Statistically-Aided Codebook-Based Hybrid Precoding for Millimeter Wave Channels

AHMED WAGDY SHABAN1, (Student Member, IEEE), OUSSAMA DAMEN1, (Senior Member, IEEE), YAN XIN2, AND EDWARD AU2, (Senior Member, IEEE)
1Electrical and Computer Engineering Department, University of Waterloo, Waterloo, ON N2L 3G1, Canada
2Ottawa Research & Development Centre, Huawei Technologies Canada, Ottawa, ON K2K 3J1, Canada
Corresponding author: Ahmed Wagdy Shaban (ahmed.w.shaban@uwaterloo.ca)

This work was supported in part by Huawei Technologies Company, in part by the University of Waterloo, and in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

ABSTRACT In this paper, we propose practical yet effective statistically-aided codebook-based hybrid precoding schemes for massive multiple-input multiple-output systems in millimeter wave bands. Particularly, we develop novel low-overhead hybrid precoding algorithms for selecting the baseband digital and radio frequency analog precoders from statistically skewed DFT-based codebooks. The proposed algorithms aim at maximizing the spectral efficiency based on minimizing the chordal distance between the optimal unconstrained precoder and the hybrid beamformer and maximizing the signal to the interference noise ratio for the single-user and multi-user cases, respectively. We investigate the performance of the proposed algorithms by considering the mutual information of the analog beamforming procedure (the common stage among the proposed algorithms) as a performance evaluation metric. We derive lower and upper bounds on the mutual information of the channel given the proposed algorithms. Moreover, we show that the performance gap between the lower and upper bounds depends heavily on how many DFT columns are aligned to the largest eigenvectors of the transmit antenna array response of the millimeter wave channel or equivalently the transmit channel covariance matrix when only statistical channel knowledge is available at the transmitter. Then, we show that the proposed algorithms are asymptotically optimal as the number of transmit antennas $M$ goes to infinity and the millimeter wave channel has a limited number of paths $P$, i.e., $P < M$. Further, we verify the performance of the proposed algorithms numerically where results illustrate that the spectral efficiency of the proposed algorithms can approach that of the optimal precoder in certain scenarios. Furthermore, these results show that the proposed hybrid precoding schemes have superior spectral efficiency performance while requiring lower (or at most comparable) channel feedback overhead in comparison with the prior art.

INDEX TERMS Codebook-based precoding, hybrid beamforming, massive MIMO, mmWave channels, statistical precoding.

I. INTRODUCTION

Hybrid (analog/digital) beamforming is the low-cost and energy-saving solution to achieve high spectral efficiency performance for most massive multiple-input multiple-output (MIMO) systems operating in millimeter wave (mmWave) bands [1]. Extensive research works have demonstrated the effectiveness of hybrid beamforming in achieving a near-optimal tradeoff between the high spectral efficiency performance and the low operational cost under the perfect channel state information at the transmitter (CSIT) assumption (e.g., [2], [3] and references therein). However, in massive MIMO frequency-division duplex (FDD) systems,1 perfect CSIT is an impractical assumption where such channel information is hard to be estimated and acquired at the transmitter.2 This is mainly due to the large downlink training, where the required training dimensions scales linearly with the number

---

1The FDD architecture is the most deployed architecture in the current wireless communication systems.

2Obtaining perfect CSIT in FDD is a two-stage process where transmitter sends downlink training signals (pilot symbols) so that the receiver estimates the channel and then feeds it back to the transmitter.
of antenna elements, and the corresponding huge channel feedback overhead [1], [4], [5]. This, in turn, may impact the downlink and the uplink system capacity if no countermeasures are considered.

Huge research efforts have been exerted to alleviate the channel feedback overhead by relaxing the perfect CSIT assumption. In one direction, many works have adopted the statistical channel information, particularly the spatial channel covariance matrix, to design the analog precoder, and the perfect channel information to design the digital precoder (e.g., [4]–[7] and references therein). This lessens the problem by reducing the required training and the corresponding feedback overhead since (i) the channel statistics of most wireless applications are slowly varying and (ii) the dimensions of the effective channel (the channel after the analog precoding) are significantly reduced. However, this direction requires optimistic conditions on the channel such as spatial user grouping, very small angular spread, uniformity of the covariance matrix across users and/or subcarriers, and the stationarity of the angles of arrival and departure in order to achieve near-optimal spectral efficiency performance. In another direction, the limited feedback hybrid precoding has been considered where the channel or the analog precoder is selected from a predefined finite set of codewords (e.g., [2], [8]–[17] and references therein). This reduces the feedback overhead dramatically where only a small number of bits are required to indicate the preferred codeword. However, existing research works in this approach invoke computational intensive beamforming algorithms that either perform exhaustive search [2] or require complex iterative processing such as coordinate descent algorithm [12], Tabu search [10] and cross-entropy optimization method [11], and utilize inefficient codebooks such as Hadamard codebooks [13] and fixed parts of DFT matrices [15].

To address these shortcomings, we propose a novel approach to design the hybrid precoder based on leveraging second-order statistics and propagation properties of the mmWave channel aiming mainly at decreasing the feedback overhead in FDD systems. The proposed approach is embodied in two main hybrid precoding schemes, i.e., Algorithm 1 and Algorithm 2, developed for single and multi-user cases, respectively. Moreover, we provide different variants of Algorithm 1 and Algorithm 2 which lend themselves to various channel knowledge frameworks such as limited feedback channel without statistical information and mixed-CSIT. The contributions of this work compared to prior art are summarized as follows:

- For the single-user case, we present a simple hybrid precoding design which is based on minimizing the chordal distance between the fully-digital precoder and the hybrid one. In contrast to the exhaustive search [2], complex gradient decent-based algorithms [10]–[12] and iterative greedy algorithms [8], [16], [17] developed in prior works, we propose low complexity and non-iterative hybrid beamforming algorithms (Algorithm 1, and its variants). Moreover, opposed to the channel independent (fixed) codebooks such as beamsteering and Hadamard codebooks used in the majority of prior art [8], [11]–[14], [16], we utilize skewed DFT codebooks that vary with the channel statistics. This is devised in order to finely quantize the local neighborhood around the statistically preferred directions of the dominant eigenvectors and thereby enhancing the spectral efficiency performance.
- For the multi-user case, we present a hybrid precoding design which is based on maximizing the signal to interference plus noise ratio of each user aiming at maximizing the sum-rate of the network. In contrast to the prior works [4]–[7], [9], [12], [14], the analog precoder is designed in a distributive manner that lends itself to the distributed nature of multi-user networks. A distinguishing feature of Algorithm 2 from prior works in [9], [11]–[13] is that Algorithm 2 can be applied to the case when the number of assigned users is less than that of the RF chains, i.e., $K \leq N_\text{rf}$. However, the algorithms in [9], [11]–[13] are applicable only when the number of users equals to the number of RF chains where they assign each user to one RF chain.
- We derive lower and upper bounds on the average mutual information for the DFT-based hybrid precoding. The proposed bounds suggest to select DFT codewords that are aligned to the directions of the largest eigenvectors of the transmit antenna array response matrix (or the covariance matrix, depending on the type of the channel knowledge availability). Leveraging the properties of the mmWave channels, we demonstrate in Corollary 1 that the proposed schemes are asymptotically optimal. Numerical results validate the near-optimal spectral efficiency performance of the proposed schemes and their superiority over prior works.

Finally, we adopt the following notation throughout this manuscript: we use capital boldface letters for matrices and small boldface letters for vectors. $\mathbb{C}^{N \times M}$ and $\mathbb{U}^{N \times M}$ denote the space of complex matrices of size $N \times M$ and the space of complex matrices that have unit magnitude entries, respectively. Moreover, $\text{det}(A)$ is the determinant of matrix $A$, $\|A\|_F$ is its Frobenius norm, whereas $\text{diag}(A)$ denotes the diagonal entries of $A$, respectively. $A_{N_i}$ is a matrix that contains $N_i$ columns of $A$, and $I_{N_i}$ is the $N_i \times N_i$ identity matrix. The $M \times M$ DFT matrix is denoted by $D_M$. With a slight abuse of notation, we use $\arg\max\{\}$ to return the indices of the largest $X$ elements in the set $\{}$.

