Published in final edited form as:

IEEE Trans Veh Technol. 2019 ; 68: . doi:10.1109/TVT.2019.2909829.

Spectral and Energy Efficiencies of Millimeter Wave MIMO With Configurable Hybrid Precoding

Zhong Zheng [Member, IEEE],

Advanced Network Technologies Division, National Institute of Standards and Technology, Gaithersburg, MD 20899 USA. He is now with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China

Hamid Gharavi [Life Fellow, IEEE]

Advanced Network Technologies Division, National Institute of Standards and Technology, Gaithersburg, MD 20899 USA

Abstract

Hybrid precoding architectures are widely studied for millimeter wave (mmWave) massive MIMO systems. A major challenge in designing hybrid precoders is the practical constraints on the number of RF chains, which can have a direct impact on the spectral and energy efficiencies of the communication systems. In this paper, we investigate tradeoff between the two performance metrics in both static and mobile communication scenarios via closed-form expressions, when the number of active RF chains can be selected. Based on these expressions, the computational complexity to configure the hybrid precoder is reduced, which can be used to adaptively activate required RF chains for the given MIMO system and channel condition. Numerical results indicate that a certain number of RF chains should be activated in order to maximize energy efficiency at high SNRs, which is generally different from the optimal configuration to maximize spectral efficiency. Further-more, for low SNRs, we have shown that a simple analog beamforming, which uses only a single RF chain, is optimal for both spectral and energy efficiencies. In addition, the proposed mobility-aware hybrid precoding is shown to be capable of effectively achieving beamforming gain between high-speed mobile devices.

Index Terms

Millimeter wave; massive MIMO; hybrid precoding; energy efficiency; random matrix theory

I. Introduction

MILLIMETER wave (mmWave) communications and massive Multiple-Input Multiple-Output (MIMO) are the key technologies in 5th generation cellular systems [1]. To enable multi-gigabit transmissions, they will be adopted in future communication networks, including traditional cellular systems, distributed radio-over-fiber systems [2], massive ad hoc networks [3], and space-air networks [4]. Such communication systems take advantage

of large mmWave spectral bands (from 30 GHz to 300 GHz) allocated for mobile communications. Due to the very high frequency of mmWave, massive MIMO communications can be efficiently utilized to overcome more severe propagation and penetration losses as compared to sub-6 GHz transmissions [5]. Under these conditions, massive antenna arrays can be deployed for both path loss compensation and capacity improvement [6]. Such a deployment is highly feasible in mmWave communications due to the small wave-length, where hundreds of antenna elements can be compactly packed for massive MIMO implementation [7].

The traditional multi-antenna systems in sub-6 GHz bands, however, are generally implemented with fully digital architecture, where all signal processing is performed in the digital domain before transmission in the Radio Frequency (RF). To realize a fully digital architecture, the number of RF chains is equal to the number of antennas [8], which is not practical for massive MIMO systems with hundreds of antennas at each communication device. This is not only because a large number of RF chains significantly increases the hardware cost, but they also consume large amounts of energy in wireless transmission systems [9]. To cope with this issue, mmWave antenna systems adopt a hybrid analog and digital architecture [10], where the number of RF chains is much smaller than the number of antennas in order to reduce energy consumption. Accordingly, hybrid processing at transmitters, which is known as hybrid precoding, is divided into analog and digital signal processing, where the former is directly implemented with analog circuitry, such as phase shifters.

Reducing the number of RF chain can have a significant impact on both the spectral and energy efficiencies of the communication system. For instance, when the mmWave channel has a sufficient spatial degree of freedom, the number of RF chains determines the maximum number of multiplexed data streams, which can be simultaneously transmitted over wireless channels. This yields a spectral efficiency proportional to the number of multiplexed data streams at a high SNR [11]. On the other hand, RF chains contain some of the heavily energy-consuming components in a transmission system (e.g., amplifiers, frequency synthesizers, mixers, etc) [12], [13]. Using multiple RF chains substantially increases the total power consumption, hence degrading energy efficiency of the mmWave transmitter. Note that spectral and energy efficiencies are also coupled with other system parameters, such as wireless channel, Signal-to-Noise Ratio (SNR), and antenna arrays. As will be shown in the following sections, these two performance metrics display distinct behaviors under different system configurations, and optimal configuration of the RF chains has to be adapted accordingly. In order to optimize spectral efficiency and/or energy efficiency for specific channel conditions, we consider a configurable hybrid precoding scheme, which flexibly chooses the number of active RF chains for various requirements of the target spectral and energy efficiencies.

The fundamental energy efficiency of a communication system is characterized in [14] as conveyed information bits per radiated energy. RF chain management further includes hardware power consumption into the system framework, and plays an important role in optimizing the power consumption of fully-digital MIMO systems [15]–[17], where the optimal number of RF chains is found via simulations or measurements. Recently, it has also

been considered in [18] and [19] for the mmWave multi-user MIMO system, where the transmitter adopts the hybrid precoding and serves multiple single-antenna receivers. In [18], the energy efficiency is maximized by optimizing the hybrid precoder for each possible number of active RF chains. In [19], it is shown that the joint optimization of RF chain configurations and the hybrid precoder is subject to non-convex l_0 -norm constraints. By using sparsity relaxation, suboptimal solution is found via the compressed sensing technique. To the best of our knowledge, the RF chain configuration algorithms, such as those in [18] and [19], rely on numerical procedures to search for the optimal number of activated RF chains. Consequently, this incurs prohibitively high computational demands due to the non-convexity of the problem. Hence, it is impractical to apply these algorithms in mobile or low-complexity computing platforms, which would require adapting to fast varying wireless channels. To overcome this, we aim to characterize the spectral and energy efficiencies of mmWave MIMO systems with closed-form expressions. On this basis, the required number of RF chains can be efficiently calculated for given spectral and energy efficiency requirements, as well as given channel conditions, which could further enable adaptive hybrid precoding to flexibly activate RF chains on-the-fly.

Given the number of active RF chains, optimization of the hybrid precoder is a challenging issue in designing the mmWave transmitter. The prevailing design principle of the hybrid precoder is to approximate the fully digital counterpart [10], [20], [21]. Due to hardware constraints of the analog phase shifters, this is in general a difficult problem as the analog precoder is subject to non-convex constant-modulus constraints [22]. To address this issue, in [23]–[25] the phases of the constant-modulus elements in the analog precoding matrix are aligned with the optimal fully digital precoding matrix, while the digital precoder is constructed to minimize the mean squared error between the hybrid precoder and the optimal one. Such construction asymptotically approximates the optimal fully digital precoder and the approximation error is upper-bounded [26]. The distance between the fully digital precoder and the hybrid precoder can be also minimized by recasting the problem into Riemannian manifold optimization [20], [27] and iterative matrix decomposition [28]. The authors in [29] and [30] solve the non-convex hybrid precoder optimization by decomposing it into a sequence of convex sub-problems. Due to non-convex constraints, the hybrid precoder optimizations in [20], [23]–[30] do not have closed-form solutions, but require iterative search algorithms.

By leveraging the directivity and spatial sparsity of the mmWave channel, optimization of the hybrid precoder can be considerably simplified. In particular, the steering vectors of the antenna array corresponding to the propagation paths of the physical channel can be used as the spanning set of the space of hybrid precoding matrices, which is also known as directional beamforming [31]. Based on the near-orthogonality of large random matrices, the optimal analog precoder asymptotically converges to the directional beamforming scheme, as the number of antennas increases [22]. The authors in [32] show that the residue error caused by the approximation [22] can be further suppressed by the linear least square estimator, implemented as the digital precoder. In [10], the hybrid precoder was optimized via the orthogonal matching pursuit, which yields the optimal linear combination of the steering vectors.

Motivated by the efficient designs of hybrid precoding as in [10], [22], [31], [32], we consider configurable hybrid precoding together with directional beamforming, where the number of active RF chains can be selected. Therein, the analog precoder is configured as the steering vectors of the antenna array, and the digital precoder is chosen to maximize the mutual information of the equivalent channel, combining the effects of the physical channel and the analog precoding. To efficiently determine the required number of active RF chains, we characterize spectral and energy efficiencies of the mmWave MIMO channel with closed-form expressions in both static and mobile communication channels. Specifically, we apply the Gaussian-Radau quadrature rule [33] to lower bound the mutual information of static mmWave channels, which is constructed with the first two moments of the equivalent channel. The derived lower bound is especially accurate in the low SNR regime. In the high SNR regime, we apply the assumption of large antenna arrays and obtain an accurate approximation of the mutual information. In addition, assuming a high-speed mobile communication scenario, a mobility-aware hybrid precoding is proposed, which leverages multiple predicted beams to compensate for performance loss caused by the movement of a communication device. This also includes deriving the corresponding mutual information of the proposed mobility-aware hybrid precoding. Based on these closed-form expressions, the number of activated RF chains can be efficiently calculated to optimize spectral and/or energy efficiencies. We have the following observations on the optimal RF chain configurations:

1. In the high SNR regime, when the propagation paths have similar path gains and when the transmit power is relatively large, there exists a certain optimal number of RF chains to be activated in order to maximize the energy efficiency. As the strength of a single path dominates the others, the optimal number of RF chains decreases and eventually becomes one (i.e., using only analog beamforming [34]). When the transmit power is small, the analog beamforming is optimal towards maximizing energy efficiency. On the other hand, activating more RF chains always improve spectral efficiency. As a result, the spectral and energy efficiencies cannot be optimized simultaneously. Therefore, depending on the performance requirement, there exists a tradeoff between the two performance metrics.
2. On the contrary, in the low SNR regime, both the spectral and energy efficiencies can be optimally achieved by activating a single RF chain with analog beamforming. As the strength of the dominant path increases, analog beamforming becomes more effective and achieves higher spectral and energy efficiencies. This is in line with the behavior of low-SNR multi-antenna systems, where the antenna array is used to exploit the diversity gain of the MIMO channels. In other words, hybrid precoding reduces to a simple analog beamforming at a low SNR, which greatly simplifies configuration of the mmWave transmitter.
3. In a high-speed communication scenario, the proposed mobility-aware hybrid precoding can effectively improve the spectral efficiency of the communications as compared to the case without mobility-aware design. Numerical results show that mobility-aware hybrid precoding can yield a nearly constant information rate

when the transmitter moves at 120 km/h. Indeed, the performance improvement is due to activating additional RF chains and beamforming in some predicted directions of future steering vectors to be seen by the moving device. Therefore, the spectral efficiency increases at the expense of greater power being consumed by the transceiver. A trade-off between spectral and energy efficiencies can be established to configure the mobility-aware hybrid precoding, which is also characterized by closed-form expressions.

The rest of the paper is organized as follows: Section II outlines the mmWave channel model, the configurable hybrid precoding scheme, and the power model of transceivers. In Sections III and IV, spectral and energy efficiencies of the mmWave MIMO systems are characterized in high and low SNR regimes, respectively. In Section V, the mobility-aware hybrid precoding is presented and the corresponding performance metrics are derived. Numerical results are provided in Section VI. We conclude the main findings of this paper in Section VII.

We use the following notations throughout the paper: Bold lower-case letter $\mathbf{a}$ denotes a vector and bold upper-case letter $\mathbf{A}$ denotes a matrix. The element of $\mathbf{A}$ on the i -th row and j -th column is represented as $A_{ij} = \mathbf{A}(i, j)$, and the sub-matrix formed by columns of $\mathbf{A}$ is denoted as $\mathbf{A}[a_1, \dots, a_n]$, where $a_1, \dots, a_n$ are the indexes of the columns. $\text{Tr}(\mathbf{A})$, $\mathbf{A}^T$, and $\mathbf{A}^\dagger$ are the trace, transpose, and the conjugate transpose of $\mathbf{A}$. We use notation $\mathcal{CN}(\mu, \Sigma)$ to denote a complex circularly symmetric Gaussian random vector with mean μ and covariance matrix Σ , and $\mathbf{I}_n$ denotes an $n \times n$ identity matrix.

II. System Model

Consider mmWave communication with n_T transmit antennas and n_R receive antennas. The transfer function between the transmitter and the receiver is denoted as

$$\mathbf{y} = \sqrt{g}\mathbf{H}\mathbf{x} + \mathbf{n}, \quad (1)$$

where the complex vectors $\mathbf{x} \in \mathbb{C}^{n_T \times 1}$ and $\mathbf{y} \in \mathbb{C}^{n_R \times 1}$ are the transmit and receive signals, respectively. The additive thermal noise $\mathbf{n}$ is modeled as the complex white Gaussian vector with variance P_n , i.e., $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, P_n \mathbf{I}_{m_R})$. The noise power is calculated as $P_n = BN_0$, where B denotes the communication bandwidth and N_0 denotes the power spectral density of the thermal noise. The average channel gain, including the effect of the distance-dependent path loss, is denoted as g .

