Channel-Covariance and Angle-of-Departure Aided Hybrid Precoding for WideBand Multiuser Millimeter Wave MIMO Systems

Yun Chen, Da Chen, Tao Jiang, *Fellow, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

Abstract—Hybrid precoding is essential for millimeter wave (mmWave) multiple-input multiple output (MIMO) systems due to its inherent advantage of a high gain, whilst alleviating the high cost of hardware. However, most of the existing literature considered either the narrowband or wideband single-user mmWave MIMO scenarios. Hence in this paper we focus our attention on the more challenging design of hybrid Transmit Precoding (TPC) for wideband multiuser mmWave MIMO systems by exploiting the long-term channel's covariance matrix and the angle of departure (AoD) information. Specifically, in the analog TPC designed, we firstly derive the channel's covariance matrix for each user. Then the analog TPC matrix having an infinite angle resolution is constructed based on the channel's covariance matrix with the aid of phase extraction. Next, by exploiting the AoD information, we also propose a non-uniform quantization codebook based analog TPC having finite angle resolution. As for the design of the digital TPC, a three-stage scheme is proposed to compensate for the deleterious effects of beam squint, to cancel the inter-user interference and to attain multiplexing gains. We study the effects of various parameters on the achievable sum rate and demonstrate with the aid of our simulation results that the proposed hybrid TPC is capable of achieving a similar performance to the excessive-complexity fully digital TPC.

Index Terms—millimeter wave communication, hybrid TPC, multiuser, wideband, channel's covariance matrix.

I. INTRODUCTION

MILLIMETER wave (mmWave) multiple-input multiple output (MIMO) communication has been considered as one of the most promising techniques for next-generation systems as a benefit of significantly improving the achievable rate [1]. The small wavelength of the millimeter wave signals (0.1 mm∼1 mm) makes it feasible to pack a large number of antennas into small form factors and hence to make massive MIMO practical for wireless communications. In return, the large number of antennas using MIMO signal processing techniques like precoding is capable of providing substantial spatial multiplexing gains to guarantee sufficient signal power at the receiver [2]. However, the high cost and power consumption of hardware components at the ten-fold increase of the mmWave carrier frequency prohibit the use of the traditional fully digital Transmit Precoding (TPC) relying on a dedicated radio frequency (RF) chain for each antenna

Y. Chen, D. Chen and T. Jiang are with Wuhan National Laboratory for Optoelectronics, School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: chen_yun@hust.edu.cn; chenda@hust.edu.cn; tao.jiang@ieee.org).

L. Hanzo is with School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: lh@ecs.soton.ac.uk).

[3]–[7]. To circumvent this problem, the concept of hybrid TPC was proposed, where the TPC is partitioned into a high-dimensional analog TPC and a low-dimensional digital TPC [8]–[11].

A. Prior Work

Most prior work on the design of the hybrid TPC focuses on the narrowband single user (SU) scenario [10]–[21]. The orthogonal matching pursuit (OMP) based hybrid TPC scheme was the first one proposed for mmWave wave MIMO systems, which formulates a sparse reconstruction problem in order to obtain the hybrid TPC having a transmission profile 'closest' to the optimal unconstrained digital TPC [10]. Based on [10], diverse methods have been proposed for near-optimal TPC design, such as the alternating minimization (AM), matrix decomposition and iterative search methods [12]–[15]. Alternatively, low-complexity hybrid TPCs may be designed by exploiting the sparse Channel Impulse Response (CIR) of the mmWave link [16]–[20]. Moreover, research efforts have also been dedicated to designing energy-efficient hybrid TPC architectures [11], [21].

Naturally, the design of the hybrid TPC in narrowband multiuser (MU) scenarios is more challenging than that of its SU counterpart, because the inter-user interference gravely degrades the achievable sum rate [22]–[25]. In [22], a hybrid TPC scheme based on imperfect channel state information (CSI) (the CSI with limited feedback) was proposed for the MU scenario, where the analog TPC of each user relies on a uniform-quantization (UQ) based codebook and the digital TPC is designed using the zero-forcing (ZF) solution. In [23], a MU hybrid TPC scheme was proposed by Dai and Clerckx based on the channel's covariance matrix for the frequency-flat mmWave channel. They conceived a so-called rate splitting (RS) transmission strategy for mitigating the inter-user interference. Furthermore, a block diagonalization (BD) based hybrid TPC has also been conceived, where the digital TPC of one user is designed to operate within the null space of other users' channels [24], [25].

More recently, the design of hybrid TPC schemes for the wideband mmWave MIMO system has attracted more interests, where typically orthogonal frequency-division multiplexing (OFDM) is utilized for overcoming the multipath fading [12], [26]–[30]. In general, the schemes designed for narrowband mmWave MIMO systems cannot be directly extended to the wideband scenario. The main impediment is that the analog

TPC is frequency-flat and it is operated in the time domain, whilst the digital TPC operates in the digital domain. Hybrid TPC schemes have been designed for the wideband SU scenario in [12], [26]–[28]. In [26], a near-optimal hybrid TPC weight vector was developed in closed form for both the fully-connected and sub-array-connected OFDM-based SU mmWave systems. An AM based hybrid TPC scheme has been conceived for wideband SU mmWave systems in [12], where the analog TPC weights are constructed by searching through the solutions on the Riemannian manifold. Another codebook based wideband SU hybrid TPC scheme was proposed in [27], which is based on Gram-Schmidt orthogonalization relying on realistic imperfect CSI. Finally, in [28] a MIMO-OFDM hybrid TPC scheme was designed without any instantaneous CSI feedback, where the analog TPC was constructed based on both the long-term mean angles and on the angular spread.

Only a small number of papers considered the design of hybrid TPCs for wideband MU scenarios [29], [30]. In [29], the authors only determined the number of RF chains for hybrid TPCs with the goal of achieving a similar performance as the high-complexity digital TPC and evaluated the hybrid TPC solution for the high signal-to-noise ratio (SNR) regime. In [30], the uplink digital TPC was designed based on several criteria, such as the mean-squared-error (MSE) criterion, and a pair of hybrid TPCs were constructed using the factorization and AM techniques, respectively.

Furthermore, most prior work on the design of hybrid TPCs assumes having perfect and instantaneous CSI at the transmitter, when designing the analog TPC. However, obtaining full-dimensional instantaneous CSI at the transmitter equipped with a large number of antennas is still an open challenge in reality, since the number of RF chains is much lower than the number of antennas. Therefore, the long-term channel's covariance matrix is more attractive using for the analog TPC, as shown in the recent contributions [23], [28], [31]–[33].

Moreover, in wideband MIMO systems especially in mmWave wideband MIMO systems having a substantial bandwidth (2-5 GHz), beam squint (representing the beam's changing spatial direction as a function of frequency [34]) will also dramatically affect the performance of TPCs [35]. The effect of the beam squint has been considered in mmWave MIMO channel estimation using hybrid TPC relying on lens antenna arrays for achieving a higher accuracy in [35]. However, there is no paper that has taken the beam squint into account for the design of hybrid TPC in wideband MU mmWave MIMO systems associated with phased arrays.

B. Contributions

In this paper, we consider the design of hybrid TPCs for wideband MU mmWave MIMO systems aided by the long-term channel's covariance matrix and on the angle of departure (AoD) information. The main contributions of this paper are summarized as follows.

- We design an analog TPC scheme relying on the channel's covariance matrix and on the AoD information. Specifically, we firstly derive the covariance matrix associated with the channel matrix for each user on each

TABLE I
CONTRASTING OUR NOVEL CONTRIBUTIONS TO THE STATE-OF-THE-ART.

	[9]	[27]	[28]	[30]	[32]	Our Scheme
Single User TPC	✓	✓	✓			
Multiuser TPC				✓	✓	✓
Narrowband Scenario	✓		✓	✓		
Widewband Scenario		✓	✓	✓		✓
Instantaneous CSI	✓	✓		✓		
Channel-Covariance			✓		✓	✓
Limited Feedback	✓	✓				✓
Beam Squint						✓

subcarrier, which shows that the covariance matrix is mainly determined by the correlation matrix of the array response matrix at the transmitter. Then, the unconstrained analog TPC is designed by applying the classic singular value decomposition (SVD) to the correlation matrix. Finally, the analog TPC satisfying the constant modulus constraint is given by simple phase extraction.

- We propose a non-uniform quantization (NUQ) codebook based analog TPC scheme having a finite resolution for each user by exploiting the sparse CIR of the mmWave channel in the AoD domain. This scheme is capable of quantizing the physical steering angles more accurately than the traditional UQ codebook based schemes so as to increase the achievable sum rate without increasing the feedback overhead.

- We propose a three-stage digital TPC for jointly handling the beam squint, the inter-user interference and the spatial multiplexing for improving the achievable sum rate. Specifically, the first-stage digital TPC is designed to alleviate the effect of beam squint by providing phase compensation (PC) for the frequency-flat analog TPC, where the compensation matrix is constructed with the assistance of the AoD information. The second-stage digital TPC is designed for canceling the inter-user interference, which is mainly based on the idea of BD. Furthermore, the third-stage digital TPC is constructed for attaining the optimal multiplexing gain based on the low-dimensional effective CSI.

- We study the effects of various transmission parameters, hybrid TPC architectural parameters and channel parameters on the achievable sum rate with the aid of our numerical experiments. Our simulation results demonstrate that the proposed hybrid TPC scheme is capable of compensating the beam squint, of cancelling the inter-user interference and of achieving a similar performance to the high-complexity fully digital TPC.

To clarify our contributions clearly, we contrast the proposed hybrid TPC scheme against the state-of-the-art in Table I.

The rest of the paper is organized as follows. Section II presents the system model, channel model, problem formulation and some of our assumptions. The derivation of the long-

term channel covariance matrix and the design of the analog TPC are introduced in Section III. In Section IV, our three-stage digital TPC scheme is proposed. Section V discusses our simulation results. Finally, we conclude in Section VI.

We use the following notations in this paper: a is a scalar, $\mathbf{a}$ is a vector, $\mathbf{A}$ is a matrix and $\mathcal{A}$ is a set. $[\mathbf{A}]_{i,j}$ is the $(i,j)^{th}$ element of $\mathbf{A}$ and $\angle(\mathbf{A})$ is a matrix whose $(i,j)^{th}$ element equals the phase of the $(i,j)^{th}$ element in $\mathbf{A}$. $\mathbf{A}^T, \mathbf{A}^*, \mathbf{A}^{-1}$, $|\mathbf{A}|$ and $\|\mathbf{A}\|_F$ denote the transpose, conjugate transpose, inverse, determinant, and Frobenius norm of $\mathbf{A}$, respectively. $|\mathbf{a}|$ is the modulus of $\mathbf{a}$. $\mathbf{I}_N$ denotes a $N \times N$ identity matrix. $\mathcal{CN}(\mathbf{a}, \mathbf{A})$ is a complex Gaussian vector with mean $\mathbf{a}$ and covariance matrix $\mathbf{A}$. $\mathbb{E}[\mathbf{A}]$ is the expectation of $\mathbf{A}$. $\mathrm{Span}(\mathbf{A})$ denotes the linear subspace generated by columns of $\mathbf{A}$ and $\mathrm{Span}^{\perp}(\mathbf{A})$ is the orthogonal complement of $\mathrm{Span}(\mathbf{A})$.

II. SYSTEM MODEL, CHANNEL MODEL AND PROBLEM FORMULATION

In this section, we introduce the system model, the channel model, the problem formulation and some of our assumptions.

A. System Model

Consider the wideband multiuser downlink mmWave MIMO system of Fig. 1, where a base station (BS) equipped with N_t antennas and N_{RF}^t RF chains transmits its signals to U mobile stations (MSs) having $N_{r,u}$ antennas and $N_{RF,u}^r$ RF chains. The fully-connected hybrid architecture is adopted in this paper, where each RF chain is connected to all antennas utilizing phase shifters. Due to the limitation of the number of RF chains, we have $N_s \leq N_{RF}^t \leq N_t$, $N_{s,u} \leq N_{RF,u}^r \leq N_{r,u}$, and $\sum_{u=1}^{U} N_{RF,u}^r \leq N_{RF}^t$, where N_s denotes the total number of data streams, $N_{s,u}$ is the number of independent data streams of the uth user and $N_s = \sum_{u=1}^{U} N_{s,u}$. Moreover, we assume that the MSs can perform optimal nearest neighbor decoding based on the $N_{r,u}$-dimensional received signal relying on fully digital hardware, and focus our attention on the design of hybrid TPCs at the BS, similarly to [27].

