 2014.

[3] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[4] L. Wei, R. Q. Hu, Y. Qian, and G. Wu, "Key elements to enable millimeter wave communications for 5G wireless systems," *IEEE Wireless Commun.*, vol. 21, no. 6, pp. 136–143, Dec. 2014.

[5] R. W. Heath Jr., N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

[6] A. Alkhateeb, O. E. 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.

[7] S. Han, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 186–194, Jan. 2015.

[8] O. E. 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.

[9] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485–500, Apr. 2016.

[10] 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.

[11] 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.

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

[13] D. H. N. Nguyen, L. B. Le, T. Le-Ngoc, and R. W. Heath Jr., "Hybrid MMSE precoding and combining designs for mmWave multiuser systems," *IEEE Access*, vol. 5, pp. 19167–19181, 2017.

[14] W. Utschick, C. Stöckle, M. Joham, and J. Luo, "Hybrid LISA precoding for multiuser millimeter-wave communications," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 752–765, Feb. 2018.

[15] T. Lin, J. Cong, Y. Zhu, J. Zhang, and K. Ben Letaief, "Hybrid beamforming for millimeter wave systems using the MMSE criterion," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3693–3708, May 2019.

[16] J. Cong, T. Lin, and Y. Zhu, "Hybrid MMSE beamforming for multiuser millimeter-wave communication systems," *IEEE Commun. Lett.*, vol. 22, no. 11, pp. 2390–2393, Nov. 2018.

[17] X. Cheng, C. Tang, and Z. Zhang, "Accurate channel estimation for millimeter-wave MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 5159–5163, May 2019.

[18] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, Feb. 2017.

[19] J. Joung and Y. H. Lee, "Regularized channel diagonalization for multiuser MIMO downlink using a modified MMSE criterion," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1573–1579, Apr. 2007.