II. SYSTEM MODEL

We consider a massive MIMO FDD system with a limited number of RF chains. In this system, we present two communication scenarios, namely, single-user and multi-user massive MIMO scenarios. In the single-user scenario, the transmitter and the receiver are equipped with $M$-transmit and $N$-receive antenna elements. This point-to-point massive MIMO model is of interest due to its wide-range applicability
in many recent applications such as high definition video streaming, virtual-reality/augmented-reality, connected cars and links between base stations [18]. The transmitter sends $N_t$ independent data streams to the receiver with the help of $N_{ff} \ll \min(M, N)$ RF chains. Due to the use of a limited number of RF chains, the hybrid beamforming structure is considered. In particular, we assume that the data vector, $\mathbf{s} \in \mathbb{C}^{N_t \times 1}$, is pre-processed by two different precoding matrices. As a result, the received signal is given by:

$$\mathbf{r} = \sqrt{\rho} \mathbf{H}_{RF} \mathbf{F}_{BB} \mathbf{s} + \mathbf{n},$$

(1)

where $\mathbf{F}_{BB} \in \mathbb{C}^{N_t \times N_t}$ is the baseband precoder and $\mathbf{F}_{RF} \in \mathbb{C}^{M \times N_{rf}}$ is the RF one. $\rho$ is the average signal to noise ratio, and $\mathbf{n}$ is the additive white Gaussian noise vector, with i.i.d entries $\sim \mathcal{CN}(0, \sigma_n^2)$. Generally, the fading channel matrix, i.e., $\mathbf{H} \in \mathbb{C}^{N \times M}$ in mmWave bands, is spatially correlated due to the usage of large-scale and densely-packed phased antenna arrays at both sides [2], [4], [6], [9] and sparse in the angular domain due to the limited scattering nature of mmWave bands [2], [19]. This is modeled mathematically in literature on mmWave channels by either the Kronecker or the clustered channel models. The Kronecker correlation model describes the stochastic spatial correlation evolution of each channel realization. As a result, the channel is given by [4], [5]:

$$\mathbf{H}_t \overset{\text{def}}{=} \mathbf{R}_t^{1/2} \mathbf{R}_t^{1/2},$$

(2)

where $\mathbf{R}_t = \mathbb{E}[\mathbf{H}_t \mathbf{H}_t^\dagger]$ is the receive correlation matrix, $\mathbf{R}_t = \mathbb{E}[\mathbf{H}_t^\dagger \mathbf{H}_t]$ is the transmit correlation matrix and $\mathbf{H}_t \in \mathbb{C}^{N \times M}$ has entries are i.i.d. Gaussian distributed. On the other hand, the clustered channel model expresses the channel as a function of its spatial parameters where the channel matrix is given as [2], [9]:

$$\mathbf{H}_{CI} = \sqrt{\frac{MN}{N_c N_{ray}}} \sum_{i,l} \alpha_{il} \mathbf{a}_t(\theta_{il}^t) \mathbf{a}_r^H(\theta_{il}^r),$$

(3)

$$= \mathbf{A}_t \mathbf{G} \mathbf{A}_r^H,$$

(4)

where $\alpha_{il}$ denotes the complex gains of the $l$th ray in the $i$th cluster, with $N_c$ clusters, each contributing $N_{ray}$ rays such that the total number of rays/paths is $P = N_c N_{ray}$. The vector $\mathbf{a}_t(\theta_{il}^t)$ is the transmit antenna array response vector of length $M$ for a given angle of departure $\theta_{il}^t$, and $\mathbf{a}_r(\theta_{il}^r)$ is the receive antenna array response vector of length $N$ for a given angle of arrival $\theta_{il}^r$. Eq. (4) is the augmented matrix representation of the clustered channel model where $\mathbf{A}_t$ and $\mathbf{A}_r$ are the augmented transmit and receive antenna array response matrices and $\mathbf{G}$ is a diagonal channel contains the normalized complex gains $\alpha_{il} \sqrt{\frac{MN}{N_c N_{ray}}}$. Although the two models are different, they bear some similar statistical properties under certain propagation assumptions. Mainly, the transmit and the receive correlation matrices are separable where $\mathbb{E}[\mathbf{H}_t^\dagger \mathbf{H}_t] = \mathbf{R}_t$ and $\mathbb{E}[\mathbf{H}_r^\dagger \mathbf{H}_r^\dagger] = \mathbf{R}_r$ and $\mathbb{E}[\mathbf{H}_t^\dagger \mathbf{H}_t] = \mathbf{A}_t \mathbf{A}_r^H$ and $\mathbb{E}[\mathbf{H}_{CI}^\dagger \mathbf{H}_{CI}] = \mathbf{A}_t \mathbf{A}_r^H$ for both the Kronecker and the clustered channel models, respectively under the uncorrelated scattering (channel gain paths, angles of departure and arrival are mutually independent) and equal power of channel paths assumptions (typical assumptions on the literature of mmWave channels [2], [4]–[6]). Given these assumptions, both models are being used interchangeably in hybrid precoding over mmWave channels literature where the Kronecker model is preferred in theoretical analysis (especially statistical ones) and the clustered model is utilized in numerical simulations by setting $\mathbf{R}_t = \mathbb{E}_A \{ \mathbf{A}_t \mathbf{A}_r^H \}$ (see [5], [6]). We follow the same approach and hereafter we drop the subscript of $\mathbf{H}$ for the sake of notation simplicity. The total power is normalized such that $\mathbb{E}[\mathbf{s}_t^\dagger \mathbf{s}_r] = \frac{1}{N_t} \mathbf{I}_{N_t}$ and $\mathbb{E}[\mathbf{H}_t^\dagger \mathbf{H}_r] = \mathbf{I}_N$. Similarly, the receiver processes the received vector $\mathbf{r}$ by two different combining matrices:

$$\tilde{\mathbf{s}} = \sqrt{\rho} \mathbf{W}_{BB}^H \mathbf{H}_{RF} \mathbf{F}_{RF} \mathbf{F}_{BB} \mathbf{s} + \mathbf{n},$$

(5)

where $\mathbf{W}_{RF} \in \mathbb{C}^{N \times N_t}$ and $\mathbf{W}_{BB} \in \mathbb{C}^{N_t \times N_t}$ are the RF and baseband combiners, respectively. We note here that both RF (analog) precoder and combiner are implemented by analog phase shifters with constant amplitude amplifiers, therefore, their entries have a constant norm.

The multi-user scenario is obtained directly from the point-to-point one by considering minor modifications. We assume that the base station communicates with $K$ single-antenna users and sends an independent data stream to each user. Therefore, the hybrid beamforming structure is implemented at the transmitter side only. The vector of the received signals is given by (1) with a minor change in the dimensions since $N = N s = K$. Further, the augmented channel matrix is given by $\mathbf{H}_r = [\mathbf{h}_1, \cdots, \mathbf{h}_K]$, where $\mathbf{h}_k$ is the channel vector of the $k$th user.

We assume that the channel is known at the receiver(s), i.e., CSI. CSI is commonly adopted in all the current wireless standards, and it will also be implemented in mmWave-based standards (e.g., IEEE 802.11ad). The CSI is not only utilized by the proposed schemes but it is also an essential requirement for many signal reception processes such as signal detection and hybrid (or digital) combining at the receiver(s) [20]. Owing to the intrinsic sparsity of millimeter wave channels, many efficient two-stage CSI estimation algorithms have been developed based on compressed sensing techniques. This is in contrast to the single-stage techniques which are utilized in massive MIMO [21]–[23]. We consider two different types of short-term CSIT:

- **Limited feedback CSIT:** The transmitter has a finite rate (quantized) knowledge about the channel throughout a limited capacity feedback channel. Particularly, we assume that both the transmitter and the receiver agree on two predefined codebooks, $C_{BB}$ and $C_{RF}$, one for the digital stage and the other for the analog one.

- **Mixed (partial) CSIT:** The transmitter has two types of CSIT: a finite rate feedback knowledge for the massive MIMO channel and a perfect knowledge for the effective channel, i.e., $\mathbf{H}_{RF}$. This assumption is practical when the communication channel is quasi-static and
the number of transmitted stream is small. When the channel varies slowly with respect to the transmission rate, this allows for estimating the channel accurately at the receiver and feeding it back to the transmitter with a negligible rate overhead compared to the information rate. This assumption is widely considered in both multi-user and single-user cases (e.g., [4], [5], [7]–[9], [12], [15], [24]).

In addition to these two types of CSI which provide the base station with short-term channel updates, we assume, as in many other prior works [4], [6], [7], that the long-term second-order channel statistics are available at terminals. We note that the fading channel of many applications are locally wide sense stationary over time where the channel statistics remain constant for a very long period of time [25]. Hence, the channel covariance can be estimated very accurately at the base station without requiring frequent training. For instance, the channel covariance matrices can be known to the base station through efficient covariance estimation techniques for the hybrid structure such as the ones based on compressed sensing techniques (e.g., [26] and references therein). However, covariance estimation is out of the scope of this paper. We point out that, the combination of long and short-term channel state information creates different channel knowledge availability scenarios. We refer to the ones that have long-term channel knowledge as statistically-aided scenarios. We note that \( F_{RF} \) is designed based on either only limited feedback CSIT or limited feedback CSIT aided with statistical information such as \( R_f \). This is different from prior work [2] where \( F_{RF} \) is designed based on the perfect channel realization, [4]–[7] where it depends only on the second-order channel statistics and [2], [8]–[13], [15], [16] where it is constructed based on fixed predefined codebooks that do not change with the long-term channel statistics.

### III. SINGLE-USER PROBLEM FORMULATION AND PROPOSED ALGORITHMS

The canonical limited feedback problem in the literature of the fully-digital beamforming is to design a codebook that match a certain selection criteria in order to minimize an average distortion measure [27]. Similarly, the limited feedback hybrid precoding problem in mmWave literature is a codebook-based problem but with additional constraints on the analog precoder [2], [8]–[13], [15], [16]. As a result, it boils down to how to define a codeword selection criterion and to design codebooks that naturally fit these constraints while utilizing the characteristics of mmWave channels. Here, we formalize the limited feedback precoding as a codebook-based subspace approximation problem while exploiting the sparsity and the spatial correlation of mmWave massive MIMO channels. Our design and problem formulation are described as follows. We start our design by discussing the selection metric.