OFDM modulation having K subcarriers and a sufficiently long cyclic prefix (CP) is utilized to mitigate the dispersion of the wideband millimeter wave channel. At the transmitter, the $N_{s,u}$ data streams conveyed by each subcarrier for each user are first precoded using the $(N_{RF}^t \times N_{s,u})$-element digital TPC matrix $\mathbf{F}_{BB,u}[k]$, $k = 1, 2, ..., K, u = 1, 2, ..., U$. Then, after the K-point IFFT and adding the CP, the data streams are precoded by the $(N_t \times N_{RF}^t)$-element analog TPC matrix $\mathbf{F}_{RF}$. It is important to emphasize that the analog TPC $\mathbf{F}_{RF}$ is applied in the time domain, which is the same for all sub-carriers. By contrast, the digital TPC $\mathbf{F}_{BB,u}[k]$ is performed in the digital domain and it differs for each subcarrier. This is one of the major distinguishing features of the wideband hybrid TPC compared to the conventional high-complexity fully-digital TPC.

At the receiver, assuming perfect carrier and frequency offset synchronization, first the CP is removed. After K-point FFT, the received signals are transformed back to the frequency domain. The discrete-time received signals of the uth user at the kth subcarrier can be written as

$$\mathbf{y}_u[k] = \mathbf{H}_u[k] \sum_{l=1}^{U} \mathbf{F}_{RF} \mathbf{F}_{BB,l}[k] \mathbf{s}_l[k] + \mathbf{n}_u[k], \quad (1)$$

where $\mathbf{s}_l[k]$ is the $(N_{s,l} \times 1)$ transmit signal vector of the lth user at the kth subcarrier, where we have $\mathbb{E}[\mathbf{s}_l[k]] = 0$, $\mathbb{E}[\mathbf{s}_l[k]\mathbf{s}_l^*[k]] = \mathbf{I}_{N_{s,l}}$ and $\mathbb{E}[\mathbf{s}_l[k]\mathbf{s}_p^*[k]] = \mathbf{0}$ for $p \neq l$. Furthermore, $\mathbf{H}_u[k]$ is the channel matrix of the uth user at the kth subcarrier, and $\mathbf{n}_u[k] \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I}_{N_{r,u}})$ is the additive white Gaussian noise vector. Similar to [30], an equal power allocation is assumed for each MS and subcarrier, so that we have $\|\mathbf{F}_{RF}\mathbf{F}_{BB,l}[k]\|_F^2 = \frac{P_t}{UN_{s,l}}$, where P_t is the total transmit power. Furthermore, since the analog TPC is implemented using analog phase shifters, the constant modulus constraint is imposed on its entries, i.e., $|[\mathbf{F}_{RF}]_{i,j}|^2 = \frac{1}{N_t}$.

B. Channel Model

Given its limited scattering characteristics, the Saleh-Valenzuela model is adopted for modelling the wideband millimeter wave channel. The frequency domain (FD) channel matrix of the uth user at the kth subcarrier can be formulated as [2], [25], [35]

$$\mathbf{H}_u[k] = \gamma \sum_{m=1}^{N_{cl}} \sum_{n=1}^{N_{ray}} \alpha_{u,m,n} \beta_{u,m,n,k} \mathbf{a}_r(\phi_{u,m,n,k}) \mathbf{a}_t^*(\varphi_{u,m,n,k}), \quad (2)$$

where $\gamma = \sqrt{N_t N_r / N_{cl} N_{ray}}$ is the normalization factor, N_{cl} is the number of clusters, each of which contributes N_{ray} propagation paths, $\alpha_{u,m,n}$ represents the path gains, $\beta_{u,m,n,k} = e^{-j2\pi\tau_{u,m,n}f_k}$ denotes the delay component, in which $\tau_{u,m,n}$ is the delay of the nth path in the mth cluster. Furthermore, $\phi_{u,m,n,k}$ and $\varphi_{u,m,n,k}$ are the spatial AoA and AoD of the nth path in the mth cluster at the kth subcarrier, respectively, which can be written as [35]

$$\phi_{u,m,n,k} = \frac{f_k}{c} d\sin(\theta_{u,m,n}), \quad (3)$$

$$\varphi_{u,m,n,k} = \frac{f_k}{c} d\sin(\vartheta_{u,m,n}), \quad (4)$$

where $f_k = f_c + \frac{f_s}{K}(k - 1 + \frac{K-1}{2})$ is the frequency at the kth subcarrier, f_c and f_s are the carrier frequency and bandwidth, respectively, c is the speed of light, $d = \frac{c}{2f_c} = \frac{\lambda}{2}$ denotes the aperture domain sample spacing, namely the antenna-element spacing. Still referring to (3) and (4), λ is the signal wavelength, $\theta_{u,m,n} \in [0, 2\pi)$ and $\vartheta_{u,m,n} \in [0, 2\pi)$ are the corresponding physical steering angles at the receiver and the transmitter, respectively. Note that in the narrowband mmWave MIMO system, it is assumed that $f_k = f_c$, which means that the spatial angles are frequency-independent. However, in the wideband mmWave MIMO system this assumption is not true and will dramatically affect the achievable rate, when considering that the large number of antennas are used at the BS. Moreover, the uniform linear array (ULA) is considered in the paper, but the proposed scheme is applicable to arbitrary
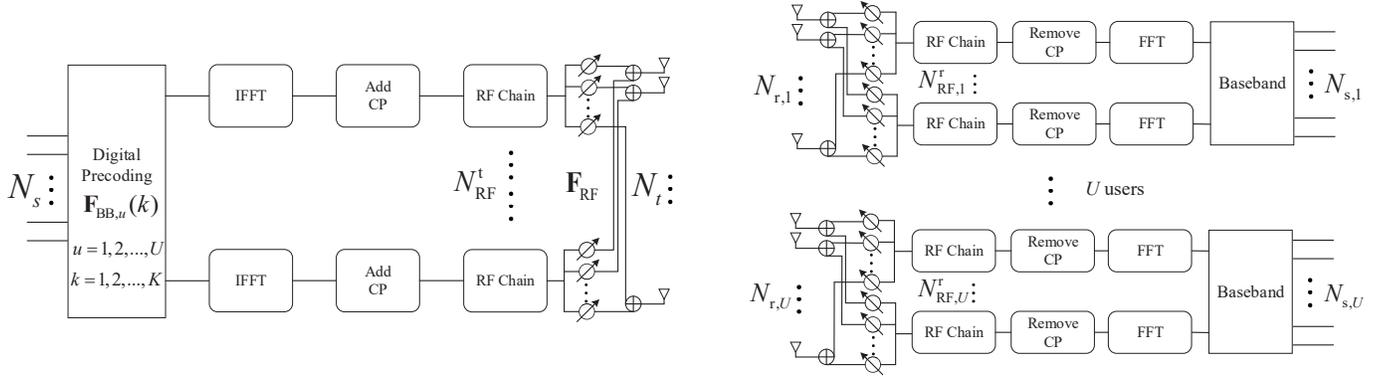
Fig. 1. The block diagram of the wideband multiuser mmWave MIMO system model that employs the hybrid TPC.

antenna arrays. Therefore, the array response vectors can be written as

$$\mathbf{a}_{\mathrm{r}}(\phi_{u,m,n,k}) = \frac{1}{\sqrt{N_{\mathrm{r}}}}\Big[1, e^{j2\pi\phi_{u,m,n,k}}, ..., e^{j(N_{\mathrm{r}}-1)2\pi\phi_{u,m,n,k}}\Big]^{T}, \tag{5}$$

and

$$\mathbf{a}_{\mathrm{t}}(\varphi_{u,m,n,k}) = \frac{1}{\sqrt{N_{\mathrm{t}}}}\Big[1, e^{j2\pi\varphi_{u,m,n,k}}, ..., e^{j(N_{\mathrm{t}}-1)2\pi\varphi_{u,m,n,k}}\Big]^{T}. \tag{6}$$

Nota that (2) can be written in a more compact form as

$$\mathbf{H}_{u}[k] = \gamma\mathbf{A}_{\mathrm{r},u}[k]\mathbf{\Delta}_{u}[k]\mathbf{A}_{\mathrm{t},u}^{*}[k], \tag{7}$$

where we have:

$$\mathbf{\Delta}_{u}[k] = \begin{bmatrix} \alpha_{u,1,1}\beta_{u,1,1,k} & & & \\ & \alpha_{u,1,2}\beta_{u,1,2,k} & & \\ & & \ddots & \\ & & & \alpha_{u,N_{\mathrm{cl}},N_{\mathrm{ray}}}\beta_{u,N_{\mathrm{cl}},N_{\mathrm{ray}},k} \end{bmatrix}, \tag{8}$$

and

$$\mathbf{A}_{\mathrm{r},u}[k] = \big[\mathbf{a}_{\mathrm{r}}(\phi_{u,1,1,k}), ..., \mathbf{a}_{\mathrm{r}}(\phi_{u,1,N_{\mathrm{ray}},k}), ..., \mathbf{a}_{\mathrm{r}}(\phi_{u,N_{\mathrm{cl}},N_{\mathrm{ray}},k})\big], \tag{9}$$

as well as

$$\mathbf{A}_{\mathrm{t},u}[k] = \big[\mathbf{a}_{\mathrm{t}}(\varphi_{u,1,1,k}), ..., \mathbf{a}_{\mathrm{t}}(\varphi_{u,1,N_{\mathrm{ray}},k}), ..., \mathbf{a}_{\mathrm{t}}(\varphi_{u,N_{\mathrm{cl}},N_{\mathrm{ray}},k})\big] \tag{10}$$

in (7) contain the array response vectors corresponding to the MS and BS, respectively.

C. Problem Formulation

The objective of designing the hybrid TPC is to maximize the achievable sum rate on all subcarriers and users achieved by assuming Gaussian signalling over the wideband mmWave channel, where the achievable ergodic sum rate is given by [9]

$$R_{\mathrm{sum}} = \frac{1}{K}\sum_{u=1}^{U}\sum_{k=1}^{K}\log_{2}\Big(\Big|\mathbf{I}_{N_{\mathrm{r},u}} + \mathbf{R}_{u}^{-1}[k] \\ \times \mathbf{H}_{u}[k]\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB},u}[k]\mathbf{F}_{\mathrm{BB},u}^{*}[k]\mathbf{F}_{\mathrm{RF}}^{*}\mathbf{H}_{u}^{*}[k]\Big|\Big), \tag{11}$$

where $\mathbf{R}_{u}[k] = \sum_{p=1,p\neq u}^{U}\mathbf{H}_{u}[k]\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB},p}[k]\mathbf{F}_{\mathrm{BB},p}^{*}[k]\mathbf{F}_{\mathrm{RF}}^{*}\mathbf{H}_{u}^{*}[k] + \sigma_{n}^{2}\mathbf{I}_{N_{\mathrm{r},u}})$ is the covariance matrix of the aggregate inter-user

interference and noise. Therefore, the optimization problem can be written as

$$(\mathbf{F}_{\mathrm{RF}}^{\mathrm{opt}}, \mathbf{F}_{\mathrm{BB},u}^{\mathrm{opt}}[k]) = \arg\max R_{\mathrm{sum}}, \forall u, k, \\ \text{s.t.} \quad |[\mathbf{F}_{\mathrm{RF}}]_{i,j}|^{2} = \frac{1}{N_{\mathrm{t}}}, \forall i, j, \\ \|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB},u}[k]\|_{F}^{2} = \frac{P_{\mathrm{t}}}{UN_{\mathrm{s},u}}. \tag{12}$$

Under our power constraint, the globally optimal solution to (12) should be determined by the joint optimization of the analog and digital TPCs. However, unfortunately this is intractable even upon neglecting the constant modulus constraint imposed on the analog TPC. According to state-of-the-art in hybrid TPC design [9], [13], [22], [23], separating the analog TPC and digital TPC achieves satisfactory performance without the need for iterative procedures. Therefore, to facilitate a low-complexity design, the TPC matrices $\mathbf{F}_{\mathrm{RF}}$ and $\mathbf{F}_{\mathrm{BB},u}[k]$ are determined in a decoupled manner.

Furthermore, our scheme is based on the following assumptions concerning the channel state information:

- *The design of analog TPC depends only on the long-term channel's covariance matrix, but not on the instantaneous CSI, while the digital TPC is designed based on the short-term low-dimensional effective CSI, similar to [32].* Naturally, obtaining the perfect instantaneous CSI at the BS is not realistic in reality since numerous antennas are used both at the BS and MSs. Moreover, in the hybrid TPC architecture, the number of RF chains is much less than the number of antennas, which exacerbates the challenge of obtaining the instantaneous CSI. More importantly, feeding back the instantaneous CSI from the MSs to the BS imposes an extremely high overhead. However, for the digital TPC, the required short-term effective CSI is of low dimension since the number of RF chains is limited, which imposes a moderate overhead.
- *The AoDs are known at the BS.* According to [20], [36], [37], the path angles vary much slower than the path gains. Specifically, the new concept of angular coherence time was defined in [36] as a comparably long time, during which the AoDs can be regarded as near-constant. Va *et al.* showed by their measurements that the angular coherence time can be an order of magnitude longer than

the traditional channel coherence time. Therefore, again, we assume that the BS knows the AoDs.

