Diversity Analysis of Millimeter-Wave Massive MIMO Systems

Dian-Wu Yue, Shuai Xu, and Ha H. Nguyen

Abstract—This paper is concerned with asymptotic diversity analysis for millimeter-wave (mmWave) massive MIMO systems. First, for a single-user mmWave system employing distributed antenna subarray architecture in which the transmitter and receiver consist of K_t and K_r subarrays, respectively, a diversity gain theorem is established when the numbers of antennas at subarrays go to infinity. Specifically, assuming that all subchannels have the same number of propagation paths L, the theorem states that by employing such a distributed antenna-subarray architecture, a diversity gain of $K_r K_t L - N_s + 1$ can be achieved, where N_s is the number of data streams. This result means that compared to the co-located antenna architecture, using the distributed antenna-subarray architecture can scale up the diversity gain or multiplexing gain proportionally to $K_r K_t$. The diversity gain analysis is then extended to the multiuser scenario as well as the scenario with conventional partially-connected RF structure in the literature. Simulation results obtained with the hybrid analog/digital processing corroborate the analysis results and show that the distributed subarray architecture indeed yields significantly better diversity performance than the co-located antenna architectures.

Index Terms—Millimeter-wave communications, massive MIMO, diversity gain, multiplexing gain, diversity-multiplexing tradeoff, distributed antenna-subarrays, hybrid precoding.

I. INTRODUCTION

Recently, millimeter-wave (mmWave) communication has gained considerable attention as a candidate technology for 5G mobile communication systems and beyond [1]–[3]. The main reason for this is the availability of vast spectrum in the mmWave band (typically 30-300 GHz) that is very attractive for high data rate communications. However, compared to communication systems operating at lower microwave frequencies (such as those currently used for 4G mobile communications), propagation loss in mmWave frequencies is much higher, in orders of magnitude. Fortunately, given the much smaller carrier wavelengths, mmWave communication systems can make use of compact massive antenna arrays to compensate for the increased propagation loss.

Nevertheless, the large-scale antenna arrays together with high cost and large power consumption of the mixed ana-

Dian-Wu Yue is with the College of Information Science and Technology, Dalian Maritime University, Dalian, Liaoning 116026, China (e-mail: dwyue@dlmu.edu.cn), and also with the Department of Electrical and Computer Engineering, University of Saskatchewan, 57 Campus Drive, Saskatoon, SK, Canada S7N 5A9.

Shuai Xu is with the College of Information Science and Technology, Dalian Maritime University, Dalian, Liaoning 116026, China (e-mail: xu_shuai@dlmu.edu.cn).

Ha H. Nguyen is with the Department of Electrical and Computer Engineering, University of Saskatchewan, 57 Campus Drive, Saskatoon, SK, Canada S7N 5A9 (e-mail: ha.nguyen@usask.ca).

log/digital signal components make it difficult to equip a separate radio-frequency (RF) chain for each antenna element and perform all the signal processing in the baseband. Therefore, research on *hybrid* analog-digital processing of precoder and combiner for mmWave communication systems has attracted very strong interests from both academia and industry [4] − [16]. In particular, a large body of work has been performed to address challenges in using a limited number of RF chains for massive antenna arrays. For example, the authors in [4] considered single-user precoding in mmWave massive MIMO systems and established the optimality of beam steering for both single-stream and multi-stream transmission scenarios. In [7], the authors showed that the hybrid processing can realize any fully digital processing exactly if the number of RF chains is twice the number of data streams.

Two architectures for connecting the RF chains in the hybrid processing that have been studied in the literature are *full-connected* and *partially-connected*. In the former, each RF chain is connected to all the antenna elements, while only a subset of antenna elements is connected to each RF chain in the later. The partially-connected architecture is more energy-efficient and implementation-friendly since it can reduce the number of required phase shifters without significant performance loss. In the conventional partially-connected architecture [8]–[12] the antenna array is partitioned into a number of smaller disjoint subarrays, each of which is driven by a single transmission chain. More recently, a more general partially-connected architecture, referred to as hybridly-connected in [13] and overlapped subarray-based in [14], has been proposed. In such a hybridly-connected structure, each sub-array is connected to multiple RF chains, and each RF chain is connected to all the antennas corresponding to the sub-array in question. In particular, the authors in [13] demonstrate that the spectral efficiency of the hybridly-connected structure is better than that of the partially-connected structure and that its spectral efficiency can approach that of the fully-connected structure with the increase in the number of RF chains.

Nevertheless, due to the facts that the antenna arrays in the above-mentioned RF architectures are co-located and mmWave signal propagation has an important feature of multipath sparsity in both the temporal and spatial domains [17], [18], it is expected that the potentially available diversity and multiplexing gains are not large for the co-located antenna deployment. In order to enlarge diversity/multiplexing gains in mmWave massive MIMO communication systems, this paper considers a more general antenna array architecture, called *distributed antenna subarray architecture*, which includes co-located array architecture as special cases. It is pointed

out that, deploying distributed antennas has been shown a promising technique to increase spectral efficiency and expand coverage of wireless communication networks [19]–[23]. As such, it is of great interest to consider distributed antenna deployment in the context of mmWave massive MIMO systems.

The diversity-multiplexing tradeoff (DMT) is a compact and convenient framework to compare different MIMO systems in terms of the two main and related system indicators: data rate and error performance [24]–[27]. This tradeoff was originally characterized in [24] for MIMO communication systems operating over independent and identically distributed (i.i.d.) Rayleigh fading channels. The framework has then ignited a lot of interests in analyzing various communication systems and under different channel models. For a mmWave massive MIMO system, how to quantify the diversity performance and characterize its DMT is a fundamental and open research problem. In particular, to the best of our knowledge, until now there is no unified diversity gain analysis for mmWave massive MIMO systems that is applicable to both co-located and distributed antenna array architectures.

To fill this gap, this paper investigates the diversity performance of mmWave massive MIMO systems with the proposed distributed subarray architecture (the multiplexing performance will be investigated in another paper). The focus is on the asymptotical diversity gain analysis in order to find out the potential diversity advantage provided by multiple distributed antenna arrays. The obtained analysis can be used conveniently to compare various mmWave massive MIMO systems with different distributed antenna array structures. The main contributions of this paper are summarized as follows: First, for a single-user system employing the proposed distributed subarray architecture, a diversity gain expression is obtained when the number of antennas at each subarray increases without bound. This expression clearly indicates that one can obtain a large diversity gain and/or multiplexing gain by employing the proposed distributed subarray architecture. Second, the diversity gain analysis is extended to the multiuser scenario with downlink and uplink transmission, as well as the single-user system employing the conventional partially-connected RF structure based on the distributed subarrays. Simulation results are provided to corroborate the analysis results and show that the distributed subarray architecture yields significantly better diversity performance than the co-located single-array architecture.

The remainder of this paper is organized as follows. Section II describes the massive MIMO system model and hybrid processing with the distributed subarray architecture in mmWave fading channels. Section III provides the asymptotical diversity analysis for the single-user mmWave system. In Sections IV and V, the diversity gain analysis is extended to the multiuser scenario and the scenario with the conventional partially-connected RF architecture, respectively. Numerical results are presented in Section VI. Section VII concludes the paper.

Throughout this paper, the following notations are used. Boldface upper and lower case letters denote matrices and column vectors, respectively. The superscripts $(\cdot)^T$ and $(\cdot)^H$ stand for transpose and conjugate-transpose, respectively. $\mathrm{diag}\{a_1, a_2, \ldots, a_N\}$ stands for a diagonal matrix with di-

agonal elements $\{a_1, a_2, \ldots, a_N\}$. The expectation operator is denoted by $\mathbb{E}()$. $[\mathbf{A}]_{ij}$ gives the (i, j)th entry of matrix $\mathbf{A}$. $\mathbf{A} \bigotimes \mathbf{B}$ is the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$. We write a function $a(x)$ of x as $o(x)$ if $\lim_{x \to 0} a(x)/x = 0$. Finally, $\mathcal{CN}(0, 1)$ denotes a circularly symmetric complex Gaussian random variable with zero mean and unit variance.

II. SYSTEM MODEL

Consider a single-user mmWave massive MIMO system as shown in Fig. 1. The transmitter is equipped with a distributed antenna array to send N_s data streams to a receiver, which is also equipped with a distributed antenna array. Here, a distributed antenna array means an array consisting of several remote antenna units (RAUs) (i.e., antenna subarrays) that are distributively located, as depicted in Fig. 2. Specifically, the antenna array at the transmitter consists of K_t RAUs, each of which has N_t antennas and is connected to a baseband processing unit (BPU) by fiber. Likewise, the distributed antenna array at the receiver consists of K_r RAUs, each having N_r antennas and also being connected to a BPU by fibers. Such a MIMO system shall be referred to as a (K_t, N_t, K_r, N_r) distributed MIMO (D-MIMO) system. When $K_t = K_r = 1$, the system reduces to a conventional co-located MIMO (C-MIMO) system.