A. Millimeter Wave Channel Model

Channel $\mathbf{H}$ admits the geometrical channel model with L propagation paths [35] and is given as:

$$\begin{aligned}\mathbf{H} &= \sqrt{n_T n_R} \sum_{l=1}^L \psi_l \mathbf{a}_R(\theta_l) \mathbf{a}_T(\phi_l)^\dagger \quad (2) \\ &= \sqrt{n_T n_R} \mathbf{A}_R \mathbf{\Psi} \mathbf{A}_T^\dagger,\end{aligned}$$

where ψ_l denotes the complex channel gain of the l -th propagation path.¹ Accordingly, θ_l and ϕ_l are the angle-of-arrival and the angle-of-departure corresponding to the l -th path. The antenna array responses at the receiver and transmitter are denoted as vectors $\mathbf{a}_R(\theta_l)$ and $\mathbf{a}_T(\phi_l)$, respectively. In (2), $\mathbf{\Psi} = \text{diag}(\psi_1, \dots, \psi_L)$ is a diagonal matrix with the l -th diagonal entry being ψ_l , the matrices $\mathbf{A}_R = [\mathbf{a}_R(\theta_1), \dots, \mathbf{a}_R(\theta_L)]$ and $\mathbf{A}_T = [\mathbf{a}_T(\phi_1), \dots, \mathbf{a}_T(\phi_L)]$.

The complex channel gains $\boldsymbol{\psi} = [\psi_1, \dots, \psi_L]^\text{T}$ are modeled as independent complex Gaussian random variables $\boldsymbol{\psi} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R})$, where $\mathbf{R} = \mathbb{E}[\boldsymbol{\psi} \boldsymbol{\psi}^\dagger] = \text{diag}(r_1, \dots, r_L)$. Here, r_l denotes the fraction of the average received power contributed by the l -th propagation path, and $\{r_l\}_{l=1}^L$ are normalized as $\sum_{l=1}^L r_l = 1$. Without loss of generality, $\{r_l\}_{l=1}^L$ are arranged to be ordered:

$$r_1 \geq \dots \geq r_L. \quad (3)$$

The angle-of-arrivals $\{\theta_l\}_{l=1}^L$ and the angle-of-departures $\{\phi_l\}_{l=1}^L$ are independently and uniformly distributed within the interval $[-\pi, \pi)$. The array responses depend on the geometry of the antenna array. In particular, we consider the Uniform Linear Arrays (ULAs) and their array responses are given by

$$\mathbf{a}_R(\theta_l) = \frac{1}{\sqrt{n_R}} \left[1, e^{i\theta_l}, \dots, e^{i(n_R-1)\theta_l} \right]^\text{T}, \quad (4)$$

$$\mathbf{a}_T(\phi_l) = \frac{1}{\sqrt{n_T}} \left[1, e^{i\phi_l}, \dots, e^{i(n_T-1)\phi_l} \right]^\text{T}. \quad (5)$$

Note that matrices $\mathbf{A}_T$ and $\mathbf{A}_R$ are in the form of rectangular Vandermonde matrices [36]. Channel $\mathbf{H}$ is normalized as $\mathbb{E}[\text{Tr}(\mathbf{H}\mathbf{H}^\dagger)] = n_T n_R$.

¹In [35, Eq. (4)], a prefactor $1/\sqrt{L}$ is used to normalize the complex channel gains $\{\psi_l\}_{l=1}^L$.

B. Configurable Hybrid Precoder

The traditional multi-antenna transmitter with fully digital transceivers is equipped with n_T RF chains, and each is connected to one individual antenna element. However, the antenna arrays of typical mmWave transmitters usually have hundreds of antenna elements. Due to high energy consumption and hardware cost of RF chains, it is difficult to be realized as fully digital architecture. Instead, the mmWave transmitter adopts the hybrid analog and digital architecture, and signal processing is separately implemented as the analog and digital precoding.

In the considered hybrid architecture, the transmitter is equipped with $m_{\max}$ RF chains, $m_{\max} \leq n_T$. Each RF chain can be configured as activated or deactivated, and the inactive RF chain will be turned off or put into sleep mode to reduce the total power consumption. As will be shown in (12) and (16), these performance metrics of the considered hybrid precoding scheme heavily depend on the number of activated RF chains, which can be flexibly configured to meet the targeting spectral and energy efficiency requirements.

Given $m \leq m_{\max}$ RF chains activated, the data streams are digitally precoded and fed to the transmit antennas via m RF chains. In order to perform analog precoding, the outputs of the m RF chains undergo a network of analog phase shifters before feeding to the antenna elements. We consider fully-connected phase shifters, where the analog signal between each pair of RF chain and antenna element can be individually phase-shifted. In this case, the signal vector, after analog and digital precoding, is given by

$$\mathbf{x} = \mathbf{F}_A \mathbf{F}_D \mathbf{z}, \quad (6)$$

where $\mathbf{z} \in \mathbb{C}^s$ denotes the information symbols from the s data streams, $s \leq m$. We adopt the Gaussian signaling such that $\mathbf{z}$ is a complex Gaussian random vector with $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, P_t/s \mathbf{I}_s)$, where P_t is the transmit power. Matrix $\mathbf{F}_D \in \mathbb{C}^{m \times s}$ denotes the digital precoder, which maps the s data streams to the m RF chains. Matrix $\mathbf{F}_A \in \mathbb{C}^{n_T \times m}$ denotes the phase shifts of the incoming data streams and each entry of $\mathbf{F}_A$ satisfies

$$|\mathbf{F}_A(i, j)|^2 = n_T^{-1}, \quad 1 \leq i \leq n_T, \quad 1 \leq j \leq m, \quad (7)$$

where $|\cdot|$ denotes the modulus of complex number. The hybrid analog and digital precoder is normalized as

$$\|\mathbf{F}_A \mathbf{F}_D\|_F^2 = s, \quad (8)$$

where $\|\mathbf{A}\|_F = \sqrt{\text{Tr}(\mathbf{A}\mathbf{A}^\dagger)}$ denotes the Frobenius norm of a matrix $\mathbf{A}$.

Define the Singular Value Decomposition (SVD) of $\mathbf{H}$ as $\mathbf{H} = \mathbf{U}^{1/2} \mathbf{V}^\dagger$ and the SVD of the matrix $\mathbf{V}^{1/2} \mathbf{V}^\dagger \mathbf{F}_A (\mathbf{F}_A^\dagger \mathbf{F}_A)^{-1/2} = \overline{\mathbf{U}} \overline{\mathbf{\Sigma}}^{1/2} \overline{\mathbf{V}}^\dagger$. In [37], given a certain analog precoder $\mathbf{F}_A$, the corresponding optimal digital precoder is expressed as

$$\mathbf{F}_D^{\text{opt}} = (\mathbf{F}_A^\dagger \mathbf{F}_A)^{-1/2} \overline{\mathbf{V}}[1, \dots, s] \mathbf{\Lambda}^{1/2}, \quad (9)$$

where $(\cdot)^{1/2}$ denotes the square root of matrix and $\overline{\mathbf{V}}[1, \dots, s]$ denotes the first s columns of $\overline{\mathbf{V}}$. The $s \times s$ diagonal matrix $\mathbf{\Lambda}$ is obtained via the water-filling algorithm across the eigen-channels $\{\overline{\Sigma}(i, i)\}_{1 \leq i \leq s}$ with the constraint $\sum_{i=1}^s \Lambda(i, i) \leq s$.

On the other hand, the design of the optimal analog precoder $\mathbf{F}_A$ is rather complicated due to the non-convex constraint (7). In this work, we apply the widely adopted directional beamforming to design the analog precoder. Specifically, given that m RF chains are activated, the analog precoder $\mathbf{F}_A$ is constructed by the steering vectors corresponding to the first m propagation paths having the strongest average path gains $r_1 \dots r_m$ i.e.,

$$\mathbf{F}_A = [\mathbf{a}_T(\phi_1), \dots, \mathbf{a}_T(\phi_m)] = \mathbf{A}_T[1, \dots, m]. \quad (10)$$

For ease of exposition, we denote $\mathbf{A}_{T1} = \mathbf{A}_T[1, \dots, m]$ and $\mathbf{A}_{T2} = \mathbf{A}_T[m+1, \dots, L]$. In other words, by choosing the analog precoder as in (10), transmissions are beamformed towards the directions of the first m strongest propagation paths, which efficiently leverages the geometrical property of the mmWave channel (2). Note that although directional beamforming is suboptimal in general, it significantly simplifies the design of an analog precoder, while becoming asymptotically optimal as the number of antennas goes to infinity [22].

To enable a tractable analysis on mmWaveMIMO systems, we adopt the following practical approximations to further simplify the structure of the hybrid precoder:

The water-filling power allocation $\mathbf{\Lambda}$ can be approximated as an identity matrix. This approximation is valid when the eigenvalues $\{\overline{\Sigma}(i, i)\}_{1 \leq i \leq s}$ are relatively large. This condition holds in typical mmWave MIMO channels with a large number of antennas and relatively small number of RF chains [30], [37].

The number of data streams equals the number of activated RF chains, i.e., $s = m$. This transmitter setting has been also utilized in [37], which uses the full multiplexing gain of the effective channel $\mathbf{H}\mathbf{F}_A$. This is optimal when the effective channel has rank m and the eigenvalues $\{\overline{\Sigma}(i, i)\}_{1 \leq i \leq m}$ are sufficiently large, as has already been assumed in (A1). This assumption also requires the number of activated RF chains does not exceed the number of propagation paths, i.e., $m \leq L$.

Inserting (10) into (9) and applying approximations (A1) and (A2), we obtain

$$\mathbf{F}_D = (\mathbf{A}_{T1}^\dagger \mathbf{A}_{T1})^{-1/2} \bar{\mathbf{V}}. \quad (11)$$

With the analog precoder as in (10) and the digital precoder as in (11), the corresponding spectral efficiency of channel (1), in term of nats/s/Hz, is given by the mutual information of the Gaussian MIMO channel [38], and expressed as

$$\begin{aligned} \mathcal{S}(\gamma) &= \mathbb{E} \left[\log \det \left(\mathbf{I}_{n_R} + \frac{\gamma}{m} \mathbf{H} \mathbf{F}_A \mathbf{F}_D \mathbf{F}_D^\dagger \mathbf{F}_A^\dagger \mathbf{H}^\dagger \right) \right] \quad (12) \\ &= \mathbb{E} \left[\log \det \left(\mathbf{I}_L + \frac{n_T n_R}{m} \gamma \mathbf{W} \right) \right], \end{aligned}$$

where $\gamma = P_t g / P_n$ denotes the Signal-to-Noise-Ratio (SNR). The matrix $\mathbf{W} = \mathbf{\Psi} \mathbf{P} \mathbf{\Psi}^\dagger \mathbf{Q}$, and

$$\mathbf{Q} = \mathbf{A}_R^\dagger \mathbf{A}_R, \quad (13)$$

$$\mathbf{P} = \mathbf{A}_T^\dagger \mathbf{A}_{T1} (\mathbf{A}_{T1}^\dagger \mathbf{A}_{T1})^{-1} \mathbf{A}_{T1}^\dagger \mathbf{A}_T. \quad (14)$$

C. Power Model

To characterize the power consumption of the transmitter using configurable hybrid precoder, we adopt the power model of the mmWave transceiver as in [27]. Therein, the total power consumed by the transceiver system is modeled as

$$P_{\text{total}} = m P_{\text{RF}} + \frac{P_t}{\alpha} + n_T m P_{\text{shift}}, \quad (15)$$

where P_{RF} and P_{shift} denote the power consumption of a single RF chain and a phase shifter, respectively. The first term of (15) is the total power consumption of the m active RF chains, which converts the digital signal into an analog signal. The analog signals are then amplified to the transmit power P_t and fed to the phase shifter network. The second term of (15) accounts for the total power consumed by the amplifiers, where α is the efficiency factor. As the fully-connected network of phase shifters is assumed for the hybrid precoder, there are in total $n_T m$ phase shifters between the RF chains and the transmit antenna array. The total power of the phase shifter network is given as the third term of (15).

The energy efficiency of the transmitter is defined as the transmitted information bits per unit energy (bits per Joule), and is given by the ratio between the achieved data rate and the total power consumption. By evaluating the communication data rate via the spectral efficiency $\mathcal{S}(\gamma)$, the energy efficiency is given by

$$\mathcal{E}(\gamma) = \frac{B\mathcal{S}(\gamma)}{\log 2 P_{\text{total}}}, \quad (16)$$

where B is the bandwidth of the communication channel. In Sections IV and III, we will derive closed-form expressions of $\mathcal{S}(\gamma)$ in the low and high SNR regimes, respectively. Based on these expressions, the behavior of spectral efficiency and energy efficiency can be understood. The optimal settings of the configurable hybrid precoder can be determined depending on the mmWave MIMO channel conditions.