### A. SELECTION METRIC

Since the massive MIMO and mmWave technologies are primarily meant to dramatically increase the capacity of 5G networks, the ergodic capacity, spectral efficiency or mutual information are reasonable selection metrics. However, these selection metrics have high computational complexity where they require computationally intensive calculations such as matrix determinant and inverse. This makes them impractical in high-dimensional applications. Exploiting the sparsity of mmWave channels, it has been shown that the loss in the mutual information due to approximating the optimal precoder, i.e., \( F_{opt} \), by the hybrid precoder, \( F_{RF}F_{BB} \) is dictated by the squared chordal distance between them, i.e.,

\[
\| F_{opt}F_{BB}^{-1}F_{RF}F_{BB}F_{opt} - F_{RF}F_{BB}F_{opt}F_{BB}^{-1}F_{RF} \|^2_F [2].
\]

The chordal distance is a subspace distance that measures the geodesic distance between two subspaces on the Grassmannian manifold [28]. We choose the chordal distance as a selection metric for two reasons. First, minimizing the chordal distance between the hybrid and optimal precoders directly minimizes the mutual information loss due to the hybrid structure [2]. Second, since it is a subspace distance, it suits the codebook-based precoding problem formulation as a subspace approximation problem [27]. The optimal precoder is given by the largest \( N_f \) right singular vectors of \( H \), i.e., \( F_{opt} = V_{N_f} \) where \( H = U_H \Sigma_H V_H^H \) is the singular value decomposition (SVD) of the channel matrix.

### B. CODEBOOK DESIGN

The design of the codebook should account for two main properties of mmWave channels, i.e., the spatial correlation and the angular sparsity. Contrary to the spatially independent channel where its eigenvectors are isotropically distributed, the dominant eigenvectors of the spatially correlated channel point to certain preferred directions [29]. Moreover, the angular sparsity makes the channel tends to have a few numbers of dominant eigenvectors [2]. Considering both observations, we design the codebooks as follows. We consider a DFT-based codebook where its bases are drawn from a DFT matrix of size \( M \times M \). The \( M \times M \) DFT matrix \( D_M = \)

\[
\begin{bmatrix}
1 & 1 & 1 & \cdots & 1 \\
1 & \omega & \omega^2 & \cdots & \omega^{(M-1)} \\
1 & \omega^2 & \omega^4 & \cdots & \omega^{2(M-1)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & \omega^{(M-1)} & \omega^{2(M-1)} & \cdots & \omega^{(M-1)(M-1)}
\end{bmatrix}
\]

where \( \omega = e^{-j \frac{2 \pi}{M}} \). Considering all the combinations of \( N_{sf} \) columns of the DFT matrix in (6), we can construct the RF codebook, \( C_{RF} \subseteq D(M, N_{sf}) \), where \( D(M, N_{sf}) \) is the set of \( M \times N_{sf} \) matrices, which has \( \binom{M}{N_{sf}} \) matrices of size \( M \times N_{sf} \). We note here that each matrix in \( D(M, N_{sf}) \) represents an \( N_{sf} \)-dimensional subspace of \( \mathbb{C}^{M \times N_{sf}} \) where \( \mathbb{C}^{M \times N_{sf}} \) denotes
the space of $M \times N_f$ matrices that have constant magnitudes entries.

This codebook choice suits the angular sparsity of the mmWave channel since it divides the angular space into $M$ orthogonal beams which have $M$ distinctive angular directions. However, this does not account for the effect of the spatial correlation on the directivity of the subspaces of the eigenvectors of the channel. The main function of the RF codebook is to efficiently approximate these subspaces, therefore, the RF codebooks have to be tilted towards the subspace of the dominant eigenvectors of the channel, i.e., the column space of the optimal precoder. This is realized by multiplying $i$th codeword in $D(M, N_f)$, i.e., $\{D(M, N_f)\}_i$, $1 \leq i \leq (M/N_f)$, from the left by the transmit correlation matrix $R_t$. This results in the skewed RF codebook, i.e., $C_{RF,\text{skewed}}$

$$
C_{RF,\text{skewed}} = \{ R_t \{D(M, N_f)\}_1 \} \times \{ R_t \{D(M, N_f)\}_2 \} \times \cdots
$$

This linear transformation makes the subspace of each codeword of the skewed RF codebook lives in the subspace of the optimal precoder. This, in turn, allows the skewed RF codebook to finely quantize the local neighborhood around the statistically preferred directions of the dominant eigenvectors, and thereby, leveraging efficiently the spatial correlation of mmWave channels.

We point out that directly designing codebooks under constant magnitude entries constraint is extremely difficult and results in intractable optimization problems. This conclusion has been established in both fully-digital [30] and hybrid beamforming [31] literature. Instead, research works resort to either use predefined codebooks that have phases only entries (adhere to the RF hardware constraints) such as the beamsteering codebooks [2], [9]–[12] and Hadamard codebooks [13] or quantizing sub-optimal solutions which are obtained by imposing these constraints on the optimal unconstrained precoder as in [3], [32], [33]. However, majority of research works follow the former approach due to its superior performance compared to the latter one. In comparison to Hadamard codebooks, the DFT codebooks provide finer quantization, and thereby, having better performance. Moreover, compared to beamsteering codebooks, the columns of DFT codebooks are orthogonal. The orthogonality between columns is favorable in transmitting multiple streams (multiplexing mode) and in dividing spanned spaces into orthogonal subspaces. On the contrary, the beamsteering codebooks are obtained by uniformly quantizing the angle of departure of the transmit antenna array response vector, i.e., $a_i(\phi)$. For instance, for uniform linear array, $a_i(\phi) = \frac{1}{\sqrt{M}} \left[ 1, e^{jkd \sin \phi}, \ldots, e^{jkd(M-1) \sin \phi} \right]^H$, where $d$ is distance between two consecutive antenna elements, and $k = \frac{2\pi}{\lambda}$ is the wave number. Quantizing the angle of departure uniformly using $B = 2^q$ points, i.e., $\phi = \frac{2\pi u}{B}$, $u \in \{0, 1, \ldots, B - 1\}$, leads to a codebook given by: $a_i\left( \frac{2\pi u}{B} \right) = \frac{1}{\sqrt{M}} \left[ 1, e^{jkd \sin \left( \frac{2\pi}{B} \right)}, \ldots, e^{jkd(M-1) \sin \left( \frac{2\pi}{B} \right)} \right]^H$.

When $B \leq M$, the columns of the beamsteering codebooks are not orthogonal due to the periodicity of the sine function. Moreover, the columns of the beamsteering codebook are not necessarily asymptotically orthogonal. In particular, for some $u$ and $k \in \{0, 1, \ldots, B - 1\}$, $\sin\left( \frac{2\pi u}{B} \right) - \sin\left( \frac{2\pi k}{B} \right) = \frac{1}{B}$; checking for the orthogonality between the columns of an $M \times M$ beamsteering matrix, i.e., $[a_i^H(\frac{2\pi u}{B}) a_i(\frac{2\pi k}{B})]$, one has

$$
\frac{1}{M} |a_i^H(\frac{2\pi u}{B}) a_i(\frac{2\pi k}{B})| = \left| \frac{1 - e^{j\pi \sin(\frac{2\pi u}{B}) - \sin(\frac{2\pi k}{B})}}{M} \right|
$$

In practice, $M$ is large but still finite and thus there is a non-negligible probability that there are $u$ and $k$ such that $\sin\left( \frac{2\pi u}{B} \right) - \sin\left( \frac{2\pi k}{B} \right) \leq \frac{1}{B}$. This directly makes the performance of the DFT codebook (utilized in the proposed algorithms) better than the beamsteering codebooks (used in [2], [8], [9], [12], [16]) as shown in the numerical results section.

In contrast to the RF precoding matrix, the baseband precoder $F_{BB}$ does not have hardware constraints, i.e., $F_{BB} \in \mathbb{C}^{N_r \times N_t}$ and it can be implemented digitally. Moreover, it has relatively small dimensions of the same order of regular MIMO systems. This gives higher degrees of freedom in designing the digital codebook, i.e., $C_{BB}$. Such codebook design is well-studied in the literature on limited feedback regular MIMO [30]. The optimal yet theoretical codebooks for quantizing $F_{BB}$ are the ones that are based on Grassmannian codebooks. However, constructing an optimal Grassmannian codebook is a challenging problem [30]. In practice, it is preferable to use easily constructed and structured codebooks such as Hadamard [13], QPSK alphabet-based [34], and DFT codebooks. We consider the DFT codebooks in quantizing $F_{BB}$ since they have nested structures, are easily constructed, and provide finer quantization than Hadamard and QPSK alphabet-based codebooks.