III. ANALOG TPC DESIGN

In this section, we firstly introduce the channel's covariance matrix for each user on each subcarrier. Then, the analog TPC wirh infinite resolution is constructed based on the covariance matrix. Furthermore, we also consider analog TPC relying on realistic quantized angles and propose a NUQ codebook based analog TPC scheme aided by the known AoDs.

A. Channel's Covariance Matrix

Most prior contributions on the design of analog TPCs assume that the instantaneous CSI is perfectly known at the BS, which is not realistic. By contrast, the channel's covariance matrix varies only over a longer time scale than the instantaneous CSI, which is hence easier to estimate. Therefore, we firstly derive the covariance matrix corresponding to the channel matrix of each user on each subcarrier.

Lemma 1: Using the compact form of the channel in (7), its covariance matrix is given by

$$
\begin{aligned}
\mathbf{R}_u[k] &= \mathbb{E}[\mathbf{H}_u^*[k]\mathbf{H}_u[k]] \\
&= \gamma^2 \zeta_{u,k} \mathbf{A}_{\mathrm{t},u}[k] \mathbf{A}_{\mathrm{t},u}^*[k],
\end{aligned}
\tag{13}
$$

where $\zeta_{u,k} = \sigma^2_{\alpha_u, \beta_{u,k}} = \sigma^2_{\alpha_u} \sigma^2_{\beta_{u,k}}$, in which $\sigma^2_{\alpha_u}$ and $\sigma^2_{\beta_{u,k}}$ are the variances of α_u and $\beta_{u,k}$, respectively. Note that the expectation is estimated using the sample average during the angular coherence time.

Proof: Since the path gains $\alpha_{u,m,n}$, the delay component $\beta_{u,m,n,k}$, as well as AoAs and AoDs are independent, we have

$$
\begin{aligned}
\mathbf{R}_u[k] &= \mathbb{E}[\mathbf{H}_u^*[k]\mathbf{H}_u[k]] \\
&= \gamma^2 \mathbb{E}[\mathbf{A}_{\mathrm{t},u}[k]\boldsymbol{\Delta}_u^*[k]\mathbf{A}_{\mathrm{r},u}^*[k]\mathbf{A}_{\mathrm{r},u}[k]\boldsymbol{\Delta}_u[k]\mathbf{A}_{\mathrm{t},u}^*[k]] \\
&= \gamma^2 \mathbb{E}[\mathbf{A}_{\mathrm{t},u}[k]\mathbf{R}_{\boldsymbol{\Delta},\mathrm{r}}\mathbf{A}_{\mathrm{t},u}^*[k]] \\
&= \gamma^2 \mathbf{A}_{\mathrm{t},u}[k]\mathbf{R}_{\boldsymbol{\Delta},\mathrm{r}}\mathbf{A}_{\mathrm{t},u}^*[k],
\end{aligned}
\tag{14}
$$

where

$$
\begin{aligned}
\mathbf{R}_{\boldsymbol{\Delta},\mathrm{r}} &= \mathbb{E}[\boldsymbol{\Delta}_u^*[k]\mathbf{A}_{\mathrm{r},u}^*[k]\mathbf{A}_{\mathrm{r},u}[k]\boldsymbol{\Delta}_u[k]] \\
&= \mathbb{E}[\boldsymbol{\Delta}_u^*[k]\mathbb{E}[\mathbf{A}_{\mathrm{r},u}^*[k]\mathbf{A}_{\mathrm{r},u}[k]]\boldsymbol{\Delta}_u[k]]
\end{aligned}
\tag{15}
$$

is an $(N_{\mathrm{cl}}N_{\mathrm{ray}} \times N_{\mathrm{cl}}N_{\mathrm{ray}})$-element diagonal matrix, since $\alpha_{u,m,n}$ are independent with zero means. Note that each column in $\mathbf{A}_{\mathrm{r},u}[k]$ is of unit norm, and the diagonal elements of $\mathbb{E}[\mathbf{A}_{\mathrm{r},u}^*[k]\mathbf{A}_{\mathrm{r},u}[k]]$ are one. Thus the diagonal elements of $\mathbf{R}_{\boldsymbol{\Delta},\mathrm{r}}$ are equal to $\sigma^2_{\alpha_u,\beta_{u,k}}$, i.e., $\mathbf{R}_{\boldsymbol{\Delta},\mathrm{r}} = \sigma^2_{\alpha_u,\beta_{u,k}}\mathbf{I}_{N_{\mathrm{cl}}N_{\mathrm{ray}}}$. Furthermore, we have $\mathbb{E}[\mathbf{A}_{\mathrm{t},u}[k]\mathbf{A}_{\mathrm{t},u}^*[k]] = \mathbf{A}_{\mathrm{t},u}[k]\mathbf{A}_{\mathrm{t},u}^*[k]$ since the angles can be regarded as constant during the angular coherence time. Upon substituting $\mathbf{R}_{\boldsymbol{\Delta},\mathrm{r}}$ into (14) we obtain (13). Moreover, since $\alpha_{u,m,n}$ and $\beta_{u,m,n,k}$ are independent, we have $\sigma^2_{\alpha_u,\beta_{u,k}} = \sigma^2_{\alpha_u}\sigma^2_{\beta_{u,k}}$. ∎

Lemma 1 indicates that the covariance matrix of the channel $\mathbf{H}_u[k]$ is mainly determined by $\mathbf{A}_{\mathrm{t},u}[k]\mathbf{A}_{\mathrm{t},u}^*[k]$, while $\zeta_{u,k}$ is a constant for each user at each subcarrier. Our the analog TPC of each user is designed in the following subsection by exploiting this property.

B. Analog TPC with Infinite Resolution

According to (13), the covariance matrix of the channel $\mathbf{H}_u[k]$ is frequency-dependent. However, for maintaining a low hardware complexity, our analog TPCs operate in the time domain and are frequency-flat. Moreover, the inter-user interference is temporarily ignored so that the analog TPCs of different users can be designed separately.

To design the common frequency-flat analog TPC for all subcarriers of each user, we approximate $\mathbf{R}_u[k]$ as

$$
\mathbf{R}_u[k] \approx \gamma^2 \zeta_{u,k} \mathbf{A}_{\mathrm{t},u,\mathrm{c}} \mathbf{A}_{\mathrm{t},u,\mathrm{c}}^*,
\tag{16}
$$

where we have

$$
\mathbf{A}_{\mathrm{t},u,\mathrm{c}} = \left[\mathbf{a}_{\mathrm{t}}(\varphi_{u,1,1,\mathrm{c}}),...,\mathbf{a}_{\mathrm{t}}(\varphi_{u,1,N_{\mathrm{ray}},\mathrm{c}}),...,\mathbf{a}_{\mathrm{t}}(\varphi_{u,N_{\mathrm{cl}},N_{\mathrm{ray}},\mathrm{c}})\right],
\tag{17}
$$

in which $\mathbf{a}_{\mathrm{t}}(\varphi_{u,m,n,\mathrm{c}})$ denotes the array response vector of the central spatial AoD $\varphi_{u,m,n,\mathrm{c}}$ formulated as

$$
\mathbf{a}_{\mathrm{t}}(\varphi_{u,m,n,\mathrm{c}}) = \frac{1}{\sqrt{N_{\mathrm{t}}}}\left[1, e^{j2\pi\varphi_{u,m,n,\mathrm{c}}},...,e^{j(N_{\mathrm{t}}-1)2\pi\varphi_{u,m,n,\mathrm{c}}}\right]^T,
\tag{18}
$$

while the central spatial AoD $\varphi_{u,m,n,\mathrm{c}}$ is defined by

$$
\varphi_{u,m,n,\mathrm{c}} = \frac{f_{\mathrm{c}}}{\mathrm{c}}d\sin(\vartheta_{u,m,n}) = \varphi_{u,m,n,k} - \triangle_f d\sin(\vartheta_{u,m,n}),
\tag{19}
$$

where $\triangle_f = \frac{f_{\mathrm{s}}}{K}(k-1+\frac{K-1}{2})$ represents the spatial angle offset.

According to the approximate covariance matrix of the channel in (16), we apply SVD of $\mathbb{E}[\mathbf{A}_{\mathrm{t},u,\mathrm{c}}\mathbf{A}_{\mathrm{t},u,\mathrm{c}}^*]$ to approximate the SVD of $\mathbf{R}_u[k]$, similar to [28]. Thus we have

$$
\mathbf{A}_{\mathrm{t},u,\mathrm{c}}\mathbf{A}_{\mathrm{t},u,\mathrm{c}}^* = \mathbf{V}_{\mathrm{t},u,\mathrm{c}}\boldsymbol{\Sigma}_{\mathrm{t},u,\mathrm{c}}\mathbf{V}_{\mathrm{t},u,\mathrm{c}}^*.
\tag{20}
$$

Accordingly, the unconstrained analog TPC for the uth user can be constructed by the $N_{\mathrm{RF}}^{\mathrm{t}}/U$ columns of $\mathbf{V}_{\mathrm{t},u,\mathrm{c}}$ corresponding to the largest $N_{\mathrm{RF}}^{\mathrm{t}}/U$ singular values, i.e.,

$$
\widetilde{\mathbf{F}}_{\mathrm{RF},u} = \mathbf{V}_{\mathrm{t},u,\mathrm{c}}(:,1:N_{\mathrm{RF}}^{\mathrm{t}}/U),
\tag{21}
$$

and the unconstrained total analog TPC can be readily written as

$$
\widetilde{\mathbf{F}}_{\mathrm{RF}} = [\widetilde{\mathbf{F}}_{\mathrm{RF},1},\widetilde{\mathbf{F}}_{\mathrm{RF},2},...,\widetilde{\mathbf{F}}_{\mathrm{RF},U}].
\tag{22}
$$

Note that we assume that all users have the same priority and are assigned the same number of RF chains.

Furthermore, to meet the constant modulus constraint, the total analog TPC is designed to be as close to the unconstrained total analog TPC as it can be, which is formulated as

$$
\min_{\mathbf{F}_{\mathrm{RF}},|[\mathbf{F}_{\mathrm{RF}}]_{i,j}|^2=\frac{1}{N_{\mathrm{t}}}}\left\|\mathbf{F}_{\mathrm{RF}} - \widetilde{\mathbf{F}}_{\mathrm{RF}}\right\|_F^2.
\tag{23}
$$

The solution of (23) has been found by Zhang *et al.* in [8] to be given by $[\mathbf{F}_{\mathrm{RF}}]_{i,j} = \frac{1}{\sqrt{N_{\mathrm{t}}}}e^{j\angle([\widetilde{\mathbf{F}}_{\mathrm{RF}}]_{i,j})}$.

C. Analog TPC with Quantized Angles

The analog TPC obtained in Section III.B is of continuous angles. However, constructing a phase shifter having infinite resolution is either infeasible or excessively costly. Therefore, in this subsection, we consider the design of analog TPCs having realistic finite resolution, where the angles of the analog

TPC obey a quantized angls set and hence the analog TPC vectors can only assume certain values.

