The final version of record is available at http://dx.doi.org/10.1109/LCOMM.2020.2983018

Hybrid beamforming in MU-MIMO using partial interfering beam feedback

Silpa S. Nair, Srikrishna Bhashyam

***Abstract*—We propose a hybrid beamforming scheme with partial interfering beam feedback for a codebook-based multi-user multiple-input multiple-output (MU-MIMO) system, where users feed back information only about the top-p transmit beams. For the analog part of the precoding, we consider two codebooks, the conventional Discrete Fourier Transform (DFT) codebook with uniform amplitude beamforming vectors and the Taylor codebook with non-uniform amplitude beamforming vectors. For the digital precoding part, the effective channel matrix is approximated and used for zero-forcing (ZF). We also propose a beam pairing algorithm that results in reduced inter-beam interference and simplifies beam and user selection in MU-MIMO. When p is equal to the number of beams in the transmit codebook, the proposed scheme includes an existing scheme with full effective channel matrix feedback as a special case. Numerical results show that the proposed hybrid beamforming performs better than an existing hybrid precoding scheme based on channel reconstruction.**

I. Introduction

Millimeter wave (mmWave) systems [1] use a large number of antennas to overcome path loss and increase range. To reduce complexity and power consumption, hybrid precoding that combines analog and digital precoding with a reduced number of RF chains is widely studied [2–8]. Under full channel state information (CSI), hybrid precoding and combining designs exploiting the sparse nature of the mmWave channels are proposed in [2]. Hybrid precoding in a large multi-input multi-output (MIMO) setting with full CSI is studied in [3]. Practical systems obtain CSI using codebook-based training and limited feedback [4, 5]. Recent studies have focused on hybrid beamforming for limited feedback multi-user (MU)-MIMO systems [6–8]. In [6], the authors proposed a hybrid beamforming technique in which the analog precoder and combiners are designed using the best codewords from beamtraining, and digital precoder is zero-forcing (ZF) on the corresponding effective channel that includes the analog precoder, channel and the analog combiners. This method assumed full feedback of the effective channel matrix, which requires each user to feedback information about all beams in the transmit codebook. Directional precoding based on feedback only about the best transmit-receive beam pair from each user was proposed in [7]. In [8], this was generalized to use feedback about the top-q beam pairs for each user.

In this letter, we: (1) propose hybrid beamforming based on partial interfering beam feedback, *i.e.*, partial knowledge of the effective channel matrix, (2) consider the use of the Taylor codebook (non-uniform magnitude) [9, 10] for the analog part of the hybrid precoder, and (3) propose a beam grouping algorithm to choose the beams that can be used simultaneously in MU-MIMO for a given hybrid beamforming design. In the proposed system, each user sends feedback about the top-p transmitter beams for its best receiver beam. When p is equal to the number of beams in the transmit codebook, this scheme reduces to the scheme in [6]. Unlike [8], where the top-q transmit-receive beam pair information is used to obtain a rank-q reconstruction of the channel, we use the top-p beam information to approximate the effective channel matrix directly. The analog part of the hybrid precoder is usually designed using only phase shifters [2, 4, 5] using the Discrete Fourier Transform (DFT) codebook (uniform magnitude). But, the possibility of practical implementation of non-uniform magnitudes in analog precoder has been recently studied for mmWave systems [8, 10]. We consider the Taylor codebook. Once we have the hybrid precoder design, the proposed beam pairing algorithm helps in: (1) reducing inter-beam interference since the partial feedback does not fully describe the interference as in [6], and (2) simplifying the beam and user selection problem. Finally, we compare the proposed scheme with [6] and [8]. Simulation results show that the proposed hybrid precoding scheme: (1) bridges the gap between directional beamforming (corresponding to $p = 1$) based on best beam feedback and the scheme in [6] with feedback about all transmit codebook beams, and (2) outperforms the scheme in [8] when the Taylor codebook is used.

II. System model

We consider a downlink MU-MIMO transmission scenario with one base station (BS) having N_t transmit antennas and U users each with N_r receive antennas. During each time slot, the BS serves K out of the U users simultaneously sending K streams of data, one for each of the K users (See Fig. 1). The received vector $\mathbf{y}_k \in \mathbb{C}^{N_r \times 1}$ at the k^{th} user is given by: $\mathbf{y}_k = \mathbf{H}_k \mathbf{F}_A \mathbf{F}_D \mathbf{s} + \mathbf{n}_k$, where $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix corresponding to the k^{th} user, $\mathbf{F}_A \in \mathbb{C}^{N_t \times K}$ and $\mathbf{F}_D \in \mathbb{C}^{K \times K}$ are the analog and digital precoding matrices, $\mathbf{s} = [s_1, s_2, \ldots, s_K]^T$ is the Gaussian input signal which satisfies $\mathbb{E}[s_k] = 0$, and $\mathbb{E}[|s_k|^2] = 1, \ \forall\ k$, and $\mathbf{n}_k \in \mathbb{C}^{N_r \times 1}$ is the complex white Gaussian noise vector with zero mean and variance $\mathbb{E}[\mathbf{n}_k \mathbf{n}_k^H] = \mathbf{I}_{N_r}$. We assume that all streams use equal power $P_k = P_t/K, \ \forall\ k$, where P_t is the total power at the BS. Suppose $\mathbf{F_A F_D} = [\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_K]$, the total input vector from the BS is $\mathbf{x} = \sum_{k=1}^{K} \mathbf{f}_k s_k$ with $\|\mathbf{f}_k\|_2^2 = P_t/K \ \forall\ k$. If the k^{th} user uses the combiner $\mathbf{g}_k \in \mathbb{C}^{N_r \times 1}$ to get $\hat{s}_k$, the achievable sum-rate R_{sum} is