C. PROBLEM FORMULATION AND PROPOSED ALGORITHMS

We consider minimizing the chordal distance between the unconstrained optimal precoder, given by the $N_t$ right singular vectors of the channel matrix, i.e., $F_{opt} = \tilde{V} N_t$, and the hybrid beamforming matrix, i.e., $F_{RF} F_{BB}$, such that these matrices are selected from $C_{RF}$ and $C_{BB}$, respectively. As a result, the limited feedback precoding problem for the mmWave channel with hybrid precoding structure is:

$$
\begin{equation}
\begin{aligned}
\arg \min_{F_{RF},F_{BB}} & \quad \| F_{opt} F_{opt}^H - F_{RF} F_{BB} F_{RF}^H \|_F^2 \\
\text{subject to} & \quad 1 - \| F_{RF} \|_{C_{RF,\text{skewed}}} \\
& \quad 2 - \| F_{BB} \|_{C_{BB}}
\end{aligned}
\end{equation}
$$

101504
Unfortunately, the optimization problem in (7) is non-convex as a result of the combinatorial nature of the constraints. Hence, finding its global optimal solution requires prohibitive complexity, and thereby, in practice an efficient sub-optimal solution is preferred specially in high dimension applications. Contrary to the traditional way of the exhaustive search that has been considered in [2], [35] and complex optimization algorithms [10]–[12] and inspired by the greedy selections algorithms developed in [8], [13], [16], we solve the optimization problem in (7) algorithmically in two steps. These two steps are summarized in the following.

Algorithm 1: Statistically-Aided Maximum Projection

1. **Input:** \( F_{opt}, D_M, D_{N_f}, R_f \)
2. **Indices\( _{RF} \) = arg max \( \{ \text{diag}(YY^H) \} \)
3. \( F_{RF} = (R_f D_M) \) \( F_{opt} \)
4. \( Z = (D_{N_f}^H F_{RF} D_{N_g}^H) - (F_{RF} D_{N_g})^H F_{opt} \)
5. \( Indices_{BB} = \text{arg max} \{ \text{diag}(ZZ^H) \} \)
6. \( F_{BB} = D_{N_g}(., Indices_{BB}) \)
7. **Return:** \( Indices_{RF}, Indices_{BB} \)

We point out that all the calculations of Algorithm 1 and its variants are performed at the receiver side aiming mainly to reduce the feedback overhead while exploiting the available CSIR (a prerequisite for other reception processing). Moreover, calculating \( F_{opt} \), which is based on the SVD of \( H \), does not require overhead of calculations since computing the SVD of \( H \) is a prerequisite for most of fully-digital or hybrid combining techniques [2]. We note that Algorithm 1 requires (i) limited feedback channel and (ii) statistical information about the spatial correlation. Moreover, Algorithm 1 can be modified in order to fit into different frameworks, such as the unavailability of statistical correlation matrix at the transmitter, and the presence of quasi-static channels.

**Algorithm 1 Variant 1:** When the statistical correlation information is not available at the transmitter or the channel is statistically uncorrelated, Algorithm 1 is easily modified by just replacing \( R_f \) by the identity matrix. This allows the constructed \( F_{RF} \) to have the constant magnitude entries and thereby there is no need to use \( Z \) operator in Step 4 in Algorithm 1 where \( F_{RF} = D_{M}(., Indices_{RF}) \). Moreover, it preserves the orthogonality between the DFT columns, and hence, there is no need for the matrix inverse operation of Step 5 in Algorithm 1, i.e., \( Z = (F_{RF} D_{N_g})^H F_{opt} \).

**Algorithm 1 Variant 2:** The second modification is based on utilizing the mixed CSIT instead of the limited one in order to improve the performance of Algorithm 1. The mixed CSIT assumes that the receiver sends the baseband precoder to the transmitter instantaneously and with infinite precision while the analog precoder is available to the transmitter with finite-precision (few bits). In particular, one selects the columns of \( F_{RF} \) from a statistically skewed DFT codebook and solves for \( F_{BB} \) as the least square solution such that selected columns have maximum projections on the subspace spanned by the optimal unconstrained (fully-digital) precoder. This permits the digital precoder to have entries with variable magnitude and phase, which improves the performance at the cost of increasing the feedback overhead. The hybrid beamforming procedure based on mixed CSIT is summarized in the following two steps. First, constructing the analog precoder using the same maximum projection procedure of Algorithm 1 (from line 1 to line 4). Second, given the analog precoder, and instead of executing the last part of Algorithm 1, one obtains \( F_{BB} \) by minimizing the chordal distance, in (7), while relaxing the second constraint. As a result, the baseband precoder is \( F_{BB} = (\frac{F_{RF}^H F_{RF}}{||F_{RF}||^2}) F_{RF} \). Finally, when Variant 1 and Variant 2 are combined, i.e., in the lack of statistical channel information and the availability of mixed CSIT, we call it Algorithm 1 Variant 3.

**IV. MULTI-USER PROBLEM FORMULATION AND PROPOSED ALGORITHMS**

In this section, we consider the design of the codebook-based hybrid precoding for the downlink multi-user (MU)
multiple-input single-output (MISO) broadcast (BC) system in which the base station is equipped with an $M$-element antenna array and a limited number of RF chains, and serves $K$ single-antenna users where $M \geq N_{rf} \geq K$. Before we proceed with the problem formulation, we highlight the main differences between MU MISO and SU MIMO systems that drive us to treat the problem differently. First, maximizing the sum-rate of the $K$-user MISO BC requires sophisticated and computationally intensive non-linear operations such as dirty paper coding [37] and its optimal solution does not adhere to the RF constraints on the analog precoder. Second, since all users (receivers) are separated, where no joint processing of their signals can be done at receivers, approximating sub-optimal linear precoding schemes, such as zero forcing [38], minimum mean square error and generalized eigenvector beamforming [39], is not applicable in our framework. This is mainly because all these solutions require either global beamforming [39], is not applicable in our framework. This is mainly because all these solutions require either global

where SINR $k$ is the signal to interference noise ratio (SINR) for user $k$. The optimal solution is typically not achievable in practice, especially in scenarios with large numbers of users and simultaneous transmissions. Therefore, various approximations and suboptimal solutions have been proposed to find near-optimal solutions.

The received signal of the single-antenna user $k$ is:

$$ y_k = h_k^H f_{BB,k} \tilde{s}_k + \sum_{i \neq k} h_i^H f_{BB,i} s_i + n_k, \quad (8) $$

where $f_{BB,k}$ is the baseband precoder vector for user $k$, i.e., the $k^{th}$ column of $F_{BB}$ and $h_k^H$ is the channel row vector of length $M \times 1$ for user $k$ where $h_k = R_{i,k}^{1/2} h_{w,k}$ and $h_{w,k}$ has i.i.d Gaussian distributed entries. Moreover, we assume that the users have different spatial correlation matrices $R_{i,k} = \mathbb{E}[h_i h_k^H]$. Considering an equal power normalized transmission strategy, the achievable rate of user $k$ is:

$$ r_k = \log_2 (1 + \text{SINR}_k) \quad (9) $$

$$ \text{SINR}_k = \frac{|h_k^H f_{BB,k}|^2}{\sum_{i \neq k} |h_i^H f_{BB,i}|^2 + K \sigma_n^2}, \quad (10) $$

where SINR $k$ is the signal to interference noise ratio of user $k$. Therefore, the sum-rate of the $K$-user broadcast channel is given as $R_{sum} = \sum_k r_k$.

Optimizing the sum-rate and other sophisticated performance metrics such as energy efficiency and bit error rate while considering the RF hardware constraints results in notorious non-convex and sparse problem formulations where a series of convex relaxations and approximations are performed to secure satisfactory sub-optimal solutions [10]–[12]. These solutions are obtained based on a series of sophisticated tangled iterative algorithms that require either modern or classic convex solvers such as MOSEK and interior point, respectively, where their convergence depends highly on the initial point.

Given the intractability and impracticality associated with applying the prior art of fully-digital [37] and hybrid beamforming [10]–[12] methodologies in massive MIMO mmWave systems, we tackle the design problem differently. Particularly, the proposed scheme is developed based on maximizing the signal to interference plus noise ratio (SINR) over two successive stages using non-iterative and low-complexity procedure; for instance, one iteration of the algorithm in [12] has a computational complexity of $O(M^6 + 64K^2 + 6K^2M^2)$ [12] while Algorithm 2 has a computational complexity of $O(M^3K + 8MK + 9K^3 + 2MK^2)$ where its complexity is dominated by the Moore-Penrose pseudo-inverse [40]. We show that this approach achieves higher sum-rate than prior art in [10]–[12] and comparable to fully-digital scheme with perfect CSIT (see Fig. 4, 5 and 6). This is based on leveraging the property of spatially correlated channels that the channel vectors of different users exist in different subspaces identified by the statistical covariance matrices and point to specific preferred directions. This can be utilized to relieve the global perfect channel knowledge assumption, consolidate the separability of users in the space, and enhance the performance of the hybrid precoding schemes. Particularly, exploiting the directivity of subspaces of users’ channels, each user designs its analog precoder vector(s) selfishly to maximize its signal strength while ignoring the interference. Its aim is to decrease the feedback overhead and for CSIT where this step requires only local statistical knowledge of its own channel vector at the receiver and limited feedback channel. In the second stage, the digital precoder is designed based on the effective channel, i.e., $H^H f_{RF}$, in order to cancel out the inter-user interference.

