Randomly-Directional Beamforming in Millimeter-Wave Multiuser MISO Downlink

Gilwon Lee, *Student Member, IEEE*, Youngchul Sung, *Senior Member, IEEE*, and Junyeong Seo, *Student Member, IEEE*

***Abstract*—In this paper, the performance of opportunistic random beamforming (RBF) and the multiuser (MU) gain in millimeter-wave (mm-wave) MU multiple-input single-output (MISO) downlink systems are analyzed based on the uniform random single-path (UR-SP) channel model suitable for highly directional mm-wave radio propagation channels. It is shown that under the UR-SP channel model, RBF achieves linear sum rate scaling with respect to (w.r.t.) the number of transmit antennas and, furthermore, yields optimal sum rate performance when the number of transmit antennas is large, if the number of users increases linearly w.r.t. the number of transmit antennas. Several beam training and user selection methods are investigated to yield insights into the most effective beamforming and scheduling choice for mm-wave MU-MISO in various operating conditions. Simulation results validate our analysis based on asymptotic techniques for finite cases.**

***Index Terms*—Millimeter-Wave, Multi-User MIMO, Massive MIMO, Opportunistic Random Beamforming, Randomly-Directional Beamforming.**

I. Introduction

Motivation: Recently, mm-wave multiple-input multiple-output (MIMO) operating in the band of 30-300 GHz is considered as a promising technology to attain high data rates for 5G wireless communications. Radio propagation in the mm-wave band has several intrinsic properties; the propagation in the mm-wave band is highly directional with large path loss and very few multi-paths. To compensate for the large path loss in the mm-wave band, highly directional beamforming is required based on large antenna arrays which can easily be implemented in the mm-wave band due to small wavelength. To perform highly directional downlink beamforming to a user in the cell, accurate channel state information (CSI) is required at the base station (BS). However, the channel is sparse in the arrival angle domain and downlink channel estimation is difficult [4]–[6]. That is, it is difficult to identify the sparse propagation angle and gain between the BS and an arbitrary receiver in the cell, and identifying the sparse channel in the angle domain requires sophisticated algorithms and heavy training overhead [4]–[7]. However, the focus of the existing channel estimation methods is single-user mm-wave MIMO systems which do not have MU diversity. Suppose directional downlink beamforming with a large uniform linear array (ULA) of transmit antennas at the BS. Although the downlink beam is highly directional, it still has some beam width because the number of transmit antennas is finite in practice. Thus, one might ask what happens if there are many users in the cell and the BS just selects the transmission beam direction randomly in the angle domain and looks for a receiver that happens to be in the beam width of the selected beam of the BS. Of course, if there exists only a single receiver in the cell, such *randomly-directional beamforming (RDB)* with a narrow beam width will not perform well because it will miss the receiver in most cases. However, if there exist more than one receivers randomly located in the cell, the RDB scheme may perform reasonably well with a sufficient number of users in the cell. Then, a natural question is "how many users in the cell are enough for reasonable performance of such simple RDB and RDB with multiple beams, so-called RBF [8], in the mm-wave band?" In this paper, we investigate the performance of RDB and the associated MU gain in the mm-wave band to answer the above question.

Channel model for mm-wave MIMO systems : Since the performance of RDB (or RBF) depends on the channel model, answering the above question should be based on a meaningful channel model. In conventional lower band MIMO communication, many MU gain analyses were performed with the assumption of rich scattering, i.e., mostly under the independent and identically distributed (i.i.d.) Rayleigh fading channel model or its variants such as correlated fading or one-ring channel model [8]–[17]. However, the propagation in the mm-wave band is quite different from that in the lower band; propagation in the mm-wave band is highly directional and there are very few multi-paths in propagation channels [4], [6], [7], [18]. To model wireless channels in the mm-wave band, *the UR-SP channel model* was proposed in [19], [20]. The UR-SP channel model well captures the highly directional propagation in the mm-wave band and is still analytically tractable [19], [20]. Such single-path channel models were widely-used for performance analysis of mm-wave systems [19]–[25]. Under the

Manuscript received May 1, 2015; revised August 3, 2015 and September 14, 2015; accepted September 24, 2015. Date of publication September 28, 2015; date of current version February 8, 2016. This work was supported in part by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant 2013R1A1A2A10060852, in part by the Cross-Ministry Giga KOREA Project of The Ministry of Science, ICT and Future Planning, Korea (GK14N0100, 5G mobile communication system development based on mmWave). A related paper of this work was presented at IEEE SPAWC, Stockholm, Sweden, June 2015 [1] and an extended version of this work was submitted to the IEEE Journal on Selected Topics in Signal Processing [2], and a portion of this work was presented at IEEE ITW, Jeju, Korea, October 2015 [3]. The associate editor coordinating the review of this paper and approving it for publication was Bala Natarajan.

The authors are with Department of Electrical Engineering, KAIST, Daejeon 305-701, South Korea (e-mail: gwlee@kaist.ac.kr; ycsung@kaist.ac.kr; and jyseo@kaist.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2015.2483493

UR-SP channel model, the channel vector of each user in the cell has a single path component with a random direction (or angle) and a random path gain. Since there is only one path in each user's channel under the UR-SP channel model, the UR-SP channel model is a simplified channel model capturing the dominant path in directional propagation environments. To gain insights into random beamforming in the highly-directional mm-wave band and make performance analysis tractable, we adopt the UR-SP channel model in this paper even though the actual channel may lie somewhere between the UR-SP channel model and the i.i.d. Rayleigh fading channel model.

Summary of Results: The MU gain under rich scattering environments has been investigated extensively during the last decade [8]–[17]. However, not much work has been done yet regarding the MU gain in mm-wave MU-MISO/MIMO systems. Recently, in [20], Ngo *et al.* simplified the UR-SP channel model as an urns-and-balls model and numerically showed that user scheduling can improve the worst-user performance. This work provides an intuitive and insightful observation regarding the MU gain in mm-wave MU-MISO, but the urns-and-balls channel model seems a bit oversimplified compared to the UR-SP channel model since the urns-and-balls model does not consider non-orthogonal regions of UR-SP and ignores inter-beam interference. In this paper, we rigorously analyze the RDB scheme, the associated MU gain, and user scheduling in mm-wave MU-MISO in an asymptotic regime in which the number of transmit antennas tends to infinity, under the UR-SP channel model capturing inter-beam interference and the assumption of an ULA at the BS, and provide guidelines for optimal operation in highly directional mm-wave MU-MISO systems. The main result of this paper is that under the UR-SP channel model, RBF achieves linear sum rate scaling w.r.t. the number of transmit antennas as the number of users increases linearly w.r.t. the number of transmit antennas, and furthermore it yields optimal sum rate performance for MU-MISO downlink with a large number of transmit antennas only with a number of users in the cell that is in the same order of the number of transmit antennas. This result is contrary to the existing result in rich scattering environments that opportunistic random beamforming does not provide a gain for MU-MISO in the regime of a large number of transmit antennas under rich scattering environments [8], [10], [12]. The fundamental reason for the performance difference of RBF is that the degree-of-freedom in the UR-SP channel model is *one* regardless of the number of transmit antennas. Thus, in this model, the orthogonality of the multiple transmit beams can be attained by simply dividing the 1-dimensional line of the normalized angle by line segments. Hence, the dimension of the search space for orthogonality is reduced from the number of transmit antennas to one when considering the UR-SP channel model instead of the Rayleigh fading channel model.