The authors are with Indian Institute of Technology Madras, Chennai, India.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/LCOMM.2020.2983018

$$R_{sum} = \sum_{k=1}^{K} \log\left(1 + \frac{|\mathbf{g}_k^H \mathbf{H}_k \mathbf{f}_k|^2}{1 + \sum_{i\neq k} |\mathbf{g}_k^H \mathbf{H}_k \mathbf{f}_i|^2}\right). \quad (1)$$

The BS has $N_t = N_{t_h} \times N_{t_v}$ antennas, and each user has $N_r = N_{r_h} \times N_{r_v}$ antennas placed in the xz-plane as 2D antenna arrays. The antennas are equally spaced with distances d_h and d_v in the horizontal and vertical directions, respectively.

A. Codebook for analog beamforming

We have two pre-defined codebooks: the transmitter codebook $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, .., \mathbf{c}_N\}$ consisting of N beamforming vectors of size $N_t \times 1$ at the BS and the receive codebook $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, .., \mathbf{d}_M\}$ consisting of M combining vectors of size $N_r \times 1$ at each user. The vector $\mathbf{c}_n$ (and, similarly $\mathbf{d}_n$) is the Kronecker product of its corresponding horizontal and vertical beamforming vectors, *i.e.,* $\mathbf{c}_n = \mathbf{c}_{h_n} \otimes \mathbf{c}_{v_n}$, where

$$\mathbf{c}_{h_n} = [\mathbf{c}_{h_n}(1), \mathbf{c}_{h_n}(2)e^{j\Omega_{h_n}}, .., \mathbf{c}_{h_n}(N_{t_h})e^{j(N_{t_h}-1)\Omega_{h_n}}]^T,$$
$$\mathbf{c}_{v_n} = [\mathbf{c}_{v_n}(1), \mathbf{c}_{v_n}(2)e^{j\Omega_{v_n}}, .., \mathbf{c}_{v_n}(N_{t_v})e^{j(N_{t_v}-1)\Omega_{v_n}}]^T.$$

Here, $\Omega_{h_n} = \frac{2\pi}{\lambda} d_h \cos(\phi_n)\sin(\theta_n)$ and $\Omega_{v_n} = \frac{2\pi}{\lambda} d_v \cos(\theta_n)$. Each codeword is steered towards a particular direction, specified by Ω_{h_n} and Ω_{v_n}.

In our performance comparison in Section IV, we use two types of codebooks: the DFT codebook and the Taylor codebook. Typically, mmWave systems use the standard DFT codebook [11] given by analog beamforming vectors $\mathbf{c}_n$ that have equal magnitude components, i.e., $\mathbf{c}_n(i) = 1/\sqrt{N_t}$, and vectors corresponding to equally spaced $(\Omega_{h_n}, \Omega_{v_n})$ values. However, these DFT codewords generate beams with high side lobe levels, which causes high interference in MU-MIMO. Non-uniform amplitude beamforming vectors allow control over the side lobes, making them a good choice in the case of MU-MIMO [8, 10]. We consider the Taylor codebook that uses amplitude tapering as in [9, p.1156], i.e., $\mathbf{c}_{h_n}(i) = J_o\big(j\pi B\sqrt{1-(2i/(N_{t_h}-1))^2}\big)$, and $\mathbf{c}_{v_n}(i) = J_o\big(j\pi B\sqrt{1-(2i/(N_{t_v}-1))^2}\big)$, where J_0 is the Bessel function of zeroth order, and B is a constant. We also enforce that the maximum per-antenna power for the Taylor codebook should be less than or equal to that of the DFT codebook. Thus, the transmit power using the Taylor codebook is less than the transmit power for the DFT codebook.