The transmitter accepts as its input N_s data streams and is equipped with $N_t^{(\mathrm{rf})}$ RF chains, where $N_s \leq N_t^{(\mathrm{rf})} \leq N_t K_t$. Given $N_t^{(\mathrm{rf})}$ transmit RF chains, the transmitter can apply a low-dimension $N_t^{(\mathrm{rf})} \times N_s$ baseband precoder, $\mathbf{W}_t$, followed by a high-dimension $K_t N_t \times N_t^{(\mathrm{rf})}$ RF precoder, $\mathbf{F}_t$. Note that amplitude and phase modifications are feasible for the baseband precoder $\mathbf{W}_t$, while only phase changes can be made by the RF precoder $\mathbf{F}_t$ through the use of variable phase shifters and combiners. The transmitted signal vector can be written as

$$\mathbf{x} = \mathbf{F}_t \mathbf{W}_t \mathbf{s}, \tag{1}$$

where $\mathbf{s}$ is the $N_s \times 1$ symbol vector such that $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \frac{P}{N_s}\mathbf{I}_{N_s}$. Thus P represents the average total input power. Considering a narrowband block fading channel, the $K_r N_r \times 1$ received signal vector is

$$\mathbf{y} = \mathbf{H}\mathbf{F}_t \mathbf{W}_t \mathbf{s} + \mathbf{n} \tag{2}$$

where $\mathbf{H}$ is $K_r N_r \times K_t N_t$ channel matrix and $\mathbf{n}$ is a $K_r N_r \times 1$ vector consisting of i.i.d. $\mathcal{CN}(0, 1)$ noise samples. Throughout this paper, $\mathbf{H}$ is assumed known to both the transmitter and receiver. Given that $N_r^{(\mathrm{rf})}$ RF chains (where $N_s \leq N_r^{(\mathrm{rf})} \leq N_r K_r$) are used at the receiver to detect the N_s data streams, the processed signal is given by

$$\mathbf{z} = \mathbf{W}_r^H \mathbf{F}_r^H \mathbf{H}\mathbf{F}_t \mathbf{W}_t \mathbf{s} + \mathbf{W}_r^H \mathbf{F}_r^H \mathbf{n} \tag{3}$$

where $\mathbf{F}_r$ is the $K_r N_r \times N_r^{(\mathrm{rf})}$ RF combining matrix, and $\mathbf{W}_r$ is the $N_r^{(\mathrm{rf})} \times N_s$ baseband combining matrix.

Furthermore, according to the architecture of RAUs at the transmitting and receiving ends, $\mathbf{H}$ can be written as

$$\mathbf{H} = \begin{bmatrix} \sqrt{g_{11}}\mathbf{H}_{11} & \cdots & \sqrt{g_{1K_t}}\mathbf{H}_{1K_t} \\ \vdots & \ddots & \vdots \\ \sqrt{g_{K_r1}}\mathbf{H}_{K_r1} & \cdots & \sqrt{g_{K_rK_t}}\mathbf{H}_{K_rK_t} \end{bmatrix}. \tag{4}$$
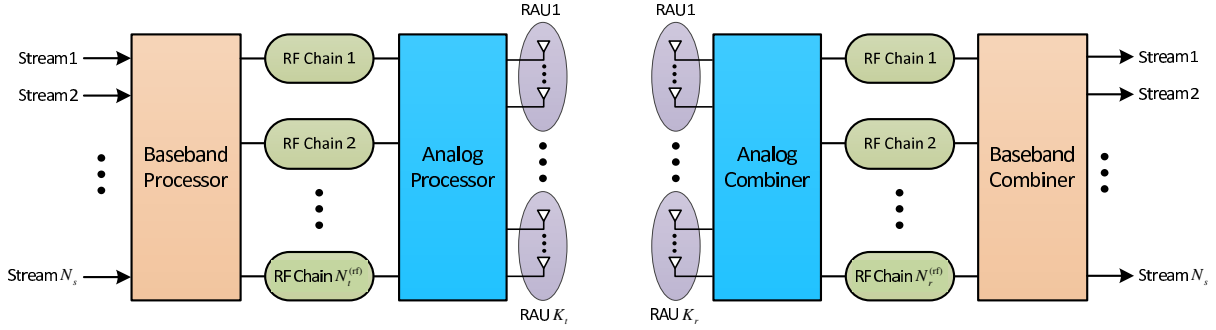
Fig. 1. Block diagram of a mmWave massive MIMO system with distributed antenna arrays.
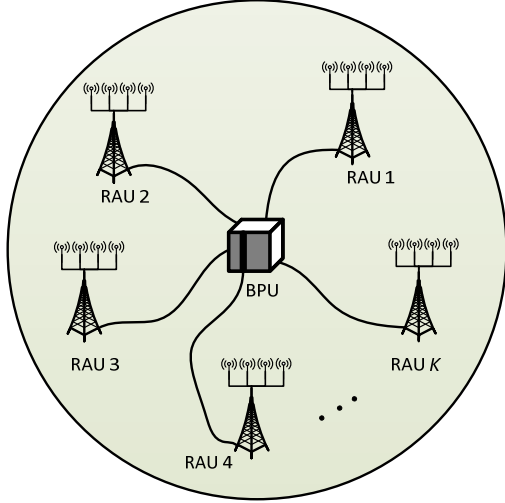

Fig. 2. Illustration of distributed antenna array deployment.

In the above expression, g_{ij} represents the large scale fading effect between the ith RAU at the receiver and the jth RAU at the transmitter, which is assumed to be constant over many coherence-time intervals. The normalized subchannel matrix $\mathbf{H}_{ij}$ represents the MIMO channel between the jth RAU at the transmitter and the ith RAU at the receiver.

A clustered channel model based on the extended Saleh-Valenzuela model is often used in mmWave channel modeling and standardization [4], [12], [13] and it is also adopted in this paper. For simplicity of exposition, each scattering cluster is assumed to contribute a single propagation path.[1] Using this model, the subchannel matrix $\mathbf{H}_{ij}$ is given by

$$\mathbf{H}_{ij} = \sqrt{\frac{N_t N_r}{L_{ij}}} \sum_{l=1}^{L_{ij}} \alpha_{ij}^l \mathbf{a}_r(\phi_{ij}^{rl}, \theta_{ij}^{rl}) \mathbf{a}_t^H(\phi_{ij}^{tl}, \theta_{ij}^{tl}), \quad (5)$$

where L_{ij} is the number of propagation paths, α_{ij}^l is the complex gain of the lth ray, and ϕ_{ij}^{rl} (θ_{ij}^{rl}) and ϕ_{ij}^{tl} (θ_{ij}^{tl}) are its random azimuth (elevation) angles of arrival and departure, respectively. Without loss of generality, the complex gains α_{ij}^l

[1]This assumption can be relaxed to account for clusters with finite angular spreads and the results obtained in this paper can be readily extended for such a case.

are assumed to be $\mathcal{CN}(0,1)$. [2] The vectors $\mathbf{a}_r(\phi_{ij}^{rl}, \theta_{ij}^{rl})$ and $\mathbf{a}_t(\phi_{ij}^{tl}, \theta_{ij}^{tl})$ are the normalized receive/transmit array response vectors at the corresponding angles of arrival/departure. For an N-element uniform linear array (ULA) , the array response vector is

$$\mathbf{a}^{\text{ULA}}(\phi) = \frac{1}{\sqrt{N}} \left[1, e^{j2\pi \frac{d}{\lambda} \sin(\phi)}, \ldots, e^{j2\pi(N-1)\frac{d}{\lambda} \sin(\phi)} \right]^T \quad (6)$$

where λ is the wavelength of the carrier and d is the inter-element spacing. It is pointed out that the angle θ is not included in the argument of $\mathbf{a}^{\text{ULA}}$ since the response for an ULA is independent of the elevation angle. In contrast, for a uniform planar array (UPA), which is composed of N_h and N_v antenna elements in the horizontal and vertical directions, respectively, the array response vector is represented by

$$\mathbf{a}^{\text{UPA}}(\phi, \theta) = \mathbf{a}_h^{\text{ULA}}(\phi) \otimes \mathbf{a}_v^{\text{ULA}}(\theta), \quad (7)$$

where

$$\mathbf{a}_h^{\text{ULA}}(\phi) = \frac{1}{\sqrt{N_h}} \left[1, e^{j2\pi \frac{d_h}{\lambda} \sin(\phi)}, \ldots, e^{j2\pi(N_h-1)\frac{d_h}{\lambda} \sin(\phi)} \right]^T \quad (8)$$

and

$$\mathbf{a}_v^{\text{ULA}}(\theta) = \frac{1}{\sqrt{N_v}} \left[1, e^{j2\pi \frac{d_v}{\lambda} \sin(\theta)}, \ldots, e^{j2\pi(N^v-1)\frac{d_v}{\lambda} \sin(\theta)} \right]^T . \quad (9)$$

III. DIVERSITY GAIN ANALYSIS

The most common performance metric of a digital communication system is the error probability, which can be defined either as the probability of symbol error or the probability of bit error (i.e., the bit error rate (BER)). When communicating over a fading channel, errors obviously depend on specific channel realizations. As such, the random behavior of a fading channel needs to be taken into account, leading to the concept of average error probabilities [28]. Determining exact expressions for the average error probabilities for a digital communication system operating over a certain fading channel is usually tedious and might not give a clear insight about the system behavior. As such, there is a need to characterize the performance of a communication system in a simple and insightful way. A popular approach is to shift the focus

[2]The different variances of α_{ij}^l can easily accounted for by absorbing into the large scale fading coefficients g_{ij}.

from exact performance analysis to asymptotic performance analysis, i.e., analyzing performance at the high signal-to-noise (SNR) region, as done in [29]. This is a reasonable approach since the performance of practical interest is in the high SNR region and in such a region, good approximation can be made on the exact analysis.