Notations and Organization: Vectors and matrices are written in boldface with matrices in capitals. For a matrix $\mathbf{A}$, $\mathbf{A}^T$, $\mathbf{A}^H$, and $\text{tr}(\mathbf{A})$ indicate the transpose, conjugate transpose, and trace of $\mathbf{A}$, respectively. $\mathbf{I}_n$ stands for the identity matrix of size n. (The subscript will be omitted if unnecessary.) The notation $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ means that $\mathbf{x}$ is complex Gaussian distributed with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, and $\theta \sim \text{Unif}[a, b]$ means that θ is uniformly distributed over the range $[a, b]$. $\mathbb{E}[\cdot]$ denotes the expectation. $\iota := \sqrt{-1}$ and $\mathbb{Z}$ is the set of integers. $a \uparrow b$ indicates that a converges to b from the below.

The remainder of this paper is organized as follows. In Section II, the system model and preliminaries are described. In Section III, the considered RDB scheme is explained. The asymptotic performance is analyzed for the single beam case in Section IV and for the multiple beam case with single user selection or multiple user selection in Section V. Numerical results are provided in Section VI, followed by conclusions in Section VII.

II. System Model and Preliminaries

We consider a single-cell mm-wave MU-MISO downlink system in which a BS equipped with an ULA of M transmit antennas communicates with K single-antenna users. The received signal at user k is then given by

$$y_k = \mathbf{h}_k^H \mathbf{x} + n_k, \quad k = 1, 2, \cdots, K, \tag{1}$$

where $\mathbf{h}_k = [h_{k,1}, h_{k,2}, \cdots, h_{k,M}]^T$ is the channel vector of user k, $\mathbf{x}$ is the transmitted signal vector subject to a power constraint $\text{tr}(\mathbb{E}\{\mathbf{x}\mathbf{x}^H\}) \le P_t$, and $n_k \sim \mathcal{CN}(0, 1)$ is the additive noise at user k.

A. Channel Model

For a typical mm-wave channel, there exist very few multipaths due to the highly directional and quasi-optical nature of electromagnetic wave propagation in the mm-wave band. A general mm-wave channel is composed of a line-of-sight (LoS) propagation component and a set of few single-bounce non-LoS (NLoS) components, and hence the mm-wave channel for ULA systems can be modeled as

$$\mathbf{h}_k = \alpha_{k,LOS}\sqrt{M}\mathbf{a}\left(\theta_{k,LOS}\right) + \sum_i \alpha_{k,NLOS}^{(i)}\sqrt{M}\mathbf{a}\left(\theta_{k,NLOS}^{(i)}\right), \tag{2}$$

for $k = 1, \cdots, K$, where $\alpha_{k,LOS}$ and $\theta_{k,LOS}$ are the complex gain and normalized direction of the LoS path for user k, $\{\alpha_{k,NLOS}^{(i)}\}$ and $\{\theta_{k,NLOS}^{(i)}\}$ represent the complex gains and normalized directions of NLoS paths for user k, and $\mathbf{a}(\theta)$ is the array steering vector given by

$$\mathbf{a}(\theta) = \frac{1}{\sqrt{M}}\left[1, e^{-\iota\pi\theta}, \cdots, e^{-\iota\pi(M-1)\theta}\right]^T. \tag{3}$$

Here, the normalized direction θ is connected with the physical angle of departure $\phi \in [-\pi/2, \pi/2]$ as $\theta = \frac{2d\sin(\phi)}{\lambda}$, where d and λ are the distance between two adjacent antennas and the carrier wavelength, respectively. We assume the critically-sampled environment, i.e., $\frac{d}{\lambda} = \frac{1}{2}$ in this paper. Note that the array steering vector in (3) has unit norm and thus the normalization factor $\sqrt{M}$ is included in (2).

For mm-wave channels with LoS links, the effect of NLoS links is marginal since the path loss of NLoS components is much larger than that of the LoS component; the power $|\alpha_{k,i}|^2$ associated with NLoS paths is typically 20 dB weaker than the LoS component $|\alpha_k|^2$ [18]. However, even when there is no LoS path, it is suggested that communication is feasible through a NLoS path with highly directional beamforming [18]. In both cases, there exist very few paths and hence, in many works for mm-wave MIMO [19]–[26] researchers adopted the single-path (SP) model considering one dominant path which is the LoS path if a LoS path exists or the dominant NLoS path if a LoS does not exist. Here, to make analysis tractable and gain insights into the MU gain in highly-directional mm-wave MIMO, we also adopt the SP model considering the dominant path. We assume that the dominant path gain is Gaussian-distributed, i.e., $\alpha_k \stackrel{i.i.d.}{\sim} \mathcal{CN}(0, 1)$ (this captures that the dominant path can be a LoS path or a NLoS path) and that the normalized direction θ_k for each user k is independent and identically distributed (i.i.d.) with $\theta_k \stackrel{i.i.d.}{\sim}$ Unif$[-1, 1]$. From the above assumptions, the considered mm-wave channel model is expressed as

$$\mathbf{h}_k = \alpha_k \sqrt{M} \mathbf{a}(\theta_k), \quad \text{for } k = 1, \cdots, K, \tag{4}$$

where α_k and θ_k are the path gain and angle of the dominant path for user k. This channel model is the UR-SP model considered in [19], [20]. In this paper, we adopt this channel model. Note that the power of the UR-SP channel model (4) is given by $\mathbb{E}\{\|\mathbf{h}_k\|^2\} = M$. Thus, the channel power linearly increases w.r.t. M as in the i.i.d. Rayleigh channel model $\mathbf{h}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. This means that the power radiated in the space is collected by the receiver antenna.