III. Spectral Efficiency and Energy Efficiency in High SNR Regime

In this section, we derive the spectral efficiency and energy efficiency of mmWave MIMO assuming the received SNR is large. In the existing high-SNR analysis for MIMO capacity, prevailing techniques include the Gauthier-Grant's lower bound as in [39], [40], and Minkowski's inequality based lower bound as in [41]. Both approaches yield closed-form expressions for the mutual information of rich scattering MIMO channels, which are asymptotically tight in the high-SNR regime. However, these approaches require evaluation of $\mathbb{E}[\log \det(\mathbf{W})]$, which trivially equals zero as $\mathbf{W}$ is rank-deficient in the considered hybrid precoding scheme.

To address this issue, we utilize the large antenna array assumption to approximate spectral efficiency (12) in the high SNR regime. In particular, when the number of antennas n_T and n_R are much larger than the number of propagation paths L , the following approximations are accurate for large Vandermonde matrices [22]:

$$\begin{aligned} \mathbf{Q} &= \mathbf{A}_R^\dagger \mathbf{A}_R \approx \mathbf{I}_L, \mathbf{A}_{T1}^\dagger \mathbf{A}_{T1} \approx \mathbf{I}_m, \\ \mathbf{A}_{T1}^\dagger \mathbf{A}_{T2} &\approx \mathbf{0}_{m \times (L-m)}. \end{aligned} \quad (17)$$

Note that we do not directly apply the approximations (17) in (13) and (14), which yields an overly simplified approximation and does not reveal the impacts due to steering matrices $\mathbf{A}_T$ and $\mathbf{A}_R$. Instead, we use (17) in a way that the approximation error due to (17) is sufficiently small as the SNR $\gamma \rightarrow \infty$, while the expression of spectral efficiency $\mathcal{S}(\gamma)$ still depends on $\mathbf{A}_T$ and $\mathbf{A}_R$.

For ease of exposition, we denote

$$\rho = \frac{n_T n_R}{m} \gamma \quad (18)$$

and spectral efficiency $\mathcal{S}(\gamma)$ can be approximated as

$$\begin{aligned}\mathcal{S}(\gamma) &= \mathbb{E} \left[\log \det(\mathbf{I}_L - \mathbf{Q} + \mathbf{Q} + \rho \mathbf{\Psi} \mathbf{P} \mathbf{\Psi}^\dagger \mathbf{Q}) \right] \quad (19) \\ &\approx \mathcal{V}_R + \mathbb{E} \left[\log \det(\mathbf{I}_L + \rho \mathbf{\Psi} \mathbf{P} \mathbf{\Psi}^\dagger) \right],\end{aligned}$$

where $\mathcal{V}_R = \mathbb{E}[\log \det(\mathbf{Q})]$. The expression (19) is obtained by using the approximation $\mathbf{Q} \approx \mathbf{I}_L$ and therefore, $\mathbf{I}_L - \mathbf{Q}$ is sufficiently small compared to $\mathbf{Q} + \rho \mathbf{\Psi} \mathbf{P} \mathbf{\Psi}^\dagger \mathbf{Q}$ assuming large SNR, a.k.a. large ρ . Let $\mathbf{\Psi}_1$ be the upper-left $m \times m$ submatrix of $\mathbf{\Psi}$ and $\mathbf{\Psi}_2$ the lower-right $(L - m) \times (L - m)$ submatrix. Substituting $\mathbf{P}$ with (14), the second term on the Right-Hand-Side (RHS) of (19) can be rewritten as in (20)–(22), shown at the bottom of this page, where (21) is due to the determinant identity [42] for a block matrix and is given in (22). Next, we

$$\begin{aligned}\mathbb{E} \left[\log \det(\mathbf{I}_L + \rho \mathbf{\Psi} \mathbf{P} \mathbf{\Psi}^\dagger) \right] &= \mathbb{E} \left[\log \det \right. \quad (20) \\ &\quad \left. \begin{bmatrix} \mathbf{I}_m + \rho \mathbf{\Psi}_1 \mathbf{A}_{T1}^\dagger \mathbf{A}_{T1} \mathbf{\Psi}_1^\dagger & \rho \mathbf{\Psi}_1 \mathbf{A}_{T1}^\dagger \mathbf{A}_{T2} \mathbf{\Psi}_2^\dagger \\ \rho \mathbf{\Psi}_2 \mathbf{A}_{T2}^\dagger \mathbf{A}_{T1} \mathbf{\Psi}_1^\dagger & \mathbf{I}_{L-m} + \rho \mathbf{\Psi}_2 \mathbf{A}_{T2}^\dagger \mathbf{A}_{T1} \mathbf{A}_{T1}^\dagger \mathbf{A}_{T1} \mathbf{\Psi}_2^\dagger \end{bmatrix} \right]\end{aligned}$$

$$= \mathbb{E} \left[\log \det(\mathbf{I}_m + \rho \mathbf{\Psi}_1 \mathbf{A}_{T1}^\dagger \mathbf{A}_{T1} \mathbf{\Psi}_1^\dagger) \right] + \mathbb{E} \left[\log \det(\mathbf{I}_{L-m} + \Delta) \right] \quad (21)$$

$$\begin{aligned}\Delta &= \rho \mathbf{\Psi}_2 \mathbf{A}_{T2}^\dagger \mathbf{A}_{T1} (\mathbf{A}_{T1}^\dagger \mathbf{A}_{T1})^{-1} \mathbf{A}_{T1}^\dagger \mathbf{A}_{T2} \mathbf{\Psi}_2^\dagger - \rho^2 \mathbf{\Psi}_2 \mathbf{A}_{T2}^\dagger \mathbf{A}_{T1} \mathbf{\Psi}_1^\dagger (\mathbf{I}_m + \rho \mathbf{\Psi}_1 \mathbf{A}_{T1}^\dagger \mathbf{A}_{T1} \mathbf{\Psi}_1^\dagger)^{-1} \quad (22) \\ &\quad \times \mathbf{\Psi}_1 \mathbf{A}_{T1}^\dagger \mathbf{A}_{T2} \mathbf{\Psi}_2^\dagger\end{aligned}$$

show that $\log \det(\mathbf{I}_{L-m} + \Delta)$ in (21) is sufficiently small compared to other terms and can be ignored. In the high SNR regime with large ρ , $(\mathbf{I}_m + \rho \mathbf{\Psi}_1 \mathbf{A}_{T1}^\dagger \mathbf{A}_{T1} \mathbf{\Psi}_1^\dagger)^{-1}$ can be expanded as a power series in terms of ρ^{-1} as

$$(\rho \mathbf{\Psi}_1 \mathbf{A}_{T1}^\dagger \mathbf{A}_{T1} \mathbf{\Psi}_1^\dagger)^{-1} + (\rho \mathbf{\Psi}_1 \mathbf{A}_{T1}^\dagger \mathbf{A}_{T1} \mathbf{\Psi}_1^\dagger)^{-2} + \mathcal{O}(\rho^{-3}). \quad (23)$$

Inserting (23) into (22), is expressed as

$$\mathbf{\Psi}_2 \mathbf{A}_{T2}^\dagger \mathbf{A}_{T1} \left(\mathbf{\Psi}_1 (\mathbf{A}_{T1}^\dagger \mathbf{A}_{T1})^2 \mathbf{\Psi}_1^\dagger \right)^{-1} \mathbf{A}_{T1}^\dagger \mathbf{A}_{T2} \mathbf{\Psi}_2^\dagger + \mathcal{O}(\rho^{-1}).$$

By using approximations $\mathbf{A}_{T1}^\dagger \mathbf{A}_{T1} \approx \mathbf{I}_m$ and $\mathbf{A}_{T1}^\dagger \mathbf{A}_{T2} \approx \mathbf{0}_{m \times (L-m)}$, the elements of the matrix , and eventually $\log \det(\mathbf{I}_{L-m} + \dots)$ become vanishingly small as $\rho \rightarrow \infty$. Finally, we approximate the first term of (21) as

$$\mathbb{E}[\log \det(\mathbf{I}_m + \rho \mathbf{\Psi}_1 \mathbf{A}_{T1}^\dagger \mathbf{A}_{T1} \mathbf{\Psi}_1^\dagger)] \approx \mathcal{V}_T + \mathbb{E}[\log \det(\mathbf{I}_m + \rho \mathbf{\Psi}_1^\dagger \mathbf{\Psi}_1)], \quad (24)$$

where $\mathcal{V}_T = \mathbb{E}[\log \det(\mathbf{A}_{T1}^\dagger \mathbf{A}_{T1})]$ and we have applied $\mathbf{A}_{T1}^\dagger \mathbf{A}_{T1} \approx \mathbf{I}_m$.

To sum up steps (19)–(24), spectral efficiency $\mathcal{S}(\gamma)$ is now approximated in a favorable form in the high-SNR regime as

$$\mathcal{S}(\gamma) \approx \mathcal{V}_\Psi + \mathcal{V}_R + \mathcal{V}_T, \quad (25)$$

where $\mathcal{V}_\Psi = \mathbb{E}[\log \det(\mathbf{I}_m + \rho \mathbf{\Psi}_1^\dagger \mathbf{\Psi}_1)]$. The approximation (25) decouples channel amplitudes $\mathbf{\Psi}$, and the steering matrices $\mathbf{A}_R$ and $\mathbf{A}_{T1}$ when calculating spectral efficiency $\mathcal{S}(\gamma)$, where each part can be obtained analytically.

First, we notice that $\mathcal{V}_\Psi$ is the mutual information of m parallel channels, with complex channel amplitudes $\{\sqrt{\rho}\psi_i\}_{1 \leq i \leq m}$. Recalling that the complex channel $\psi_i \sim \mathcal{CN}(0, r_i)$, we have

$$\mathcal{V}_\Psi = \sum_{i=1}^m \mathbb{E}[\log(1 + \rho|\psi_i|^2)] \quad (26)$$

$$= \sum_{i=1}^m \frac{1}{r_i} \int_0^\infty \log(1 + \rho x) e^{-\frac{x}{r_i}} dx$$

$$= \sum_{i=1}^m \exp\left(\frac{1}{r_i \rho}\right) E_1\left(\frac{1}{r_i \rho}\right), \quad (27)$$

where $E_1(z) = \int_z^\infty t^{-1} e^{-t} dt$ denotes the exponential integral [43, pp. XXXV] and the last equality is due to [43, Eq. (4.337.2)]. Next, we obtain an asymptotic approximation of $\mathcal{V}_R$ by using the well-known matrix identity $\log \det(\mathbf{D}) = \text{Tr} \log(\mathbf{D})$ for square matrix $\mathbf{D}$. In particular, $\mathcal{V}_R$ can be immediately rewritten as

$$\mathcal{V}_R = \mathbb{E}[\text{Tr} \log(\mathbf{A}_R^\dagger \mathbf{A}_R)]. \quad (28)$$

The RHS of (28) can be expanded as a Maclaurin series in terms of $\mathbf{A}_R^\dagger \mathbf{A}_R - \mathbf{I}_L$ as

$$\begin{aligned} \mathcal{V}_R &= \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \mathbb{E} \left[\text{Tr} \left((\mathbf{A}_R^\dagger \mathbf{A}_R - \mathbf{I}_L)^k \right) \right] \quad (29) \\ &= L \sum_{k=1}^{\infty} \sum_{j=0}^k \frac{(-1)^{j+1}}{k} \binom{k}{j} \mu_j^{\mathbf{A}}, \end{aligned}$$

where $\mu_j^{\mathbf{A}} = \frac{1}{L} \mathbb{E} \left[\text{Tr} \left((\mathbf{A}_R^\dagger \mathbf{A}_R)^j \right) \right]$ and the second equality is due to the binomial expansion. The normalized moment $\mu_j^{\mathbf{A}}$ is rather complicated as its computation involves the so-called sum over set partitions [36]. It is a combinatorial object that requires one to enumerate a total of B_{nR} summands [44], where B_n denotes the Bell number corresponding to a finite set with n elements and it grows superexponentially as n increases. However, as shown in [36, Thm. 2], the first 4 moments $\mu_1^{\mathbf{A}}, \dots, \mu_4^{\mathbf{A}}$ can be exactly calculated as follows:

$$\mu_1^{\mathbf{A}} = 1 \quad (30)$$

$$\mu_2^{\mathbf{A}} = 1 + \frac{L-1}{n_R}$$

$$\mu_3^{\mathbf{A}} = 1 - \frac{3}{n_R} + \frac{2}{n_R^2} + \left(1 - \frac{1}{n_R}\right) \frac{3L}{n_R} + \left(\frac{L}{n_R}\right)^2$$

$$\mu_4^{\mathbf{A}} = 1 - \frac{20}{3n_R} + \frac{12}{n_R^2} - \frac{19}{3n_R^3} + \left(\frac{20}{3} - \frac{18}{n_R} + \frac{34}{3n_R^2}\right) \frac{L}{n_R} + 6 \left(1 - \frac{1}{n_R}\right) \left(\frac{L}{n_R}\right)^2 + \left(\frac{L}{n_R}\right)^3.$$

Since $\mathbf{A}_R^\dagger \mathbf{A}_R \approx \mathbf{I}_L$, the series expansion (29) converges fast and we truncate the infinite summation at $k=4$, which only requires the first 4 moments given by (30). Similarly, we can also obtain the asymptotic approximation of $\mathcal{V}_T$ as in (29)–(30) by replacing n_R with n_T and L with m , respectively.