In the high-SNR region, the average BER function can be approximated in most cases as [29]

$$\overline{\text{BER}} \approx (G_c \cdot \bar{\gamma})^{-G_d} \qquad (10)$$

where G_d and G_c are referred to as the diversity and coding gains, respectively, and $\bar{\gamma}$ is the average receive SNR. The diversity gain determines the slope of the BER curve versus $\bar{\gamma}$ at high SNR in a log-log scale, whereas the coding gain determines how the curve is shifted along the horizontal axis with respect to a benchmark BER curve $\bar{\gamma}^{-G_d}$. This yields a simple parameterized average BER characterization for high SNR, which can provide meaningful insights on the system performance behavior.

In this section, the diversity gain is first examined for a *generalized selection combining*. The main result is then invoked in the diversity analysis of the distributed mmWave massive MIMO system studied in this paper.

A. Diversity Gain of Generalized Selection Combining

Selection combining (SC) is the most popular low-complexity combining scheme. With selection combining, the receiver estimates the SNRs of all available diversity branches and then select the one with the highest SNR for detection. For *generalized* selection combining (GSC) considered here, the receiver also estimates the SNRs of all available diversity branches. However, instead of selecting the branch with the highest SNR, it selects a branch with the lth highest SNR for detection. It is pointed out that, while such a GSC scheme has no practical interest in its own right, its diversity analysis can be used in performance analysis of the mmWave massive MIMO system considered in this paper.

Lemma 1: Consider a GSC system with L receive antennas operating over i.i.d. Rayleigh fading channels. If the receiver selects the branch with the lth highest SNR for detection then the system achieves diversity gain

$$G_d = L - l + 1. \qquad (11)$$

Proof: Let $F(\gamma)$ and $f(\gamma)$ be the probability density function (PDF) and cumulative distribution function (CDF) of the instantaneous SNRs in all branches, respectively. Let $\bar{\gamma}$ denote the average receive SNR of each branch. With Rayleigh fading, it follows from [29] that $F(\gamma)$ and $f(\gamma)$ can be written as

$$F(\gamma) = 1 - e^{-\frac{\gamma}{\bar{\gamma}}} = \frac{\gamma}{\bar{\gamma}} + o\left(\frac{\gamma}{\bar{\gamma}}\right) \qquad (12)$$

and

$$f(\gamma) = \frac{1}{\bar{\gamma}}e^{-\frac{\gamma}{\bar{\gamma}}} = \frac{1}{\bar{\gamma}} - \frac{\gamma}{\bar{\gamma}^2} + o\left(\frac{\gamma}{\bar{\gamma}}\right). \qquad (13)$$

If the receiver selects the branch with the lth highest SNR for detection, then based on the theory of order statistics [30], the

PDF of the instantaneous receive SNR at the receiver, denoted γ_l, is given by

$$f_{l:L}(\gamma_l) = \frac{L!}{(L-l)!(l-1)!}[F(\gamma_l)]^{L-l}[1 - F(\gamma_l)]^{l-1}f(\gamma_l)$$

$$= \frac{L!}{(L-l)!(l-1)!}\frac{1}{\bar{\gamma}}\left(\frac{\gamma_l}{\bar{\gamma}}\right)^{L-l} + o\left(\left(\frac{\gamma_l}{\bar{\gamma}}\right)^{L-l}\right). \qquad (14)$$

Applying the above PDF in Proposition 1 in [29] leads to the desired result. $\qquad \square$

Lemma 1 can be extended to the case of independent but not identically distributed (i.n.i.d.) Rayleigh fading channels and the result is stated in the next lemma.

Lemma 2: Suppose that the GSC system with L receive antennas operates over the i.n.i.d. Rayleigh fading channels. If it selects the path with the lth highest SNR for detection, then it can achieve diversity gain

$$G_d = L - l + 1. \qquad (15)$$

Proof: Let $\bar{\gamma}_{\min}$ and $\bar{\gamma}_{\max}$ denote the maximum and minimum values of the average receive SNRs of all these L diversity paths, respectively. Furthermore, let $\mathcal{A}$ and $\mathcal{B}$ denote two GSC systems, each equipped with L receive antennas and operating over i.i.d. Rayleigh fading channels such that the average receive SNRs equal to $\bar{\gamma}_{\max}$ and $\bar{\gamma}_{\min}$, respectively. It is known from Lemma 1 that the diversity gains of these two systems are the same and equal to $L - l + 1$ if both systems select the branch with the lth highest instantaneous SNR for detection. Furthermore, since the GSC system under consideration cannot have better diversity performance than System $\mathcal{A}$ and cannot have worse diversity performance than System $\mathcal{B}$, it can then be concluded that the i.n.i.d. system must also achieve the diversity gain of $L - l + 1$. $\qquad \square$

B. Diversity Gain Analysis of the Distributed mmWave Massive MIMO System

From the structure and definition of the channel matrix $\mathbf{H}$ in Section II, there is a total of $L_s = \sum_{i=1}^{K_r}\sum_{j=1}^{K_t}L_{ij}$ propagation paths. Naturally, $\mathbf{H}$ can be decomposed into a sum of L_s rank-one matrices, each corresponding to one propagation path. Specifically, $\mathbf{H}$ can be rewritten as

$$\mathbf{H} = \sum_{i=1}^{K_r}\sum_{j=1}^{K_t}\sum_{l=1}^{L_{ij}}\tilde{\alpha}_{ij}^l\tilde{\mathbf{a}}_r(\phi_{ij}^{rl}, \theta_{ij}^{rl})\tilde{\mathbf{a}}_t^H(\phi_{ij}^{tl}, \theta_{ij}^{tl}), \qquad (16)$$

where

$$\tilde{\alpha}_{ij}^l = \sqrt{g_{ij}\frac{N_tN_r}{L_{ij}}}\alpha_{ij}^l, \qquad (17)$$

$\tilde{\mathbf{a}}_r(\phi_{ij}^{rl}, \theta_{ij}^{rl})$ is a $K_rN_r \times 1$ vector whose bth entry is defined as

$$[\tilde{\mathbf{a}}_r(\phi_{ij}^{rl}, \theta_{ij}^{rl})]_b = \begin{cases} [\mathbf{a}_r(\phi_{ij}^{rl}, \theta_{ij}^{rl})]_{b-(i-1)N_r}, & b \in Q_i^r \\ 0, & b \notin Q_i^r \end{cases} \qquad (18)$$

where $Q_i^r = ((i-1)N_r, iN_r]$. And $\tilde{\mathbf{a}}_t(\phi_{ij}^{tl}, \theta_{ij}^{tl})$ is a $K_tN_t \times 1$ vector whose bth entry is defined as

$$[\tilde{\mathbf{a}}_t(\phi_{ij}^{tl}, \theta_{ij}^{tl})]_b = \begin{cases} [\mathbf{a}_t(\phi_{ij}^{tl}, \theta_{ij}^{tl})]_{b-(j-1)N_t}, & b \in Q_j^t \\ 0, & b \notin Q_j^t. \end{cases} \qquad (19)$$

where $Q_j^t = ((j-1)N_t, jN_t]$.

Lemma 3: Suppose that the antenna configurations at all RAUs are either ULA or UPA. Then all L_s vectors $\{\tilde{\mathbf{a}}_r(\phi_{ij}^{rl}, \theta_{ij}^{rl})\}$ are orthogonal to each other when $N_r \to \infty$. Likewise, all L_s vectors $\{\tilde{\mathbf{a}}_t(\phi_{ij}^{tl}, \theta_{ij}^{tl})\}$ are orthogonal to each other when $N_t \to \infty$.