We firstly introduce the UQ codebook based scheme, which is considered by most of the prior literature [9], [17], [22]. Given the total number of quantization bits b, the entire angular domain can be partitioned into 2^b segments:

$$\mathcal{F}_{\mathrm{UQ}} = \{\varphi_{\mathrm{UQ},1}, ..., \varphi_{\mathrm{UQ},2^b}\}, \qquad (24)$$

where $\varphi_{\mathrm{UQ},q} = \frac{1}{2}\sin(\frac{2\pi(q-1)}{2^b})$ and $\frac{2\pi}{2^b}$ is the quantization accuracy. The UQ codebook can then be written as

$$\mathbf{F}_{\mathrm{UQ}} = [\mathbf{a}_{\mathrm{t}}(\varphi_{\mathrm{UQ},1}), \mathbf{a}_{\mathrm{t}}(\varphi_{\mathrm{UQ},2}), ..., \mathbf{a}_{\mathrm{t}}(\varphi_{\mathrm{UQ},2^b})]. \qquad (25)$$

Although the UQ codebook is of relatively small size, it provides good quantization performance but does not fully exploit the CIR's sparsity in the angular domain. According to the recent measurement results of NYU WIRELESS, the AOAs/AODs of the paths in the millimeter wave channel may be grouped into distinct spatial lobes [38]–[40], where the physical steering angles of the paths in different spatial lobes are distinctly separated. Again, we assume that the spatial lobes are the same as the clusters due to the limited angular spread of each cluster and the angles of the paths in different clusters are distinctly separated with a high probability. Moreover, the angles of the paths for all users are assumed to be known at the BS within the angular coherence time. Therefore, we are able to construct an NUQ codebook similar to [20]. Note that the quantized entries in the NUQ codebook are the physical steering angles, which are the same for all subcarriers and we use f_{c} to approximate f_k so as to construct the NUQ codebook for all subcarriers simultaneously. Given the AoD, the entire effective angular coverage for the uth user is

$$\mathcal{CV}_u = \bigcup_{m=1,2,...,N_{\mathrm{cl}}} \mathcal{CV}_{u,m}, \qquad (26)$$

where $\mathcal{CV}u, m = \bigcup_{n=1,2,...,N_{\mathrm{ray}}} \mathcal{CV}[\mathbf{a}_{\mathrm{t}}(\varphi_{u,m,n,\mathrm{c}})]$ is the effective angular coverage of the mth cluster for the uth user and $\mathcal{CV}[\mathbf{a}_{\mathrm{t}}(\varphi_{u,m,n,\mathrm{c}})] \subset [\underline{\vartheta_{u,m,\mathrm{c}}}, \overline{\vartheta_{u,m,\mathrm{c}}}]$, where $\underline{\vartheta_{u,m,\mathrm{c}}}$ and $\overline{\vartheta_{u,m,\mathrm{c}}}$ are the corresponding lower and upper bounds of the path angles for the mth cluster, respectively. The lower and upper angular bounds $\underline{\vartheta_{u,m,\mathrm{c}}}$ and $\overline{\vartheta_{u,m,\mathrm{c}}}$ are set a little lower and higher than the actual smallest and largest physical angles in the mth cluster to guarantee that the NUQ codebook becomes capable of quantizing all the path angles. We use $angle(\mathcal{CV}_u)$ to represent the whole effective angular range of the uth user, i.e., $\sum_{m=1}^{N_{\mathrm{cl}}} (\overline{\vartheta_{u,m,\mathrm{c}}} - \underline{\vartheta_{u,m,\mathrm{c}}})$. Naturally, we have $angle(\mathcal{CV}_u) \leq 2\pi$ due to the CIR's sparsity in the angular domain. Then the total number of quantization bits b is appropriately mapped to the effective angular coverage, yielding the quantization accuracy of $\frac{angle(\mathcal{CV}_u)}{2^b}$. Note that there is no angular overlap between different clusters, hence the set of quantized spatial angles for the mth cluster of the uth user can be written as

$$\mathcal{F}_{\mathrm{NUQ},u,m} = \{\varphi_{\mathrm{NUQ},u,m,1}, ..., \varphi_{\mathrm{NUQ},u,m,G_m}\}, \qquad (27)$$

where $\varphi_{\mathrm{NUQ},u,q} = \frac{1}{2}\sin(\underline{\vartheta_{u,m,\mathrm{c}}} + (q-1)\frac{angle(\mathcal{CV}_u)}{2^b})$, $G_m = \frac{\overline{\vartheta_{u,m,\mathrm{c}}} - \underline{\vartheta_{u,m,\mathrm{c}}}}{angle(\mathcal{CV}_u)/2^b}$, and the total set of the quantized spatial angles of the uth user is

$$\mathcal{F}_{\mathrm{NUQ},u} = \{\mathcal{F}_{\mathrm{NUQ},u,1}, \mathcal{F}_{\mathrm{NUQ},u,2}, ..., \mathcal{F}_{\mathrm{NUQ},u,N_{\mathrm{cl}}}\}. \qquad (28)$$

Therefore, the NUQ codebook can be constructed as:

$$\begin{aligned} \mathbf{F}_{\mathrm{NUQ},u}(q,:) &= \mathbf{a}_{\mathrm{t}}(\varphi_{\mathrm{NUQ},u,q}), \\ \varphi_{\mathrm{NUQ},q} &\in \mathcal{F}_{\mathrm{NUQ},u}, q = 1, 2, ..., 2^b. \end{aligned} \qquad (29)$$

Having constructed the NUQ codebooks, the analog TPC of the uth user having quantized phases can be designed by

$$\min_{\mathbf{F}_{\mathrm{RF},u}^{\mathrm{quant}}, \mathbf{F}_{\mathrm{RF},u}^{\mathrm{quant}} \in \mathbf{F}_{\mathrm{NUQ},u}} \left\| \mathbf{F}_{\mathrm{RF},u}^{\mathrm{quant}} - \widetilde{\mathbf{F}}_{\mathrm{RF},u} \right\|_F^2, \qquad (30)$$

which could be readily solved by finding the $N_{\mathrm{RF},u}^{\mathrm{r}}$ vectors in $\mathbf{F}_{\mathrm{NUQ},u}$ along which $\widetilde{\mathbf{F}}_{\mathrm{RF},u}$ has the largest $N_{\mathrm{RF},u}^{\mathrm{r}}$ projections. More details about the NUQ codebook can be found in [20].

IV. DIGITAL TPC DESIGN

Section III introduced the analog TPC which is identical for all subcarriers of each users by using the channel's covariance matrix. However, the differences between different subcarriers (caused by the beam squint) in the wideband mmWave channel and the inter-user interference have been tentatively ignored to simplify the hardware. Again, the analog TPC is designed to be frequency-flat across all subcarriers, while the digital TPC is capable of performing dedicated precoding for each subcarrier. Therefore, in this section, we propose a three-stage digital TPC scheme for jointly handling the beam squint, the inter-user interference and spatial multiplexing in order to improve the achievable sum rate.

A. Beam Squint Compensation

The first-stage digital TPC is designed to alleviate the effect of beam squint on the achievable sum rate by refining the frequency-flat analog TPC constructed. We will commence by introducing the concept of beam squint. Then the operation of beam squint compensation will be carried out with the aid of the AoD information.

Beam (angle) squint refers to the physical phenomenon, where the AoD deviates from the perpendicular direction with respect to the plane of the antenna, as a function of the frequency of a wideband transmit signal. In simple terms, the AoD of a wideband signal becomes frequency-dependent. More explicitly, as we have shown, the spatial angular offset is given by $\triangle_f = \frac{f_{\mathrm{s}}}{K}(k - 1 + \frac{K-1}{2})$, which means that the larger the bandwidth, the higher the angular offset. Naturally, a wideband mmWave MIMO system is likely to have high bandwidth (may be up to 4 GHz [1]) to achieve high-rate transmission and it is usually equipped with a large phased array. Therefore, the beam directions of different subcarriers substantially deviate, which will dramatically affect the system performance attained.

Note that within the angular coherence time, the AoDs are assumed to be known at the BS. Armed with the AoD, the BS becomes capable of constructing the accurate array response

vectors $\mathbf{A}_{t,u}^*[k]$. Note that $\mathbf{R}_u[k]$ is mainly determined by $\mathbf{A}_{t,u}[k]\mathbf{A}_{t,u}^*[k]$ and the singular vectors of $\mathbf{R}_u[k]$ are the same as those of $\mathbf{A}_{t,u}[k]\mathbf{A}_{t,u}^*[k]$. Upon applying the SVD to $\mathbf{A}_{t,u}[k]\mathbf{A}_{t,u}^*[k]$, we have

$$\mathbf{A}_{t,u}[k]\mathbf{A}_{t,u}^*[k] = \mathbf{V}_{t,u}[k]\mathbf{\Sigma}_{t,u}[k]\mathbf{V}_{t,u}^*[k]. \qquad (31)$$

Therefore, the "accurate analog TPC" of the kth subcarrier can be formulated as:

$$\widehat{\mathbf{F}}_{\mathrm{RF},u}[k] = \mathbf{V}_{t,u,}[k](:, 1:N_{\mathrm{RF}}^t/U). \qquad (32)$$

The remaining refining operations are to provide phase compensation to the common analog TPC $\mathbf{F}_{\mathrm{RF}}$ according to (32).

1) Direct Compensation: A straightforward technique is to directly apply phase rotations to each element of $\mathbf{F}_{\mathrm{RF}}$ in (23) for each subcarrier, as facilitated by the constant modular constraint. The elements of the direct compensation (DC) matrix for the kth subcarrier can be readily expressed as

$$[\mathbf{F}_{\mathrm{DC},u}[k]]_{i,j} = \frac{1}{\sqrt{N_t}}(e^{j\angle([\widehat{\mathbf{F}}_{\mathrm{RF},u}[k]]_{i,j})} - e^{j\angle([\mathbf{V}_{t,u,c}(:,1:\frac{N_{\mathrm{RF}}^t}{U})]_{i,j})}). \qquad (33)$$

Following this the phase compensation, the elements of the extended "analog TPC" can be written as

$$[\mathbf{F}_{\mathrm{RF},\mathrm{PC}}[k]]_{i,j} = \sqrt{N_t}[\mathbf{F}_{\mathrm{RF}}]_{i,j}[\mathbf{F}_{\mathrm{DC},u}[k]]_{i,j}. \qquad (34)$$

Note that the direct compensation still only depends on the long-term channel's covariance matrix and on the AoD information, which means that we do not have to update the phase compensation matrix within the anular coherence time.

2) Indirect Compensation: Beaning in mind that the dimension of the analog TPC is large, and the last two digital TPCs are both designed based on the effective CSI or the effective channel covariance matrix, the PC operation can be equivalently applied to the low-dimensional effective channel covariance matrix. The $N_{r,u} \times N_{\mathrm{RF}}^t$ effective CSI and the short-term $N_{\mathrm{RF}}^t \times N_{\mathrm{RF}}^t$ effective channel covariance matrix after the analog TPC are

$$\mathbf{H}_{\mathrm{eff},u}[k] = \mathbf{H}_u[k]\mathbf{F}_{\mathrm{RF}}, \qquad (35)$$

and

$$\mathbf{R}_{\mathrm{eff},u}[k] = \mathbf{F}_{\mathrm{RF}}^*\mathbf{H}_u^*[k]\mathbf{H}_u[k]\mathbf{F}_{\mathrm{RF}}, \qquad (36)$$

respectively. According to (32), we can calculate the accurate effective channel covariance matrix as

$$\widehat{\mathbf{R}}_{\mathrm{eff},u}[k] = \widehat{\mathbf{F}}_{\mathrm{RF},u}[k]^*\mathbf{H}_u^*[k]\mathbf{H}_u[k]\widehat{\mathbf{F}}_{\mathrm{RF},u}[k]. \qquad (37)$$

Therefore, the element of the indirect compensation (IDC) matrix conditioned for the kth subcarrier can be readily expressed as

$$[\mathbf{F}_{\mathrm{IDC},u}[k]]_{i,j} = (e^{j\angle([\widehat{\mathbf{R}}_{\mathrm{RF},u}[k]]_{i,j})} - e^{j\angle([\mathbf{R}_{\mathrm{eff},u}[k]]_{i,j})}). \qquad (38)$$

The extended "analog TPC" specifically conditioned for each subcarrier and each user is the same for both the direct compensation and indirect compensation,, i.e., $\mathbf{F}_{\mathrm{RF},\mathrm{PC}}[k]$. The power constraint now becomes

$$\|\mathbf{F}_{\mathrm{RF},\mathrm{PC}}[k]\mathbf{F}_{\mathrm{BB},u}[k]\|_F^2 = \frac{P_t}{UN_{s,u}}, \qquad (39)$$

where $\mathbf{F}_{\mathrm{BB},u}[k]$ consists of the next two digital TPCs stages, which will be introduced in the following sections.

B. Block Diagonalization

Note that in the design of the analog TPC and the first-stage digital TPC the inter-user interference has been tentatively ignored in order to obtain the TPCs of all users simultaneously. However, in multiuser wideband systems, the inter-user interference is quite a severe problem, hence neglecting it would dramatically degrade the system's performance, especially when serving a large number of users and at high SNRs. Therefore, the second-stage digital TPC is specifically designed for canceling the inter-user interference, which is mainly based on the classic idea of block diagonalization. In this process, the perfect low-dimensional short-term effective CSI is assumed to be known at the BS.