B. Review of Opportunistic Random Beamforming in Rich Scattering Environments

Before introducing the considered RDB for large mm-wave MIMO systems with the UR-SP channel model, we briefly review the RBF scheme in [8] and its performance in rich scattering environments under which each element $h_{k,j}$ in the channel vector $\mathbf{h}_k$ has an i.i.d. Rayleigh fading:

$$h_{k,j} \stackrel{i.i.d.}{\sim} \mathcal{CN}\left(0, \sigma_h^2\right) \quad \text{for } j = 1, \cdots, M. \tag{5}$$

In the RBF scheme, the BS constructs a set of S random orthonormal beam vectors $\{\mathbf{u}_1, \cdots, \mathbf{u}_S\}$ and transmits each beam sequentially to the K users in the cell during the training period. Then, each user k computes the signal-to-interference-plus-noise ratio (SINR) for each beam direction at the end of the training period, given by $\text{SINR}_{k,i} = \frac{\frac{P_t}{S}|\mathbf{h}_k^H \mathbf{u}_i|^2}{1+\frac{P_t}{S}\sum_{j\neq i}|\mathbf{h}_k^H \mathbf{u}_j|^2}$ for $i = 1, \cdots, S$. After the training period, each user k feeds back its maximum SINR value, i.e., $\max_{1\le i\le S} \text{SINR}_{k,i}$, and the beam index i at which the SINR is maximum. Then, after the feedback the BS assigns each beam i to the user $k'(i)$ with the highest SINR for beam i, i.e., $k'(i) = \arg\max_{1\le k\le K} \text{SINR}_{k,i}$, and transmits S data streams to the selected S users. In [8], Sharif and Hassibi derived several scaling laws of this RBF scheme in the case of $S = M$ with the *small-scale*[1] MIMO in mind, i.e., $M \ll K$, as $K \to \infty$. Specifically, they showed

$$\mathcal{R}_{RBF} \sim_K \begin{cases} M \log\log K, & \text{for fixed } M, \\ cM, & \text{for } M = O(\log K), \end{cases} \tag{6}$$

where $\mathcal{R}_{RBF} = \mathbb{E}[\sum_{i=1}^{M} \log(1 + \max_{1\le k\le K} \text{SINR}_{k,i})]$ and c is a positive constant. (Here, $x \sim_K y$ indicates that $\lim_{K\to\infty} x/y = 1$.) Furthermore, they showed that [8]

$$\lim_{K\to\infty} \frac{\mathcal{R}_{RBF}}{M} = 0, \tag{7}$$

if $\lim_{K\to\infty} \frac{M}{\log K} = \infty$ ($\lim_{K\to\infty} \frac{M}{\log K} = \infty$ is equivalent to $\lim_{K\to\infty} \frac{\log K}{M} = 0$). The above scaling laws state that the sum rate of the RBF scheme maintains linear scaling w.r.t. the number M of transmit antennas when M grows no faster than $\log K$ as $K \to \infty$, but this linear scaling w.r.t. M is not achieved when M grows faster than $\log K$ as $K \to \infty$. That is, the RBF scheme performs well, i.e., the RBF data rate grows linearly w.r.t. the number M of antennas in small-scale MIMO systems with a large number of users in the cell, but does not show linear scaling rate w.r.t. M in massive MIMO situations under rich scattering environments.

Now consider the case of mm-wave MIMO. Due to large path loss in the mm-wave band, highly directional beamforming is required to compensate for the large path loss. This means a large antenna array at the BS, i.e., M is very large. In the following sections, we investigate the performance of RBF under the UR-SP channel model in a progressive manner from one single random beam and single user selection to multiple random (asymptotically-orthogonal) beams and multiple user selection under a massive MIMO asymptote in which M tends to infinity.

III. Randomly-Directional Beamforming in Massive mm-Wave MISO

First, we consider the RDB strategy in the single beam downlink transmission case. In this case, during the training period, the BS chooses a normalized direction ϑ randomly and transmits the beam $\mathbf{x}$ in (1) given by

$$\mathbf{x} = \mathbf{a}(\vartheta) \tag{8}$$

where $\vartheta \sim$ Unif$[-1, 1]$ and $\mathbf{a}(\theta)$ is given by (3). (We set $P_t = 1$ for simplicity here.) Then, each user k in the cell composed of K users feeds back the average received power[2] $|\bar{y}_k|^2$ ($\approx |\mathbf{h}_k \mathbf{x}|^2 + \frac{1}{N_s}$) to the BS, where $|\mathbf{h}_k^H \mathbf{x}|^2 = |\alpha_k|^2 \cdot M|\mathbf{a}(\theta_k)^H \mathbf{a}(\vartheta)|^2$. After the feedback period is over, the BS selects the user that has maximum signal power and transmits a data stream with the beamforming vector $\mathbf{x}$ in (8) to the user. Then, the expected rate $\mathcal{R}_1$ of the RDB scheme is given by

$$\mathcal{R}_1 = \mathbb{E}\left[\log\left(1 + \max_{1\le k\le K} |\alpha_k|^2 M |\mathbf{a}(\theta_k)^H \mathbf{a}(\vartheta)|^2\right)\right], \tag{9}$$

[1] In small-scale MIMO systems, M is small and K is relatively large. Hence, the authors of [8] focused on the asymptotic scenario in which K grows to infinity with fixed M or M growing much slower than K. Note that $K = \Theta(e^M)$ for K as a function of M for the scaling of $M = \Theta(\log K)$ considered in [8].

[2] To average out the noise effect, each user can have multiple time samples $y_k(i)$ during the training period and average the multiple samples for the feedback value $|\bar{y}_k|^2 = |\frac{1}{N_s}\sum_{i=1}^{N_s} y_k(i)|^2 \stackrel{(a)}{=} |\mathbf{h}_k^H \mathbf{x}|^2 + \frac{1}{N_s}$. We assume that sufficient sample average is done and will ignore possible error in step (a) in this paper.

where the expectation is over $\mathbf{h}_k$ and $\mathbf{x}$.

Consider the case of $K = 1$, i.e., only one user in the cell. In this case, we have an upper bound on $\mathcal{R}_1$ from Jensen's inequality as

$$\begin{aligned}\mathcal{R}_1 &= \mathbb{E}\left[\log\left(1 + |\alpha_1|^2 M|\mathbf{a}(\theta_1)^H \mathbf{a}(\vartheta)|^2\right)\right]\\ &\le \log\left(1 + \mathbb{E}\left[|\alpha_1|^2 M|\mathbf{a}(\theta_1)^H \mathbf{a}(\vartheta)|^2\right]\right)\\ &= \log\left(1 + \mathbb{E}\left[|\alpha_1|^2\right]\mathbb{E}\left[M|\mathbf{a}(\theta_1)^H \mathbf{a}(\vartheta)|^2\right]\right) = \log 2.\end{aligned}$$

The last equality holds from $\mathbb{E}[|\alpha_1|^2] = 1$ because $|\alpha_1|^2$ has a chi-square distribution with degree-of-freedom two, i.e., $|\alpha_1|^2 \sim \chi^2(2)$ and from

$$\begin{aligned}\mathbb{E}[M|\mathbf{a}(\theta_1)^H \mathbf{a}(\vartheta)|^2] &= \frac{1}{M}\mathbb{E}\left[\left|\sum_{n=0}^{M-1} e^{-\iota\pi n(\vartheta-\theta_1)}\right|^2\right]\\ &= \frac{1}{M}\mathbb{E}\left[M + \sum_{\substack{n,m\\ n\neq m}} e^{-\iota\pi(m-n)(\vartheta-\theta_1)}\right]\\ &\overset{(a)}{=} 1,\end{aligned}$$

where step (a) holds because $\mathbb{E}[e^{-j\pi(m-n)(\vartheta-\theta_1)}] = \frac{1}{2}\int_{-1}^{1} e^{-j\pi(m-n)\tilde{\theta}_k} d\tilde{\theta}_k = \frac{\sin\pi(m-n)}{\pi(m-n)} = 0$ for any $(m-n) \in \mathbb{Z}\backslash\{0\}$ [20].[3] Furthermore, it will be shown in the next section that $\mathcal{R}_1$ actually goes to zero as $M \to \infty$. Thus, the rate of the RDB scheme for $K = 1$ is insignificant. In this case, it is imperative to obtain the CSI of the single user to achieve the attainable rate of $\log(1 + |\alpha_1|^2 M) \sim_M \log M$ [4]–[6].