Proof: It follows immediately from (18) and (19) that if $u \neq v$, then vectors $\{\tilde{\mathbf{a}}_r(\phi_{up}^{rl}, \theta_{up}^{rl})\}$ and $\{\tilde{\mathbf{a}}_r(\phi_{vq}^{rl}, \theta_{vq}^{rl})\}$ are orthogonal. On the other hand, when $u = v$ and $p \neq q$, it is known from Lemma 1 and Corollary 2 in [4] (also see [31]) that vectors $\{\tilde{\mathbf{a}}_r(\phi_{up}^{rl}, \theta_{up}^{rl})\}$ and $\{\tilde{\mathbf{a}}_r(\phi_{vq}^{rl}, \theta_{vq}^{rl})\}$ are orthogonal. The proof that $\{\tilde{\mathbf{a}}_t(\phi_{ij}^{tl}, \theta_{ij}^{tl})\}$ is a set of orthogonal vectors can be shown similarly. $\square$

Theorem 1: Suppose that both sets $\{\tilde{\mathbf{a}}_r(\phi_{ij}^{rl}, \theta_{ij}^{rl})\}$ and $\{\tilde{\mathbf{a}}_t(\phi_{ij}^{tl}, \theta_{ij}^{tl})\}$ are orthogonal vector sets when $N_r \to \infty$ and $N_t \to \infty$. Let $N_s \leq L_s$. Then the distributed massive MIMO system with large N_r and N_t can achieve a diversity gain of

$$G_d = L_s - N_s + 1. \tag{20}$$

Proof: The distributed massive MIMO system can be considered as a co-located massive MIMO system with L_s paths that have complex gains $\{\tilde{\alpha}_{ij}^l\}$, receive array response vectors $\{\tilde{\mathbf{a}}_r(\phi_{ij}^{rl}, \theta_{ij}^{rl})\}$ and transmit response vectors $\{\tilde{\mathbf{a}}_t(\phi_{ij}^{tl}, \theta_{ij}^{tl})\}$. Furthermore, order all paths in a decreasing order of the absolute values of the complex gains $\{\tilde{\alpha}_{ij}^l\}$. Then the channel matrix can be written as

$$\mathbf{H} = \sum_{l=1}^{L_s} \tilde{\alpha}^l \tilde{\mathbf{a}}_r(\phi^{rl}, \theta^{rl}) \tilde{\mathbf{a}}_t(\phi^{tl}, \theta^{tl})^H, \tag{21}$$

where $\tilde{\alpha}^1 \geq \tilde{\alpha}^2 \geq \cdots \geq \tilde{\alpha}^{L_s}$. One can rewrite $\mathbf{H}$ in a matrix form as

$$\mathbf{H} = \mathbf{A}_r \mathbf{D} \mathbf{A}_t^H \tag{22}$$

where $\mathbf{D}$ is a $L_s \times L_s$ diagonal matrix with $[\mathbf{D}]_{ll} = \tilde{\alpha}^l$, and $\mathbf{A}_r$ and $\mathbf{A}_t$ are defined as follows:

$$\mathbf{A}_r = [\tilde{\mathbf{a}}_r(\phi^{r1}, \theta^{r1}), \tilde{\mathbf{a}}_r(\phi^{r2}, \theta^{r2}), \ldots, \tilde{\mathbf{a}}_r(\phi^{rL_s}, \theta^{rL_s})] \tag{23}$$

and

$$\mathbf{A}_t = [\tilde{\mathbf{a}}_t(\phi^{t1}, \theta^{t1}), \tilde{\mathbf{a}}_t(\phi^{t2}, \theta^{t2}), \ldots, \tilde{\mathbf{a}}_t(\phi^{tL_s}, \theta^{tL_s})]. \tag{24}$$

Since both $\{\tilde{\mathbf{a}}_r(\phi^{rl}, \theta^{rl})\}$ and $\{\tilde{\mathbf{a}}_t(\phi^{tl}, \theta^{tl})\}$ are orthogonal vector sets when $N_r \to \infty$ and $N_t \to \infty$, $\mathbf{A}_r$ and $\mathbf{A}_t$ are asymptotically unitary matrices. Then one can form a singular value decomposition (SVD) of matrix $\mathbf{H}$ as

$$\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H = [\mathbf{A}_r | \mathbf{A}_r^\perp] \mathbf{\Sigma} [\tilde{\mathbf{A}}_t | \tilde{\mathbf{A}}_t^\perp]^H \tag{25}$$

where $\mathbf{\Sigma}$ is a diagonal matrix containing all singular values on its diagonal, i.e.,

$$[\mathbf{\Sigma}]_{ll} = \begin{cases} |\tilde{\alpha}^l|, & \text{for } 1 \leq l \leq L_s \\ 0, & \text{for } l > L_s \end{cases} \tag{26}$$

and the submatrix $\tilde{\mathbf{A}}_t$ is defined as

$$\tilde{\mathbf{A}}_t = [e^{-j\psi_1}\tilde{\mathbf{a}}_t(\phi^{t1}, \theta^{t1}), \ldots, e^{-j\psi_{L_s}}\tilde{\mathbf{a}}_t(\phi^{tL_s}, \theta^{tL_s})] \tag{27}$$

where ψ_l is the phase of complex gain $\tilde{\alpha}^l$ corresponding to the lth path.

Based on (25), the optimal precoder and combiner are chosen, respectively, as

$$[\mathbf{F}_t\mathbf{W}_t]_{\text{opt}} = [e^{-j\psi_1}\tilde{\mathbf{a}}_t(\phi^{t1}, \ldots, e^{-j\psi_{L_s}}\tilde{\mathbf{a}}_t(\phi^{tN_s}, \theta^{tN_s})] \tag{28}$$

and

$$[\mathbf{F}_r\mathbf{W}_r]_{\text{opt}} = [\tilde{\mathbf{a}}_r(\phi^{r1}, \ldots, \tilde{\mathbf{a}}_r(\phi^{rN_s}, \theta^{rN_s})]. \tag{29}$$

To summarize, when N_t and N_r are large enough, the massive MIMO system can employ the optimal precoder and combiner given in (28) and (29), respectively. Then it follows from the above SVD analysis that the instantaneous SNR of the lth data stream is given by

$$\text{SNR}_l = \frac{P}{N_s}|\tilde{\alpha}^l|^2, \quad l = 1, 2, \ldots, N_s. \tag{30}$$

Now the detection of the lth data stream is equivalent to the detection in a generalized selection combining system, which selects the path with the lth highest SNR for detection. Therefore, it follows from Lemma 2 that the detection performance of the lth data stream has a diversity gain $L_s - l + 1$. Since the overall BER is the arithmetic mean of individual BERs, i.e., $\overline{\text{BER}} = \frac{1}{N_s}\sum_{l=1}^{N_s} \overline{\text{BER}}(l)$, the system's diversity gain equals to the diversity gain in detecting the N_sth data stream, which is the worst among all data streams. Therefore, the result in (20) is obtained. $\square$

Remark 1: When N_t and N_r are large enough, (25) indicates that the system multiplexing gain is at most equal to L_s. This is reasonable since there exist only L_s effective singular values in the channel matrix $\mathbf{H}$. Theorem 1 provides a simple diversity-multiplexing tradeoff of a mmWave massive MIMO system: adding one data stream to the system decreases the diversity gain by one, whereas removing one data stream increases the diversity gain by one. Such a tradeoff is useful in designing a system to meet requirements on both data rate and error performance.

Remark 2: Under the case where N_t and N_r are large enough, it can be deduced from the proof of Theorem 1 that the diversity performance of the mmWave massive MIMO system only depends on the singular value set $\{\tilde{\alpha}^l\}$ and is not influenced by how sub-matrices $\{\sqrt{g_{ij}}\mathbf{H}_{ij}\}$ are placed in the channel matrix $\mathbf{H}$ (see further discussion of Fig. 10 on this point).

Corollary 1: Consider the scenario that the antenna configuration at each RAU is ULA. Also assume that $L_{ij} = L$ for any i and j. Let $N_s \leq K_r K_t L$. When both N_t and N_r are very large, the distributed massive MIMO system can achieve a diversity gain

$$G_d = K_r K_t L - N_s + 1. \tag{31}$$

In particular, when $K_r = K_t = 1$, the massive MIMO system with co-located antennas arrays can achieve a diversity gain

$$G_d = L - N_s + 1 \tag{32}$$

Remark 3: Corollary 1 implies that for a mmWave co-located massive MIMO system, its diversity gain and multiplexing gain are limited and at most equal to the number of paths L. However, these gains can be increased by employing the distributed antenna architecture and can be scaled up proportionally to $K_r K_t$.