IV. Spectral Efficiency and Energy Efficiency in Low SNR Regime

In this section, we will derive the lower bound for spectral efficiency $\mathcal{S}(\gamma)$ and energy efficiency $\mathcal{E}(\gamma)$ in low SNR regime. These performance bounds are obtained based on the Gauss-Radau quadrature rule [33], which requires the first few moments of the random matrix $\mathbf{W}$ [45], defined as $\mu_i^{\mathbf{W}} = \mathbb{E} [\text{Tr}(\mathbf{W}^i)]$, $i = 1$. The simplest form of the Gauss-Radau quadrature rule involves the first two moments $\mu_1^{\mathbf{W}}$ and $\mu_2^{\mathbf{W}}$. Recall that $\{r_i\}_{i=1}^L$ denote the average gains of the L propagation paths. In the next proposition, we derive the exact expression of $\mu_1^{\mathbf{W}}$, and the asymptotic expression of $\mu_2^{\mathbf{W}}$ as the number of antennas n_T and n_R approaches infinity.

Proposition 1: The expectation $\mu_1^{\mathbf{W}} = \mathbb{E}[\text{Tr}(\mathbf{W})]$ is given by

$$\mu_1^{\mathbf{W}} = \sum_{a=1}^m r_a + \frac{m}{n_T} \sum_{a=m+1}^L r_a. \quad (31)$$

As n_T and n_R approach infinity, the asymptotic expression of $\mu_2^{\mathbf{W}}$ is given by

$$\begin{aligned} \mu_2^{\mathbf{W}} = & 2 \sum_{a=1}^m r_a^2 + \left(\frac{1}{n_R} + \frac{1}{n_T} \right) \sum_{a_1=1}^m \sum_{\substack{a_2=1 \\ a_2 \neq a_1}}^m r_{a_1} r_{a_2} + \frac{2}{n_T} \sum_{a_1=1}^m \sum_{a_2=m+1}^L r_{a_1} r_{a_2} \\ & + \mathcal{O} \left(\frac{1}{n_T n_R} + \frac{1}{n_T^2} + \frac{1}{n_R^2} \right). \end{aligned} \quad (32)$$

Proof: The proof of Proposition 1 is in Appendix A. ■

Using $\mu_1^{\mathbf{W}}$ and $\mu_2^{\mathbf{W}}$ calculated in Proposition 1, the lower bound of spectral efficiency $\mathcal{S}(\gamma)$ can be derived using the Gauss-Radau quadrature rule and is presented in the next proposition.

Proposition 2: The spectral efficiency is lower bounded as

$$\mathcal{S}(\gamma) \geq \mathcal{S}_{\text{LB}}(\gamma) = G_{\text{mx}} \log(1 + G_{\text{array}} \gamma), \quad (33)$$

where G_{array} and G_{mx} are the array gain and the spatial-multiplexing gain of MIMO systems and are given as

$$G_{\text{array}} = \frac{n_T n_R \mu_2^{\mathbf{W}}}{m \mu_1^{\mathbf{W}}}, \quad (34)$$

$$G_{\text{mx}} = \frac{(\mu_1^{\mathbf{W}})^2}{\mu_2^{\mathbf{W}}}. \quad (35)$$

Proof: The proof of Proposition 2 is a direct application of Bai and Golub's inequality [46, Eq. (10)]. Therein, a lower bound λ_{LB} of the smallest eigenvalue of matrix $\mathbf{I}_L + (n_R n_R / m) \gamma \mathbf{W}$ is needed. A choice of such a lower bound is $\lambda_{\text{LB}} = 1$. ■

As will be shown in Section VI, the lower bound given in (33) is especially tight when the SNR is relatively small. In (33), G_{mx} and G_{array} can have a significant impact on the performance of MIMO systems [47] as both are functions of the number of activated RF chains, m . In generic system settings, an adequate number of RF chains can be determined by evaluating (33)–(35) for a range of m , which can be done efficiently by using these closed-form expressions. In addition, for some specific channels, as will be discussed below, (34) and (35) provide intuitive indications on the adequate value of m . Recall that the average gains of propagation paths are normalized as $\sum_{i=1}^L r_i = 1$. We consider a channel structure having one path with channel gain r_1 and $L - 1$ paths with common channel gain $r_2 = \dots = r_L = r_c = (1 - r_1)/(L - 1)$. When $r_1 = r_c$, i.e. all channels have the same gain, G_{mx} and G_{array} can be approximated as

$$G_{\text{mx}} = \frac{(\mu_1^{\mathbf{W}})^2}{\mu_2^{\mathbf{W}}} \approx \frac{m^2 r_c^2}{2mr_c} = \frac{m}{2}, \quad (36)$$

$$G_{\text{array}} = \frac{n_T n_R}{m} \frac{\mu_2^{\mathbf{W}}}{\mu_1^{\mathbf{W}}} \approx 2 \frac{n_T n_R}{m} r_c, \quad (37)$$

where we have applied the limits $n_T, n_R \rightarrow \infty$ to obtain the RHS of (36) and (37). Equation (36) shows that the spectral efficiency may be improved by increasing the number of RF chains as the multiplexing gain G_{mx} is linearly proportional to m . However, the array gain G_{array} is inverse proportional to m , which reduces the effective SNR as m increases. As a result, when the number of activated RF chains is relatively large, the improvement of spectral efficiency is only marginal while severely impacting the power consumption.

When $r_c \ll r_1 \approx 1$, i.e., there is a dominant propagation path, G_{mx} is approximately calculated as

$$\begin{aligned} G_{\text{mx}} &\approx \frac{(r_1 + (m-1)r_c)^2}{2r_1^2 + 2(m-1)r_c^2} \\ &= \frac{r_1^2 + 2(m-1)r_1 r_c + (m-1)^2 r_c^2}{2r_1^2 + 2(m-1)r_c^2} \\ &\approx \frac{1}{2} + \frac{m-1}{L-1} \frac{1-r_1}{r_1}, \end{aligned} \quad (38)$$

where the second approximation is obtained by ignoring the small r_c^2 terms. Under the same condition, the array gain G_{array} is calculated as

$$G_{\text{array}} \approx \frac{n_T n_R}{m} \frac{2r_1^2 + 2(m-1)r_c^2}{r_1 + (m-1)r_c} \quad (39)$$

$$\approx 2r_1 \frac{n_T n_R}{m} \left(1 - (m-1) \frac{r_c}{r_1}\right),$$

where the first approximation is obtained by letting $n_T, n_R \rightarrow \infty$, and the second approximation is obtained by ignoring the r_c^2 term and using the Taylor expansion of $1/(1 + (m-1)r_c/r_1)$. In this case, as $r_1 \approx 1$, the multiplexing gain G_{mx} in (38) is approximately 1/2 regardless of the number of activated RF chains. However, the array gain G_{array} in (39) maximizes when $m = 1$. Therefore, when there is a dominant propagation path with $r_1 \approx 1$, it is optimal to activate a single RF chains.

The Gauss-Radau quadrature rule can be extended to incorporate higher order moments of $\mathbf{W}$, which could improve accuracy of the lower bound (33). However, the derivations of $\mu_i^{\mathbf{W}}$, $i \geq 3$, involve complicated combinatorial structures due to the Gaussian moment theorem [48]. We leave studies of higher order moments of $\mathbf{W}$ and the corresponding improved lower bound of the spectral efficiency to future works. By substituting $\mathcal{S}_{\text{LB}}(\gamma)$ into (16), the energy efficiency is lower bounded as

$$\mathcal{E}(\gamma) \geq \mathcal{E}_{\text{LB}}(\gamma) = \frac{B \mathcal{S}_{\text{LB}}(\gamma)}{\log 2 P_{\text{total}}}. \quad (40)$$

V. Spectral Efficiency and Energy Efficiency of Mobility-Aware Hybrid Precoding

In this section, we apply the directional beamforming in mobile communications, where the transmitter is a moving vehicle, while the propagation environment and the receiver are fixed. In this circumstance, the steering matrix $\mathbf{A}_T$ of the channel (2) changes with the movement of the vehicle and therefore depends on time, speed, and the moving direction. To illustrate such dependence, with a bit abuse of the notations in Section II, we first present the time-varying channel model $\mathbf{H}[t]$. Then, we apply a Mobility-Aware Beamforming (MAB) technique to construct the hybrid precoder at the transmitter, which sends information symbols in the directions of predicted steering vectors as the channel $\mathbf{H}[t]$ changes. In particular, we hereafter adopt an important assumption that the hybrid precoder is updated periodically. Accordingly, the design of the hybrid precoder takes into account the variations of the channel between two consecutive update events. This assumption is relevant for practical mobile communication systems, where the update interval is selected to trade off between performance and operational complexity. For example, the hybrid precoder may be assumed to be updated at the beginning of an LTE frame, which spans 10 milliseconds in time domain [49].

As shown in Fig. 1, we consider a single-path propagation channel with a scatterer located at the origin of the Euclidean plane. The antenna elements of an ULA transmitter are aligned along the antenna axis, which has angle φ_A relative to the horizontal axis. The transmitter has initial location $\xi[0]$, moves at a velocity $\mathbf{v}$, and arrives at $\xi[t] = \xi[0] + \mathbf{v}t$ after time t , where $0 \leq t \leq t_{\text{update}}$ and t_{update} denotes the time duration between two updates of the hybrid precoder. The physical angle between the antenna axis and the vector $\xi[t]$ is calculated as $\phi_P[t] = \phi_A - \phi_\xi[t] + \pi$, where $\phi_\xi[t]$ denotes the angle of the vector $\xi[t]$. By applying [47, Eq. (7.25)], the steering vector of transmit array at time t is written as

$$\mathbf{a}_T(\phi[t]) = \frac{1}{\sqrt{n_T}} \begin{bmatrix} 1 \\ e^{i\pi \cos(\varphi_A - \varphi_\xi[t])} \\ \vdots \\ e^{i\pi(n_T-1)\cos(\varphi_A - \varphi_\xi[t])} \end{bmatrix}, \quad (41)$$

where $\phi[t] = \pi \cos(\varphi_A - \varphi_\xi[t])$ by comparing the RHS of (41) with (5). Since the scatterer and the receiver are assumed to be fixed, the steering vectors of the receiver array is not time-dependent and therefore, the single-path channel can be written as²

$$\mathbf{H}[t] = \sqrt{n_T n_R} \psi \mathbf{a}_R(\theta) \mathbf{a}_T(\phi[t])^\dagger. \quad (42)$$

As discussed above, the hybrid precoder is presumably updated at the beginning of every transmission frames, while keeps fixed for the current frame. The analog precoder chosen for one instantaneous steering vector, such as the one given by (10), causes significant performance degradation when the steering vector mismatches the analog precoder due to mobility of the transmitter. To address this issue, we hereafter propose the MAB scheme, where the current and some of the predicted steering vectors to be seen by the transmit array are chosen as the analog precoder. As shown in Fig. 2, the transmitter sends a direct beam to the instantaneous direction of the scatterer and three predicted beams. As the transmitter moves upwards, the scatterer will be illuminated by these partially overlapped beams, and the total beamwidth of the transmission is increased compared to a single beam. Of course, the number of beams should be properly adjusted to trade off between the spectral and energy efficiencies as the power consumption of the transceiver is also increased due to the additional beams and RF chains. Next, given a certain number of beams, we illustrate how to choose the directions of the predicted beams to avoid fluctuations in the received SNR. Then, we derive the closed-form expression of the corresponding spectral efficiency, where the MAB is applied to construct the hybrid precoder. The closed-form expression facilitates to determine the numbers of beams and active RF chains, which optimally trades off the spectral and energy efficiencies.