B. Partial interfering beam feedback

During the codebook-based training phase, each user k estimates the SNR for each transmit-receive beam pair $\{\mathbf{c}_n, \mathbf{d}_m\}$:

$$\text{SNR}_{n,m}^{(k)} = |\mathbf{d}_m^H \mathbf{H}_k \mathbf{c}_n|^2, \quad (2)$$

and the quantities $S_{n,m}^{(k)} \triangleq |\mathbf{d}_m^H \mathbf{H}_k \mathbf{c}_n|$ and $A_{n,m}^{(k)} \triangleq \angle(\mathbf{d}_m^H \mathbf{H}_k \mathbf{c}_n)$. Let m^k be the index of the best combining vector of user k, *i.e.*, $m^k = \arg\max_m \max_n \text{SNR}_{n,m}^{(k)}$. In this work, we consider a limited feedback system where the users can send back information only about p beams. Each user determines the best transmit-receive beam pair and sends feedback about the top-p transmitter beams corresponding to the best receive beam, *i.e.*, each user k feeds back information about its best transmit beam and the top $p-1$ interfering beams for the best transmit-receive beam pair. Thus, the feedback from the k^{th} user are the following $3 \times p$ quantities:

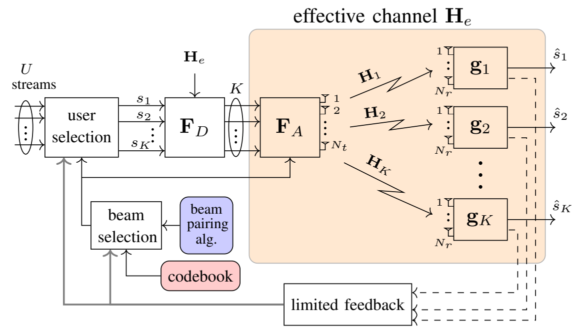

Fig. 1: Downlink limited feedback hybrid beamforming system

$$\mathcal{N}^{(k)} \triangleq \{n_1^k, n_2^k, \ldots, n_p^k\}, \quad (3a)$$
$$\mathcal{S}^{(k)} \triangleq \{S_{n_1^k,m^k}^{(k)}, S_{n_2^k,m^k}^{(k)}, \ldots, S_{n_p^k,m^k}^{(k)}\}, \quad (3b)$$
$$\mathcal{A}^{(k)} \triangleq \{A_{n_1^k,m^k}^{(k)}, A_{n_2^k,m^k}^{(k)}, \ldots, A_{n_p^k,m^k}^{(k)}\}, \quad (3c)$$

where n_l^k is the index of the l^{th} best beamforming vector when beam m^k is used at the user k, satisfying $\text{SNR}_{n_1^k,m^k}^{(k)} \geq \text{SNR}_{n_2^k,m^k}^{(k)} \geq \cdots \geq \text{SNR}_{n_p^k,m^k}^{(k)} \cdots \geq \text{SNR}_{n_N^k,m^k}^{(k)}$. Based on this feedback from all the U users, the BS performs beam selection and user selection, followed by beamforming, as explained in Section III. When $p = N$, our scheme reduces to the scheme in [6] where the full effective channel $\mathbf{H}_e$ in Fig. 1 is fed back. We will also compare with another limited feedback scheme in [8] based on channel-reconstruction using the top-q transmit-receive beam pairs for each user.

III. Beamforming Design and User Selection

In this section, we first present the proposed hybrid beamforming for a given choice of K users and their feedback using the partial interfering beam feedback scheme. Then, we present a beam grouping scheme to choose K beams out of N. Finally, we discuss beam group and user selection.

A. Proposed Hybrid Beamforming

For a given selection of K users, $\mathbf{F}_A$ and $\mathbf{g}_k$'s can be chosen as the best beamfoming vectors in $\mathcal{C}$ and the best combining vectors in $\mathcal{D}$, respectively, [6]:

$$\mathbf{F}_A \triangleq [\tilde{\mathbf{f}}_1, \tilde{\mathbf{f}}_2, \ldots, \tilde{\mathbf{f}}_K] = [\mathbf{c}_{n_1^1}, \mathbf{c}_{n_1^2}, \ldots, \mathbf{c}_{n_1^K}] \quad (4)$$
$$\mathbf{g}_k = \mathbf{d}_{m^k}, \quad k = 1, \ldots, K. \quad (5)$$

Now, we write $\hat{\mathbf{s}} \triangleq [\hat{s}_1, \hat{s}_2, \ldots, \hat{s}_K] = \mathbf{H}_e \mathbf{F}_D \mathbf{s} + \tilde{\mathbf{n}}$, where $\mathbf{H}_e = [\mathbf{g}_1^H \mathbf{H}_1 \mathbf{F}_A, \mathbf{g}_2^H \mathbf{H}_2 \mathbf{F}_A, \ldots, \mathbf{g}_K^H \mathbf{H}_K \mathbf{F}_A]^T$ and $\tilde{\mathbf{n}} = [\mathbf{g}_1^H \mathbf{n}_1, \mathbf{g}_2^H \mathbf{n}_2, \ldots, \mathbf{g}_K^H \mathbf{n}_K]^T$. Then, in [6, Alg. 1], $\mathbf{F}_D$ is the ZF digital precoder for the effective channel:

$$\mathbf{F}_D = \mathbf{H}_e^H (\mathbf{H}_e \mathbf{H}_e^H)^{-1}. \quad (6)$$

Finally, the columns of $\mathbf{F}_A \mathbf{F}_D$ have to be normalized to satisfy the power constraints. In order to implement this scheme in