After the analog TPC and the first-stage digital TPC, the effective channel for the uth user on the kth subcarrier can be written as

$$\mathbf{H}_{\mathrm{BD},u}[k] = \mathbf{H}_u[k]\mathbf{F}_{\mathrm{RF},\mathrm{PC}}[k], \qquad (40)$$

and the $N_{\mathrm{RF}}^t \times N_{\mathrm{RF}}^t$ correlation matrix of the effective channel can be accordingly expressed as

$$\mathbf{R}_{\mathrm{BD},u}[k] = \mathbf{F}_{\mathrm{RF},\mathrm{PC}}^*[k]\mathbf{H}_u^*[k]\mathbf{H}_u[k]\mathbf{F}_{\mathrm{RF},\mathrm{PC}}[k]. \qquad (41)$$

In order to obtain the second-stage digital TPC $\mathbf{F}_{\mathrm{BD}}[k] = [\mathbf{F}_{\mathrm{BD},1}[k], \mathbf{F}_{\mathrm{BD},2}[k], ..., \mathbf{F}_{\mathrm{BD},U}[k]]$ on the kth subcarrier, where $\mathbf{F}_{\mathrm{BD},u}[k]$ contains the second-stage digital TPC vectors for the uth users, we first define $\widehat{\mathbf{R}}_{\mathrm{BD},u}[k]$ as

$$\widehat{\mathbf{R}}_{\mathrm{BD},u}[k] = [\mathbf{R}_{\mathrm{BD},1}^T[k], ..., \mathbf{R}_{\mathrm{BD},u-1}^T[k], \mathbf{R}_{\mathrm{BD},u+1}^T[k], ..., \mathbf{R}_{\mathrm{BD},U}^T[k]]^T. \qquad (42)$$

Our goal is to design $\mathbf{F}_{\mathrm{BD},u}[k]$ for ensuring that it satisfies

$$\mathbf{R}_{\mathrm{BD},p}[k]\mathbf{F}_{\mathrm{BD},u}[k] = \mathbf{0}, p \neq u. \qquad (43)$$

The equality in (43) can be enforced exactly if $\mathbf{F}_{\mathrm{BD},u}[k]$ lies in the null space of $\widehat{\mathbf{R}}_{\mathrm{BD},u}[k]$. Moreover, it should be satisfied that

$$\mathrm{dim}(\mathrm{Span}(\mathbf{R}_{\mathrm{BD},u}[k]) \cap \mathrm{Span}^{\perp}(\{\mathbf{R}_{\mathrm{BD},p}[k] : p \neq u\})) \geq N_{s,u} \qquad (44)$$

in order to obtain the $N_{s,u}$ spatial multiplexing gains (i.e., the number of interference-free data streams) in the baseband. Let us denote the rank of $\mathbf{R}_{\mathrm{BD},u}[k]$ as $r_u, u = 1, 2, ..., U$. When the ranks r_u are excessive, enforcing the exact BD operation as in (43) would result in spatial multiplexing gains lower than $N_{s,u}$. Moreover, since the number of RF chains N_{RF}^t is lower than the number of antennas, the degree of freedom in the null space is limited, which will also restrict the spatial multiplexing gains. Therefore, the approximate BD approach is adopted for circumventing this problem, by selecting $\widetilde{r}_u$ dominant eigenmodes of $\mathbf{R}_{\mathrm{BD},u}[k]$ for all users. Applying the SVD to the correlation matrix yields,

$$\mathbf{R}_{\mathrm{BD},u}[k] = \mathbf{V}_{\mathrm{BD},u}[k]\mathbf{\Sigma}_{\mathrm{BD},u}[k]\mathbf{V}_{\mathrm{BD},u}^*[k], \qquad (45)$$

where the dominant eigenmodes may be set as $\mathbf{V}_{\mathrm{BD},u}^*[k](:, 1 : \widetilde{r}_u)$. Therefore (43) is relaxed to

$$\mathbf{V}_{\mathrm{BD},u}^*[k](:, 1 : \widetilde{r}_p)\mathbf{F}_{\mathrm{BD},u}[k] = \mathbf{0}, p \neq u, \qquad (46)$$

and $\widehat{\mathbf{R}}_{\mathrm{BD},u}[k]$ is modified to

$$\widehat{\mathbf{R}}_{\mathrm{BD},u}[k] = \Big[\mathbf{V}^*_{\mathrm{BD},1}[k](:,1:\widetilde{r}_1); ...; \mathbf{V}^*_{\mathrm{BD},u-1}[k](:,1:\widetilde{r}_{u-1}); \\ \mathbf{V}^*_{\mathrm{BD},u+1}[k](:,1:\widetilde{r}_{u+1}); ...; \mathbf{V}^*_{\mathrm{BD},U}[k](:,1:\widetilde{r}_U)\Big], \tag{47}$$

which has a dimension of $\sum_{p\neq u}\widetilde{r}_p \times N^{\mathrm{t}}_{\mathrm{RF}}$ and rank $\sum_{p\neq u}\widetilde{r}_p$. Then the SVD of $\widehat{\mathbf{R}}_{\mathrm{BD},u}[k]$ is given by

$$\widehat{\mathbf{R}}_{\mathrm{BD},u}[k] = \widehat{\mathbf{V}}_{\mathrm{BD},1}[k]\widehat{\mathbf{\Sigma}}_{\mathrm{BD},u}[k]\Big[\widehat{\mathbf{V}}^{(1)}_{\mathrm{BD},u}[k], \widehat{\mathbf{V}}^{(0)}_{\mathrm{BD},u}[k]\Big]^*, \tag{48}$$

where $\widehat{\mathbf{V}}^{(0)}_{\mathrm{BD},u}[k]$ is of size $N^{\mathrm{t}}_{\mathrm{RF}} \times (N^{\mathrm{t}}_{\mathrm{RF}} - \sum_{p\neq u}\widetilde{r}_p)$ and forms a unitary basis for the orthogonal complement of $\mathrm{Span}^{\perp}(\{\widehat{\mathbf{R}}_{\mathrm{BD},p}[k] : p \neq u\})$, i.e., $\mathrm{Span}(\widehat{\mathbf{V}}^{(0)}_{\mathrm{BD},u}[k]) = \mathrm{Span}^{\perp}(\{\widehat{\mathbf{R}}_{\mathrm{BD},p}[k] : p \neq u\})$. Furthermore, the second-stage digital TPC is obtained by concatenating its projection onto $\mathrm{Span}(\widehat{\mathbf{V}}^{(0)}_{\mathrm{BD},u}[k])$ with eigen-beamforming along the dominant singular vectors of $\widehat{\widehat{\mathbf{R}}}_{\mathrm{BD},u}[k]$, which is given by

$$\begin{aligned}\widehat{\widehat{\mathbf{R}}}_{\mathrm{BD},u}[k] &= (\widehat{\mathbf{V}}^{(0)}_{\mathrm{BD},u}[k])^*\mathbf{R}_{\mathrm{BD},u}[k]\widehat{\mathbf{V}}^{(0)}_{\mathrm{BD},u}[k] \\ &= (\widehat{\mathbf{V}}^{(0)}_{\mathrm{BD},u}[k])^*\mathbf{V}_{\mathrm{BD},u}[k]\mathbf{\Sigma}_{\mathrm{BD},u}[k]\mathbf{V}^*_{\mathrm{BD},u}[k]\widehat{\mathbf{V}}^{(0)}_{\mathrm{BD},u}[k] \\ &= \widehat{\widehat{\mathbf{V}}}_{\mathrm{BD},u}[k]\widehat{\widehat{\mathbf{\Sigma}}}_{\mathrm{BD},u}[k]\widehat{\widehat{\mathbf{V}}}^*_{\mathrm{BD},u}[k],\end{aligned} \tag{49}$$

where $\widehat{\widehat{\mathbf{V}}}_{\mathrm{BD},u}[k] = \Big[\widehat{\widehat{\mathbf{V}}}^{(1)}_{\mathrm{BD},u}[k], \widehat{\widehat{\mathbf{V}}}^{(0)}_{\mathrm{BD},u}[k]\Big]$ and $\widehat{\widehat{\mathbf{V}}}^{(1)}_{\mathrm{BD},u}[k]$ consists of the dominant b_u singular vectors of $\widehat{\widehat{\mathbf{R}}}_{\mathrm{BD},u}[k]$. Finally, the second-stage digital TPC can be obtained by

$$\mathbf{F}_{\mathrm{BD},u}[k] = \widehat{\mathbf{V}}^{(0)}_{\mathrm{BD},u}[k]\widehat{\widehat{\mathbf{V}}}^{(1)}_{\mathrm{BD},u}[k] \in \mathbb{C}^{N^{\mathrm{t}}_{\mathrm{RF}} \times b_u}, \tag{50}$$

which is orthogonal to the dominant eigenmodes of $\widehat{\mathbf{R}}_{\mathrm{BD},u}[k]$ and matched to the b_u singular vectors of $\mathbf{R}_{\mathrm{BD},u}[k]$. Note that we have $b_u \leq \min\{r_u, N^{\mathrm{t}}_{\mathrm{RF}} - \sum_{p\neq u}\widetilde{r}_p\}$, if $\sum_p^U r_p \leq N^{\mathrm{t}}_{\mathrm{RF}}$, and we can choose $b_u = \widetilde{r}_u = r_u$ to obtain the second-stage digital TPC using exact BD. However, considering that the number of RF chains $N^{\mathrm{t}}_{\mathrm{RF}}$ is quite limited, b_u and $\widetilde{r}_u$ can be set as $b_u = \widetilde{r}_u = N^{\mathrm{r}}_{\mathrm{RF},u}$ to guarantee a good performance.

C. Spatial Multiplexing

Since b_u may be higher than the required spatial multiplexing gains $N_{\mathrm{s},u}$ and since $N_{\mathrm{s},u}$ is usually more than one for each user, i.e., $b_u \geq N_{\mathrm{s},u} > 1$, the third-stage digital TPC is specifically constructed for maximizing each user's achievable rate on each subcarrier based on the aforementioned analog TPC and the first two stage's of the digital TPC. This time the effective channel for the uth user on the kth subcarrier can be written as

$$\mathbf{H}_{\mathrm{SM},u}[k] = \mathbf{H}_u[k]\mathbf{F}_{\mathrm{RF,PC}}[k]\mathbf{F}_{\mathrm{BD},u}[k], \tag{51}$$

and the corresponding $b_u \times b_u$ correlation matrix is given by

$$\mathbf{R}_{\mathrm{SM},u}[k] = \mathbf{F}^*_{\mathrm{BD},u}[k]\mathbf{F}^*_{\mathrm{RF,PC}}[k]\mathbf{H}^*_u[k]\mathbf{H}_u[k]\mathbf{F}_{\mathrm{RF,PC}}[k]\mathbf{F}_{\mathrm{BD},u}[k]. \tag{52}$$

Since the inter-user interference has now been effectively canceled by the second-stage digital TPC, the achievable sum rate (11) can be approximately simplified to

$$\begin{aligned}\widetilde{R}_{\mathrm{sum}} = \frac{1}{K}\sum_{u=1}^{U}\sum_{k=1}^{K}\log_2\Big(\Big|\mathbf{I}_{N_{\mathrm{r},u}} + \frac{1}{\sigma^2_n}(\mathbf{H}_u[k]\mathbf{F}_{\mathrm{RF,PC}}[k]) \\ \times \mathbf{F}_{\mathrm{BD},u}[k]\mathbf{F}_{\mathrm{SM},u}[k]\mathbf{F}^*_{\mathrm{SM},u}[k]\mathbf{F}^*_{\mathrm{BB},u}[k]\mathbf{F}^*_{\mathrm{RF,PC}}[k]\mathbf{H}^*_u[k]\Big|\Big).\end{aligned} \tag{53}$$

Therefore, the problem of constructing the third-stage digital TPC for further achieve improving the achievable rate of the uth user on the kth subcarrier can be formulated as

$$\begin{aligned}(\mathbf{F}^{\mathrm{opt}}_{\mathrm{SM},u}[k]) &= \arg\max \widetilde{R}_{\mathrm{sum}}, \forall u, k, \\ \text{s.t.} \quad & \|\mathbf{F}_{\mathrm{RF,PC}}[k]\mathbf{F}_{\mathrm{BB},u}[k]\|^2_F = \frac{P_{\mathrm{t}}}{UN_{\mathrm{s},u}},\end{aligned} \tag{54}$$

where $\mathbf{F}_{\mathrm{BB},u}[k] = \mathbf{F}_{\mathrm{BD},u}[k]\mathbf{F}_{\mathrm{SM},u}[k]$. The problem (54) is difficult to solve due to the coupling between the third-stage digital TPC $\mathbf{F}_{\mathrm{SM},u}[k]$ and $\mathbf{F}_{\mathrm{RF,PC}}[k]\mathbf{F}_{\mathrm{BD},u}[k]$ in the power constraint. We will formulate the optimal solution of (54) in the following lemma.