²The propagation delay and the Doppler effect due to the movement cause phase changes of the received signal, which are included in the complex channel gain ψ .

Let $\mathbf{f}_{A,0} = \mathbf{a}_T(\phi[0])$ denote the direct beam and $\{\mathbf{f}_{A,k}\}_{k=1}^{m-1}$ denote the $m-1$ predicted beams. The analog precoder applying the MAB is constructed as

$$\mathbf{F}_A^{\text{MAB}} = [\mathbf{f}_{A,0}, \dots, \mathbf{f}_{A,m-1}]. \quad (43)$$

Similar to (11), the corresponding optimal digital precoder is selected as

$$\mathbf{F}_D^{\text{MAB}} = \left((\mathbf{F}_A^{\text{MAB}})^\dagger \mathbf{F}_A^{\text{MAB}} \right)^{-1/2} \mathbf{V}, \quad (44)$$

which shows that the number of activated RF chains is the same as the number of transmitted beams. Consider the projection of an arbitrary steering vector $\mathbf{f} = \mathbf{a}_T(\phi)$ onto $\mathbf{f}_{A,0}$ calculated as

$$|\mathbf{f}^\dagger \mathbf{f}_{A,0}| = \left| \frac{1}{n_T} \sum_{i=1}^{n_T} e^{i(c-1)(\phi[0] - \phi)} \right| = \left| \frac{\sin\left(\frac{n_T}{2}(\phi[0] - \phi)\right)}{n_T \sin\left(\frac{1}{2}(\phi[0] - \phi)\right)} \right| \stackrel{n_T \rightarrow \infty}{=} \left| \text{sinc}\left(\frac{n_T}{2\pi}(\phi[0] - \phi)\right) \right|, \quad (45)$$

where $\text{sinc}(x) = \sin(\pi x)/(\pi x)$ is the sinc function and the last equality is obtained by taking the limit $n_T \rightarrow \infty$ using [47, Eq. (7.41)]. As $\mathbf{f}_{A,0} = \mathbf{a}_T(\phi[0])$ is the steering vector of the channel at time $t=0$, the projection (45) represents how the beamforming gain degrades as the direction-of-departure ϕ changes due to, e.g., mobility of the transmitter. In particular, we notice that the sinc function becomes zero, when

$$\phi = \phi[0] - \frac{2\pi l}{n_T}, \quad (46)$$

where l is positive integer. In other words, when the direction-of-departure ϕ changes to the RHS of (46), the beamforming gain vanishes. In order to overcome this problem, the predicted beams are transmitted in the first few directions of (46), i.e.,

$$\mathbf{f}_{A,1} = \mathbf{a}_T\left(\phi[0] - \frac{2\pi}{n_T}\right), \dots, \mathbf{f}_{A,m-1} = \mathbf{a}_T\left(\phi[0] - \frac{2\pi(m-1)}{n_T}\right),$$

which compensates the performance loss when ϕ changes from $\phi[0]$ to $\phi[0] - \frac{2\pi(m-1)}{n_T}$.

By applying the analog precoder $\mathbf{F}_A^{\text{MAB}}$ as in (43) and the digital precoder $\mathbf{F}_D^{\text{MAB}}$ as in (44), the spectral efficiency is calculated as

$$\begin{aligned}
\mathcal{S}(\gamma) &= \mathbb{E} \left[\log \det \left(\mathbf{I}_{n_R} + \frac{\gamma}{m} \mathbf{H}[t] \mathbf{F}_A^{\text{MAB}} \mathbf{F}_D^{\text{MAB}} (\mathbf{F}_D^{\text{MAB}})^\dagger (\mathbf{F}_A^{\text{MAB}})^\dagger \mathbf{H}[t]^\dagger \right) \right] \quad (47) \\
&= \mathbb{E} \left[\log \left(1 + \frac{n_T n_R \gamma}{m} \left| \psi \right|^2 p[t] q \right) \right] \\
&= \exp \left(\frac{m}{n_T n_R p[t] q \gamma} \right) \mathbb{E}_1 \left(\frac{m}{n_T n_R p[t] q \gamma} \right),
\end{aligned}$$

where $p[t]$ and q are given by³

$$\begin{aligned}
p[t] &= \mathbf{a}_T(\phi[t])^\dagger \mathbf{F}_A^{\text{MAB}} \left((\mathbf{F}_A^{\text{MAB}})^\dagger \mathbf{F}_A^{\text{MAB}} \right)^{-1} \times (\mathbf{F}_A^{\text{MAB}})^\dagger \mathbf{a}_T(\phi[t]), \\
q &= \mathbf{a}_R(\theta)^\dagger \mathbf{a}_R(\theta).
\end{aligned}$$

The third equality of (47) is due to (27). Again, the energy efficiency $\mathcal{E}(\gamma)$ is calculated as in (16). As will be shown in Section VI, a trade-off between these two performance metrics exists to configure the mobility-aware hybrid precoder, where the trade-off depends on the velocity of the moving device.

VI. Numerical Results

First, we illustrate the accuracy of the high-SNR approximation (25) and the low-SNR bound (33) for the spectral efficiency of mmWave MIMO using hybrid precoder. In Fig. 3, we assume that the received SNR γ is relatively large,⁴ and set the number of transmit and receive antennas to $n_T = 128$ and $n_R = 64$, respectively. The transmitted signals propagate via $L = 10$ distinct paths, where $r_1 = 0.1$ or 0.9 while $r_2 = \dots = r_L = (1 - r_1)/(L - 1)$ as we discussed in Section IV. Fig. 3 shows that the approximation given in (25) agrees with simulations for the considered range of SNRs and the system settings. At large SNRs, the achieved spectral efficiency improves by activating additional RF chains, which allows more data streams to be transmitted simultaneously. This is in line with the high-SNR MIMO transmission strategy that utilizes the multiplexing gain of MIMO channels.

Fig. 4 illustrates the accuracy of the lower bound of the spectral efficiency by comparing $\mathcal{S}_{\text{LB}}(\gamma)$ calculated in (33) and the numerical simulation results. The number of antennas at the transmitter and receiver are set to $n_T = 128$ and $n_R = 16$, and the number of RF chains is $m = 1$ or $m = 4$. We adopt similar channel configurations with $L = 10$ paths, where the average gain of the dominant path is $r_1 = 0.1$ or $r_1 = 0.9$. Fig. 4 shows that the corresponding lower bound (33) captures the behavior of the spectral efficiency well, which is especially accurate when γ is small. When the mmWave MIMO channel has an equal path gain with $r_1 = 0.1$, the spectral efficiency is higher with activated 4 RF chains. On the other hand, when a

³Recall that $\mathbf{a}_T(\phi[t])$ is given by (41).

⁴Note that the values of γ in Fig. 3 are still small as γ denotes the pre-processed SNR without the antenna array gain. Instead, the parameter ρ in (18) takes into account the effect of antenna array and justifies the high SNR assumption with large n_T and n_R .

dominant propagation path exists with $r_1 = 0.9$, the achieved spectral efficiency is substantially higher compared to $r_1 = 0.1$. In this case, the optimal transmission strategy is to activate only a single RF chain and beamform towards the strongest propagation path (i.e., analog beamforming without digital precoding). Results in Fig. 4 are in line with our prediction in Section IV.

Next, we study the impact of the proposed configurable hybrid precoding scheme on both of the energy and spectral efficiency of mmWave MIMO communications. Specifically, we illustrate the energy efficiency and spectral efficiency as functions of the activated number of RF chains at macro and femto base stations, under different physical channels at high, medium, and low SNR settings. In the numerical simulations, we assume there are $n_T=128$ transmit antennas, $n_R = 64$ receive antennas, and $L = 10$ propagation paths. Other key system parameters are summarized in Table I. The subscripts “macro” and “femto” refer to the parameters of the macro and femto base stations, respectively.

In Fig. 5, the energy efficiency and spectral efficiency of the mmWave MIMO channels are plotted at a relatively high SNR with $\gamma = 0$ dB. Here, we apply the high-SNR approximation (25) to compute $\mathcal{E}(\gamma)$ and $\mathcal{S}(\gamma)$, and the results are then compared with the numerical simulations, where $r_1 = 0.1, 0.7$, or 0.9 . Note that the spectral efficiency of the macro and femto base stations is identical as $\mathcal{S}(\gamma)$ does not depend on a specific power model (15). As shown in Fig. 5 (a), when $r_1 = 0.1$, in order to maximize the energy efficiency of the macro base station, the optimal number of RF chains to be activated is 5, and the hybrid precoder is configured to beamform towards the corresponding strongest 5 propagation paths. As r_1 increases to 0.7 and 0.9, there exists a dominating propagation path between transmitter and receiver, and the optimal hybrid precoding is to activate two and one RF chains, respectively. In these two cases, activating more RF chains reduces the energy efficiency due to high power consumption of the RF chains and the phase shifting network (15). In the case of the femto base station, a single RF chain is optimal for all three channels, as shown in Fig. 5 (b). This is because the RF chain of the femto base station consumes much less power compared to the macro base station, which substantially decreases the total power P_{total} in energy efficiency (16). On the other hand, Fig. 5 (c) shows that the spectral efficiency always improves when more RF chains are activated. This is in line with the behavior of high-SNR MIMO capacity, where the capacity improves linearly to the number of multiplexed data streams.

In Fig. 6, we plot the energy and spectral efficiencies when the SNR $\gamma = -15$ dB, where the same high-SNR approximation (25) is used to obtain the analytical results. Figs. 6 (a) and (b) show that the energy efficiency of macro and femto base stations decreases approximately by half compared to the corresponding high SNR counterparts, as shown in Figs. 5 (a) and (b), respectively. Meanwhile, as the SNR decreases, it becomes less spectrally efficient to activate more RF chains. This is illustrated in Fig. 6, where the spectral efficiency curves become flattened or even decreasing as the number of RF chains increases. Therefore, the optimal number of RF chains required at macro base stations is less for all the three channels, resulting in 2 RF chains when $r_1 = 0.1$, and one RF chain when $r_1 = 0.7$ and 0.9 . Results in Figs. 5 and 6 also show that there exists a tradeoff between spectral efficiency and energy efficiency for most of the channel conditions, especially when the propagation

paths have similar path gains $\{\tau_i\}_{1 \leq i \leq L}$. In this case, we cannot maximize the spectral efficiency and the energy efficiency of the hybrid precoding simultaneously. Instead, we have to trade off one performance metric for another.

In Fig. 7, when SNR $\gamma = -30$ dB, the behavior of the spectral and energy efficiencies is investigated via lower-bounds (33) and (40), respectively. Here, the dominant propagation path has the path gain $\tau_1 = 0.1, 0.4$, or 0.9 . As shown in Figs. 7 (a) and (b), the optimal energy efficiency of both macro and femto base stations is achieved when a single RF chain is activated. In addition, Fig. 7 (c) shows that the spectral efficiency has only marginal improvement when $\tau_1 = 0.1$ and decreases drastically under other channels. Therefore, using a single RF chain maximizes both the energy efficiency and the spectral efficiency, which is in contrast to the high SNR cases, where a tradeoff between the two performance metrics exists. In other words, the optimal hybrid precoder reduces to the analog beamforming in the low SNR regime, which simplifies the design of analog and digital precoders.

Finally, we study the spectral efficiency of the mobility-aware hybrid precoding, where a transmitter is a high-speed moving vehicle. When $t = 0$, we assume that the transmitter is located 5 meters away from the scatterer and moves perpendicularly towards it. The velocity of the transmitter is set at 120 km/h. Along the trajectory of the transmitter, the spectral efficiency is calculated using (47), which is also confirmed by numerical simulations denoted by markers (i.e., see Fig. 8). The number of transmitted beams, as well as the activated RF chains, varies from $m = 1$ to 6. When no predicted beam is used, i.e., $m = 1$, Fig. 8 shows that the spectral efficiency quickly decreases due to movement of the transmitter. The variation of the spectral efficiency, from 9 to 2 nats/s/Hz, also causes difficulties in designing the encoder at the transmitter as the encoder has to track the instantaneous state of the channel for rate adaptation. With $m = 2$, peak spectral efficiency (which is achieved at the initial transmitter's location) decreases compared to the case $m = 1$. Yet, a fixed information rate can be maintained even with the movement of the transmitter. As m increases, the information rate remains unchanged for a longer time period (i.e., up to 10-millisecond), especially for $m = 5$ and 6.