Therefore, the analog precoding problems, while abstracting the digital processing, are given by:

$$ \arg\max_{f_{RF,k}} |h_k^H f_{RF,k}|^2, \quad 1 \leq k \leq K $$

subject to $f_{RF,k} \in C_{RF,skewed,k}, \quad 1 \leq k \leq K, \quad (11) $
RF precoder matrix by selecting the preferred $K$ different codewords out of the total $K \times L$ codewords received from all $K$ receivers. Each of the preferred $K$ codewords (vectors) is corresponding to one respective user. Particularly, each user’s receiver transmits a set of indices indicating the best $L$ columns in descending order (or ascending order) such that the first (or last) element in the set of indices indicates the column along which the channel vector has the maximum projection. In the case where two or more receivers share one or more codewords, the base station selects the next best codeword(s) out of the $L$ codewords (i.e., using the next entry in the set of indices), such that all the columns of $\mathbf{F}_{RF}$ are different codewords. Hence, the base station can select $K$ DFT columns corresponding to respective $K$ MU receivers, to form the RF precoder matrix as $\mathbf{F}_{RF} = [\mathbf{f}_{RF,1}, \mathbf{f}_{RF,2}, \cdots, \mathbf{f}_{RF,K}]$. This selection strategy is devised to enforce the full rank constraint on $\mathbf{F}_{RF}$ which is required to achieve the highest multiplexing gain offered by the network. We note here that $L$ is identified empirically since it is a function of the number of common clusters/scatters between users and the location of users.

We also note that this precoding strategy can be extended to the case where there are more RF chains than the number of users. When $2K \geq N_{rf} > K$, the base station selects the columns of the analog precoder in two rounds. First, it constructs the first $K$ columns of $\mathbf{F}_{RF}$, similar to the previous strategy, by selecting the first entries of all users’ sets, i.e., the indices of the columns that have the largest projection, such that they are different. Then, for the remaining $N_{rf} - K$ columns of $\mathbf{F}_{RF}$, the base station selects the second entries of only $N_{rf} - K$ user’s sets.

On the other hand, the digital precoder problem is given based on the effective channel as:

$$\arg\max_{\mathbf{F}_{BB}} \frac{|\mathbf{h}_{\text{eff},k}^{H} \mathbf{f}_{BB,k}|^2}{\sum_{l \neq k} |\mathbf{h}_{\text{eff},l}^{H} \mathbf{f}_{BB,l}|^2 + K\sigma_n^2}, \quad 1 \leq k \leq K,$$

subject to $||\mathbf{F}_{RF}\mathbf{F}_{BB}||_F^2 \leq K$ (12)

Contrarily to the analog precoding, the digital precoder has an optimal solution given by zero-forcing [41]. Having determined the RF precoder matrix $\mathbf{F}_{RF}$, the base station then determines the baseband precoder matrix $\mathbf{F}_{BB}$ based on the effective channel after applying the RF precoder matrix. Particularly, each user’s receiver feeds back its estimated effective channel vector $\mathbf{h}_{\text{eff}}$. The baseband precoder matrix $\mathbf{F}_{BB}$ is determined by the base station to be the zero-forcing solution based on these effective channel vectors, represented by $\mathbf{H}_{\text{eff}} = \mathbf{H}_{RF}$, such that $\mathbf{F}_{BB} = \mathbf{H}_{\text{eff}}^{H} (\mathbf{H}_{\text{eff}} \mathbf{H}_{\text{eff}}^{H})^{-1}$. We note here that the effective channel has small dimensions in comparison to the massive MIMO channel. This significantly reduces the channel training and feedback required for estimating and acquiring the effective channel. These two steps are summarized in Algorithm 2.

**Algorithm 2 Distributed Statistically-Aided Maximum Projection Multi-User**

1: **Input:** $\mathbf{H}, \{\mathbf{R}_1, \cdots, \mathbf{R}_K\}, \mathbf{D}_M$
2: While $k \leq K$
3: $\mathbf{Y}_k = (\mathbf{R}_k \mathbf{D}_M)^{H} \mathbf{h}_k$
4: $\text{Indices}_{RF,k} = \arg\max_{L} (\text{diag}(\mathbf{Y}_k \mathbf{Y}_k^{H}))$
5: End while
6: $\mathbf{F}_{BB} = \mathbf{F}_{RF}^{H} \mathbf{H} (\mathbf{H}^{H} \mathbf{F}_{RF} \mathbf{F}_{RF}^{H})^{-1}$
7: $\mathbf{F}_{BB} = \sqrt{\frac{K}{M}} \mathbf{F}_{BB}$
8: Return $\mathbf{F}_{RF}$ and $\mathbf{F}_{BB}$

We note that Algorithm 2 requires mixed CSIT; specifically (i) limited feedback channels and (ii) statistical information about the spatial correlation for constructing the analog precoder and (iii) perfect knowledge about the low dimensional effective channel. Similar to Algorithm 1 Variant 1, Algorithm 2 can be modified in order to accommodate for the unavailability of the transmit covariance matrix at the transmitter. We refer to this modification as Algorithm 2 Variant 1.

**V. BOUNDS AND ASYMPTOTIC ANALYSIS OF THE ACHIEVABLE RATE OF THE PROPOSED SCHEMES**

In order to evaluate the performance of Algorithm 1, Algorithm 2 and their variants, we consider the mutual information of the analog precoding stage (the common stage among the proposed schemes) as a performance metric. This is made possible by abstracting the digital precoding and receiver side processing while considering an equal power transmission strategy. We start by providing lower and upper bounds on the mutual information of any DFT codebook-based analog precoding strategy. Then, we show that the proposed schemes are asymptotically optimal as the number of transmit antennas $M$ goes to infinity and the millimeter wave channel has a limited number of paths, i.e., $P < M$.

**Theorem 1:** The mutual information of the mmWave channel with hybrid precoding structure at the transmitter, where the analog precoder is selected from an $M \times M$ DFT matrix, i.e., $\log \det (\mathbf{I}_{N_{rf}} + \frac{P}{\sigma_n^2} \mathbf{F}_{RF}^{H} \mathbf{H} \mathbf{H}^{H} \mathbf{F}_{RF})$, is bounded by (13) shown at the bottom of the next page. In Eq. (13), $\lambda_1(\mathbf{A}) \geq \cdots \geq \lambda_M(\mathbf{A})$ are the eigenvalues of the matrix $\mathbf{AA}^{H}$ in a descending order.
Proof: We start with the mutual information of the channel when using the analog beamformer at the transmitter:

\[ R = \log \det \left( I_{N_{rf}} + \frac{\rho}{\sigma_n^2 N_s} F_{RF}^H H^H H F_{RF} \right) \]

(a) \[ \implies \log \det \left( I_{N_{rf}} + \frac{\rho}{\sigma_n^2 N_s} F_{RF}^H A_t G H A_t^H A_t G A_t^H F_{RF} \right) \]

(b) \[ \implies \log \det \left( I_{N_{rf}} + \frac{\rho}{\sigma_n^2 N_s} U_X \Sigma_X V_X^H G H A_t G V_X \Sigma_X^T U_X^H \right) \]

(c) \[ \implies \log \det \left( I_{N_{rf}} + \frac{\rho}{\sigma_n^2 N_s} \Sigma_X G H A_t^H A_t G \Sigma_X^T \right) \]

(d) \[ \implies \log \det \left( I_{N_{rf}} + \frac{\rho}{\sigma_n^2 N_s} \Sigma_X \Sigma_Y \Sigma_Y^T \right) \]

(e) \[ \sum_{i=1}^{N_{rf}} \log \left( 1 + \frac{\rho}{\sigma_n^2 N_s} \Sigma_X^2 \Sigma_Y^2 \right) \]

where (a) is due to the use of the augmented matrix representation of the channel in (4) and (b) is based on considering the SVD of \( X = F_{RF}^H A_t = U_X \Sigma_X V_X^H \) and \( Y = A_t G = U_Y \Sigma_Y V_Y^H \). Eq. (c) and (d) are due to the invariance of the mutual information formula under unitary transformation, i.e., multiplication by \( U_X, V_X, U_Y, V_Y \), and \( \det(A^H) = \det(A) \). Eq. (e) comes directly from (d) because \( I_{N_{rf}}, \Sigma_X, \) and \( \Sigma_Y \) are diagonal matrices where \( \det(\Sigma) = \prod_{i=1}^{N_{rf}} |\Sigma_{i,i}| \) and \( \log(\prod) = \sum \log(\cdot) \). Bounding (e) requires bounding the non-negative elements of \( \Sigma_X \) where the effect of precoding is only captured in \( \Sigma_X \). The elements of \( \Sigma_X^2 \) correspond to the eigenvalue decomposition of \( F_{RF}^H A_t A_t^H F_{RF} \) where \( F_{RF} \) is a part of an \( M \times M \) DFT matrix, i.e., \( F_{RF} = D_{N_{rf}} \) and \( D_{N_{rf}} \) denotes a collection of \( N_{rf} \) columns selected from an \( M \times M \) DFT matrix. Since the \( D_{N_{rf}} \) is a semi-unitary matrix, i.e., \( D_{N_{rf}}^H D_{N_{rf}} = I_{N_{rf}} \) and using Lemma 3.3.1 in [42], the eigenvalues of the product, i.e., \( \lambda_i(D_{N_{rf}}^H A_t) \), can be bounded by:

\[ \lambda_{M-N_{rf}+i}(A_t) \leq \lambda_i(D_{N_{rf}}^H A_t) \leq \lambda_i(A_t), \quad (15) \]

where \( \lambda_i(D_{N_{rf}}^H A_t) = |\Sigma_X_{i,i}| \). Substituting these bounds in (e), (13) readily follows. \( \square \)

Similar bounds can be obtained as a function of the transmit correlation matrix \( R_t \) instead of \( A_t \) by using (2) instead of (4).