Lemma 2: Let us define the dummy effective channel matrix $\widehat{\mathbf{H}}_{\mathrm{SM},u}[k]$ as

$$\begin{aligned}\widehat{\mathbf{H}}_{\mathrm{SM},u}[k] &= \mathbf{H}_{\mathrm{SM},u}[k]\Big(\mathbf{F}^*_{\mathrm{BD},u}[k]\mathbf{F}^*_{\mathrm{RF,PC}}[k]\mathbf{F}_{\mathrm{RF,PC}}[k]\mathbf{F}_{\mathrm{BD},u}[k]\Big)^{-\frac{1}{2}} \\ &= \mathbf{H}_u[k]\mathbf{F}_{\mathrm{RF,PC}}[k]\mathbf{F}_{\mathrm{BD},u}[k] \\ &\quad \times \Big(\mathbf{F}^*_{\mathrm{BD},u}[k]\mathbf{F}^*_{\mathrm{RF,PC}}[k]\mathbf{F}_{\mathrm{RF,PC}}[k]\mathbf{F}_{\mathrm{BD},u}[k]\Big)^{-\frac{1}{2}},\end{aligned} \tag{55}$$

and defining the SVD decomposition of the corresponding dummy correlation matrix $\widehat{\mathbf{R}}_{\mathrm{SM},u}[k]$ as

$$\begin{aligned}\widehat{\mathbf{R}}_{\mathrm{SM},u}[k] &= \widehat{\mathbf{H}}^*_{\mathrm{SM},u}[k]\widehat{\mathbf{H}}_{\mathrm{SM},u}[k] \\ &= \mathbf{V}_{\mathrm{SM},u}[k]\mathbf{\Sigma}_{\mathrm{SM},u}[k]\mathbf{V}^*_{\mathrm{SM},u}[k].\end{aligned} \tag{56}$$

Then, the optimal third-stage digital TPC for the uth user on the kth subcarrier $\mathbf{F}_{\mathrm{SM,u}}[k]$ can be formulated as

$$\begin{aligned}\mathbf{F}_{\mathrm{SM,u}}[k] &= \Big(\mathbf{F}^*_{\mathrm{BD},u}[k]\mathbf{F}^*_{\mathrm{RF,PC}}[k]\mathbf{F}_{\mathrm{RF,PC}}[k]\mathbf{F}_{\mathrm{BD},u}[k]\Big)^{-\frac{1}{2}} \\ &\quad \times \mathbf{V}_{\mathrm{SM},u}[k](:,1:N_{\mathrm{s},u}).\end{aligned} \tag{57}$$

Proof: To circumvent the problem that there is a power-coupling between the different TPCs, we first let $\mathbf{F}_{\mathrm{SM,u}}[k] = \Big(\mathbf{F}^*_{\mathrm{BD},u}[k]\mathbf{F}^*_{\mathrm{RF,PC}}[k]\mathbf{F}_{\mathrm{RF,PC}}[k]\mathbf{F}_{\mathrm{BD},u}[k]\Big)^{-\frac{1}{2}}\widehat{\mathbf{F}}_{\mathrm{SM,u}}[k]$. Then the problem (54) becomes equivalent to

$$\begin{aligned}(\mathbf{F}^{\mathrm{opt}}_{\mathrm{SM},u}[k]) &= \arg\max \widetilde{R}_{\mathrm{sum}}, \forall u, k, \\ \text{s.t.} \quad & \Big\|\widehat{\mathbf{F}}_{\mathrm{SM,u}}[k]\Big\|^2_F = \frac{P_{\mathrm{t}}}{UN_{\mathrm{s},u}},\end{aligned} \tag{58}$$

and $\widetilde{R}_{\mathrm{sum}}$ can be written as

$$\begin{aligned}\widetilde{R}_{\mathrm{sum}} = \frac{1}{K}\sum_{u=1}^{U}\sum_{k=1}^{K}\log_2\Big(\Big|\mathbf{I}_{N_{\mathrm{r},u}} + \frac{1}{\sigma^2_n}(\widehat{\mathbf{H}}_{\mathrm{SM},u}[k]\widehat{\mathbf{F}}_{\mathrm{SM,u}}[k]) \\ \times \widehat{\mathbf{F}}^*_{\mathrm{SM,u}}[k]\widehat{\mathbf{H}}^*_{\mathrm{SM,u}}[k]\Big|\Big).\end{aligned} \tag{59}$$

Since $\widehat{\mathbf{H}}_{\mathrm{SM},u}[k]$ is known at the BS after the second-stage digital TPC and the power is now decoupled, the optimal $\widehat{\mathbf{F}}_{\mathrm{SM},u}[k]$ can be readily found by the SVD decomposition of the $b_u \times b_u$ dummy correlation matrix $\widehat{\mathbf{R}}_{\mathrm{SM},u}[k]$ i.e., $\widehat{\mathbf{F}}_{\mathrm{SM},u}[k] = \mathbf{V}_{\mathrm{SM},u}[k](:, 1:N_{\mathrm{s},u})$ and we can obtain the result shown in (57). $\blacksquare$

V. SIMULATION RESULTS

In this section, we evaluate the performances of the proposed hybrid TPC scheme using numerical simulations in our multiuser wideband mmWave MIMO system. We characterize four implement actions of the hybrid TPC scheme explicitly, 'Hybrid-PC' is the scheme utilizing all the three-stages digital TPC, 'Hybrid-PC w/o BD' is the scheme where the BD digital TPC is removed, 'Hybrid' is the scheme that does not apply phase compensation (the first-stage digital TPC) but contains the BD, and finally 'Hybrid w/o BD' represents the corresponding scheme that only carries out the spatial multiplexing operation. Moreover, the hybrid TPC schemes whose analog TPCs use quantized phases selected from the NUQ codebook and the UQ codebook are marked as 'Hybrid quantization-NUQ', and 'Hybrid quantization-UQ', respectively. We assume that the ULAs at both the BS and MSs use $\lambda/2$ element-spacing. The channel parameters for each user are set as $N_{\mathrm{cl}} = 3$ and $N_{\mathrm{ray}} = 2$ for all figures, except for Fig. (9). The azimuth AOAs and AODs obey the Laplacian distribution with uniformly distributed mean angles within $(0, 2\pi]$ and angular spreads of $10°$ for all figures, except for Fig. (10). The paths gains obey $\alpha_{u,m,n} \sim \mathcal{CN}(0, \sigma_{\alpha_u}^2)$ with $\sigma_{\alpha_u}^2 = 1$, and the delay assume $\tau_{u,m,n} \sim \mathcal{U}(0, 20\mathrm{ns})$ and $\max\limits_{u,m,n} \tau_{u,m,n} = 20\mathrm{ns}$. The carrier frequency is $f_{\mathrm{c}} = 28$ GHz, the number of subcarriers is set to $K = 128$, the bandwidth is $f_{\mathrm{s}} = 4$ GHz for all figures except for Fig. (6), and the number of quantization bits is set to $b = 7$. Finally, the SNR is defined as $\frac{P_{\mathrm{t}}}{U\sigma_n^2}$.

Benchmark. The fully digital TPC schemes with short-term CSI and AoD information are adopted as the benchmark. We consider two fully digital implement actions, i.e., fully digital with BD ('Digital') and fully digital without BD ('Digital w/o BD'). The one using BD is designed based on the $N_{\mathrm{t}} \times N_{\mathrm{t}}$ correlation matrix for each user on each subcarrier $\mathbf{R}_{\mathrm{Full},u}[k] = \mathbf{H}_u^*[k]\mathbf{H}_u[k] = \mathbf{V}_u[k]\boldsymbol{\Sigma}_u[k]\mathbf{V}_u^*[k]$ and the remaining operations of BD for the fully digital TPC are similar to those discussed in Section IV.B. The corresponding steps to obtain the spatial multiplexing gains are also similar to (57), upon replacing $\mathbf{F}_{\mathrm{RF,PC}}[k]\mathbf{F}_{\mathrm{BD},u}[k]$ by $\mathbf{F}_{\mathrm{BD,Full},u}[k]$, where $\mathbf{F}_{\mathrm{BD,Full},u}[k]$ is the BD matrix obtained for the fully digital TPC scheme. Note that the 'Digital' scheme has higher degrees of freedom than the hybrid TPC schemes in order to cancel the inter-user interference, since we have $N_{\mathrm{t}} \gg N_{\mathrm{RF}}^t$.

A. Achievable sum rate vs transmission parameters (SNR, U, N_{s} and f_{s})

This subsection evaluates the achievable sum rate of the proposed 'Hybrid-PC', 'Hybrid-PC w/o BD', 'Hybrid', 'Hybrid w/o BD' schemes, as well as the 'Digital' and 'Digital w/o BD' schemes for different transmission parameters.
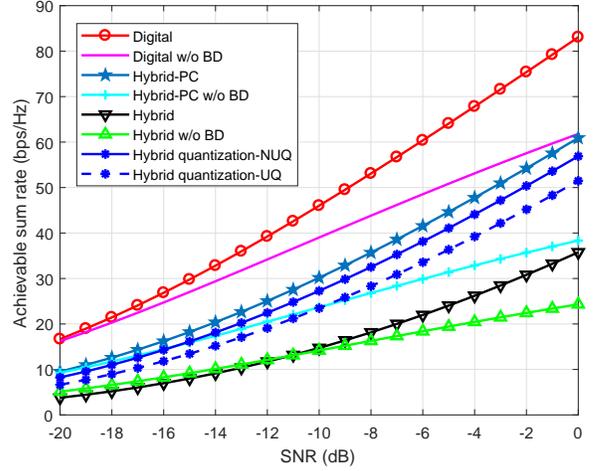
Fig. 2. The performance of the proposed hybrid TPC scheme, the proposed hybrid TPC schemes based on the codebooks, and the corresponding fully digital TPC schemes for different SNRs, where $N_{\mathrm{t}} = 128$, $N_{\mathrm{r},u} = 16$, $U = 6$, $N_{\mathrm{RF},u}^{\mathrm{r}} = N_{\mathrm{s},u} = 2$, and $N_{\mathrm{RF}}^{\mathrm{t}} = UN_{\mathrm{s},u}$.
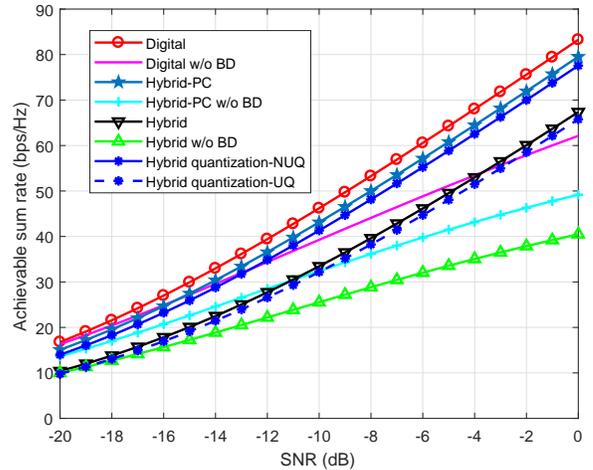
Fig. 3. The performance of the proposed hybrid TPC scheme, the proposed hybrid TPC schemes based on the codebooks, and the corresponding fully digital TPC schemes for different SNRs, where $N_{\mathrm{t}} = 128$, $N_{\mathrm{r},u} = 16$, $U = 6$, $N_{\mathrm{RF},u}^{\mathrm{r}} = N_{\mathrm{s},u} = 2$, and $N_{\mathrm{RF}}^{\mathrm{t}} = 3UN_{\mathrm{s},u}$.