In Fig. 9, we plot the averaged spectral and energy efficiencies as functions of the number of beams m (or activated RF chains). In these experiments, the average is taken over a 10-millisecond time frame. When the transmitter moves at 40 km/h, spectral efficiency is maximized with two beams, but drops when using more beams, which is due to a reduction in the radiated transmission power within each beam. The results in Fig. 9 also indicate that as the velocity increases to 80 and 120 km/h, spectral efficiencies drops, but reaches the maximum with $m = 4$ and $m = 5$, respectively. In all these cases, however, energy efficiency is maximized only when one beam is transmitted. Therefore, similar to the results in Figs. 5 and 6, there exists a tradeoff between the spectral and energy efficiencies in designing the mobility-aware hybrid precoding.

VII. Conclusions

Spectral efficiency and energy efficiency are key physical-layer performance metrics when designing a wireless communication system. In this work, we consider configurations of a

flexible hybrid precoding scheme for millimeter wave MIMO communications, where the number of active RF chains can be adjusted to achieve optimal, yet different, settings for the two performance metrics. To characterize the achieved performance, we derive lower bounds for the spectral and energy efficiencies using the Gaussian-Radau quadrature rule, which is especially tight in the low SNR regime. In the high SNR regime, closed-form approximations are obtained using the assumption of large antenna arrays. In addition, a mobility-aware hybrid precoding is proposed for mobile vehicular communications and its performance metrics are derived. Based on these explicit expressions, the proper number of active RF chains can be efficiently determined, so as to facilitate adaptive hybrid precoding depending on varying channel conditions.

Numerical results justify the accuracy of the analytical bounds/approximations, and show that there exists a tradeoff between spectral efficiency and energy efficiency by varying the number of activated RF chains. In other words, optimal settings to maximize these two performance metrics are generally different and one has to configure the hybrid precoder according to the specific requirements. On the contrary, in the low SNR regime, both the spectral and energy efficiencies maximize when hybrid precoding reduces to analog beamforming using a single RF chain. Under such channel conditions, configuration of the hybrid precoding becomes simple and analog beamforming is both spectral and energy efficient. Moreover, by transmitting additional predicted beams, the proposed mobility-aware hybrid precoding is capable of maintaining a communication link with a stable information rate, which is suitable for high-speed vehicular communications. Depending on the velocity of the communication device, numerical results show that the number of beams and the corresponding RF chains ranging from 2–5 would be able to maximize the spectral efficiency. Such configurations would also have a reasonable power consumption for high-speed mobile communications.

Appendix A

Proof of Proposition 1

Denote the Kronecker symbol $\delta_{a,b} = 1$ when $a = b$. Otherwise $\delta_{a,b} = 0$. The proof relies on the integral identities of complex Gaussian random variables and random Haar unitary matrices, which are stated as below.

Lemma 1. (Wick's Lemma [48]): Let $Z_1, \dots, Z_t$ denote *i.i.d.* complex Gaussian random variables with zero mean and unit variance. For $1 \leq a_1, a_2, b_1, b_2 \leq t$, the following identities hold:

$$\mathbb{E}[Z_{a_1} Z_{b_1}^*] = \delta_{a_1, b_1},$$

$$\mathbb{E}[Z_{a_1} Z_{a_2} Z_{b_1}^* Z_{b_2}^*] = \delta_{a_1, b_1} \delta_{a_2, b_2} + \delta_{a_1, b_2} \delta_{a_2, b_1}.$$

Denote $\mathcal{U}(n)$ as the set of $n \times n$ Haar unitary matrices. Let $\mathbf{U} \in \mathcal{U}(n)$, and denote $\mathbf{U} = \{U(p, q)\}_{1 \leq p, q \leq n}$ and $d\mathbf{U}$ as the normalized Haar measure on the unitary group $\mathcal{U}(n)$.

Lemma 2. (Integral identities of Haar unitary matrices [50]): The following integral identity of Haar unitary matrices holds:

$$\int_{\mathcal{U}(n)} U(a_1, b_1) U(a_2, b_2) U(c_1, d_1)^* U(c_2, d_2)^* d\mathbf{U} = \frac{\delta_{a_1, c_1} \delta_{b_1, d_1} \delta_{a_2, c_2} \delta_{b_2, d_2} + \delta_{a_2, c_1} \delta_{b_2, d_1} \delta_{a_1, c_2} \delta_{b_1, d_2}}{n^2 - 1} - \frac{\delta_{a_1, c_1} \delta_{b_2, d_1} \delta_{a_2, c_2} \delta_{b_1, d_2} + \delta_{a_2, c_1} \delta_{b_1, d_1} \delta_{a_1, c_2} \delta_{b_2, d_2}}{n(n^2 - 1)}.$$

We are now ready to prove Proposition 1. Inserting (4) into (13), the entries of $\mathbf{Q}$ are given by

$$Q_{i,j} = \mathbf{Q}(i, j) = \frac{1}{n_R} \sum_{c=1}^{n_R} e^{i(c-1)(\theta_j - \theta_i)}. \quad (48)$$

Denote the sets of indices $\mathcal{A}_1$, $\mathcal{A}_2$, and $\mathcal{A}_3$ as

$$\mathcal{A}_1 = \{1 \leq i, j \leq m\}, \quad (49)$$

$$\mathcal{A}_2 = \{m+1 \leq i \leq L, 1 \leq j \leq m\} \cup \{1 \leq i \leq m, m+1 \leq j \leq L\}, \quad (50)$$

$$\mathcal{A}_3 = \{m+1 \leq i, j \leq L\}. \quad (51)$$

The entries of the matrix $\mathbf{P}$, with the indices $(i, j) \in \mathcal{A}_1 \cup \mathcal{A}_2$, are given by inserting (5) into (14) as

$$P_{i,j} = \mathbf{P}(i, j) = \frac{1}{n_T} \sum_{c=1}^{n_T} e^{i(c-1)(\phi_j - \phi_i)}. \quad (52)$$

Denote the SVD of $\mathbf{A}_{T1}$ as $\mathbf{A}_{T1} = \mathbf{U}_{T1} \mathbf{\Sigma}_{T1}^{1/2} \mathbf{V}_{T1}^\dagger$. As $n_T > m$, the rectangular diagonal matrix $\mathbf{\Sigma}_{T1}$ is of the form $\mathbf{\Sigma}_{T1} = \begin{bmatrix} \underline{\mathbf{\Sigma}_{T1}}, \mathbf{0}_{m \times (n_T - m)} \end{bmatrix}^T$ and the diagonal elements of $\underline{\mathbf{\Sigma}_{T1}}$ are the non-

zero eigenvalues of $\mathbf{A}_{T1}\mathbf{A}_{T1}^\dagger$. The matrix block $\mathbf{A}_{T2}^\dagger\mathbf{A}_{T1}(\mathbf{A}_{T1}^\dagger\mathbf{A}_{T1})^{-1}\mathbf{A}_{T1}^\dagger\mathbf{A}_{T2}$ can be rewritten as

$$\mathbf{A}_{T2}^\dagger\mathbf{U}_{T1}\left[\frac{\underline{\Sigma}_{T1}}{\mathbf{0}}\right]^{1/2}\left[\underline{\Sigma}_{T1}^{-1}[\underline{\Sigma}_{T1}^{1/2}\mathbf{0}]\mathbf{U}_{T1}^\dagger\mathbf{A}_{T2} = \mathbf{A}_{T2}^\dagger\mathbf{U}_{T1}[1, \dots, m]\mathbf{U}_{T1}[1, \dots, m]^\dagger\mathbf{A}_{T2}. \quad (53)$$

The entries of $\mathbf{P}$, with the indices $(i, j) \in \mathcal{A}_3$, are therefore expressed as

$$P_{i,j} = \mathbf{P}(i, j) = \frac{1}{n_T} \sum_{b=1}^m \sum_{c=1}^{n_T} \sum_{d=1}^{n_T} U_{T1}(c, b)U_{T1}(d, b)^* \times e^{i(d-1)\phi_j - i(c-1)\phi_i}. \quad (54)$$

The first moment $\mu_1^{\mathbf{W}} = \mathbb{E}[\text{Tr}(\mathbf{W})]$ is derived as

$$\begin{aligned} \mu_1^{\mathbf{W}} &= \mathbb{E}\left[\sum_{a=1}^L \sum_{b=1}^L \psi_a P_{a,b} \psi_b^* Q_{b,a}\right] \quad (55) \\ &\stackrel{(a)}{=} \sum_{a=1}^L \sum_{b=1}^L \mathbb{E}[\psi_a \psi_b^*] \mathbb{E}[P_{a,b}] \mathbb{E}[Q_{b,a}] \\ &\stackrel{(b)}{=} \sum_{a=1}^L \sum_{b=1}^L \delta_{a,b} \sqrt{r_a r_b} \mathbb{E}[P_{a,b}] \mathbb{E}[Q_{b,a}] \\ &= \sum_{a=1}^L r_a \mathbb{E}[P_{a,a}] \mathbb{E}[Q_{a,a}], \end{aligned}$$

where (a) is due to the independence among Ψ , $\mathbf{A}_T$, and $\mathbf{A}_R$, (b) is due to Lemma 1. Replacing (48), (52), and (54) into (55), we obtain

$$\begin{aligned} \mu_1^{\mathbf{W}} &= \frac{1}{n_T} \sum_{a=m+1}^L r_a \sum_{b=1}^m \sum_{c,d=1}^{n_T} \mathbb{E}[U_{T1}(c, b)U_{T1}(d, b)^*] \times \mathbb{E}\left[e^{i(d-c)\phi_a}\right] + \sum_{a=1}^m r_a \quad (56) \\ &\stackrel{(c)}{=} \frac{1}{n_T} \sum_{a=m+1}^L r_a \sum_{b=1}^m \sum_{c=1}^{n_T} \mathbb{E}[|U_{T1}(c, b)|^2] + \sum_{a=1}^m r_a \\ &\stackrel{(d)}{=} \sum_{a=1}^m r_a + \frac{m}{n_T} \sum_{a=m+1}^L r_a, \end{aligned}$$

where (c) is due to the expectation $\mathbb{E}\left[e^{i(d-c)\phi_a}\right] = \delta_{c,d}$ over the uniformly distributed random variables ϕ_a , and (d) is due to the orthonormal property of the columns of the unitary matrix $\mathbf{U}_{T1}$, i.e., $\sum_{c=1}^{n_T} |U_{T1}(c, b)|^2 = 1$.

The second moment $\mu_2^{\mathbf{W}}$ is derived in (57) as shown at the bottom of this page. Then, we calculate the RHS of (57) in (58)–(60), shown at the bottom of the next page, when the indices a_1 and a_2 of the summands fall within the sets $\mathcal{A}_1$, $\mathcal{A}_2$, and $\mathcal{A}_3$, respectively. Note that the exact expressions of the summations over $\mathcal{A}_1$ and $\mathcal{A}_2$ can be obtained analytically. However, for $(a_1, a_2) \in \mathcal{A}_3$, the three terms on the RHS of (60) have to be

$$\begin{aligned}
 \mu_2^{\mathbf{W}} &= \sum_{a_1, a_2, b_1, b_2=1}^L \mathbb{E} \left[\psi_{a_1} \psi_{a_2} \psi_{b_1}^* \psi_{b_2}^* \right] \mathbb{E} \left[P_{a_1, b_1} P_{a_2, b_2} \right] \mathbb{E} \left[Q_{b_1, a_2} Q_{b_2, a_1} \right] \\
 &= \sum_{a_1, a_2, b_1, b_2=1}^L \sqrt{r_{a_1} r_{a_2} r_{b_1} r_{b_2}} \left(\delta_{a_1, b_1} \delta_{a_2, b_2} + \delta_{a_1, b_2} \delta_{a_2, b_1} \right) \mathbb{E} \left[P_{a_1, b_1} P_{a_2, b_2} \right] \mathbb{E} \left[Q_{b_1, a_2} Q_{b_2, a_1} \right] \\
 &= \sum_{a_1, a_2=1}^L r_{a_1} r_{a_2} \left(\mathbb{E} \left[P_{a_1, a_1} P_{a_2, a_2} \right] \mathbb{E} \left[Q_{a_1, a_2} Q_{a_2, a_1} \right] + \mathbb{E} \left[P_{a_1, a_2} P_{a_2, a_1} \right] \right)
 \end{aligned} \tag{57}$$