However, the situation becomes different as K becomes large. To see this, we need to represent the key term $|\mathbf{a}(\theta_k)^H \mathbf{a}(\vartheta)|$ in this paper in an explicit form. Under the UR-SP channel model, this inner product $|\mathbf{a}(\theta_k)^H \mathbf{a}(\vartheta)|$ between the transmission beam $\mathbf{a}(\vartheta)$ with the direction angle ϑ and the channel of user k located at angle θ_k plays a key role and defines the beam pattern associated with transmission beam $\mathbf{a}(\vartheta)$. The inner product $|\mathbf{a}(\theta_k)^H \mathbf{a}(\vartheta)|$ is explicitly given by

$$\begin{aligned}|\mathbf{a}(\theta_k)^H \mathbf{a}(\vartheta)| &= \frac{1}{M}\left|\sum_{n=0}^{M-1} e^{-\iota\pi n(\vartheta-\theta_k)}\right| = \frac{1}{M}\left|\frac{1-e^{-\iota\pi(\vartheta-\theta_k)M}}{1-e^{-\iota\pi(\vartheta-\theta_k)}}\right|\\ &= \frac{1}{M}\left|\frac{\sin\frac{\pi(\vartheta-\theta_k)M}{2}}{\sin\frac{\pi(\vartheta-\theta_k)}{2}}\right| =: F_M(\vartheta-\theta_k), \qquad (10)\end{aligned}$$

which is called the *Fejér kernel* $F_M(\cdot)$ of order M [27]. Fig. 1 shows the value of (10) versus $\vartheta - \theta_k$. Note that the beam pattern ripples in a roughly diminishing manner as θ_k goes away from the beam center angle ϑ, and the pattern has nulls at the angles that are multiples of $2/M$. The Fejér kernel beam pattern is the standard beam pattern of ULAs [28, Chs. 2 and 3]. Typically, the non-zero region around the beam center angle ϑ is called the main lobe and the non-zero regions between nulls are called side-lobes. From (10), we have $|\mathbf{a}(\theta_k)^H \mathbf{a}(\vartheta)| \to 0$ as $M \to \infty$ for fixed ϑ and θ_k. On the other hand, we have $|\mathbf{a}(\theta_k)^H \mathbf{a}(\vartheta)| \to \left|\frac{2\sin\frac{\pi\Delta}{2}}{\pi\Delta}\right|$ as $M \to \infty$, if $\vartheta - \theta_k = \frac{\Delta}{M}$ for some $\Delta > 0$ [20]. This is because

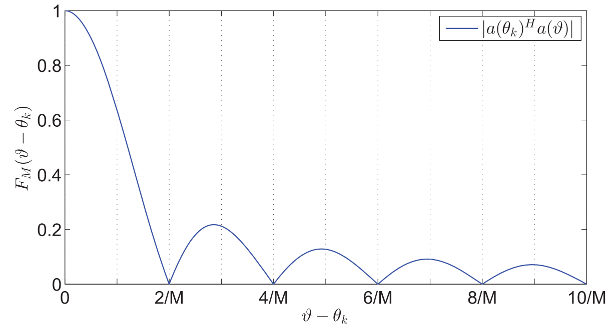

Fig. 1. $F_M(\vartheta - \theta_k)$ in (10) when $M = 100$.

$$\frac{1}{M}\left|\frac{\sin\frac{\pi(\vartheta-\theta_k)M}{2}}{\sin\frac{\pi(\vartheta-\theta_k)}{2}}\right| \overset{(a)}{\approx} \frac{1}{M}\left|\frac{\sin\frac{\pi\Delta}{2}}{\frac{\pi\Delta}{2M}}\right| \to \left|\frac{2\sin\frac{\pi\Delta}{2}}{\pi\Delta}\right| \qquad (11)$$

where (a) holds from $\sin\epsilon \approx \epsilon$ for small $\epsilon > 0$. That is, the asymptotic value of $|\mathbf{a}(\theta_k)^H \mathbf{a}(\vartheta)|$ may not be zero if $(\vartheta - \theta_k)$ becomes sufficiently small in the order of $O\left(\frac{1}{M}\right)$ (i.e., within the half of the main lobe width) as $M \to \infty$. Suppose that we can find a user k such that $|\vartheta - \theta_k| < \frac{1}{M}$ almost surely due to MU diversity. Then, the rate $\mathcal{R}_1$ of the RDB scheme is lower bounded by

$$\mathcal{R}_1 \ge \mathbb{E}\left[\log\left(1 + |\alpha_k|^2 M\frac{4}{\pi^2}\right)\right] \sim_M \log M, \qquad (12)$$

as $M \to \infty$. In other words, if the number K of users as a function of M is sufficiently large such that there exists a user k for whom $|\vartheta - \theta_k|$ is sufficiently small in the order of $O\left(\frac{1}{M}\right)$ with high probability, the RDB scheme has asymptotically good performance.

IV. Asymptotic Analysis of the RDB Rate: The Single Beam Case

With the intuition gained in the previous section, we first rigorously analyze the asymptotic performance of the RDB scheme in the single downlink beam case in this section. Direct computation of $\mathcal{R}_1$ in (9) is difficult since the integral in (9) does not have a closed-form expression. To circumvent this difficulty, we use several techniques to bound $\mathcal{R}_1$ by first assuming that $\alpha_k = 1$ for all k and focusing on the term $Z_k := M|\mathbf{a}(\theta_k)^H \mathbf{a}(\vartheta)|^2$ in (9). Then, we will include the term $\alpha_k \sim \mathcal{CN}(0, 1)$ in the performance analysis later.

Lemma 1: For any constant $p \in (-1, 1)$ and sufficiently large M, we have

$$|\tilde{\theta}_k| < \frac{1}{\frac{\pi}{4}M^{(1+p)/2}} \qquad (13)$$

under the event $\{Z_k > M^p\}$, and furthermore

$$\frac{1}{2\pi M^{(1+p)/2}} < \Pr\{Z_k > M^p\} < \frac{1}{\frac{\pi}{4}M^{(1+p)/2}}, \qquad (14)$$

[3] We can regard $\tilde{\theta}_k := \vartheta - \theta_k \sim \mathrm{Unif}[-1, 1]$ in case that $\vartheta - \theta_k$ appears as $e^{\iota\pi l(\vartheta-\theta_k)}$ for any integer l due to the periodicity of period two. See Appendix A.

where $\tilde{\theta}_k = \vartheta - \theta_k$ and $Z_k = M|\mathbf{a}(\theta_k)^H \mathbf{a}(\vartheta)|^2$.