Remark 1 (SNR Gap): Indeed, there is a wide SNR gap between the upper and lower bounds in (13) and it is more pronounced at small values of \( N_{rf} \) as shown in Fig. 1 subplots (a) and (b). This is expected since the proposed bounds are valid for any DFT-based hybrid precoding scheme. However, this SNR gap reduces gradually with increasing \( N_{rf} \) till reaching to zero, i.e., both bounds coincide, at \( N_{rf} = M \) as shown in Fig. 1 (c) (subplot (c) shows the SNR gaps between different schemes and the upper bound). The wide SNR gap between the upper and lower bounds at small values of \( N_{rf} \) suggests that the DFT columns of \( F_{RF} \) should be carefully selected to achieve the upper bound while the small SNR gap at large values of \( N_{RF} \) hints that no sophisticated selection is required and just fixed or random selection is sufficient.

Remark 2 (Design Insight): The upper bound in (13) is achieved when all the selected DFT columns are in the same directions of the largest \( N_{rf} \) eigenvectors of \( A_t \). On the other hand, the lower bound in (13) is achieved when none of the selected DFT columns are pointing toward any of the largest \( N_{rf} \) eigenvectors’ directions of \( A_t \). This suggests that efficient DFT-based hybrid precoding schemes should consider aligning the DFT columns of \( F_{RF} \) to the eigenvectors of \( A_t \) and its efficiency is measured by how close its mutual information curve to the upper bound. On the other hand, inefficient DFT-based hybrid precoding schemes do not consider any alignment technique, and thereby, their performance is close to the lower bound. For instance, comparing subplots (a) and (b) in Fig. 1, one finds the mutual information curves of the proposed analog precoder, denoted by (Maximum Projection), outperform the ones of the prior art scheme developed in [15] for different values of \( N_{rf} \). This is because, in the proposed analog strategy, the \( l \)-th DFT column is selected to be aligned to the \( l \)-th eigenvector of \( R_t \) whereas, in [15], the columns of \( F_{RF} \) are either randomly selected from an \( M \times M \) DFT matrix, or are represented as the first \( N_{rf} \) columns of an \( M \times M \) DFT matrix.
Corollary 1: For a large uniform linear antenna array where $M \to \infty$ in a limited scattering mmWave channel, defined in (3), and $P < M$, the proposed analog precoding strategy (the common procedure among the proposed algorithms) achieves the upper bound on the mutual information in (13).

Proof: For large $M$ ($M \to \infty$ and $P < M$), the columns of $A_t$ are asymptotically orthogonal and have unit norms [43]. Moreover, $A_t$ and the optimal precoder, given by the largest right singular vectors of the channel matrix, i.e., $F_{opt} = \overline{V}_N$, span the same subspace. Therefore, the channel matrix representation in (4) converges to its SVD [43, Lemma 2], i.e., $A_t$ and $A_r$ converge to $V_H$ and $U_H$, respectively. Considering ULA at the base station, the columns $A_t$ have the same structure as the DFT matrix; hence the optimal precoder, i.e., $F_{opt}$, has a DFT structure as well [4]. Consequently, the dot product between the $M \times M$ DFT matrix and the optimal digital precoder is sufficient to select the DFT columns that are perfectly aligned to the optimal precoder vectors, i.e., the largest eigenvectors of $A_t$. Since $F_{RF}$ in all the proposed algorithms is constructed based on these $N_f$ DFT columns, its columns are asymptotically aligned to the eigenvectors of $A_t$ as well. Consequently, and based on [42, Lemma 3.3.1], the proposed analog procedure asymptotically achieves the upper bound in (13).

VI. SIMULATION RESULTS

In this section, we evaluate the performance of the proposed algorithms, Algorithm 1 and Algorithm 2, and their variants. All these hybrid beamforming schemes are compared with the prominent prior works, mentioned in Section I, in terms of spectral efficiency over mmWave bands. Since, we are mainly concerned with decreasing the channel feedback overhead, we limit the application of the proposed algorithms to the transmitter side. For the single-user case, the hybrid precoder is obtained by either Algorithm 1 or its variants (depending on the available CSIT) while the hybrid combiner, i.e., $W_{RF} W_{BB}$, is obtained by approximating the dominant $N_f$ left singular vectors of the channel, i.e., $U_N$, using the procedure described in [2, equation (16)-(18)]. On the other hand, for multi-user case, since all users are equipped with single antennas, there is no combining available at the users’ sides.

A. SIMULATION SETUPS

We consider the clustered channel model, described in (3) where the complex channel gains, i.e., $a_{kl}$, are i.i.d. $CN(0, \sigma^2_{h,k,l})$ and $\sigma^2_{h,k,l}$ are randomly generated from an exponential distribution and normalized such that $\sum_k \sigma^2_{h,k,l} = 1$ [2], [5], [6]. Moreover, we assume that the total number of paths/trays, $P$, is sufficiently larger than the number of transmitted streams per user, i.e., $P = N_{ray} N_{cl} > N_t$ and $P > 1$ for single and multi-user cases, respectively. For the single-user case, the transmitter is equipped with an antenna array with 64 elements, and the receiver is equipped with an antenna array with 16 elements. On the other hand, for the multi-user case the base station is equipped with 64 or 32 elements and the number of user varies between 4 and 20. Moreover, we assume that the number of users $K$ is smaller than the number of antenna elements $M$. Thus, there is no need to consider any users’ scheduling or opportunistic selection strategy and instead, we serve all users. In both cases, we consider ULA where the transmit/receive antenna array responses of ULA with half wave length element spacing and $M$-element is given by:

$$a_{t/r}(\theta^{\alpha}_{t/r}) = \frac{1}{\sqrt{M}} [1, e^{j \pi \sin(\theta^1_{t/r})}, \ldots, e^{j(M-1)\pi \sin(\theta^M_{t/r})}]^H$$

(16)

The angles of departure, $\theta^1_t$, and arrival $\theta^1_r$, are drawn from Laplace distributions with means $\bar{\theta}^1_t$ and $\bar{\theta}^1_r$ and angular spread $\Delta_\theta$ of $7.5^\circ$ [2], [6]. Accordingly, the $(m,n)$ entry of the transmit correlation matrix of user $k$, i.e., $R_{t,k}$, for ULA with half wave length element spacing is [4]–[6]:

$$[R_{t,k}]_{m,n} = \frac{1}{N_{cl}} \sum_{l=1}^{N_{cl}} 1 \int e^{-\frac{1}{2} \pi |\psi - \bar{\theta}^1_t - j\pi (m-n)\sin\psi|^{2}} d\psi,$$

(17)

B. SPECTRAL EFFICIENCY PERFORMANCE EVALUATION

From the rich literature on hybrid beamforming, we choose the most prominent and relevant prior art schemes as benchmarks in order to evaluate the efficacy of the proposed schemes. The first benchmark solves the hybrid beamforming problem algorithmically based on an efficient compressed sensing algorithm, namely, orthogonal matching pursuit (OMP) [2] where the analog precoder is represented as linear combinations of the columns of $A_t$. This scheme requires perfect CSIT since $A_t$ changes with each channel realization. Therefore, this benchmark works as an upper bound to its exhaustive search-based limited feedback version in [2] and all hybrid beamforming schemes that utilize mixed or limited feedback CSIT. We call this benchmark by spatially sparse OMP. The second also utilized the OMP algorithm, however, the analog beamformer is obtained based on the beamsteering codebook whereas the digital precoder is designed to eliminate the inter-user interference based on the zero-forcing approach [8], [16]. We call this benchmark by codebook-based spatially sparse [8], [16]. This scheme requires mixed CSIT and it represents the prior works that utilizes beamsteering codebooks and compressed sensing greedy algorithms [9], [13], [14], [16], [17]. We note here that the schemes in [14], [17] are very similar to the ones in [8], [16] expect the digital precoders of the former are designed to minimize the mean squared error while those of the latter are designed to maximize the spectral efficiency. The third is also a codebook-based hybrid beamforming scheme where the analog precoder is selected from beamsteering codebooks using a MOSEK-based algorithm where its selection metric is based on maximizing the sum-rate. We refer to this scheme as codebook-based sum-rate.
maximization [12]. This scheme also requires mixed CSIT and it represents the prior works that utilize complex iterative processing [10], [11]. The fourth is developed in [6] where the analog precoders are jointly designed at the base station to maximize the signal to leakage and noise ratio (SLNR) based on the second-order statistical channel knowledge of all users; specifically, the analog precoder is given as the largest $K$ eigenvectors of the sum of the transmit covariance matrices of the users. Moreover, the digital precoder is designed, based on perfect effective channel knowledge, to minimize the SINR. We refer to this benchmark as SLNR-based statistical beamforming [6]. The fifth is widely known as the joint spatial division multiplexing scheme (JSDM) [4], [5], [7]). The basic idea of JSDM is to partition the users into groups where users with similar covariance matrices are grouped together while maintaining orthogonality between the groups. The analog RF beamforming is designed to reduce the inter-group interference by employing the well-known block diagonalization technique using only statistical knowledge where the analog beamforming is the augmented matrix of the largest eigenvectors of the covariance matrix of the users’ groups. On the other hand, the digital baseband precoding is designed to eliminate the inter-user interference between users in the same group using linear precoders based on perfect channel knowledge.