IV. DIVERSITY GAIN ANALYSIS WITH THE CONVENTIONAL PARTIALLY-CONNECTED STRUCTURE

The previous section has analyzed the diversity gain for the massive MIMO system with the general fully-connected RF architecture. This section focuses on a massive MIMO system employing the conventional partially-connected RF architecture as illustrated in Fig. 3. Here the transmitter equipped with K_t RF chains sends N_s data streams to the receiver equipped with K_r RF chains. Each RF chain at the transmitter or receiver is connected to only one RAU. It is assumed that $N_s \leq \min\{K_t, K_r\}$. The numbers of antennas per each RAU at the transmitter and receiver are fixed as N_t and N_r, respectively. Note that $N_t \gg N_s$ and $N_r \gg N_s$. Both the transmitter and receiver employ very small digital processors and very large analog processors, represented respectively by $\mathbf{W}_t$ and $\mathbf{F}_t$ for the transmitter, and $\mathbf{W}_r$ and $\mathbf{F}_r$ for the receiver.

As before, denote by $\mathbf{s}$ the transmitted symbol vector, by $\mathbf{H}$ the fading channel matrix, and by $\mathbf{n}$ the noise vector. Then at the receiver the processed signal vector $\mathbf{z}$ is given by (3), whereas $\mathbf{H}$ is described as in (4). Due to the partially-connected RF architecture, the analog processors $\mathbf{F}_t$ and $\mathbf{F}_r$ are block diagonal matrices, expressed as

$$\mathbf{F}_t = \text{diag}\{\mathbf{f}_{t1}, \mathbf{f}_{t2}, \ldots, \mathbf{f}_{tK_t}\} \qquad (33)$$

and

$$\mathbf{F}_r = \text{diag}\{\mathbf{f}_{r1}, \mathbf{f}_{r2}, \ldots, \mathbf{f}_{rK_r}\} \qquad (34)$$

where $\mathbf{f}_{ti}$ denotes the $N_t \times 1$ steering vector of phases for the ith RAU at the transmitter, and $\mathbf{f}_{rj}$ the $N_r \times 1$ steering vector of phases for the jth RAU at the transmitter.

Theorem 2: Consider the case that the antenna array configuration at each RAU is ULA and $L_{ij} = L$ for any i and j. In the limit of large N_t and N_r, the distributed massive MIMO system with partially-connected RF architecture can achieve a diversity gain

$$G_d = (K_t - N_s + 1)(K_r - N_s + 1)L. \qquad (35)$$

Proof: When N_t and N_r are very large, the diversity gain analysis is similar to that in Theorem 1. For the first data stream that enjoys the best path, it is simple to see that its diversity gain is the largest and equal to $K_r K_t L$. This is because the detection of the first data stream is equivalent to a selection combining system operating with $K_r K_t L$ paths. However, for the second data stream, due to the structure of $\mathbf{F}_t$ and $\mathbf{F}_r$, its detection is equivalent to a selection combining system operating with $(K_r - 1)(K_t - 1)L$ paths. Therefore, it can be inferred that its diversity gain is equal to $(K_r - 1)(K_t - 1)L$. Similarly, for the last data stream among the N_s data streams, its diversity gain is $(K_r - N_s + 1)(K_t - N_s + 1)L$. It then follows that the diversity gain of the whole system is just $(K_r - N_s + 1)(K_t - N_s + 1)L$. $\qquad \square$

Remark 4: Comparing the diversity gain given in Corollary 1 with that given in Theorem 2 reveals that when $N_s = 1$ the diversity gains with the two systems under consideration are the same. However, when $N_s > 1$, the proposed distributed antenna system with fully-connected RF architecture achieves a higher diversity gain than the system with the partially-connected architecture, and the gap between the two diversity gains is $(N_s - 1)[(K_r + K_t - N_s + 1)L - 1]$.

V. DIVERSITY GAIN ANALYSIS FOR THE MULTIUSER SCENARIO

This section considers the downlink communication in a massive multiuser MIMO system as illustrated in Fig. 4. Here the base station (BS) employs K_b RAUs with each having N_b antennas and $N_b^{(\text{rf})}$ RF chains to transmit data streams to K_u mobile stations. Each mobile station (MS) is equipped with N_u antennas and $N_u^{(\text{rf})}$ RF chains to support the reception of its own N_s data streams. This means that there is a total of $K_u N_s$ data streams transmitted by the BS. The numbers of data streams are constrained as $K_u N_s \leq N_b^{(\text{rf})} \leq K_b N_b$ for the BS, and $N_s \leq N_u^{(\text{rf})} \leq N_u$ for each MS.

At the BS, denote by $\mathbf{F}_b$ the $K_b N_b \times N_b^{(\text{rf})}$ RF precoder and by $\mathbf{W}_b$ the $N_b^{(\text{rf})} \times N_s K_u$ baseband precoder. Then under the narrowband flat fading channel model, the received signal vector at the ith MS is given by

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{F}_b \mathbf{W}_b \mathbf{s} + \mathbf{n}_i, \; i = 1, 2, \ldots, K_u \qquad (36)$$

where $\mathbf{s}$ is the signal vector for all K_u mobile stations, which satisfies $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \frac{P}{K_u N_s}\mathbf{I}_{K_u N_s}$ and P is the average transmit power. The $N_u \times 1$ vector $\mathbf{n}_i$ represents additive white Gaussian noise, whereas the $N_u \times K_b N_b$ matrix $\mathbf{H}_i$ is the channel matrix corresponding to the ith MS, whose entries $\mathbf{H}_{ij}$ are described as in Section II. Furthermore, the signal vector after combining can be expressed as

$$\mathbf{z}_i = \mathbf{W}_{ui}^H \mathbf{F}_{ui}^H \mathbf{H}_i \mathbf{F}_b \mathbf{W}_b \mathbf{s} + \mathbf{W}_{ui}^H \mathbf{F}_{ui}^H \mathbf{n}_i, \; i = 1, 2, \ldots, K_u \qquad (37)$$

where F_{ui} is the $N_u \times N_u^{(\text{rf})}$ RF combining matrix and $\mathbf{W}_{ui}$ is the $N_u^{(\text{rf})} \times N_s$ baseband combining matrix for the ith MS.

Theorem 3: Consider the case that all antenna array configurations for the downlink transmission are ULA and $L_{ij} = L$ for any i and j (i.e., all subchannels $\mathbf{H}_{ij}$ have the same number of propagation paths). In the limit of large N_b and N_u, the downlink transmission in a massive MIMO multiuser system can achieve a diversity gain

$$G_d = K_b L - N_s + 1. \qquad (38)$$

Proof: For the downlink transmission in a massive MIMO multiuser system, the overall equivalent multiuser basedband channel can be written as

$$\mathbf{H}_{\text{eq}} = \begin{bmatrix} \mathbf{F}_{u1}^H & 0 & \cdots & 0 \\ 0 & \mathbf{F}_{u2}^H & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{F}_{uK_u}^H \end{bmatrix} \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_{K_u} \end{bmatrix} \mathbf{F}_b. \qquad (39)$$

On the other hand, when both N_b and N_u are very large, both receive and transmit array response vector sets, $\{\tilde{\mathbf{a}}_r(\phi_{ij}^{rl}, \theta_{ij}^{rl})\}$ and $\{\tilde{\mathbf{a}}_t(\phi_{ij}^{tl}, \theta_{ij}^{tl})\}$, are asymptotically orthogonal. Therefore the diversity performance for the ith user depends only on the subchannel matrix $\mathbf{H}_i$ and the choices of $\mathbf{F}_{ui}$ and $\mathbf{F}_b$. The subchannel matrix $\mathbf{H}_i$ has a total of $K_b L$ propagation paths. Similar to the proof of Theorem 1, by employing the
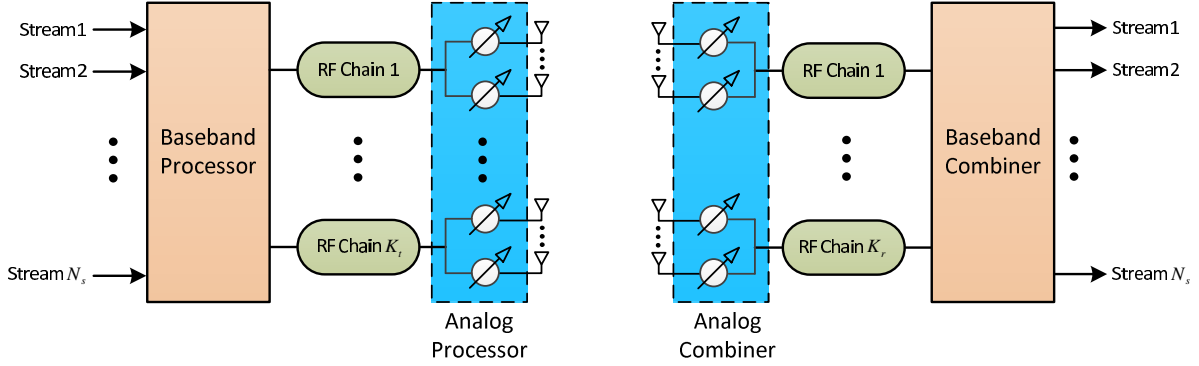
Fig. 3. Block diagram of a mmWave massive MIMO system with the conventional partially-connected RF architecture.
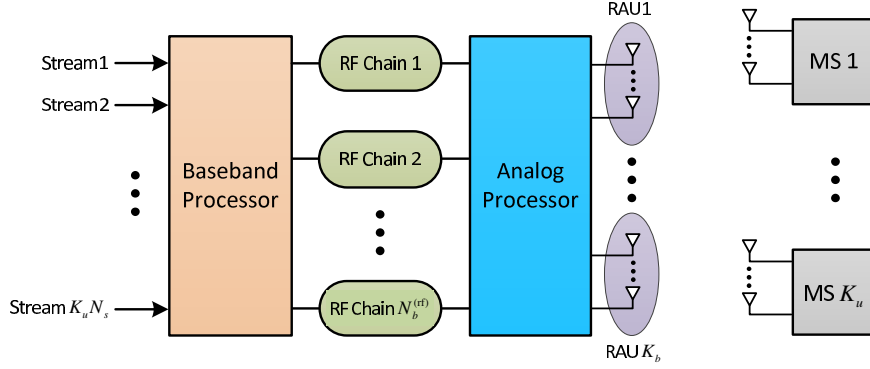