[6], $\mathbf{H}_e$ should be known at the transmitter. User k needs to feedback the k^{th} row of the effective channel $\mathbf{H}_e$ given by

$$\mathbf{H}_{e_{(k,:)}} = [\mathbf{g}_k^H \mathbf{H}_k \tilde{\mathbf{f}}_1, \mathbf{g}_k^H \mathbf{H}_k \tilde{\mathbf{f}}_2, \ldots, \mathbf{g}_k^H \mathbf{H}_k \tilde{\mathbf{f}}_K]. \quad (7)$$

Since each user is not aware of the beams to be used for MU-MIMO in advance, this can be achieved only by feeding back the information $\mathbf{g}_k^H \mathbf{H}_k \tilde{\mathbf{f}}_i$ for all $i = 1, 2, \cdots, N$ transmit beams in the codebook.

In this work, we consider a top-p beam feedback system, where the users can send back only information given in (3). For any $p \geq 1$, the diagonal entries $\mathbf{g}_k^H \mathbf{H}_k \tilde{\mathbf{f}}_k$ of $\mathbf{H}_e$ are always fed back. The off-diagonal values corresponding to the interference $\mathbf{g}_k^H \mathbf{H}_k \tilde{\mathbf{f}}_i,\ i \neq k$ is available only for i belonging to the top $p-1$ interfering beams in $\mathcal{N}^{(k)}$. For the entries of $\mathbf{H}_e$ not available in the top-p feedback, we propose to set these entries in the effective channel $\mathbf{H}_e$ to zero. This is reasonable because it is always true that $|\mathbf{g}_k^H \mathbf{H}_k \tilde{\mathbf{f}}_i| \geq |\mathbf{g}_k^H \mathbf{H}_k \tilde{\mathbf{f}}_j|$, when $n_1^i \in \mathcal{N}^{(k)}$ and $n_1^j \notin \mathcal{N}^{(k)}$. Since n_1^j is not present in the best p-beams of the k^{th} selected user, its interference on the k^{th} user will be less than the interference due to the best p-beams. Based on this observation, we propose that

$$\mathbf{g}_k^H \mathbf{H}_k \tilde{\mathbf{f}}_j \approx 0, \text{ if } n_1^j \notin \mathcal{N}^{(k)}. \quad (8)$$

When $p = 1$, the BS has only the first best beam information and assumes the off-diagonal elements of $\mathbf{H}_e$ to be zero leading to $\mathbf{F}_D = \mathbf{I}_K$, *i.e.,* analog-only beamforming. When $p = N$, we will be able to recover $\mathbf{H}_e$ fully, *i.e.*, the scheme in [6]. Thus, for different p, we now have a set of partial interference suppression schemes from analog-only beamforming to full ZF beamforming using the effective channel.

In [8], ZF beamforming is proposed based on a rank-q approximation $\hat{\mathbf{H}}_k$ of $\mathbf{H}_k$ using the top-q transmit-receive beam pairs as follows. The vector $\mathbf{g}_k^H \mathbf{H}_k$ is approximated as

$$\mathbf{g}_k^H \hat{\mathbf{H}}_k = \sum_{l=1}^{q} \left(\alpha_l^{(k)} e^{j\varphi_l^{(k)}} \right) \left(\beta_l^{(k)} e^{j\psi_l^{(k)}} \right) . \ \mathbf{c}_{\tilde{n}_l^k}^H, \quad (9)$$

where $(\tilde{n}_l^k, \tilde{m}_l^k)$ is the codebook index pair corresponding to the l^{th} highest $\text{SNR}^{(k)}$ defined in (2). So, $\alpha_l^{(k)} = \mathcal{S}_{\tilde{n}_l^k, \tilde{m}_l^k}^{(k)}$, $\varphi_l^{(k)} = \mathcal{A}_{\tilde{n}_l^k, \tilde{m}_l^k}^{(k)}$, $\beta_l^{(k)} \triangleq |\mathbf{g}_k^H \mathbf{d}_{\tilde{m}_l^k}|$ and $\psi_l^{(k)} \triangleq \angle(\mathbf{g}_k^H \mathbf{d}_{\tilde{m}_l^k})$. The feedback from each user is of size $5 \times q$ in [8] (can be reduced to $3 \times q$ if the receiver uses the best beam for combining). For the given estimated channel matrices, they proposed a ZF structure for the hybrid precoder $\mathbf{F}_A \mathbf{F}_D$ as:

$$\mathbf{F}_A \mathbf{F}_D = \tilde{\mathbf{H}}_e^H (\tilde{\mathbf{H}}_e \tilde{\mathbf{H}}_e)^{-1}, \quad (10)$$

where $\tilde{\mathbf{H}}_e = [\mathbf{g}_1^H \hat{\mathbf{H}}_1, \mathbf{g}_2^H \hat{\mathbf{H}}_2, \ldots, \mathbf{g}_K^H \hat{\mathbf{H}}_K]^T$. The constraints of hybrid precoding are later applied to the find $\mathbf{F}_A$ and $\mathbf{F}_D$. In our performance comparison, we use the solution in (10) which gives an upper bound on the performance in [8].

B. Proposed Beam Grouping for MU-MIMO

In this section, we propose a beam grouping algorithm for selecting K out of N beams ($\mathbf{c}_n$'s) for simultaneous MU-MIMO transmission by the BS. The key idea is to group the beams which result in least overall interference between them.

Algorithm 1 Beam grouping algorithm for MU-MIMO

1: Calculate $I_{i,j}\ \forall\ i, j$. Set $p = 0$, $\mathbf{G}_K = [\ ]$, $\tilde{\mathbf{G}}_K = [\ ]$.
2: **for** $n \in \{1, \ldots, N\}$, let the ${}^{N-1}C_{K-1} \times K$ matrix $\mathbf{F}$ denote all the possible ways of choosing K out of N beams such that the n^{th} beam is present.
3: **for** $m \in \{1, \ldots, {}^{N-1}C_{K-1}\}$
4: $A_m = \sum_{i \neq j} I_{i,j}$ for $i, j \in \mathbf{F}(m, 1:K)$.
5: **end**
6: **if** $\mathbf{F}[(\hat{m}, :)] \notin \tilde{\mathbf{G}}_K$, where $\hat{m} = \arg\min_m A_m$.
7: Update $\tilde{\mathbf{G}}_K = [\tilde{\mathbf{G}}_K; \mathbf{F}(\hat{m}, :)], p = p+1, \hat{A}_p = A_{\hat{m}}$
8: **end**
9: **end**
10: **for** $n \in \{1, \ldots, N\}$, $p \in \{1, \ldots, P\}$, $P =$ rows of $\tilde{\mathbf{G}}_K$.
11: **if** $\tilde{\mathbf{G}}_K(p, :)$ is the only combination in $\tilde{\mathbf{G}}_K$ such that
12: n is present, update $\mathbf{G}_K = [\mathbf{G}_K\ ;\ \tilde{\mathbf{G}}_K(p, :)]$.
13: **end**
14: **end**
15: **for** $n \in \{1, \ldots, N\}$
16: **if** $n \notin \mathbf{G}_K$, $Q = \{q : n \in \tilde{\mathbf{G}}_K(q, :), q \in \{1, \ldots, P\}\}$.
17: Update $\mathbf{G}_K = [\mathbf{G}_K; \tilde{\mathbf{G}}_K(\hat{q}, :)], \hat{q} = \arg\min_{q \in Q} \hat{A}_q$.
18: **end**
19: **end**

This algorithm is an offline algorithm that only depends on the codebook design and not on the channel realization.

Define the interference of the j^{th} beam on the i^{th} beam's region as $I_{ij} = \int_{\Omega_v^i} \int_{\Omega_h^i} \mathbf{c}_j . (\mathbf{c}_i(\Omega_h, \Omega_v))^H \, d\Omega_h d\Omega_v$, where Ω_v^i and Ω_h^i are the (Ω_h, Ω_v) region corresponding to the i^{th} beam. We evaluate this integration numerically in our algorithm. The beam pairing is done using Algorithm 1 below. The result of the beam grouping is represented by the matrix $\mathbf{G}_K$. Each row of $\mathbf{G}_K$ is a beam combination with K beams. The algorithm ensures that all beams are used in at least one group. Steps 3-9 selects one group of K beams for each n and forms the matrix $\tilde{\mathbf{G}}_K$ of groups. If a particular beam appears in *only* one of the above groups, then that group is chosen in Steps 10-14 for $\mathbf{G}_K$. If any beam has not been selected in $\mathbf{G}_K$ by Steps 10-14, then the group that results in least interference is chosen from the possible groups in $\tilde{\mathbf{G}}_K$ in Steps 15-19.