In Figs. 2 and 3, we plot the spectral efficiency of five different hybrid beamforming schemes for the single-user case, namely, spatially sparse hybrid precoding based on OMP [2], shown in red solid line with circles, Algorithm 1, shown in sky blue solid line, Algorithm 1 Variation 1 (Alg1.Var1), shown in solid blue line with squares, Algorithm 1 Variation 2 (Alg1.Var2), shown in green solid line with plus signs, Algorithm 1 Variation 3 (Alg1.Var3), shown in yellow solid line with crosses and the codebook-based spatially sparse hybrid precoding scheme in [8], shown in solid violet line with diamonds in addition to the fully-digital SVD based precoder, shown in solid black line.

Fig.2 shows that Alg1.Var2 and Alg1.Var3 have almost the same performance as the spatially sparse OMP-based scheme which requires full CSIT [2] while Alg1.Var2 and Alg1.Var3 require statistically-aided mixed CSIT and mixed CSIT, respectively. This shows the efficiency of Alg1.Var2 and Alg1.Var3 in reducing the feedback overhead (where only knowledge about the low-dimensional effective channel is required) while having marginal spectral efficiency degradation. Moreover, Fig.2 demonstrates that Algorithm 1 and Alg1.Var1 both outperform the performance of the codebook-based spatially sparse benchmark [8]. This, in turn, significantly reduces the required channel knowledge to achieve the same high spectral efficiency performance since Algorithm 1 and Alg1.Var1 require limited feedback CSIT while the scheme in [8] requires mixed CSIT. We note here that the impact of the statistical knowledge on increasing spectral efficiency is marginal where the curves of Alg1.Var2 and Alg1.Var3, and Algorithm 1 and Alg1.Var1 are on top of each other.

In order to evaluate the impact of the statistical channel knowledge on the spectral efficiency, in Fig. 3, we simulated all these schemes in the same set-up of Fig. 2 except codebooks’ lengths in this set-up are decreased from 64 to 32. Fig.3 shows that even though there is some performance loss due to the reduction in codebooks’ lengths, the statistically-aided schemes generally outperform the ones without statistical knowledge. Moreover, in comparison with Fig. 2, Fig. 3 illustrates that the spectral efficiency loss of the statistically-aided schemes, i.e., Algorithm 1 and Alg1.Var2, due to the reduction in codebooks’ lengths is less severe than loss of the ones without statistical knowledge, i.e., Alg1.Var1 and Alg1.Var3. Particularly, the required SNR gains that compensate for these losses are 2 dB and 4 dB, respectively. Further, from Fig.2, it is clear that the statistical channel knowledge has a marginal impact on the proposed hybrid beamforming schemes since there is almost no performance gap between the ones that are statistically-aided ones and the others. Thus, we can infer from Figs. 2 and 3 that the impact of the statistical knowledge on increasing the spectral efficiency performance diminishes by increasing codebooks’ lengths or equivalently the degradation in the spectral efficiency...
A. W. Shaban et al.: Statistically-Aided Codebook-Based Hybrid Precoding for Millimeter Wave Channels

FIGURE 4. Sum-rate comparison of different hybrid beamforming schemes over 10-user MISO BC mmWave channel with $M = 32$, $N_{cl} = 3$ and $N_{ray} = 2$, and $L = 2$ for Algorithm 2 and Alg2.Var1.

FIGURE 5. Sum-rate comparison of different hybrid beamforming schemes over 16-user MISO BC mmWave channel with $M = 64$, $N_{cl} = 3$ and $N_{ray} = 2$, and $L = 2$ for Algorithm 2 and Alg2.Var1.

FIGURE 6. Sum-rate comparison of different hybrid beamforming schemes over $K$-user MISO BC mmWave channel with $M = 64$, $N_{cl} = 3$ and $N_{ray} = 2$, and $L = 2$ for Algorithm 2 and Alg2.Var1.

performance due to the lack of statistical knowledge can be compensated by increasing the rate of the feedback channels.

In Figs. 4, 5, and 6, we plot the spectral efficiency of five different hybrid beamforming schemes for the multi-user case, namely, Algorithm 2, shown in blue solid line with triangles, Algorithm 2 Variation 1 (Alg2.Var1), shown in red solid line with circles, codebook-based sum-rate maximization [12], shown in violet solid line with points, SLNR-based statistical beamforming [6], shown in solid green line with crosses, and JSDM [4], [5], [7], shown in yellow solid line with diamonds in addition to the fully-digital zero-forcing precoding scheme, shown in solid black line. It is worth mentioning that all the simulated hybrid beamforming schemes require perfect knowledge of the effective channel after the analog precoding while the last two benchmarks require only second order statistical knowledge for constructing the analog precoders (none codebook-based).

Fig. 4 illustrates that the spectral efficiency performance of the proposed schemes for the $K$-user MISO broadcast channel outperforms all the benchmarks that utilize different channel knowledge. Particularly, there are SNR gaps of 3 dB, 4 dB and 8 dB between the performance of Algorithm 2 and SLNR-based statistical beamforming, codebook-based sum-rate maximization and JSDM. We note here that although JSDM has been shown to be asymptotically optimal under certain channel conditions, it suffers from severe performance degradation as depicted in Fig. 4. This is mainly because it requires a sophisticated user grouping strategy and utilizes the orthogonality between the groups of users. However, here, we consider more realistic channel models where they are characterized by multiple scattering clusters ($N_{cl} = 3$), and where these clusters may overlap.

In Fig. 5, the proposed schemes maintain their superior spectral efficiency performance compared to the prior arts even when the number of users increased from 10 to 16 while increasing the number of transmitting elements from 32 to 64 in comparison to Fig. 4. Moreover, Fig. 5 shows that the impact of the statistical knowledge on increasing the spectral efficiency of the proposed schemes (e.g., compare the curves of Algorithm 2 and Alg2.Var1) is more pronounced at higher channel dimensions. This is expected since with higher channel dimensions the statistical knowledge becomes more beneficial in directing the information towards users.

In Fig. 6, we study the effect of increasing the number of users $K$ while fixing the number of transmitting elements to 64. From Fig. 6, one can observe that the prior art schemes, which utilize only the transmit covariance matrix in the analog beamforming such as JSDM [4], [5], [7] and statistical beamforming SLNR maximization scheme [6], suffer from significant sum-rate flooring or even plunging when $K$ is increased gradually in contrast with the codebook-based schemes such as the proposed schemes and codebook-based sum-rate maximization [12]. Moreover, this illustrates that the proposed schemes maintain their superior sum-rate performance over the entire range of $K$ in comparison with benchmarks. Comparing the sum-rate curves of Algorithm 2 and Alg2.Var1, one deduces that the significance of exploiting the statistical knowledge in enhancing the
sum-rate increases with increasing the number of the users, for instance, at $K = 8$, both curves are on top of each other while, at $K = 20$, there is a sum-rate gap of 13 bits/s/Hz in favor to Algorithm 2 (the statistically-aided one). This is mainly due to the considerable role of statistical knowledge in decreasing the inter-user interference by separating users in space based on their covariance matrices.

VII. CONCLUSION

We considered single-user and multiple-user MIMO hybrid (analog/digital) precoding. Utilizing the spatial correlation and sparsity the channel of recent wireless communication systems such as massive MIMO systems working in millimeter wave bands, we developed practical and simple codebook-based hybrid precoding strategies assuming limited feedback channel or (mixed) partial channel knowledge while exploiting the statistical information of these channels. The proposed algorithms are designed efficiently to achieve high spectral efficiency while decreasing the feedback overhead by constructing the hybrid precoders based on statistically skewed DFT cookbooks. Numerical results showed that the proposed algorithms allow millimeter wave systems with much less channel knowledge and feedback overhead to approach the achievable rates of prior arts’ schemes that require perfect and mixed channel knowledge. Moreover, numerical results illustrated the potency of assisting the limited feedback systems with statistical information where, for the single-user case, it enhances the immunity of the system’s spectral efficiency against the feedback rate reduction, whereas for the multi-user cases, it consolidates the separability of the users and thereby increase the sum-rate of the network. In summary, the advantages of the proposed schemes over the state-of-the-art schemes are their simplicity, low-overhead requirement and near-optimal spectral efficiency.