Fig. 2 and Fig. 3 show the achievable sum rate of the schemes for different SNRs, where we have $N_{\mathrm{t}} = 128$, $N_{\mathrm{r},u} = 16$, $U = 6$, $N_{\mathrm{RF},u}^{\mathrm{r}} = N_{\mathrm{s},u} = 2$, while $N_{\mathrm{RF}}^{\mathrm{t}}$ is set as $N_{\mathrm{RF}}^{\mathrm{t}} = UN_{\mathrm{s},u}$ and $N_{\mathrm{RF}}^{\mathrm{t}} = 3UN_{\mathrm{s},u}$, respectively. It can be observed that

- The proposed 'Hybrid-PC' scheme always achieves better sum rate than the 'Hybrid' scheme and the 'Hybrid-PC w/o BD' is also better than the 'Hybrid w/o BD' scheme, both of which confirm the effectiveness of the first-stage digital TPC.
- The 'Digital', 'Hybrid-PC' and 'Hybrid' schemes are capable of achieving better sum rate than the 'Digital w/o BD', 'Hybrid-PC w/o BD' and 'Hybrid w/o BD' schemes, respectively, especially at high SNRs. More-over, the 'Hybrid-PC' and 'Hybrid' schemes could even
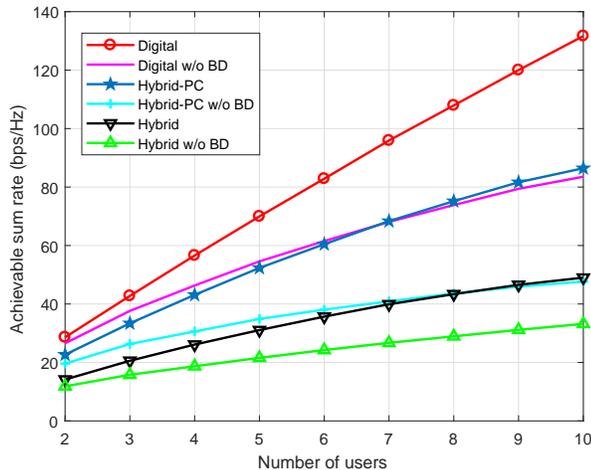
Fig. 4. The performance of the proposed hybrid TPC scheme, and the corresponding fully digital TPC schemes for different numbers of users, where $N_{\mathrm{t}} = 128$, $N_{\mathrm{r},u} = 16$, SNR=0 dB, $N_{\mathrm{RF},u}^{\mathrm{r}} = N_{\mathrm{s},u} = 2$, and $N_{\mathrm{RF}}^{\mathrm{t}} = UN_{\mathrm{s},u}$.

Fig. 5. The performance of the proposed hybrid TPC scheme, and the corresponding fully digital TPC schemes for different numbers of data streams, where $N_{\mathrm{t}} = 128$, $N_{\mathrm{r},u} = 16$, $U = 6$, SNR=0 dB, $N_{\mathrm{RF},u}^{\mathrm{r}} = N_{\mathrm{s},u}$, and $N_{\mathrm{RF}}^{\mathrm{t}} = UN_{\mathrm{s},u}$.

achieve a performance similar to (Fig. 2) or even a better performance than (Fig. 3) the 'Digital w/o BD' and 'Hybrid-PC w/o BD' schemes at high SNRs, respectively. These phenomena illustrate that the second-stage digital TPC increase the achievable sum rate.

- The proposed 'Hybrid quantization-NUQ' scheme out-performs the Hybrid quantization-UQ' scheme, which indicates that the proposed NUQ codebook is capable of better matching the optimal unconstrained analog TPC than the traditional UQ codebook.

- There are non-negligible performance gaps between the 'Digital' and the proposed 'Hybrid-PC' schemes, when $N_{\mathrm{RF}}^{\mathrm{t}} = UN_{\mathrm{s},u}$. However, when $N_{\mathrm{RF}}^{\mathrm{t}} = 3UN_{\mathrm{s},u}$, the performance gap becomes quite limited. This is because the 'Digital' scheme has higher degrees of freedom than the proposed 'Hybrid-PC' schemes for cancelling the inter-user interference and because the analog TPC of the proposed 'Hybrid-PC' scheme is based on the long-term channel's covariance matrix not on the short-term CSI.

Fig. 4 compares the achievable sum rate of the various schemes for different numbers of users, where we have $N_{\mathrm{t}} = 128$, $N_{\mathrm{r},u} = 16$, SNR=0 dB, $N_{\mathrm{RF},u}^{\mathrm{r}} = N_{\mathrm{s},u} = 2$, and $N_{\mathrm{RF}}^{\mathrm{t}} = UN_{\mathrm{s},u}$. We observe that the proposed 'Hybrid-PC' scheme and the 'Hybrid' schemes outperform the 'Hybrid-PC w/o BD' and 'Hybrid w/o BD' schemes, respectively. More-over, the performance gaps become higher as U is increased, because the inter-user interference becomes more serious for users. Furthermore, the proposed 'Hybrid-PC' scheme and the 'Hybrid' schemes always achieve similar sum-rates as the 'Digital w/o BD' and 'Hybrid-PC w/o BD' schemes, which also conforms the effectiveness of the proposed second-stage digital TPC.

In Fig. 5, the impact of the number of data streams on the achievable sum rate is presented, where we have $N_{\mathrm{t}} = 128$, $N_{\mathrm{r},u} = 16$, $U = 6$, SNR=0 dB, $N_{\mathrm{RF},u}^{\mathrm{r}} = N_{\mathrm{s},u}$, and $N_{\mathrm{RF}}^{\mathrm{t}} = UN_{\mathrm{s},u}$. It can be observed that the achievable sum rate
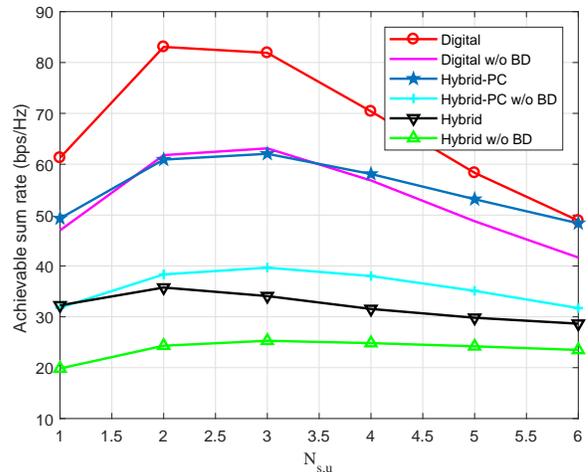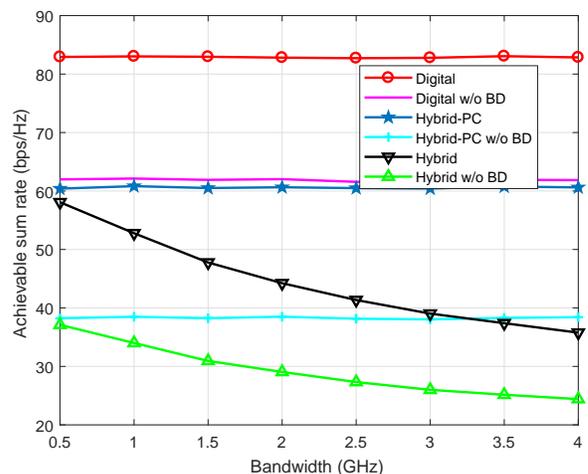
Fig. 6. The performance the proposed hybrid TPC scheme, and the corresponding fully digital TPC schemes for different bandwidths, where $N_{\mathrm{t}} = 128$, $N_{\mathrm{r},u} = 16$, $U = 6$, SNR=0 dB, $N_{\mathrm{RF},u}^{\mathrm{r}} = N_{\mathrm{s},u} = 2$, and $N_{\mathrm{RF}}^{\mathrm{t}} = UN_{\mathrm{s},u}$.

of all schemes is first increased to the point at $N_{\mathrm{s},u} = 2$ and then starts to decay, as the number of data streams increases. This downward trend is a result of equal power allocation among the different streams and due to the sparsity of the mmWave channel in the angular domain, where some of the power is allocated to delayed paths contributing reduced gains. Furthermore, Fig. 5 illustrates that the proposed 'Hybrid-PC' scheme outperforms both the 'Hybrid-PC w/o BD' and the 'Hybrid-PC' schemes, regardless of the number of data streams. Specifically, when $N_{\mathrm{s},u} = N_{\mathrm{cl}}N_{\mathrm{ray}} = 6$, there is nearly no performance gap between the proposed 'Hybrid-PC' scheme and the optimal fully digital scheme.

In Fig. 6, we evaluate the performance of four implement actions of the proposed hybrid TPC scheme benchmarked against the fully digital TPC schemes for different bandwidths,
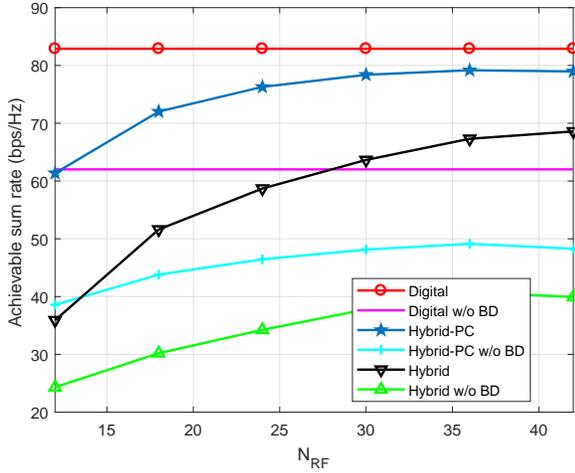
Fig. 7. The performance of the proposed hybrid TPC scheme, and the corresponding fully digital TPC schemes for different numbers of RF chains at the BS, where $N_{\mathrm{t}} = 128$, $N_{\mathrm{r},u} = 16$, $U = 6$, SNR=0 dB, and $N_{\mathrm{RF},u}^{\mathrm{r}} = N_{\mathrm{s},u} = 2$.

Fig. 8. The performance of the proposed hybrid TPC scheme, and the corresponding fully digital TPC schemes for different numbers of antennas at the BS, where $N_{\mathrm{r},u} = 16$, $U = 6$, SNR=0 dB, $N_{\mathrm{RF},u}^{\mathrm{r}} = N_{\mathrm{s},u} = 2$, and $N_{\mathrm{RF}}^{\mathrm{t}} = U N_{\mathrm{s},u}$.

where we have $N_{\mathrm{t}} = 128$, $N_{\mathrm{r},u} = 16$, $U = 6$, SNR=0 dB, and $N_{\mathrm{RF},u}^{\mathrm{r}} = N_{\mathrm{s},u} = 2$, $N_{\mathrm{RF}}^{\mathrm{t}} = U N_{\mathrm{s},u}$. It can be observed that the performance gaps between the proposed 'Hybrid-PC' scheme and the 'Hybrid' scheme, as well as between the 'Hybrid-PC w/o BD' scheme and the 'Hybrid w/o BD' scheme increase, when an increased bandwidth is adopted. This observation is consistent with our analysis in Section IV.B, indicating that the beam squint dramatically affects the system performance. Moreover, we observe that the 'Hybrid-PC w/o BD' scheme attains better performance than the 'Hybrid' scheme at high bandwidths, which indicates that the effect of the beam squint is even higher than that of the inter-user interference for wideband multi-user mmWave MIMO systems.

B. Achievable Sum-Rate vs Architectural Parameters ($N_{\mathrm{RF}}^{\mathrm{t}}$ and N_{t})

In this subsection, we evaluate the effect of the architectural parameters on the achievable sum rate of the proposed 'Hybrid-PC', 'Hybrid-PC w/o BD', 'Hybrid', 'Hybrid w/o BD' schemes, and of the corresponding fully digital TPC schemes.

Fig. 7 presents the achievable sum rate of different schemes using different numbers of RF chains, where we have $N_{\mathrm{t}} = 128$, $N_{\mathrm{r},u} = 16$, $U = 6$, SNR=0 dB, and $N_{\mathrm{RF},u}^{\mathrm{r}} = N_{\mathrm{s},u} = 2$. We observe that the achievable sum rate of the proposed 'Hybrid-PC', 'Hybrid-PC w/o BD', 'Hybrid', and 'Hybrid w/o BD' schemes increases with the number of RF chains at the BS. This is because more RF chains are capable of providing higher degrees of freedom to obtain increased multiplexing gains. Moreover, Fig. 7 also shows that the performance gap between the proposed 'Hybrid-PC' scheme and the fully digital TPC scheme 'Digital' may be compensated by increasing the number of RF chains.

Fig. 8 illustrates the achievable sum rate obtained by different schemes for different numbers of antennas at the BS, where we have $N_{\mathrm{r},u} = 16$, $U = 6$, SNR=0 dB, $N_{\mathrm{RF},u}^{\mathrm{r}} = N_{\mathrm{s},u} = 2$,
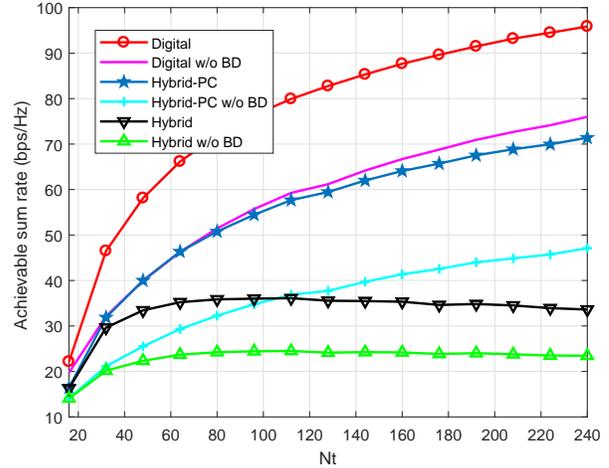
and $N_{\mathrm{RF}}^{\mathrm{t}} = U N_{\mathrm{s},u}$. It may be observed that the achievable sum rate of the proposed 'Hybrid-PC' and 'Hybrid-PC w/o BD' schemes exhibits an upward trend as the number of antennas increases. By contrast, the 'Hybrid', and 'Hybrid w/o BD' schemes have a reduce rate as the number of antennas increases, which indicates that the performance degradation imposed by the beam squint is higher than the performance increase benefits of having more antennas. Therefore, it is of great significance to impose phase compensation for the analog TPC in the wideband multi-user mmWave MIMO system.