Example: Consider $N = 16$ where the beams are in the directions of equally spaced $(\Omega_{h_n}, \Omega_{v_n})$ values covering the horizontal beamspace $(-\sqrt{3}\pi/2, \sqrt{3}\pi/2)$ and vertical beamspace $(-\pi/2, 0)$. The results of beam pairing for $K = 2,\ 3,$ and 8 for the Taylor codebook are:

$$\begin{aligned} K = 2, \mathbf{G}_2 = &[1\ 12;\ 2\ 11;\ 3\ 14;\ 4\ 13;\ 5\ 16; \\ &6\ 15;\ 7\ 16;\ 8\ 15;\ 2\ 9,\ 1\ 10], \\ K = 3, \mathbf{G}_3 = &[1\ 9\ 16;\ 2\ 10\ 15;\ 3\ 8\ 13;\ 4\ 7\ 14;\ 5\ 10\ 15; \\ &6\ 9\ 16;\ 2\ 7\ 15;\ 1\ 8\ 16;\ 6\ 11\ 16, 5\ 12\ 15], \\ K = 8, \mathbf{G}_8 = &[1\ 4\ 5\ 8\ 9\ 12\ 13\ 16;\ 2\ 3\ 6\ 7\ 10\ 11\ 14\ 15]. \end{aligned}$$

The number of groups in $\mathbf{G}_K$ satisfies $N/K \leq |\mathbf{G}_K| \leq N$.

C. Beam Group and User Selection

Optimal joint beam and user selection is hard. Therefore, we use a simpler two-step method: (1) In step 1, for K-user transmission, the BS selects one group from the set of beam groups

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/LCOMM.2020.2983018

from Algorithm 1, $\mathbf{G}_K$, based on the feedback information in (3). Let this combination of the selected beams be denoted as $\mathcal{J} = \{j_1, j_2, \ldots, j_K\}$. Let $\mathcal{S}_{j_k} = \{u : n_1^u = j_k, u = 1, \ldots, U\}$ be the set of users for whom j_k^{th} beam is the best beam. (2) In step 2, one user is selected for each selected beam j_k, from the set $\mathcal{S}_{j_k}$ based on the feedback.

For a given beam group, we consider three scheduling schemes for user selection: Round-Robin (RR), Proportional Fair (PF), and SINR-based scheduling. In RR user selection, for j_k^{th} beam, a user from the set $\mathcal{S}_{j_k}$ is chosen in round-robin fashion. In PF user selection, the user in $\mathcal{S}_{j_k}$ with maximum PF metric is selected. The PF metric for user j is $R_k[j]/T[j]$, where $R_k[j]$ is the achievable rate for user j using beam k and $T[j]$ is the average throughput so far for user j in a suitably chosen window [12]. In SINR-based scheduling, the user with best SINR in $\mathcal{S}_{j_k}$ is selected for beam j_k.

For beam group selection, we again consider RR, PF, and sum-rate (SR)-based selection. In RR selection, a group from $\mathbf{G}_K$ is selected in round-robin fashion providing best fairness among the beams without considering the channel conditions. In SR-based selection, the group in $\mathbf{G}_K$ which gives the highest sum rate for the selected users for that beam group is selected. In PF selection, the beam group in $\mathbf{G}_K$ with the highest sum of the PF metrics of the selected users is selected.

The number of metric calculations needed for exhaustive joint beam and user selection is of the order $\binom{N}{K} U_{max}^K$, where U_{max} is the maximum of $|\mathcal{S}_n|$ for $n = 1, \ldots, N$. However, for the proposed method, the number of metric calculations is reduced to $|\mathbf{G}_K| K U_{max}$, where $N/K \leq |\mathbf{G}_K| \leq N$.

IV. Performance Comparison

In this section, the performance of the proposed hybrid beamforming scheme with partial interfering beam feedback is studied and compared with the hybrid beamforming schemes in [6, 8]. We assume $N_t = 8 \times 8$ antenna elements at the BS and $N_r = 4 \times 4$ at each user, both placed in xz-plane with $d_h = d_v = \lambda/2$. The carrier frequency is 28 GHz. There are $U = 100$ users to be served by the BS in the coverage area. The 3D channel model developed by 3GPP is used to generate the channel matrices similar to the Urban Macro (UMa) scenario, considering the number of clusters and rays per cluster as 4 and 1, respectively. The detailed steps are explained in [13]. For ease of exposition, we neglected doppler effect and assumed purely vertically polarized antennas at both transmitter and receiver. The total power used at the BS is 35 dBm and the noise variance is $k_B T \Delta f = -174$ dBm $+ 10 \log_{10}(\Delta f)$, where the bandwidth Δf is assumed to be 100 MHz. The transmit codebook has size $N = 16$, covering $\phi \in (30°, 150°)$ and $\theta \in (90°, 120°)$ whereas, the receive codebook has size $M = 3$, covering $\phi \in (-150°, -30°)$ and $\theta \in (-90°, -60°)$.