Proof: From (10), the event $\{Z_k = M|\mathbf{a}(\theta_k)^H \mathbf{a}(\vartheta)|^2 > M^p\}$ is equivalent to

$$\left| \frac{\sin \frac{\pi \tilde{\theta}_k M}{2}}{\sin \frac{\pi \tilde{\theta}_k}{2}} \right| > M^{(1+p)/2}, \tag{15}$$

where $\tilde{\theta}_k \sim \text{Unif}[-1, 1]$ as shown in Appendix A. A necessary condition to satisfy (15) is that the denominator in the left-hand side (LHS) of (15) should be upper bounded as

$$\left| \sin \frac{\pi \tilde{\theta}_k}{2} \right| < \frac{1}{M^{(1+p)/2}} \tag{16}$$

since the numerator $\left| \sin \frac{\pi \tilde{\theta}_k M}{2} \right| \le 1$ and $M^{(1+p)/2} > 1$ for $p \in (-1, 1)$ and $M > 1$. For given $p \in (-1, 1)$, the value $1/M^{(1+p)/2}$ in the right-hand side (RHS) of (16) goes to zero as $M \to \infty$. Hence, by the fact that $\frac{\epsilon}{2} < \sin \epsilon$ for small $\epsilon > 0$, (16) implies that for sufficiently large M,

$$\frac{1}{2} \cdot \frac{\pi}{2} |\tilde{\theta}_k| < \frac{1}{M^{(1+p)/2}}. \tag{17}$$

Therefore, (13) holds under the event $\{Z_k > M^p\}$, and we have the upper bound part $\Pr\{Z_k > M^p\} < \frac{1}{\frac{\pi}{4} M^{(1+p)/2}}$ in (14), since (17) is a necessary condition for $\{Z_k > M^p\}$:

$$\Pr\{Z_k > M^p\} < \Pr\left\{ |\tilde{\theta}_k| < \frac{1}{\frac{\pi}{4} M^{(1+p)/2}} \right\} = \frac{1}{\frac{\pi}{4} M^{(1+p)/2}}$$

for sufficiently large M, since $\tilde{\theta}_k \sim \text{Unif}[-1, 1]$.

Now consider the lower bound part in (14). From the fact that $\sin \epsilon < \epsilon$ for $\epsilon > 0$, we have

$$\left| \sin \frac{\pi \tilde{\theta}_k}{2} \right| < \frac{1}{2M^{(1+p)/2}} \tag{18}$$

if

$$\frac{\pi}{2} |\tilde{\theta}_k| < \frac{1}{2M^{(1+p)/2}}. \tag{19}$$

If the following equation

$$\left| \sin \frac{\pi \tilde{\theta}_k M}{2} \right| \ge \frac{1}{2} \tag{20}$$

is satisfied in addition to (19) implying (18), then (15) is satisfied (i.e., the joint event of (19) and (20) is a sufficient condition for (15)). It is easy to see that the solution to (20) is

$$|\tilde{\theta}_k| \in \left\{ \left[\frac{2k}{M} + \frac{1}{3M}, \frac{2k}{M} + \frac{5}{3M} \right], k = 0, 1, 2, \cdots, \right\}. \tag{21}$$

Note that $\left| \sin \frac{\pi \tilde{\theta}_k M}{2} \right|$ in (20) has period $\frac{2}{M}$ and the length of one interval per period contained in the set (21) is $\frac{4}{3M}$. Hence, the set (21) occupies $\frac{2}{3}$ length of each period of $\frac{2}{M}$. Since the term $\frac{1}{M^{(p+1)/2}}$ for given $p \in (-1, 1)$ converges to zero slower than $\frac{1}{M}$ as $M \to \infty$, multiple discontinuous intervals in the set (21) are contained in the set defined by (19), and the length of the intersection of the sets (19) and (21) is lower bounded by $\frac{2}{3} \left(\frac{1}{\pi M^{(1+p)/2}} - \frac{2}{M} \right)$, where minus $\frac{2}{M}$ takes into account the impact of the last possibly partially overlapping interval. Hence, we have the lower bound part of (14):

$$\Pr\{Z_k > M^p\} \ge \frac{2}{3} \left(\frac{1}{\pi M^{(1+p)/2}} - \frac{2}{M} \right) > \frac{1}{2} \cdot \frac{1}{\pi M^{(1+p)/2}}.$$

for sufficiently large M. ■

Using Lemma 1 we have the following theorem.

Theorem 1: For $K = M^q$ and any given $q \in (0, 1)$,[4] we have asymptotic upper and lower bounds for $\mathcal{R}_1$ in (9) when $\alpha_k = 1$ for all k, given by

$$\log\left(1 + M^{2q-1-\epsilon}\right) \lesssim_M \mathbb{E}\left[\log(1+Z)\right] \lesssim_M \log\left(1 + M^{2q-1+\epsilon}\right) \tag{22}$$

for any sufficiently small $\epsilon > 0$, where $Z = \max_{1 \le k \le K} Z_k$ and $x \lesssim_M y$ means $\lim_{M\to\infty} x/y \le 1$.

Proof: Define c_M as

$$c_M \triangleq \frac{1}{\Pr\{Z_k > M^p\} \cdot M^{(1+p)/2}} \text{ for each } M$$

for $p \in (-1, 1)$. Then, we have

$$\Pr\{Z_k > M^p\} = \frac{1}{c_M M^{(1+p)/2}} \text{ for each } M, \tag{23}$$

and c_M is bounded between $\frac{\pi}{4}$ and 2π for sufficiently large M by Lemma 1.[5] (c_M is a bounded sequence with index M.) By (23), the probability of the event $\{Z > M^p\}$ for any $p \in (-1, 1)$ can be expressed as

$$\Pr\left\{ \max_k Z_k > M^p \right\} = 1 - \Pr\{Z_k \le M^p\}^K \tag{24}$$

$$= 1 - \left(1 - \frac{1}{c_M M^{(1+p)/2}}\right)^K. \tag{25}$$

Now consider an arbitrarily small given ϵ such that $0 < \epsilon < \min(2q, 2-2q)$ (say, if the given $q = 0.9$, then $0 < \epsilon < 0.2$).

[4] The ultimate goal of this paper is to show the sum rate optimality of RBF for large M under the UR-SP channel model as $K \uparrow M$, as shown in Corollary 2. Among several parameterizations for $K \uparrow M$ with K as a function of M, we chose the fractional power function $K = M^q$ with $0 \le q \le 1$ in this paper. This parameterization easily expresses $K \uparrow M$ by letting $q \uparrow 1$. Furthermore, when M is large as in massive MIMO, e.g. in order of hundred, the number K of active users in the cell may not be as large as M. Such an operating regime is called the *sparse user regime*[29], [30]. The parameterization $K = M^q$ with $0 \le q \le 1$ is effective to represent the number of users in the sparse user regime with a single parameter q and to analyze the performance of a scheduling algorithm depending on q, as seen in Section V-C.