Fig. 4. Block diagram of a multiuser mmWave system with distributed antenna arrays.

optimal RF precoder and combiner for the ith user, the user can achieve a maximum diversity gain $K_b L - N_s + 1$. It is then concluded that the downlink transmission can achieve a diversity gain $G_d = K_b L - N_s + 1$. $\square$

Remark 5: Theorem 3 implies that when N_b and N_u are large enough, the available diversity gain G_d does not depend on the number of mobile users K_u.

Remark 6: In a similar fashion, it is easy to prove that the uplink transmission in a massive MIMO multiuser system can also achieve a diversity gain $G_d = K_b L - N_s + 1$. Moreover, it can also be proved that when $L = 1$, the system diversity gain is equal to $G_d = K_b$ for the case $N_u = 1$, i.e., each MS has only one antenna.

VI. SIMULATION RESULTS

For all simulation results presented in this section, it is assumed that each subchannel matrix $\mathbf{H}_{ij}$ consists of $L_{ij} = L = 3$ paths, each of the large scale fading coefficients g_{ij} equals to $g = -20$ dB (except for Fig. 10), and the numbers of transmit and receive RF chains are twice the number of data streams [7] (i.e., $N_t^{(\mathrm{rf})} = N_r^{(\mathrm{rf})} = 2N_s$). It is further assumed that the variance of AWGN samples is unity and hence the input SNR is the same as the average input power P/N_s. For simplicity, only ULA array configuration with $d = 0.5$ is considered at RAUs and BPSK modulation is employed for each data stream. With such system configurations, the instantaneous BER is given by $Q(\sqrt{2\gamma})$ [32], where γ denotes the instantaneous receive SNR and

the Q-function is defined as $Q(x) = \int_x^\infty \exp\left(-\frac{y^2}{2}\right) \mathrm{d}y$. For ease of comparison and discussion, introduce the concept of *designed* SNR as $\overline{\mathrm{SNR}}_{\mathrm{dg}} = P N_r N_t / (N_s L)$. This means that $P = \overline{\mathrm{SNR}}_{\mathrm{dg}} N_s L / (N_r N_t)$ for a given designed SNR $\overline{\mathrm{SNR}}_{\mathrm{dg}}$. In fact, there exists a power scaling law for mmWave communications which states that the data transmit power P can be scaled down proportionally to $1/(N_r N_t)$ to maintain a desirable BER performance [33].

In all simulations, unless stated otherwise, there are three main steps for hybrid digital-analog processing as follows:

(a) Perform the SVD for channel matrix $\mathbf{H}$ and find the optimal overall digital precoder and combiner for N_s data streams.

(b) Form an analog precoder and an analog combiner based on the optimal overall digital precoder and combiner, respectively.

(c) Perform zero-forcing (ZF) digital detection based on the analog precoder and analog combiner and complete the data detection operation.

First, the singular values of channel matrix $\mathbf{H}$ are studied. Let $K_r = K_t = K$. It is expected that when N_t and N_r are large enough, the number of the effective singular value for the cases $K = 2$ and $K = 1$ should be equal to $L_s = 12$ and $L_s = 3$, respectively. To confirm this, Fig. 5 plots the 1st, the 12th and 13th singular values for $K = 2$, and the 1st, the 3th and 4th singular values for $K = 1$, when N_r increases from 10 to 100, It can be seen from this figure that as N_r increases, all six singular values slowly increases, but the difference at
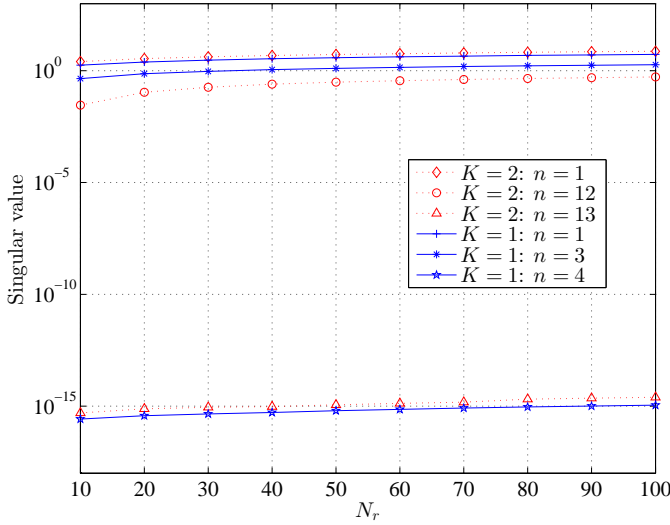
Fig. 5. Behavior of singular values of channel matrix $\mathbf{H}$ for $K = 1$ and $K = 2$.

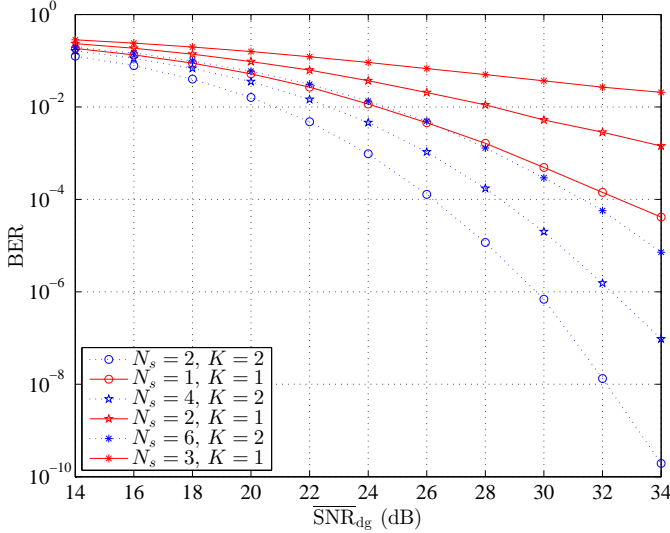Fig. 7. BER versus designed SNR: Verifying diversity gain.

Fig. 6. BER versus designed SNR: Comparison between distributed and co-located antenna array architectures.

$N_r = 10$ and $N_r = 100$ is small. The 13th singular value is very much smaller than the 12th singular value when $K = 2$ and it is almost equal to zero. Likewise, the 4th singular value is much smaller than the 3th singular value when $K = 1$ and it is almost zero. On the other hand, the 12th singular value under $K = 2$ and the 3th singular value under $K = 1$ are quite close to their corresponding largest singular values. Thus this figure verifies that the multiplexing gain is in fact at most equal to L_s as stated in Remark 1.