In Figs. 2a-2e, the CDF of the sum rates over 10000 realizations are plotted for different beamforming methods. For Figs. 2a-2b, both user and beam selections are done using PF metric. Fig. 2a shows the CDF of the sum rates of the proposed scheme using partial interfering beam feedback for 2 cases: DFT codebook and Taylor codebook. For each case, the sum rates are shown for $p = 1, 4, 8$ and 16 with $K = 8$. In both cases, as p increases (*i.e.*, the BS has more knowledge of the effective channel $\mathbf{H}_e$), the sum rates also increase. The Taylor codebook results in better performance for all cases except $p = 16$. This is because using Taylor codebook instead of DFT codebook reduces residual interference when partial interfering beam feedback is available. For $p = 16 = N$, the Taylor codebook is worse because it has lesser transmit power. Since full interference feedback is available, the ZF precoder is able to handle the interference. Note that, because of the per-antenna constraint, the total transmit power for the Taylor codebook used at the BS is 4.4 dB less than the DFT codebook. Even with lower transmit power, the Taylor outperforms the DFT codebook when $p < N$. The ZF scheme in [8] was also studied for different levels of feedback q (not shown in Fig. 2) for both the DFT and Taylor codebooks. The performance of the Taylor codebook for the scheme in [8] was found to be similar to the DFT codebook.

Fig. 2b compares the performance of the following schemes: (i) proposed hybrid beamforming using DFT codebook, (ii) proposed hybrid beamforming using Taylor codebook, (iii) ZF scheme in [8], (iv) hybrid beamforming in [6] that corresponds to the $p = N$ case of the proposed scheme. It can be observed that the sum rate achieved by the proposed hybrid beamforming using the Taylor codebook is greater than that of the DFT codebook and the ZF scheme in [8] for both $p = 1$ and $p = 8$. We can also see that the ZF in [8] gives high rates than the proposed hybrid beamforming using the DFT codebook. This is because [8] has a better approximation of the channel matrices compared to the proposed scheme under low feedback information. But, with the Taylor codebook, the proposed hybrid beamformer is able to achieve significantly high sum rates than the ZF scheme in [8]. Comparing all the 4 schemes, the hybrid beamfoming in [6] achieves the highest sum rate as it uses full feedback information ($\mathbf{H}_e$ is fully known, $p = N = 16$). The figure also shows that the Taylor codebook for $p = 8$ with same sum-power constraint (SPC) achieves better performance than the other schemes for $p = 8$.

In Fig. 2c, we note that: (i) MMSE hybrid precoding using the proposed scheme performed better than the ZF scheme, (ii) the performance of the quantized feedback scheme is almost as good as the unquantized scheme, and (iii) the relative performance of all three schemes with quantized feedback is similar to the case of unquantized feedback in Fig. 2b.

In Fig. 2d and Fig. 2e, we evaluate the performance of the proposed hybrid beamforming and the scheme in [8] for different user and beam selection schemes. Fig. 2d shows the CDF of the sum rates for different user selections under SR-based beam selection when $p = 8$ and $K = 8$. As we discussed in Section III-C, the sum rates achieved by the SINR-based scheduling is high compared to the other two scheduling. The RR scheduling gives the lowest rates as it does not depend on the channel conditions. The PF scheduling keeps a balance between the fairness and rate throughput, providing sum rates higher than the RR but less compared to the SINR-based. Further, Fig. 2e shows the CDF of the sum rates for different beam selections under SINR-based user selection when $p = 8$ and $K = 8$. Here also, it is observed that the sum rates achieved by the three schedulers follow