[5] Since Lemma 1 states $\frac{1}{2\pi M^{(1+p)/2}} < \Pr\{Z_k > M^p\} < \frac{1}{\frac{\pi}{4} M^{(1+p)/2}}$ for any constant $p \in (-1, 1)$ and sufficiently large M, such constant c_M exists for each M when M is sufficiently large.

Set $p = 2q - 1 - \epsilon$. Then, $p \in (-1, 1)$ and (25) is applicable. Compute the second term in (25) with $p = 2q - 1 - \epsilon$, given by

$$\left(1 - \frac{1}{c_M M^{(1+p)/2}}\right)^K = \left(1 - \frac{1}{c_M M^{q-\frac{\epsilon}{2}}}\right)^{M^q} \tag{26}$$

$$= e^{M^q \log\left(1 - \frac{1}{c_M M^{q-\frac{\epsilon}{2}}}\right)} \tag{27}$$

$$= e^{-\frac{1}{c_M} M^{\epsilon/2} + O\left(\frac{1}{M^{2q-\epsilon}}\right)} \tag{28}$$

$$\to 0 \text{ as } M \to \infty, \tag{29}$$

where we used the fact that $\log(1 - x) = -x + O(x^2)$ for small x in the third step. Therefore, in this case, we have

$$\Pr\{Z > M^{2q-1-\epsilon}\} \to 1, \quad \text{as } M \to \infty \tag{30}$$

and thus $\mathbb{E}[\log(1 + Z)]$ can be bounded as

$$\mathbb{E}[\log(1 + Z)] \geq \int_{M^{2q-1-\epsilon}}^{M} \log(1 + z)p(z)dz \tag{31}$$

$$\geq \log\left(1 + M^{2q-1-\epsilon}\right) \int_{M^{2q-1-\epsilon}}^{M} p(z)dz \tag{32}$$

$$\sim_M \log\left(1 + M^{2q-1-\epsilon}\right), \tag{33}$$

since $\int_{M^{2q-1-\epsilon}}^{M} p(z)dz = \Pr\{Z > M^{2q-1-\epsilon}\} \to 1$ by (30) in this case. Hence, the claim on the lower bound follows.

Now set $p = 2q - 1 + \epsilon$. Then, $p \in (-1, 1)$ and (25) is applicable. Again by using the techniques used in (26)–(29), compute the second term in (25) with $p = 2q - 1 + \epsilon$, given by

$$\left(1 - \frac{1}{c_M M^{(1+p)/2}}\right)^K = e^{-\frac{1}{c_M} M^{-\epsilon/2} + O\left(\frac{1}{M^{2q+\epsilon}}\right)} \tag{34}$$

$$\overset{(a)}{=} 1 + O\left(-\frac{1}{M^{\frac{\epsilon}{2}}} + \frac{1}{M^{2q+\epsilon}}\right) \tag{35}$$

$$= 1 - O\left(\frac{1}{M^{\frac{\epsilon}{2}}}\right), \tag{36}$$

where the step (a) holds by the identity $e^x = 1 + O(x)$ for small x. Therefore, in this case, the probability of the event $\{Z > M^{2q-1+\epsilon}\}$ is given by $O\left(\frac{1}{M^{\epsilon/2}}\right)$. Using this, we have

$$\mathbb{E}[\log(1 + Z)]$$
$$= \int_{M^{2q-1+\epsilon}}^{M} \log(1 + z)p(z)dz + \int_{0}^{M^{2q-1+\epsilon}} \log(1 + z)p(z)dz \tag{37}$$

$$\leq \log(1 + M)O\left(\frac{1}{M^{\epsilon/2}}\right) + \log\left(1 + M^{2q-1+\epsilon}\right) \tag{38}$$

$$\sim_M \log\left(1 + M^{2q-1+\epsilon}\right). \tag{39}$$

In the second step, we used $\int_{M^{2q-1+\epsilon}}^{M} p(z)dz = \Pr\{Z > M^{2q-1+\epsilon}\} = O\left(\frac{1}{M^{\epsilon/2}}\right)$. Hence, the claim on the upper bound follows. ∎

Theorem 1 states that the single-beam RDB scheme under the assumption $\alpha_k = 1, \ \forall\, k$ has asymptotically nontrivial performance, i.e., $\mathcal{R}_1 \to \infty$, as $M \to \infty$, when $K = M^q$ with $q \in (\frac{1}{2}, 1)$. On the other hand, when $K = M^q$ with $q \in (0, \frac{1}{2})$, the RDB scheme has trivial performance, i.e., $\mathcal{R}_1 \to 0$, as $M \to \infty$. Thus, $q = \frac{1}{2}$ is the performance transition point for the single-beam RDB scheme under the UR-SP channel model.

Now consider the impact of the path gain term $\alpha_k \overset{i.i.d.}{\sim} \mathcal{CN}(0, 1)$ on the single-beam RDB rate $\mathcal{R}_1$. In fact, the same is true under the assumption of $\alpha_k \overset{i.i.d.}{\sim} \mathcal{CN}(0, 1)$.

Theorem 2: For $K = M^q$ with $q \in (\frac{1}{2}, 1)$ and $\alpha_k \overset{i.i.d.}{\sim} \mathcal{CN}(0, 1)$, we have

$$\lim_{M\to\infty} \frac{\mathcal{R}_1}{\mathbb{E}\left[\log\left(1 + M \max_k |\alpha_k|^2\right)\right]} = 2q - 1, \tag{40}$$

where $\mathcal{R}_1$ is the optimal single-beam RDB rate defined in (9) considering the random path gain, and $\mathbb{E}\left[\log(1 + M \max_k |\alpha_k|^2)\right]$ is the optimal rate of exact beamforming based on perfect CSI at the BS. On the other hand, when $q \in (0, \frac{1}{2})$, $\mathcal{R}_1 \to 0$ as $M \to \infty$.

Proof: See Appendix B. ∎

Note that the ratio $2q - 1$ of the RDB rate $\mathcal{R}_1$ to the exact beamforming rate is the same for both assumptions $\alpha_k = 1$ and $\alpha_k \sim \mathcal{CN}(0, 1)$. As seen, the single-beam RDB strategy achieves $2q - 1$ fraction of the exact beamforming rate based on perfect CSI at the BS. The supremum fraction of one can be achieved arbitrarily closely when the number K of users grows almost linearly w.r.t. M, i.e., q is arbitrarily close to one. Note that in the single beam case we only have the power gain by the antenna array, as shown in the maximum rate of $\mathbb{E}\left[\log(1 + M \max_k |\alpha_k|^2)\right]$ even by perfect beamforming.