Studied next is the diversity performance of a mmWave MIMO system with distributed antenna arrays. With $N_r = N_t = N = 50$ and $K_r = K_t = K = 2$, Fig. 6 plots BER curves versus the designed SNR for different numbers of data streams, $N_s = 2, 4, 6$. For comparison, the BER curve obtained in the case of co-located antenna arrays are also plotted for $N_s = 1, 2, 3$. It can be seen that even for the
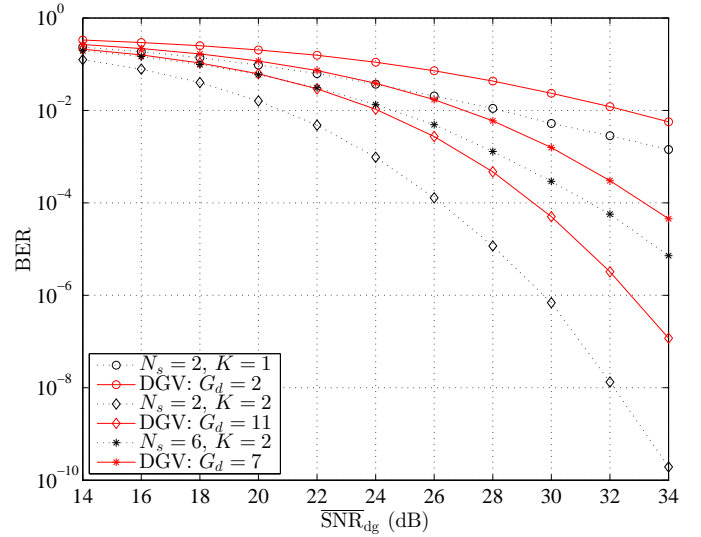
larger number of data streams, the BER performance with distributed antenna arrays is clearly better than that with co-located antenna arrays. Furthermore, as N_s decreases, the BER performance with either distributed or co-located antenna arrays is improved. These observations are expected and agree with Corollary 1, which states that using distributed antenna arrays yields higher diversity gains than using co-located antenna arrays. To verify exactly the diversity gain result given in Corollary 1, Fig. 7 plots diversity gain verifying (GDV) curves produced by simulating the generalized selection combining (GSC) systems. It can be seen that in the high SNR region, a BER curve with either distributed or co-located antenna arrays has the same slope as the corresponding GDV curve.

Illustrated in Fig. 8 is the performance with the conventional partially-connected (PC) RF architecture analyzed in Section IV. With this structure, one first carries out the SVDs for subchannel matrices $\{\mathbf{H}_{ij}\}$ rather than for the whole channel matrix $\mathbf{H}$ and then forms the analog precoder and analog combiner. Let $K_r = K_t = K$. With $N_r = N_t = N = 50$, Fig. 8 plots the BER curves for the following four cases: $(K = 1, N_s = 1)$, $(K = 2, N_s = 2)$, $(K = 3, N_s = 3)$, and $(K = 4, N_s = 4)$. It is known from Theorem 2 that the diversity gains for the four cases are identical and equal to $G_d = L = 3$. To illustrate this, a DGV curve with diversity gain $G_d = 3$ is also plotted in this figure. It can be seen that the system with the conventional PC structure for the four cases can achieve the full diversity gain 3, while the coding gain increases when both K and N_s increase. For comparison, the BER curve obtained with the general fully-connected (FC) RF structure when $N_s = 4$ and $K = 2$ is also plotted. The theoretical limit on the diversity gain in this case is 9, which agrees well with the DGV curve having $G_d = 9$. Observe that in the high SNR region the general FC structure yields significantly better diversity performance than the conventional PC structure.

Next, when $N_s = 1$, we consider the diversity performance with the multiuser downlink scenario where there are 5 or 10
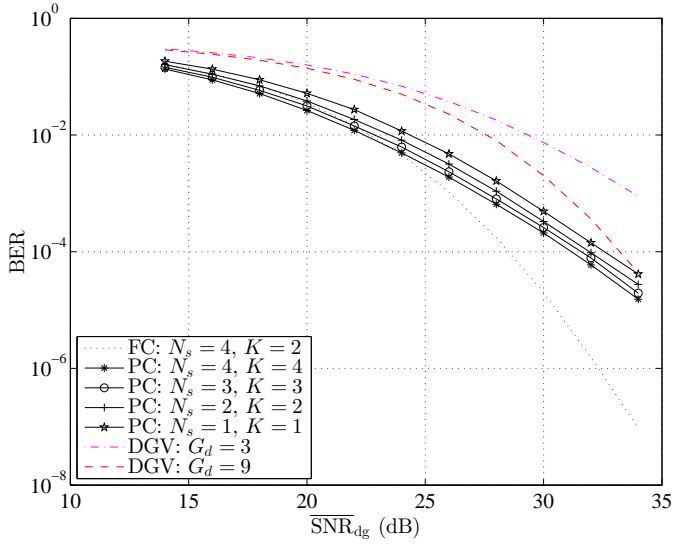
Fig. 8. BER versus designed SNR: Conventional partially-connected architectures with different numbers of data streams.

Fig. 9. BER versus designed SNR: Multiuser scenario with different numbers of subarrays.

mobile users, each having 10 antennas and each RAU at the BS is equipped with 50 antennas. Due to the fact that there is no cooperation among the users, one first carries out the SVDs for subchannel matrices $\{\mathbf{H}_i\}$ rather than for the whole channel matrix $\mathbf{H}$ and then forms the analog precoder for the BS and analog combiners for the users. Note that the BS needs to carry out ZF digital preprocessing before transmitting data. Fig. 9 plots the BER curves versus the designed SNR for different numbers of subarrays at the BS, namely $K_b = 1, 3, 5$. It can be observed from this figure that as K_b increases, the diversity performance of the multiuser system improves remarkably. This is because, as established in Theorem 3, the diversity gain becomes larger with increasing K_b. Furthermore, it can be seen from Fig. 9 that the system has the same diversity gain for different numbers of users while the coding gain increases as K_u decreases. This observation agrees with Remark 5.

Finally, the diversity performance of the single-user mmWave massive MIMO system is examined under the scenario that the distributions of large scale fading coefficients, $\{g_{ij}\}$, are inhomogeneous. To this end, let $\mathbf{G} = [g_{ij} \ (\mathrm{dB})]$ denote the large scale fading coefficient matrix. When $N_r = N_t = N = 50$ and $K_r = K_t = K = 2$, simulation is performed for the following six inhomogeneous $\mathbf{G}$:

$$\mathbf{G}_1 = \begin{bmatrix} -25 & -20 \\ -20 & -25 \end{bmatrix}, \ \mathbf{G}_2 = \begin{bmatrix} -20 & -20 \\ -25 & -25 \end{bmatrix},$$

$$\mathbf{G}_3 = \begin{bmatrix} -20 & -25 \\ -25 & -20 \end{bmatrix}, \ \mathbf{G}_4 = \begin{bmatrix} -20 & -25 \\ -20 & -25 \end{bmatrix},$$

$$\mathbf{G}_5 = \begin{bmatrix} -25 & -25 \\ -20 & -20 \end{bmatrix}, \ \mathbf{G}_6 = \begin{bmatrix} -25 & -20 \\ -25 & -20 \end{bmatrix}.$$

It can be found that the diversity performance for the six inhomogeneous cases are almost the same (see Remark 2). In order to illustrate this interesting phenomenon, Fig. 10 plots the BER curves versus the designed SNR with $\mathbf{G}_1$ and $\mathbf{G}_2$, respectively. For comparison, the two BER curves for the homogeneous distributions with $g = -20\mathrm{dB}$ and $g = -25\mathrm{dB}$
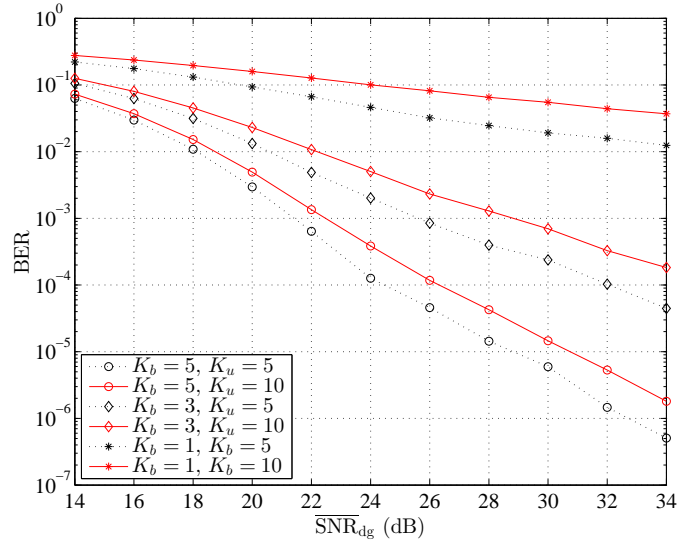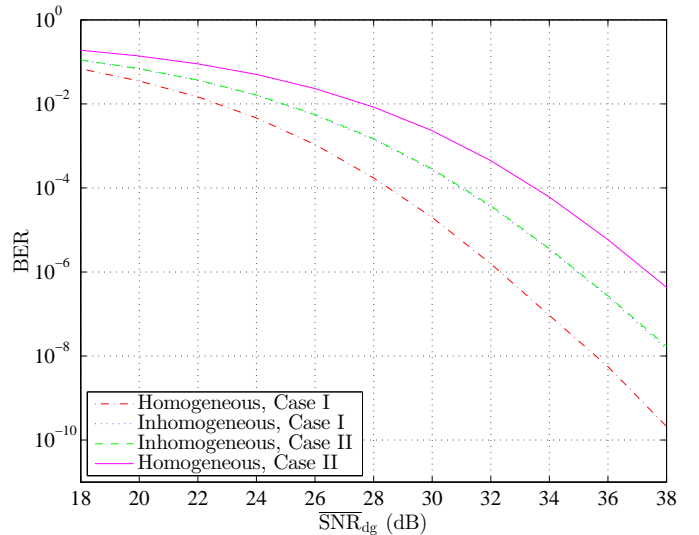

Fig. 10. BER versus designed SNR: Comparison between homogeneous distributions and inhomogeneous distributions for large scale fading coefficients $\{g_{ij}\}$.

are also plotted. As expected, the BER curves with the inhomogeneous coefficient distributions are between the two BER curves with homogeneous coefficient distributions. It can be concluded from this figure that the case of inhomogeneous coefficient distributions has the same diversity gain as in the case of homogeneous coefficient distributions.