$$\begin{aligned}
 (a_1, a_2) \in \mathcal{A}_1: \sum_{a_1, a_2=1}^m r_{a_1} r_{a_2} \left(\mathbb{E} \left[Q_{a_1, a_2} Q_{a_2, a_1} \right] + \mathbb{E} \left[P_{a_1, a_2} P_{a_2, a_1} \right] \right) &= 2 \sum_{a=1}^m r_a^2 \tag{58} \\
 + \sum_{a_1 \neq a_2} r_{a_1} r_{a_2} \left(\mathbb{E} \left[Q_{a_1, a_2} Q_{a_2, a_1} \right] + \mathbb{E} \left[P_{a_1, a_2} P_{a_2, a_1} \right] \right) \\
 &= 2 \sum_{a=1}^m r_a^2 + \sum_{a_1 \neq a_2} r_{a_1} r_{a_2} \left(\frac{1}{n_R^2} \sum_{c_1, c_2=1}^{n_R} \mathbb{E} \left[e^{i(c_1 - c_2)\theta_{a_2}} \right] \mathbb{E} \left[e^{i(c_2 - c_1)\theta_{a_1}} \right] \right. \\
 &\quad \left. + \frac{1}{n_T^2} \sum_{c_1, c_2=1}^{n_T} \mathbb{E} \left[e^{i(c_1 - c_2)\phi_{a_2}} \right] \mathbb{E} \left[e^{i(c_2 - c_1)\phi_{a_1}} \right] \right) \\
 &= 2 \sum_{a=1}^m r_a^2 + \sum_{a_1 \neq a_2} r_{a_1} r_{a_2} \left(\frac{1}{n_R^2} \sum_{c_1, c_2=1}^{n_R} \delta_{c_1, c_2} + \frac{1}{n_T^2} \sum_{c_1, c_2=1}^{n_T} \delta_{c_1, c_2} \right) = 2 \sum_{a=1}^m r_a^2 \\
 &\quad + \left(\frac{1}{n_R} + \frac{1}{n_T} \right) \sum_{a_1=1}^m \sum_{\substack{a_2=1 \\ a_2 \neq a_1}}^m r_{a_1} r_{a_2}.
 \end{aligned}$$

$$\begin{aligned}
& (a_1, a_2) \in \mathcal{A}_2; 2 \sum_{a_1=1}^m \sum_{a_2=m+1}^L r_{a_1} r_{a_2} \mathbb{E}[P_{a_2, a_2}] \mathbb{E}[Q_{a_1, a_2} Q_{a_2, a_1}] \\
& + 2 \sum_{a_1=1}^m \sum_{a_2=m+1}^L r_{a_1} r_{a_2} \mathbb{E}[P_{a_1, a_2} P_{a_2, a_1}] \\
& = 2 \sum_{a_1=1}^m \sum_{a_2=m+1}^L \frac{r_{a_1} r_{a_2}}{n_T n_R^2} \left(\sum_{b=1}^m \sum_{c=1}^{n_T} \mathbb{E}[U_{T1}(c, b) U_{T1}(d, b)^*] \mathbb{E} \left[e^{i(d-c)\phi_{a_2}} \right] \right) \\
& \left(\sum_{c_1, c_2=1}^{n_R} \mathbb{E} \left[e^{i(c_2-c_1)\theta_{a_1}} \right] \mathbb{E} \left[e^{i(c_1-c_2)\theta_{a_2}} \right] \right) + 2 \sum_{a_1=1}^m \sum_{a_2=m+1}^L \frac{r_{a_1} r_{a_2}}{n_T^2} \sum_{c_1, c_2=1}^{n_T} \mathbb{E} \left[e^{i(c_2-c_1)\phi_{a_1}} \right] \mathbb{E} \left[e^{i(c_1-c_2)\phi_{a_2}} \right] \\
& = 2 \sum_{a_1=1}^m \sum_{a_2=m+1}^L \frac{r_{a_1} r_{a_2}}{n_T n_R^2} \left(\sum_{b=1}^m \sum_{c=1}^{n_T} \mathbb{E}[|U_{T1}(c, b)|^2] \right) \left(\sum_{c_1, c_2=1}^{n_R} \delta_{c_1, c_2} \right) + 2 \sum_{a_1=1}^m \sum_{a_2=m+1}^L \frac{r_{a_1} r_{a_2}}{n_T^2} \sum_{c_1, c_2=1}^{n_T} \delta_{c_1, c_2} \\
& = \frac{2m}{n_T n_R} \sum_{a_1=1}^m \sum_{a_2=m+1}^L r_{a_1} r_{a_2} + \frac{2}{n_T} \sum_{a_1=1}^m \sum_{a_2=m+1}^L r_{a_1} r_{a_2}
\end{aligned}$$

(59)

$$\begin{aligned}
& (a_1, a_2) \in \mathcal{A}_3: \sum_{a_1, a_2 = m+1}^L r_{a_1} r_{a_2} \left(\mathbb{E} \left[P_{a_1, a_1} P_{a_2, a_2} \right] \mathbb{E} \left[Q_{a_1, a_2} Q_{a_2, a_1} \right] + \mathbb{E} \left[P_{a_1, a_2} P_{a_2, a_1} \right] \right) \\
& = 2 \sum_{a_1 = m+1}^L r_{a_1}^2 \mathbb{E} \left[P_{a_1, a_1}^2 \right] + \frac{1}{n_R} \sum_{a_1 \neq a_2} r_{a_1} r_{a_2} \mathbb{E} \left[P_{a_1, a_1} P_{a_2, a_2} \right] + \sum_{a_1 \neq a_2} r_{a_1} r_{a_2} \mathbb{E} \left[P_{a_1, a_2} P_{a_2, a_1} \right] \\
& = \sum_{a_1 \neq a_2} \frac{r_{a_1} r_{a_2}}{n_R n_T^2} \sum_{b_1, b_2 = 1}^m \sum_{c_1, c_2, d_1, d_2 = 1}^{n_T} \mathbb{E} \left[U_{T1}(c_1, b_1) U_{T1}(c_2, b_2) U_{T1}(d_1, b_1)^* U_{T1}(d_2, b_2)^* \right. \\
& \quad \left. \delta_{c_1, d_1} \delta_{c_2, d_2} + \sum_{a_1 \neq a_2} \frac{r_{a_1} r_{a_2}}{n_T^2} \sum_{b_1, b_2 = 1}^m \sum_{c_1, c_2, d_1, d_2 = 1}^{n_T} \mathbb{E} \right. \\
& \quad \left. \left[U_{T1}(c_1, b_1) U_{T1}(c_2, b_2) U_{T1}(d_1, b_1)^* U_{T1}(d_2, b_2)^* \right] \delta_{c_1, d_2} \delta_{c_2, d_1} \right. \\
& \quad \left. + 2 \sum_{a_1 = m+1}^L \frac{r_{a_1}^2}{n_T^2} \sum_{b_1, b_2 = 1}^m \sum_{c_1, c_2, d_1, d_2 = 1}^{n_T} \mathbb{E} \left[U_{T1}(c_1, b_1) U_{T1}(c_2, b_2) U_{T1}(d_1, b_1)^* U_{T1}(d_2, b_2)^* \right. \right. \\
& \quad \left. \left. \delta_{c_1 + c_2, d_1 + d_2} \right] \right)
\end{aligned} \tag{60}$$

treated differently. Specifically, by exchanging the order of summations and the expectation, the first term of (60) is calculated as

$$\sum_{a_1 \neq a_2} \frac{r_{a_1} r_{a_2}}{n_R n_T^2} \sum_{b_1, b_2 = 1}^m \mathbb{E} \left[\left(\sum_{c_1 = 1}^{n_T} |U_{T1}(c_1, b_1)|^2 \right) \times \left(\sum_{c_2 = 1}^{n_T} |U_{T1}(c_2, b_2)|^2 \right) \right] = \frac{m^2}{n_R n_T^2} \sum_{a_1 \neq a_2} r_{a_1} r_{a_2}, \tag{61}$$

where we have applied the identities $\sum_{c_1 = 1}^{n_T} |U_{T1}(c_1, b_1)|^2 = 1$ and $\sum_{c_2 = 1}^{n_T} |U_{T1}(c_2, b_2)|^2 = 1$.

Similarly, the second term of (60) is calculated as

$$\begin{aligned}
& \sum_{a_1 \neq a_2} \frac{r_{a_1} r_{a_2}}{n_T^2} \sum_{b_1, b_2=1}^m \mathbb{E} \left[\left(\sum_{c_1=1}^{n_T} U_{T1}(c_1, b_1) U_{T1}(c_1, b_2)^* \right) \times \left(\sum_{c_2=1}^{n_T} U_{T1}(c_2, b_2) U_{T1}(c_2, b_1)^* \right) \right] \\
&= \frac{m}{n_T^2} \sum_{a_1 \neq a_2} r_{a_1} r_{a_2}.
\end{aligned} \tag{62}$$

The third term of (60) is intractable as the summations over c_1 , c_2 , d_1 , and d_2 (satisfying $c_1 + c_2 = d_1 + d_2$) cannot be solved via the orthonormal property of the unitary matrix $\mathbf{U}_{T1}$, and the expectations over $\mathbf{U}_{T1}$, being the singular matrix of the Vandermonde matrix $\mathbf{A}_{T1}$, is difficult since the distribution of $\mathbf{U}_{T1}$ is unknown. Instead, we evaluate the order of magnitude of the third term of (60) by approximating the matrix $\mathbf{U}_{T1}$ with a $n_T \times n_T$ random Haar unitary matrix $\mathbf{U}$. Denoting the elements of $\mathbf{U}$ as $\{U(c, b)\}_{1 \leq c, b \leq n_T}$, we estimate each individual expectation $\mathbb{E}[U_{T1}(c_1, b_1) U_{T1}(c_2, b_2) U_{T1}(d_1, b_1)^* U_{T1}(d_2, b_2)^*]$ with the corresponding $\mathbb{E}[U(c_1, b_1) U(c_2, b_2) U(d_1, b_1)^* U(d_2, b_2)^*]$ via Lemma 2. Therefore, the third term of (60) can be approximated as

$$\begin{aligned}
& 2 \sum_{a_1=m+1}^L \frac{r_{a_1}^2}{n_T^2} \sum_{b_1, b_2=1}^m \sum_{c_1, c_2, d_1, d_2=1}^{n_T} \delta_{c_1+c_2, d_1+d_2} \\
& \times \mathbb{E}[U(c_1, b_1) U(c_2, b_2) U(d_1, b_1)^* U(d_2, b_2)^*] = \frac{2m(m+1)}{n_T(n_T+1)} \sum_{a_1=m+1}^L r_{a_1}^2.
\end{aligned} \tag{63}$$

Summarizing (58)–(63), we obtain the desired result as in (32).

Biographies

Zhong Zheng received the B.Eng. degree from the Beijing University of Technology, Beijing, China, in 2007, the M.Sc. degree from the Helsinki University of Technology, Espoo, Finland, in 2010, and the D.Sc. degree from Aalto University, Espoo, Finland, in 2015. From 2015 to 2018, he held visiting positions at the University of Texas at Dallas and National Institute of Standards and Technology. In 2019, he joined School of Information and Electronics, Beijing Institute of Technology, Beijing, China, as an Associate Professor.

His research interests include massive MIMO, secure communications, millimeter wave communications, random matrix theory, and free probability theory.

Hamid Gharavi received the Ph.D. degree from Loughborough University, Loughborough, U.K., in 1980. He joined the Visual Communication Research Department, AT&T Bell Laboratories, Holmdel, NJ, USA, in 1982. He was then transferred to Bell Communications Research (Bellcore) after the AT&T-Bell divestiture, where he became a Consultant on video technology and a Distinguished Member of Research Staff. In 1993, he joined Loughborough University as a Professor and Chair of Communication Engineering. Since September 1998, he has been with the National Institute of Standards and Technology, U.S. Department of Commerce, Gaithersburg, MD, USA. He was a Core Member of Study Group XV (Specialist Group on Coding for Visual Telephony) of the International Communications Standardization Body CCITT (ITU-T) and a member of the IEEE 2030 Standard Working Group. His research interests include smart grid, wireless multimedia, mobile communications and wireless systems, mobile ad hoc networks, and visual communications. He received the Charles Babbage Premium Award from the Institute of Electronics and Radio Engineering in 1986, the IEEE CAS Society Darlington Best Paper Award in 1989, the Washington Academy of Science Distinguished Career in Science Award for 2017. He was a Distinguished Lecturer of the IEEE Communication Society. He has been a Guest Editor for a number of Special Issues of the proceedings of the IEEE including Smart Grids, Sensor Networks & Applications, Wireless Multimedia Communications, Advanced Automobile Technologies, and Grid Resilience. He was a TPC Co-Chair for IEEE SmartGridComm in 2010 and 2012. He was a member of the Editorial Board of proceedings of the IEEE from January 2003 to December 2008. He was Editor-in-Chief of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY and IEEE WIRELESS COMMUNICATIONS.