V. Asymptotic Analysis of the RDB Rate: The Multiple Beam Case

In this section, we consider the case in which the number S of randomly-directional beams is more than one and allowed to grow to infinity as a function of M, and analyze the corresponding asymptotic performance. In the multiple beam case, the BS transmits S random beams equi-spaced in the normalized angle domain, defined as

$$\mathbf{u}_b = \mathbf{a}(\vartheta_b) = \mathbf{a}\left(\vartheta + \frac{2(b-1)}{S}\right), \text{ for } b = 1, \cdots, S, \tag{41}$$

where $\vartheta \sim \text{Unif}[-1, 1]$, to the downlink sequentially during the training period. We assume that the network is synchronized and thus each user knows the training beam index b by the corresponding training interval. Here, the difference between the normalized directions of two adjacent beams is $\frac{2}{S}$ and the offset ϑ is randomly generated on $[-1, 1]$. (Recall from (3) that $\mathbf{a}(\theta)$ is periodic in θ with period 2.) Note that the equi-spaced beams are asymptotically orthogonal to one another, i.e.,

$$\lim_{M\to\infty} |\mathbf{a}(\vartheta_{b_1})^H \mathbf{a}(\vartheta_{b_2})| = 0 \text{ for } b_1 \neq b_2, \tag{42}$$

when $S = o(M)$.

In the next subsections, we analyze the asymptotic performance of single user selection based on multiple training beams first and multiple user selection based on multiple beams later.

A. The Single User Selection Case

In the single user selection case, after the training period is over, each user reports the maximum of its received power values for the S training beams and the corresponding beam index. Then, the BS transmits a data stream to the user that has maximum received power with the corresponding beam $\mathbf{w}_b$. In this case, the rate $\mathcal{R}_S$ is given by

$$\mathcal{R}_S = \mathbb{E}\left[\log\left(1 + \max_{1\le k\le K}\max_{1\le b\le S}|\alpha_k|^2 M|\mathbf{a}(\theta_k)^H\mathbf{a}(\vartheta_b)|^2\right)\right]. \tag{43}$$

First, consider the case of $|\alpha_k| = 1$ for all $k = 1, \cdots, K$ as before. In this case, we have the following theorem:

Theorem 3: For $K = M^q$, $S = M^\ell$ and any $\ell, q \in (0, 1)$ such that $\ell + q < 1$, we have asymptotic lower and upper bounds on $\mathcal{R}_S$ in the case of $|\alpha_k| = 1$, $\forall\, k$, given by

$$\log\left(1 + M^{2q+2\ell-1-\epsilon}\right) \lesssim_M \mathbb{E}\left[\log\left(1+Z'\right)\right] \lesssim_M \log\left(1 + M^{2q+2\ell-1+\epsilon}\right) \tag{44}$$

for any sufficiently small $\epsilon > 0$, where $Z' = \max_k Z'_k$ and $Z'_k = \max_b M|\mathbf{a}(\theta_k)^H\mathbf{a}(\vartheta_b)|^2$.

Proof: Proof consists of two steps as in the proofs of Lemma 1 and Theorem 1: (i) first, we bound $\Pr\{Z'_k \le M^p\}$ and (ii) then bound $\mathbb{E}\left[\log(1+Z')\right]$ using the bounds on $\Pr\{Z'_k \le M^p\}$.

(i) First, consider

$$\Pr\{A\} := \Pr\{Z'_k = \max_b M|\mathbf{a}(\theta_k)^H\mathbf{a}(\vartheta_b)|^2 > M^p\}$$

for $p \in (-1, 1)$. Let C_i be the event that $i = \arg\max_b M|\mathbf{a}(\theta_k)^H\mathbf{a}(\vartheta_b)|^2$, i.e., C_i is the event that the i-th beam is the optimal beam for user k. Note that the distribution of $M|\mathbf{a}(\theta_k)^H\mathbf{a}(\vartheta_b)|^2 = M\cdot F_M^2\left(\vartheta_b - \theta_k\right) = M\cdot F_M^2\left(\vartheta + \frac{2(b-1)}{S} - \theta_k\right) = M\cdot F_M^2\left(\tilde{\theta}_k + \frac{2(b-1)}{S}\right)$ is independent of b since $\tilde{\theta}_k = \vartheta - \theta_k \sim \text{Unif}[-1, 1]$ and the Fejér kernel $F_M(\tilde{\theta})$ is a periodic function with period 2. Hence, the events $C_1, C_2, \cdots, C_{M^\ell}$ are equally probable as $\Pr\{C_i\} = \frac{1}{S} = \frac{1}{M^\ell}$ for every $i = 1, \cdots, M^\ell$. Furthermore, the conditional events $A|C_1, A|C_2, \cdots, A|C_{M^\ell}$ are also equally probable, i.e., $\Pr\{A|C_1\} = \cdots = \Pr\{A|C_{M^\ell}\}$ since the situation is the same for each ϑ_b due to the periodicity of $F_M(\cdot)$ of period two and $\tilde{\theta}_k \sim \text{Unif}[-1, 1]$. Hence, by the law of total probability and Bayes' rule, we have

$$\Pr\{A\} = \sum_{i=1}^{M^\ell}\Pr\{A|C_i\}\Pr\{C_i\} = \Pr\{A|C_1\} = M^\ell\cdot\Pr\{A, C_1\}.$$

Thus, to bound $\Pr\{A\}$, we need to bound $\Pr\{A, C_1\}$. In order to bound $\Pr\{A, C_1\}$, we find a sufficient condition for the event C_1. Let $\tilde{C}_1(p)$ be the event $M|\mathbf{a}(\theta_k)^H\mathbf{a}(\vartheta_1)|^2 > M^p$. Then, the event $\tilde{C}_1(p)$ with $p > 2\ell - 1$ implies

$$|\theta_k - \vartheta_1| \overset{(a)}{<} \frac{1}{\frac{\pi}{4}M^{(p+1)/2}} \overset{(b)}{=} \frac{1}{\frac{\pi}{4}M^{\ell+\delta/2}} \tag{45}$$

for sufficiently large M, where $\delta = p - (2\ell - 1) > 0$. (Here, step (a) is by (13) of Lemma 1 with $p \in (-1, 1)$, and step (b) is by the new additional condition $p > 2\ell - 1$.) Therefore, in this case, $|\theta_k - \vartheta_b| > |\frac{2}{M^\ell} - \frac{1}{\frac{\pi}{4}M^{\ell+\delta/2}}| > \frac{1}{M^\ell} = \frac{1}{2}\frac{2}{S}$ for any $b \ne 1$ and sufficiently large M, and this implies for sufficiently large M

$$\tilde{C}_1(p) \subset C_1 \text{ for } p \in (-1, 1) \text{ and } p > 2\ell - 1. \tag{46}$$