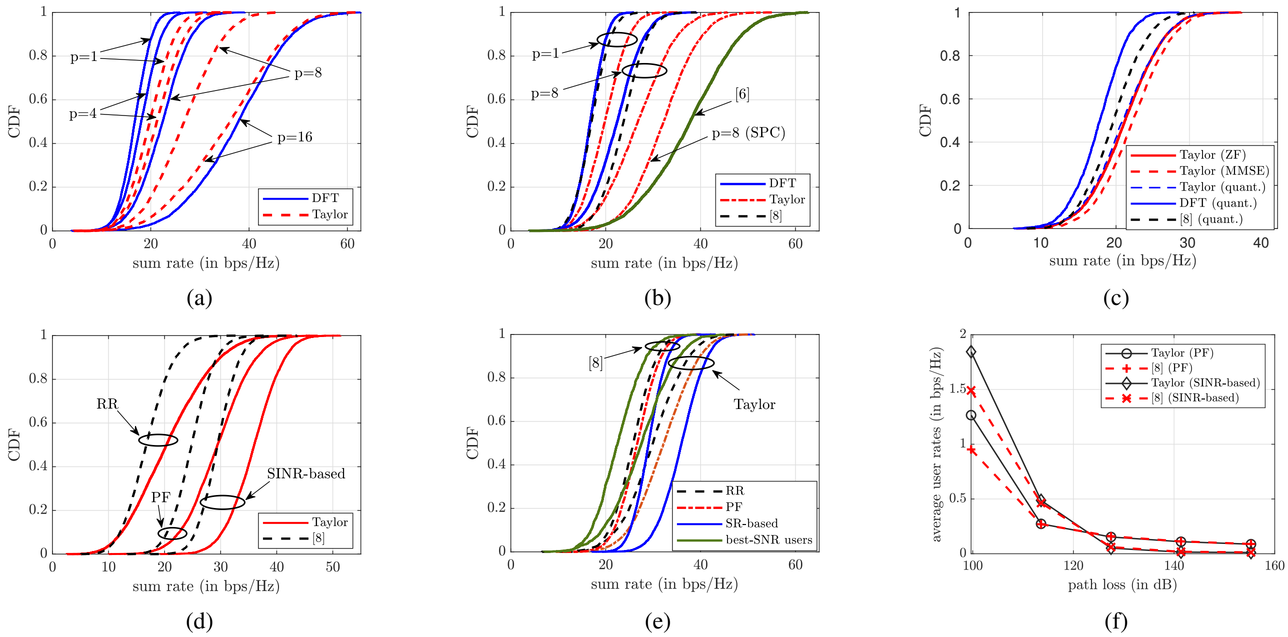

Fig. 2: (a) CDF of sum rates of the proposed hybrid beamforming using DFT and Taylor codebooks for $K = 8$. (b) Comparison of different schemes for $K = 8$. (c) Comparison with quantized feedback and comparison of Taylor MMSE scheme with Taylor ZF, $p = 4$ and $K = 8$. (d) Comparison of different user selections, $p = q = 8$ and $K = 8$. (e) Comparison of different beam selections, $p = q = 8$ and $K = 8$. (f) Average user rates vs. path loss for different user selections, $p = q = 8$ and $K = 8$.

RR<PF<SR-based. In this figure, we also show the sum rate achieved by beam selection corresponding to the top K users with the best SNR and distinct best transmit beams. We can see that the performance of this beam selection method suffers due to inter-beam interference unlike the proposed beam selection schemes that use only the beam groupings in $\mathbf{G}_K$. In both Fig. 2d and Fig. 2e, the proposed hybrid beamforming with Taylor codebook achieves better sum rates compared to [8].

Fig. 2f shows the average user rates versus path loss for the proposed scheme using the Taylor codebook and the scheme in [8] when $K = 8$ and $p = 8$, under PF beam selection. The path loss values are divided into 5 bins, and their corresponding individual user rates are averaged and plotted with respect to the median of each bin. For all the schemes, the rate decreases as path loss value increases because of the propagation loss. The proposed hybrid beamforming scheme provides higher data rates for users, which are nearer to the BS, as compared to [8]. Since the SINR-based scheduler always selects the users with high SINR, it provides high data rates than PF for nearer users. PF considers fairness and results in better rates for users that are far away from the BS than the SINR-based scheme.

V. Conclusions

The proposed hybrid beamforming using Taylor codebook provides significant improvement in sum rate over the ZF beamforming in [8] for the same amount of feedback even while transmitting lesser power. The result has been validated for different beam and user selections. We also proposed a beam grouping algorithm to reduce inter-beam interference in MU-MIMO and simplify beam and user selection. When $p = N$, our scheme reduces to the scheme in [6].

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, 2011.
[2] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, “Spatially sparse precoding in millimeter wave MIMO systems,” *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, 2014.
[3] 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, 2016.
[4] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, “Channel estimation and hybrid precoding for millimeter wave cellular systems,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, 2014.
[5] V. Raghavan, J. Cezanne, S. Subramanian, A. Sampath, and O. Koymen, “Beamforming tradeoffs for initial UE discovery in millimeter-wave MIMO systems,” *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 543–559, 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, 2015.
[7] V. Raghavan, S. Subramanian, J. Cezanne, A. Sampath, O. H. Koymen, and J. Li, “Single-user vs. multi-user precoding for millimeter wave MIMO systems,” *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1387–1401, 2017.
[8] M. R. Castellanos, V. Raghavan, J. H. Ryu, O. H. Koymen, J. Li, D. J. Love, and B. Peleato, “Channel-reconstruction-based hybrid precoding for millimeter-wave multi-user MIMO systems,” *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 2, pp. 383–398, 2018.
[9] S. J. Orfanidis, *Electromagnetic waves and antennas*, 2002. [Online]. Available: http://eceweb1.rutgers.edu/~orfanidi/ewa/
[10] Z. Briqech, A.-R. Sebak, and T. A. Denidni, “Low-cost wideband mm-wave phased array using the piezoelectric transducer for 5g applications,” *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6403–6412, 2017.
[11] D. Yang, L.-L. Yang, and L. Hanzo, “DFT-based beamforming weight-vector codebook design for spatially correlated channels in the unitary precoding aided multiuser downlink,” in *Proc. IEEE Int. Conf. Commun.*, 2010, pp. 1–5.
[12] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.
[13] 3GPP TR 38.901, “Study on channel model for frequencies from 0.5 to 100 GHz,” June 2017.