REFERENCES

[1] A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda. “Hybrid beamforming for massive MIMO: A survey,” IEEE Commun. Mag., vol. 55, no. 9, pp. 134–141, Sep. 2017.
[2] 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.
[3] 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.
[4] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, “Joint spatial division and multiplexing—The large-scale array regime,” IEEE Trans. Inf. Theory, vol. 59, no. 10, pp. 6441–6463, Oct. 2013.
[5] A. Adhikary, E. Al Safadi, M. K. Samimi, R. Wang, G. Caire, T. S. Rappaport, and A. F. Molisch, “Joint spatial division and multiplexing for mm-Wave channels,” IEEE J. Sel. Areas Commun., vol. 32, no. 6, pp. 1239–1255, Jun. 2014.
[6] S. Park, J. Park, A. Yazdan, and R. W. Heath, Jr., “Exploiting spatial channel covariance for hybrid precoding in massive MIMO systems,” IEEE Trans. Signal Process., vol. 65, no. 14, pp. 3818–3832, Jul. 2017.
[7] J. Jin, C. Xiao, W. Chen, and Y. Wu, “Channel-statistics-based hybrid precoding for millimeter-wave MIMO systems with dynamic subarrays,” IEEE Trans. Commun., vol. 67, no. 6, pp. 3991–4003, Jun. 2019.
[8] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, Jr., “Hybrid precoding for millimeter wave cellular systems with partial channel knowledge,” in Proc. Inf. Theory Appl. Workshop (ITA), Feb. 2013, pp. 1–5.
[9] 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.
[10] X. Gao, L. Dai, C. Yuen, and Z. Wang, “Turbo-like beamforming based on tabu search algorithm for millimeter-wave massive MIMO systems,” IEEE Trans. Veh. Technol., vol. 65, no. 7, pp. 5731–5737, Jul. 2016.
[11] J.-C. Chen, “Efficient codebook-based beamforming algorithm for millimeter-Wave massive MIMO systems,” IEEE Trans. Veh. Technol., vol. 66, no. 9, pp. 7809–7817, Sep. 2017.
[12] S. He, J. Wang, Y. Huang, B. Ottersten, and W. Hong, “Codebook-based hybrid precoding for millimeter wave multiuser systems,” IEEE Trans. Signal Process., vol. 65, no. 20, pp. 5289–5304, Oct. 2017.
[13] H. Seleem, A. I. Sulyman, and A. Alsayed, “Hybrid precoding-beamforming design with Hadamard RF codebook for mmWave large-scale MIMO systems,” IEEE Access, vol. 5, pp. 6813–6823, 2017.
[14] J. Mao, Z. Gao, Y. Wu, and M.-S. Alouini, “Over-sampling codebook-based hybrid minimum sum-mean-square-error precoding for millimeter-wave 3D-MIMO,” IEEE Wireless Commun. Lett., vol. 7, no. 6, pp. 938–941, Dec. 2018.
[15] W. Tan, M. Matthaiou, S. Jin, and X. Li, “Spectral efficiency of DFT-based processing hybrid architectures in massive MIMO,” IEEE Wireless Commun. Lett., vol. 6, no. 5, pp. 586–589, Oct. 2017.
[16] X. Liu, W. Zou, and S. Chen, “Joint design of analog and digital codebooks for hybrid precoding in multi-user massive MIMO systems,” IEEE Access, vol. 6, pp. 69818–69825, 2018.
[17] K. Ying, Z. Gao, S. Lyu, Y. Wu, H. Wang, and M. Alouini, “GMD-based hybrid beamforming for large reconfigurable intelligent surface assisted millimeter-wave massive MIMO,” IEEE Access, vol. 8, pp. 19530–19539, 2020.
[18] T. E. Bogale and L. B. Le, “Massive MIMO and mmWave for 5G wireless HetNet: Potential benefits and challenges,” IEEE Veh. Technol. Mag., vol. 11, no. 1, pp. 64–75, Mar. 2016.
[19] S. Hur, S. Bae, B. Kim, Y. Chang, A. F. Molisch, T. S. Rappaport, K. Haneda, and J. Park, “Proposal on millimeter-wave channel modeling for 5G cellular system,” IEEE J. Sel. Topics Signal Process., vol. 10, no. 3, pp. 454–469, Apr. 2016.
[20] M. A. Albreem, M. Juntti, and S. Shababuddin, “Massive MIMO detection techniques: A survey,” IEEE Commun. Surveys Tuts., vol. 21, no. 4, pp. 3109–3132, 4th Quart., 2019.
[21] X. Rao and V. K. N. Lau, “Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems,” IEEE Trans. Signal Process., vol. 62, no. 12, pp. 3261–3271, Jun. 2014.
[22] Z. Gao, L. Dai, Z. Wang, and S. Chen, “Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO,” IEEE Trans. Signal Process., vol. 63, no. 23, pp. 6169–6183, Dec. 2015.
[23] C. Qi, Y. Huang, S. Jin, and L. Wu, “Sparse channel estimation based on compressed sensing for massive MIMO systems,” in Proc. IEEE Int. Conf. Commun. (ICC), Jun. 2015, pp. 4558–4563.
[24] 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.
[25] T. S. Rappaport, R. W. Heath, Jr., R. C. Daniels, and J. N. Murdock, Millimeter Wave Wireless Communications. London, U.K.: Pearson, 2015.
[26] S. Park and R. W. Heath, Jr., “Spatial channel covariance estimation for the hybrid MIMO architecture: A compressive sensing-based approach,” IEEE Trans. Wireless Commun., vol. 17, no. 12, pp. 8047–8062, Dec. 2018.
[27] D. J. Love and R. W. Heath, Jr., “Limited feedback unitary precoding for spatial multiplexing systems,” IEEE Trans. Inf. Theory, vol. 51, no. 8, pp. 2967–2976, Aug. 2005.
[28] J. S. Dhillon, R. W. Heath, Jr., T. Strohmer, and J. A. Tropp, “Constructing packings in Grassmannian manifolds via alternating projection,” Experim. Math., vol. 17, no. 1, pp. 9–35, 2008, doi: 10.1080/10586458.2008.10129018.
[29] D. J. Love and R. W. Heath, Jr., “Limited feedback diversity techniques for correlated channels,” IEEE Trans. Veh. Technol., vol. 55, no. 2, pp. 718–722, Mar. 2006.
[30] D. J. Love, R. W. Heath, Jr., V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, “An overview of limited feedback in wireless communication systems,” IEEE J. Sel. Areas Commun., vol. 26, no. 8, pp. 1341–1365, Oct. 2008.
[31] R. W. Heath, N. Gonzalez-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.
AHMED WAGDY SHABAN (Student Member, IEEE) received the B.Sc. degree in electrical engineering from Alexandria University, Egypt, in 2010, and the M.Sc. degree in wireless communications and information technology form Nile University, Egypt, in 2014. He is currently pursuing the Ph.D. degree with the University of Waterloo. He was working as a Research Assistant with Qatar University, Qatar, and a Visiting Research Assistant with the Wireless Intelligent Networks Center (WINC), Nile University. His research interests include interference alignment and management, and beamforming for massive MIMO millimeter wave systems. He received the Faculty of Engineers (FOE) Award and the Best Teaching Assistantship Award from the Electrical and Computer Engineering Department, University of Waterloo, in fall 2016 and spring 2018, respectively.

OUSSAMA DAMEN (Senior Member, IEEE) received the Ph.D. degree in electronics and communications from the École Nationale Supérieure des Télécommunications (ENST) de Paris, Paris, France, in October 1999. He was a Postdoctoral Research with ENST, from November 1999 to August 2000, and with the Electrical and Computer Engineering Department, University of Minnesota, from September 2000 to March 2001. From March 2001 to June 2004, he was with the Electrical and Computer Engineering Department, University of Alberta, working as a Senior Research Associate of Alberta Informatics Circle of Research Excellence (CORE). In June 2004, he joined the Electrical and Computer Engineering Department, University of Waterloo, Waterloo, ON, Canada, where he is currently working as a Professor. He held a visiting position at Ohio State University in the summer of 2002. His current research interests include codes and modulation design for wireless communication, beamforming for massive MIMO systems, and fast decoding and detection algorithms.

YAN XIN received the Ph.D. degree from the University of Alberta, Alberta, AB, Canada, in 2002. He is currently a Principal Engineer with Huawei Technologies Canada, where he works on research and standardization of wireless communications. Prior to joining Huawei, he was a Senior Engineer with BlackBerry and Samsung Electronics. He has numerous granted U.S. patents and U.S. patent applications in pending. He has authored and coauthored a number of international refereed journals and conference papers. His research interests include physical layer development of wireless communication, signal processing, coding and modulation, and hardware implementation of communication systems.

EDWARD AU (Senior Member, IEEE) is currently a Senior Principal Engineer with Huawei Technologies Canada, where he works on product certification, test bed development, and standardization of Wi-Fi. Prior to joining Huawei, he was a Principal Engineer at Marvell Semiconductor and worked in developing foundational technologies for Bluetooth and Wi-Fi. He has actively participated in standardization organizations and industry forums. Dr. Au was awarded the IEEE Transactions on Vehicular Technology Top Editor Award in recognition of outstanding performance for articles reviews in both quantity and quality for the journal, in 2013 and 2016. He is the founding Chair of the IEEE 802.11 Task Group AY, which is aimed at developing an amendment for next-generation unlicensed millimeter wave technologies for Wi-Fi, and a Co-Editor of the IEEE 802.11 Task Group MD, which is charted to develop amendments for the IEEE 802.11 standards. He is also the Chair of Policies and Procedures of the IEEE Computer Society Standards Activities Board and a member of the IEEE Educational Activities Board/Standards Association Standards Education Committee. He also serves as a Senior Editor for the IEEE VEHICULAR TECHNOLOGY MAGAZINE and an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS. He was also an Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, a leading Guest Editor of the IEEE COMMUNICATIONS MAGAZINE on a few occasions and a Lead Track/Tutorial/Industry Forum Co-Chair of multiple IEEE-sponsored conferences.