C. Achievable Sum-Rate vs Channel Parameters (N_{c} and, angle spread)

Finally, we show the effect of channel parameters on the achievable sum rate of the proposed 'Hybrid-PC', 'Hybrid-PC w/o BD', 'Hybrid', 'Hybrid w/o BD' schemes, and on the corresponding fully digital TPC schemes.

In Fig. 9, we evaluate the performance of different schemes for different numbers of clusters, where we have $N_{\mathrm{t}} = 128$, $N_{\mathrm{r},u} = 16$, $U = 6$, SNR=0 dB, $N_{\mathrm{RF},u}^{\mathrm{r}} = N_{\mathrm{s},u} = 2$, $N_{\mathrm{RF}}^{\mathrm{t}} = U N_{\mathrm{s},u}$, and $N_{\mathrm{ray}} = 2$. The impact of the clusters on the achievable sum rate is difficult to evaluate for the different schemes from the theoretical perspective in our wideband multi-user mmWave MIMO system. On the one hand, when the CIR is very sparse, i.e. when the number of clusters is low, the multiplexing gains can benefit from larger number of clusters. However, on the other hand, more clusters would average the paths gains through the normalization factor $\gamma = \sqrt{N_{\mathrm{t}} N_{\mathrm{r}} / N_{\mathrm{cl}} N_{\mathrm{ray}}}$. Moreover, the inter-user interference will increase with the number of clusters. Nevertheless, we can still observe from Fig. 9 that the proposed 'Hybrid-PC' scheme obtains a higher sum rate than the 'Hybrid-PC w/o BD' and 'Hybrid' schemes, which validates the effectiveness of the proposed hybrid TPC scheme.

Fig. 10 compares the performance of different schemes for different angular spreads, where we have $N_{\mathrm{t}} = 128$,

Fig. 9. The performance of the proposed hybrid TPC scheme, and the corresponding fully digital TPC schemes for different numbers of clusters, where $N_{\rm t} = 128$, $N_{{\rm r},u} = 16$, $U = 6$, SNR=0 dB, $N_{{\rm RF},u}^{\rm r} = N_{{\rm s},u} = 2$, $N_{\rm RF}^{\rm t} = UN_{{\rm s},u}$, and $N_{\rm ray} = 2$.

Fig. 10. The performance of the proposed hybrid TPC scheme, and the corresponding fully digital TPC schemes for different angle spreads, where $N_{\rm t} = 128$, $N_{{\rm r},u} = 16$, $U = 6$, SNR=0 dB, $N_{{\rm RF},u}^{\rm r} = N_{{\rm s},u} = 2$, and $N_{\rm RF}^{\rm t} = UN_{{\rm s},u}$.

$N_{{\rm r},u} = 16$, $U = 6$, SNR=0 dB, $N_{{\rm RF},u}^{\rm r} = N_{{\rm s},u} = 2$, and $N_{\rm RF}^{\rm t} = UN_{{\rm s},u}$. It is also challenging to analyse the impact of the angular spread on the achievable sum rate, since an increased angular spread increases the degrees of freedom, hence increasing the multiplexing gains. However, at the same time, the total power is also spread across an increased number of less important multi-path components, which will degrade the system's performance, when the number of data streams for each user is low. Observe from Fig. 10 that the proposed 'Hybrid-PC' scheme outperforms the other hybrid TPCs in terms the achievable sum rate, when the angular spread varies from $1°$ to $15°$.

VI. CONCLUSIONS

In this paper, we considered the design of hybrid TPCs relying only on the long-term channel's covariance matrix and on the AoD information for wideband multiuser mmWave MIMO systems. We firstly derived the covariance matrix based on the mmWave channel matrix of each user, which showed that the covariance matrix is mainly determined by the array response matrix at the BS. Based on this, we constructed the analog TPC matrix by SVD and phase extraction. We also proposed a NUQ codebook based analog TPC design assisted by the AoD information to accurately quantize the physical steering angles and hence to increase the sum rate without increasing the feedback overhead. Then, a three-stage digital TPC scheme was proposed for compensating the beam squint, for cancelling the inter-user interference and for attaining the optimal multiplexing gains. Our simulation results have demonstrated that the proposed hybrid TPC scheme is capable of achieving a similar performance to the high-complexity fully digital TPC even for a very large bandwidth.

REFERENCES

[1] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar and L. Hanzo, "Millimeter-wave communications: physical channel models, design considerations, antenna constructions, and link-budget," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 870-913, 2nd Quart. 2018.

[2] R. W. Heath, N. Gonzlez-Prelcic, S. Rangan, W. Roh and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Areas Commun.,* vol. 10, no. 3, pp. 436-453, Apr. 2016.

[3] S. Qiu, Da Chen, D. Qu, K. Luo and T. Jiang, "Downlink precoding with mixed statistical and imperfect instantaneous CSI for massive MIMO systems," *IEEE Trans. Veh. Technol.,* vol. 67, no. 4, pp. 3028-3041, Apr. 2018.

[4] T. Zhang, C. Wen, S. Jin, and T. Jiang, "Mixed-ADC massive MIMO detectors: performance analysis and design optimization," *IEEE Trans. Wireless Commun.,* vol. 15, no. 11, pp. 7738-7752, Nov. 2016.

[5] P. Liu, S. Jin, T. Jiang, Q. Zhang, and M. Matthaiou, "Pilot power allocation through user grouping in multi-cell massive MIMO systems," *IEEE Trans. Commun.*, vol. 65, no. 4, pp. 1561-1574, Apr. 2017.

[6] Y. Wu, C. Xiao, Z. Ding, X. Gao and S. Jin, "Linear precoding for finite-alphabet signaling over MIMOME wiretap channels," *IEEE Trans. Veh. Technol.,* vol. 61, no. 6, pp. 2599-2612, Jul. 2012.

[7] X. Li, S. Jin, X. Gao and R. W. Heath, "Three-dimensional beamforming for large-scale FD-MIMO systems exploiting statistical channel state information," *IEEE Trans. Veh. Technol.,* vol. 65, no. 11, pp. 8992-9005, Nov. 2016.

[8] X. Zhang, A. Molisch, and S. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," *IEEE Trans. Signal Process.,* vol. 53, no. 11, pp. 4091-4103, Nov. 2005.

[9] 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.

[10] J. Zhao, F. Gao, W. Jia, S. Zhang, S. Jin and H. Lin, "Angle domain hybrid precoding and channel tracking for millimeter wave massive MIMO systems," *IEEE Trans. Wireless Commun.,* vol. 16, no. 10, pp. 6868-6880, Oct. 2017.

[11] K. Satyanarayana, M. El-Hajjar, P. H. Kuo, A. Mourad and L. Hanzo, "Dual-function hybrid beamforming and transmit diversity aided millimeter wave architecture," *IEEE Trans. Veh. Technol.,* vol. 67, no. 3, pp. 2798-2803, Mar. 2018.

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

[13] W. Ni and X. Dong, "Hybrid block diagonalization for massive multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 201-211, Jan. 2016.

[14] C. E. Chen, "An iterative hybrid transceiver design algorithm for millimeter wave MIMO systems," *IEEE Wireless Commun. Lett.*,vol.4, no. 3, pp. 285-288, Jun. 2015.

[15] 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.

[16] C. Rusu, R. Méndez-Rial, N. González-Prelcic and R. W. Heath, "Low complexity hybrid precoding strategies for millimeter wave communication systems," *IEEE Trans. Wireless Commun.,* vol. 15, no. 12, pp. 8380-8393, Dec. 2016.

[17] J. Singh and S. Ramakrishna, "On the feasibility of codebook-based beamforming in millimeter wave systems with multiple antenna arrays," *IEEE Trans. Wireless Commun.,* vol. 14, no. 5, pp. 2670-2683, May 2015.

[18] C. H. Chen, C. R. Tsai, Y. H. Liu, W. L. Hung and A. Y. Wu, "Compressive sensing (CS) assisted low-complexity beamspace hybrid precoding for millimeter-wave MIMO systems," *IEEE Trans. Signal Process.,* vol. 65, no. 6, pp. 1412-1424, Mar. 2017.

[19] Y. Chen, D. Chen, Y. Tian, and T. Jiang, "Spatial lobes division based low complexity hybrid precoding and diversity combining for mmWave IoT systems," *IEEE Internet Things J.*, 2018. DOI: 10.1109/JIOT.2018.2881171

[20] Y. Chen, D. Chen, and T. Jiang, "Non-uniform quantization codebook based hybrid precoding to reduce feedback overhead in millimeter wave MIMO systems," *IEEE Trans. Commun.*, 2018. DOI: 10.1109/TCOM-M.2018.2890227

[21] D. Zhang, Y. Wang, X. Li, W. Xiang, "Hybridly-connected structure for hybrid beamforming in mmWave massive MIMO systems," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 662-674, Feb. 2018.

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

[23] M. Dai and B. Clerckx, "Multiuser millimeter wave beamforming strategies with quantized and statistical CSIT," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7025-7038, Nov. 2017.

[24] W. Ni and X. Dong, "Hybrid block diagonalization for massive multiuser MIMO systems," *IEEE Trans. Commun.,* vol. 64, no. 1, pp. 201-211, Jan. 2016.

[25] X. Yu, J. Zhang, and K. B. Letaief, "Alternating minimization for hybrid precoding in multiuser OFDM mmWave Systems," in *Proc. IEEE Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2016.

[26] S. Park, A. Alkhateeb, and R. W. Heath, "Dynamic subarrays for hybrid precoding in wideband mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2907-2920, May 2017.

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

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

[29] T. E. Bogale, L. B. Le, A. Haghighat, and L. Vandendorpe, "On the number of RF chains and phase shifters, and scheduling design with hybrid analog-digital beamforming," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 331-3326, May 2016.

[30] J. P. González-Coma, J. Rodríguez-Fernández, N. González-Prelcic, L. Castedo and R. W. Heath, "Channel estimation and hybrid precoding for frequency selective multiuser mmWave MIMO systems," *IEEE J. Sel. Top. Signal Process.*, vol. 12, no. 2, pp. 353-367, May 2018.

[31] V. Raghavan, S. V. Hanly, and V. V. Veeravalli, "Statistical beamforming on the Grassmann manifold for the two-user broadcast channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6464-6489, Oct. 2013.

[32] S. Park, J. Park, A. Yazdan, and R. W. Heath, "Exploiting spatial channel covariance for hybrid precoding in massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 65, no. 14, pp. 3818-832, Jul. 2017.

[33] A. Adhikary, E. A. 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.

[34] J. Brady and A. M. Sayeed, "Wideband communication with high-dimensional arrays: New results and transceiver architectures," in *Proc. IEEE ICC Workshops*, Jun. 2015, pp. 1042-1047

[35] X. Gao, L. Dai, S. Zhou, A. Sayeed, and L. Hanzo, "Beamspace channel estimation for wideband millimeter-wave MIMO with lens antenna array," in *Proc. IEEE Int. Conf. Commun.*, Kansas, US, May 2018, pp. 1-6.

[36] V. Va, J. Choi, and R. W. Heath, "The impact of beamwidth on temporal channel variation in vehicular channels and its implications," *Trans. Veh. Technol.*, vol. 66, no. 6, pp. 5014-5029, Jun. 2017.

[37] W. Shen, L. Dai, B. Shim, Z. Wang, and R. W. Heath, "Channel feedback based on AoD-adaptive subspace codebook in FDD massive MIMO systems," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5235-5248, Nov. 2018.

[38] M. K. Samimi and T. S. Rappaport, "Ultra-wideband statistical channel model for non line of sight millimeter-wave urban channels," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Austin, TX, Dec. 2014, pp. 3483-3489.

[39] M. K. Samimi and T. S. Rappaport, "3-D millimeter-wave statistical channel model for 5G wireless system design," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 7, pp. 2207-2225, Jul. 2016.

[40] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3029-3056, Sep. 2015.