VII. CONCLUSIONS

This paper has provided asymptotical diversity analysis for mmWave massive MIMO systems with co-located and distributed antenna architectures when the number of antennas at each subarray goes to infinity. Theoretical analysis shows that with a co-located massive antenna array, scaling up the number of antennas of the array can increase the coding gain (array gain), but not the diversity gain. However, if the array

is built from distributed subarrays (RAUs), each having a very large number of antennas, then increasing the number of RAUs does increase the diversity gain and/or multiplexing gain. As such, the analysis leads to a novel approach to improve the diversity and multiplexing gains of mmWave massive MIMO systems. It is acknowledged that the asymptotical diversity analysis obtained in this paper is under the idealistic assumption of having perfect CSI. Performing the diversity analysis for mmWave massive MIMO systems under imperfect CSI is important and deserves further research.

REFERENCES

[1] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access,* vol. 1, pp. 335-349, May 2013.

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

[3] W. Roh *et al.*, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results," *IEEE Commun. Mag.,* vol. 52, no. 2, pp. 106-113, Feb. 2014.

[4] O. E. Ayach, R. W. Heath, S. Abu-Surra, S. Rajagopal and Z. Pi, "The capacity optimality of beam steering in large millimeter wave MIMO systems," *in Proc. IEEE 13th Intl. Workshop on Sig. Process. Advances in Wireless Commun. (SPAWC),* pp. 100-104, June 2012.

[5] L. Liang, W. Xu and X. Dong, "Low-complexity hybrid precoding in massive multiuser MIMO systems," *IEEE Wireless Commun. Letters,* vol.3, no.6, pp. 653-656, Dec. 2014.

[6] W. Ni and X. Dong, "Hybrid Block Diagonalization for Massive Multiuser MIMO Systems," *IEEE Transactions on Communications,* vol.64, no.1, pp.201-211, Jan. 2016.

[7] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE Journal of Selected Topics in Signal Processing,* vol. 10, no. 3, pp. 501-513, Apr. 2016.

[8] O. E. Ayach, R. W. Heath, S. Rajagopal, and Z. Pi, "Multimode precoding in millimeter wave MIMO transmitters with multiple antenna sub-arrays," *in Proc. IEEE Glob. Commun. Conf.,* 2013, pp. 3476-3480.

[9] J. Singh and S. Ramakrishna, "On the feasibility of beamforming in millimeter wave communication systems with multiple antenna arrays," *in Proc. IEEE GLOBECOM,* 2014, pp. 3802-3808.

[10] J. A. Zhang, X. Huang, V. Dyadyuk, and Y. J. Guo, "Massive hybrid antenna array for millimeter-wave cellular communications," *IEEE Wireless Commun.,* vol. 22, no. 1, pp. 79-87, Feb. 2015.

[11] S. He, C. Qi, Y. Wu, and Y. Huang, "Energy-efficient transceiver design for hybrid sub-array architecture MIMO systems," *IEEE Access,* vol. 4, pp. 9895-9905, 2016.

[12] N. Li, Z. Wei, H. Yang, X. Zhang, D. Yang, "Hybrid Precoding for mmWave Massive MIMO Systems With Partially Connected Structure," *IEEE Access,* vol. 5, pp. 15142-15151, 2017.

[13] D. Zhang, Y. Wang, X. Li, W. Xiang, "Hybridly-Connected Structure for Hybrid Beamforming in mmWave Massive MIMO Systems," *IEEE Transactions on Communications,* DOI 10.1109/TCOMM.2017.2756882, Sep. 2017.

[14] N. Song, T. Yang, and H. Sun, "Overlapped subarray based hybrid beamforming for millimeter wave multiuser massive MIMO," *IEEE Signal Processing Letters,* vol. 24, no. 5, pp. 550-554, May 2017.

[15] S. Kutty and D. Sen, "Beamforming for millimeter wave communications: an inclusive surey," *IEEE Communications Surveys & Tutorials,* vol. 18, no.2, pp. 949-973, Second Quarter 2016.

[16] A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda, "Hybrid beamforming for massive MIMO-a survey," *arXiv preprint arXiv:* 1609.05078, 2016.

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

[18] T. S. Rappaport, F. Gutierrez, E. Ben-Dor, J. N. Murdock, Y. Qiao, and J. I. Tamir, "Broadband millimeter-wave propagation measurements and models using adaptive-beam antennas for outdoor urban cellular communications," *IEEE Trans. Antennas Propag.,* vol. 61, no. 4, pp. 1850-1859, Apr. 2013.

[19] M. V. Clark, T. M. W. III, L. J. Greenstein, A. J. Rustako, V. Erceg, and R. S. Roman, "Distributed versus centralized antenna arrays in broadband wireless networks," *in Proc. IEEE Veh. Technology Conf. (VTC'01),* May 2001, pp. 33-37.

[20] W. Roh and A. Paulraj, "MIMO channel capacity for the distributed antenna systems," *in IEEE Veh. Technology Conf. (VTC'02),* vol. 3, Sept. 2002, pp. 1520-1524.

[21] L. Dai, "A comparative study on uplink sum capacity with co-located and distributed antennas," *IEEE J. Sel. Areas Commun.,* vol. 29, no. 6, pp. 1200-1213, June 2011.

[22] Q.Wang,D. Debbarma, A. Lo, Z. Cao, I. Niemegeers, S. H. de Groot, "Distributed antenna system for mitigating shadowing effect in 60 GHz WLAN,", *wireless Personal Communications,* vol. 82, no. 2, pp. 811-832, May 2015.

[23] S. Gimenez, D. Calabuig, S. Roger, J. F. Monserrat, and N. Cardona, "Distributed hybrid precoding for indoor deployments using millimeter wave band," *Mobile Information Systems,* vol. 2017, Article ID 5751809, 12 pages, Oct. 2017.

[24] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. Inform. Theory,* vol. 49, pp. 1073-1096, May 2003.

[25] D. Tse, P. Viswanath, and L. Zheng, "Diversity-multiplexing tradeoff in multiple access channels," *IEEE Trans. Inform. Theory,* vol. 50, pp. 1859-1874, Sep. 2004.

[26] M. Yuksel and E. Erkip, "Multiple-antenna cooperative wireless systems: A diversity-multiplexing tradeoff perspective," *IEEE Trans. Inform. Theory,* vol. 53, pp 3371-3393, Oct. 2007.

[27] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge, U.K.: Cambridge Univ. Press, 2007.

[28] L. G. Ordóñz, , D. P. Palomar, A. Pagès-Zamora, and J. R. Fonollosa, "High-SNR Analytical Performance of Spatial Multiplexing MIMO Systems With CSI," *IEEE Transactions on Signal Processing,* Vol.55, no. 11, pp. 5447-5463, Nov. 2007

[29] Z. Wang and G. B. Giannakis, "A simple and general parameterization quantifying performance in fading channels," *IEEE Trans. Signal Process.,* vol. 51, no. 8, pp. 1389-1398, Aug. 2003.

[30] H. A. David, *Order Statistics,* New York, NY: John Wiley & Sons, Inc., 1981.

[31] J. Chen, "When does asymptotic orthogonality exist for very large arrays?", *Global Commun. Conf..,* Dec. 2013, pp. 4251-4255.

[32] A. Goldsmith, *Wireless Communications,* Chapter 6, Cambrige University Press, UK, 2005.

[33] D. -W. Yue, Y. Zhang, and Y. N. Jia, "Beamforming based on specular component for massive MIMO systems in Ricean fading," *IEEE Wireless Commun. Lett.,* vol. 4, no. 2, pp. 197-200, Jan. 2015.