REFERENCES

- [1]. Busari SA, Huq KMS, Mumtaz S, Dai L, and Rodriguez J, "Millimeter-wave massive MIMO communication for future wireless-systems: A survey," *IEEE Commun. Surv. Tuts.*, vol. 20, no. 2, pp. 836–869, Apr-Jun 2018.
- [2]. Liu J, Guo H, Nishiyama H, Ujikawa H, Suzuki K, and Kato N, "New perspectives on future smart FiWi networks: Scalability, reliability, and energy efficiency," *IEEE Commun. Surv. Tuts.*, vol. 18, no. 2, pp. 1045–1072, Apr-Jun 2016.
- [3]. Liu J, Nishiyama H, Kato N, Kumagai T, and Takahara A, "Toward modeling ad hoc networks: Current situation and future direction," *IEEE Wireless Commun.*, vol. 20, no. 6, pp. 51–58, Dec. 2013.
- [4]. Liu J, Shi Y, Fadlullah Z. Md., and Kato N, "Space-air-ground integrated network: A survey," *IEEE Commun. Surv. Tuts.*, vol. 20, no. 2, pp. 2714–2741, Apr-Jun 2018.
- [5]. Pi Z and Khan F, "An introduction to millimeter-wave mobile broad-band systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, Jun. 2011.

- [6]. Sun S, Rappaport TS, Shafi M, Tang P, Zhang J, and Smith PJ, "Propagation models and performance evaluation for 5G millimeter-wave bands," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8422–8439, Sep. 2018.
- [7]. Zahir S, Gurbuz OD, Kar-Roy A, Raman S, and Rebeiz GM, "60-GHz 64- and 256-elements wafer-scale phased-array transmitters using full-reticle and subreticle stitching techniques," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 12, pp. 4701–4719, Dec. 2016.
- [8]. Heath RW, González-Prelcic N, Rangan S, Roh W, and Sayeed AM, "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.
- [9]. Björnson E, Sanguinetti L, Hoydis J, and Debbah M, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059–3075, Jun. 2015.
- [10]. Ayach OE, Rajagopal S, Abu-Surra S, Pi Z, and Heath RW, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.
- [11]. Zheng Z, Wei L, Speicher R, Müller RR, Hämäläinen J, and Corander J, "Asymptotic analysis of Rayleigh product channels: A free probability approach," *IEEE Trans. Inf. Theory*, vol. 63, no. 3, pp. 1731–1745, Mar. 2017.
- [12]. Li Y, Bakaloglu B, and Chakrabarti C, "A comprehensive energy model and energy-quality evaluation of wireless transceiver front-ends," in *Proc. IEEE Workshop Signal Process. Syst. Des. Implementation*, Athens, Greece, Nov. 2005, pp. 262–267.
- [13]. Ge X, Yang J, Gharavi H, and Sun Y, "Energy efficiency challenges of 5G small cell networks," *IEEE Commun. Mag.*, vol. 55, no. 5, pp. 184–191, 5 2017. [PubMed: 28757670]
- [14]. Verdú S, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.
- [15]. Yu H, Zhong L, and Sabharwal A, "Adaptive RF chain management for energy-efficient spatial-multiplexing MIMO transmission," in *Proc. ACM/IEEE Int. Symp. Low Power Electron. Des.*, San Francisco, CA, Aug. 2009, pp. 401–406.
- [16]. Pei Y, Pham T-H, and Liang Y, "How many RF chains are optimal for large-scale MIMO systems when circuit power is considered?" in *Proc. GLOBECOM*, Anaheim, CA, Dec. 2012, pp. 3868–3873.
- [17]. Xu Z, Yang C, Li GY, Zhang S, Chen Y, and Xu S, "Energy-efficient configuration of spatial and frequency resources in MIMO-OFDMA systems," *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 564–575, Feb. 2013.
- [18]. Zi R, Ge X, Thompson J, Wang C, Wang H, and Han T, "Energy efficiency optimization of 5G radio frequency chain systems," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 758–771, Apr. 2016.
- [19]. Ha VN, Nguyen DHN, and Frigon J, "Energy-efficient hybrid precoding for mmWave multi-user systems," in *Proc. Int. Conf. Commun.*, Kansas City, MO, 5 2018, pp. 1–6.
- [20]. Yu X, Shen JC, Zhang J, and Letaief KB, "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.
- [21]. Chen J-C, "Hybrid beamforming with discrete phase shifters for millimeter-wave massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 7604–7608, Aug. 2017.
- [22]. Ayach OE, Heath RW, Abu-Surra S, Rajagopal S, and Pi Z, "The capacity optimality of beam steering in large millimeter wave MIMO systems," in *Proc. IEEE Int. Workshop Signal Process. Adv. Wireless Commun.*, Cesme, Turkey, Jun. 2012, pp. 100–104.
- [23]. Xu Z, Han S, Pan Z, and C. L. I, "Alternating beamforming methods for hybrid analog and digital MIMO transmission," in *Proc. IEEE Int. Conf. Commun.*, London, U.K., Jun. 2015, pp. 1595–1600.
- [24]. Dai L, Gao X, Quan J, Han S, and C. L. I, "Near-optimal hybrid analog and digital precoding for downlink mmWave massive MIMO systems," in *Proc. IEEE Int. Conf. Commun.*, London, U.K., Jun. 2015, pp. 1334–1339.

- [25]. Rusu C, Mèndez-Rial R, González-Prelcic N, and Heath RW, "Low complexity hybrid precoding strategies for millimeter wave communication systems," *IEEE Trans. Wireless Commun*, vol. 15, no. 12, pp. 8380–8393, Dec. 2016.
- [26]. Rajashekar R and Hanzo L, "Iterative matrix decomposition aided block diagonalization for mm-wave multiuser MIMO systems," *IEEE Trans. Wireless Commun*, vol. 16, no. 3, pp. 1372–1384, Mar. 2017.
- [27]. Ge X, Sun Y, Gharavi H, and Thompson J, "Joint optimization of computation and communication power in multi-user massive MIMO systems," *IEEE Trans. Wireless Commun*, vol. 17, no. 6, pp. 4051–4063, Jun. 2018.
- [28]. Ni W and Dong X, "Hybrid block diagonalization for massive multiuser MIMO systems," *IEEE Trans. Commun*, vol. 64, no. 1, pp. 201–211, Jan. 2016.
- [29]. Sohrabi F and Yu W, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process*, vol. 10, no. 3, pp. 501–513, Apr. 2016.
- [30]. Sohrabi F and Yu W, "Hybrid analog and digital beamforming for mmWave OFDM large-scale antenna arrays," *IEEE J. Sel. Areas Commun*, vol. 35, no. 7, pp. 1432–1443, Jul. 2017.
- [31]. Raghavan V, Subramanian S, Cezanne J, Sampath A, Koymen OH, and Li J, "Single-user versus multi-user precoding for millimeter wave MIMO systems," *IEEE J. Sel. Areas Commun*, vol. 35, no. 6, pp. 1387–1401, Jun. 2017.
- [32]. Alkhateeb A, Ayach OE, Leus G, and Heath RW, "Hybrid precoding for millimeter wave cellular systems with partial channel knowledge," in *Proc. Inf. Theory Appl. Workshop*, San Diego, CA, Feb. 2013, pp. 1–5.
- [33]. Davis PJ and Rabinowitz P, *Methods of Numerical Integration*. New York, NY, USA: Academic, 2007.
- [34]. Raviteja P, Hong Y, and Viterbo E, "Millimeter wave analog beamforming with low resolution phase shifters for multiuser uplink," *IEEE Trans. Veh. Technol*, vol. 67, no. 4, pp. 3205–3215, Apr. 2018.
- [35]. Alkhateeb A, Leus G, and Heath RW, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Trans. Wireless Commun*, vol. 14, no. 11, pp. 6481–6494, Nov. 2015.
- [36]. Ryan O and Debbah M, "Asymptotic behavior of random Vandermonde matrices with entries on the unit circle," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3115–3147, Jul. 2009.
- [37]. Alkhateeb A and Heath RW, "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. Commun*, vol. 64, no. 5, pp. 1801–1818, 5 2016.
- [38]. Telatar E, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun*, vol. 10, no. 6, pp. 585–595, Nov-Dec 1999.
- [39]. Gauthier E, Yongacoglu A, and Chouinard JY, "Capacity of multiple antenna systems in Rayleigh fading channels," in *Proc. Can. Conf. Elect. Comput. Eng*, 2000, vol. 1, pp. 275–279.
- [40]. Grant A, "Rayleigh fading multi-antenna channels," *EURASIP J. Adv. Signal Process*, vol. 2002, no. 3, Dec. 2002, Art. no. 260208.
- [41]. Oyman O, Nabar RU, Bolcskei H, and Paulraj AJ, "Characterizing the statistical properties of mutual information in MIMO channels," *IEEE Trans. Signal Process*, vol. 51, no. 11, pp. 2784–2795, Nov. 2003.
- [42]. Silvester JR, "Determinants of block matrices," *Math. Gazette*, vol. 84, no. 501, pp. 460–467, 2000.
- [43]. Gradshteyn IS and Ryzhik IM, *Table of Integrals, Series, and Products*. Burlington, MA, USA: Academic, 2014.
- [44]. Rota G-C, "The number of partitions of a set," *Amer. Math. Monthly*, vol. 71, no. 5, pp. 498–504, 1964.
- [45]. Hiai F and Petz D, *The Semicircle Law, Free Random Variables and Entropy*. Boston, MA, USA: Amer. Math. Soc., 2006.
- [46]. Bai Z and Golub GH, "Bounds for the trace of the inverse and the determinant of symmetric positive definite matrices," *Ann. Numer. Math*, vol. 4, pp. 29–38, 1997.

- [47]. Tse DNC and Viswanath P, Fundamentals of Wireless Communications. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [48]. Reed I, "On a moment theorem for complex Gaussian processes," IRE Trans. Inf. Theory, vol. 8, no. 3, pp. 194–195, Apr. 1962.
- [49]. 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation (ver. 15.4.0)," 3rd Generation Partnership Project, Sophia Antipolis, France, Tech. Rep. TS 36.211, Jan. 2019 [Online]. Available: http://www.3gpp.org/ftp/Specs/archive/36_series/36.211/
- [50]. Forrester PJ, Log-Gases and Random Matrices. Princeton, NJ, USA: Princeton Univ. Press, 2007.

Fig. 1. Displacement of a multi-antenna transmitter. A scatterer is denoted by a circle and is located at the origin of the Euclidean plane.

Fig. 2.
Mobility-aware beamforming.

Fig. 3.

Spectral efficiency of mmWave MIMO at high SNRs with $n_T = 128$, $n_R = 64$, and the number of RF chains is $m = 1$, $m = 3$, or $m = 6$. The number of propagation paths is $L = 10$. Solid lines: the approximate spectral efficiency (25) with $r_1 = 0.1$; dashed lines: the approximate spectral efficiency (25) with $r_1 = 0.9$; markers: simulations.

Fig. 4.

Spectral efficiency of mmWave MIMO at low SNRs with $n_T = 128$, $n_R = 16$, and the number of RF chains is $m = 1$ or $m = 4$. The number of propagation paths is $L = 10$, where the dominant path has variance $r_1 = 0.1$ and 0.9 , respectively. Solid lines: lower bound of spectral efficiency (33); dashed lines: simulations.

Fig. 5.

Energy efficiency and spectral efficiency of mmWave MIMO at high SNR $\gamma = 0$ dB, where $n_T = 128$, $n_R = 64$, and $L = 10$. (a) Energy efficiency of macro base station; (b) energy efficiency of femto base station; (c) spectral efficiency.

Fig. 6. Energy efficiency and spectral efficiency of mmWave MIMO at medium SNR $\gamma = -15$ dB, where $n_T = 128$, $n_R = 64$, and $L = 10$. (a) Energy efficiency of macro base station; (b) energy efficiency of femto base station; (c) spectral efficiency.

Fig. 7. Energy efficiency and spectral efficiency of mmWave MIMO at low SNR $\gamma = -30$ dB, where $n_T = 128$, $n_R = 64$, and $L = 10$. (a) Energy efficiency of macro base station; (b) energy efficiency of femto base station; (c) spectral efficiency.

Fig. 8. Spectral efficiency of mobility-aware hybrid precoding with the pre-coder updated every 10 milliseconds. The transmitter is moving perpendicularly to the direction towards the scatterer with a speed 120 km/h.

Fig. 9. Trade-off between spectral efficiency and energy efficiency of mobility-aware hybrid precoding with different velocities.

Table I

Simulation Parameters

Parameter	Value	Parameter	Value
n_T	128	$P_{L,\text{macro}}$	38 dBm
n_R	64	$P_{t,\text{femto}}$	17 dBm
L	10	$P_{\text{RF},\text{macro}}$	6.5 W
N_0	-174 dBm/Hz	$P_{\text{RF},\text{femto}}$	0.6 W
B	1 MHz	P_{shifter}	88 mW
α_{macro}	0.228	α_{femto}	0.044