Now consider $\Pr\{A, C_1\}$

$$\begin{aligned}\Pr\{A, C_1\} &= \Pr\Big\{\max_b M|\mathbf{a}(\theta_k)^H\mathbf{a}(\vartheta_b)|^2 > M^p,\\ &\qquad 1 = \arg\max_b M|\mathbf{a}(\theta_k)^H\mathbf{a}(\vartheta_b)|^2\Big\}\\ &= \Pr\{\tilde{C}_1(p), C_1\}.\end{aligned} \tag{47}$$

By using (46) and (47), we have

$$\begin{aligned}\Pr\{A, C_1\} &\overset{(c)}{=} \Pr\{\tilde{C}_1(p), C_1\} \overset{(d)}{=} \Pr\{\tilde{C}_1(p)\}\\ &= \Pr\{M|\mathbf{a}(\theta_k)^H\mathbf{a}(\vartheta_1)|^2 > M^p\}\end{aligned} \tag{48}$$

when $p \in (-1, 1)$ and $p > 2\ell - 1$ (these two conditions are required to apply (46) for step (d), and step (c) is valid by (47)). Now by applying (14) of Lemma 1 to the last term in (48) and using $\Pr\{A\} = M^\ell\cdot\Pr\{A, C_1\}$, we have for $p \in (-1, 1)$ and $p > 2\ell - 1$,

$$\frac{1}{2\pi M^{(1+p-2\ell)/2}} < \Pr\{A\} < \frac{1}{\frac{\pi}{4}M^{(1+p-2\ell)/2}}. \tag{49}$$

(ii) Substituting $\Pr\{Z'_k \le M^p\} = 1 - \Pr\{A\}$ into $\Pr\{Z_k \le M^p\}$ in (24) of the proof in Theorem 1 and following the proof of Theorem 1, we have for $p = 2q + 2\ell - 1 - \epsilon$ with arbitrarily small $\epsilon > 0$,

$$\mathbb{E}[\log\left(1+Z'\right)] \gtrsim_M \log\left(1 + M^p\right), \tag{50}$$

and for $p = 2q + 2\ell - 1 + \epsilon$ with arbitrarily small $\epsilon > 0$,

$$\mathbb{E}\left[\log\left(1+Z'\right)\right] \lesssim_M \log\left(1 + M^p\right) \tag{51}$$

provided that $\ell + q < 1$ (this is required for the condition $p \in (-1, 1)$), where $x \gtrsim_M y$ indicates $\lim_{M\to\infty} x/y \ge 1$. Therefore, we have (44). ■

Corollary 1: For $K = M^q$, $S = M^\ell$ and any $\ell, q \in (0, 1)$ such that $\frac{1}{2} < \ell + q < 1$, we have

$$\lim_{M\to\infty}\frac{\mathbb{E}\left[\log\left(1+Z'\right)\right]}{\log(1+M)} = 2(q+\ell) - 1. \tag{52}$$

When the number S of training beams is fixed, i.e., $\ell = 0$, Corollary 1 reduces to the single beam result in (40). In the single user selection with multiple training beams, as seen in (52), the supremum of one for the achievable fraction can be achieved arbitrarily closely by the combination of multiple users q and multiple training beams ℓ. Thus, when there exist not sufficiently many users in the cell, multiple training beams can be used to enhance the RDB performance. Note that even for $q = 0$, the optimal rate can be achieved with $\mathcal{R}_S$ by making $\ell \uparrow 1$, as expected. (In fact, this case corresponds to the case considered in the previous works on channel estimation for sparse mm-wave MIMO channels, e.g. [6].) Note also that the effect of two terms is not distinguishable at least in terms of the rate during the data transmission period, although multiple training beams require more training time. It can be shown that even with consideration of the random channel gain $\alpha_k \overset{i.i.d.}{\sim} \mathcal{CN}(0, 1)$, the same result as (52) is valid.

B. The Multiple User Selection Case: Multiplexing Gain

Finally, aiming at multiplexing gain, we consider multiple user selection with RDB with multiple beams, where inter-beam interference should be considered, and investigate what can be achieved under the assumption of $|\alpha_k| = 1, \ \forall\, k$ for simplicity. The considered multi-user multi-beam scheme here is basically the conventional RBF scheme in [8] with the S random (asymptotically) orthogonal beams given by $\mathbf{u}_b$, $b = 1, \cdots, S$, defined in (41). In this scheme, we choose a user that has maximum SINR for each beam $\mathbf{u}_b$, $b = 1, \cdots, S$, and transmit S independent data streams to the S selected users. In this case, the received signal of a selected user κ_b is given by

$$y_{\kappa_b} = \sqrt{\frac{P_t}{S}}\mathbf{h}_{\kappa_b}^H \mathbf{u}_b + \sqrt{\frac{P_t}{S}} \sum_{b' \neq b} \mathbf{h}_{\kappa_b}^H \mathbf{u}_{b'} + n_{\kappa_b}, \ b = 1, \cdots, S, \tag{53}$$

where $\kappa_b = \arg\max_{1 \le k \le K} \text{SINR}_{k,b}$, $\text{SINR}_{k,b} = \frac{\rho M |\mathbf{a}(\theta_k)^H \mathbf{a}(\vartheta_b)|^2}{1 + \sum_{b' \neq b} \rho M |\mathbf{a}(\theta_k)^H \mathbf{a}(\vartheta_{b'})|^2}$, $\rho = \frac{P_t}{S}$ is the per-user power of each scheduled user, and the second term in the RHS of (53) is the inter-beam or inter-user interference. The expected sum rate of this RBF method is given by

$$\mathcal{R}_M = \sum_{b=1}^{S} \mathcal{R}_{\kappa_b}, \tag{54}$$

where the data rate of each scheduled user κ_b for beam b is given by

$$\begin{aligned}\mathcal{R}_{\kappa_b} &= \mathbb{E}\left[\log\left(1 + \max_{1 \le k \le K} \text{SINR}_{k,b}\right)\right] \\ &= \mathbb{E}\left[\log\left(1 + \frac{\rho M |\mathbf{a}(\theta_{\kappa_b})^H \mathbf{a}(\vartheta_b)|^2}{1 + \sum_{b' \neq b} \rho M |\mathbf{a}(\theta_{\kappa_b})^H \mathbf{a}(\vartheta_{b'})|^2}\right)\right].\end{aligned} \tag{55}$$

We first introduce the following lemma necessary to derive the asymptotic result regarding (54) and (55):

Lemma 2: For $|\tilde{\theta}_k| \in (0, 1]$, we have an upper bound for $F_M(\tilde{\theta})$, given by $F_M(\tilde{\theta}_k) \le \frac{1}{M|\tilde{\theta}_k|}$, where $F_M(\cdot)$ is defined in (10).

Proof: Since $F_M(\tilde{\theta}_k)$ and $\frac{1}{M|\tilde{\theta}_k|}$ are even functions, it is enough to consider $\tilde{\theta}_k \in (0, 1]$ only. From (10), we have an upper bound of $F_M(\tilde{\theta}_k)$:

$$F_M(\tilde{\theta}_k) \overset{(a)}{\le} \frac{1}{M} \frac{1}{\sin \frac{\pi \tilde{\theta}_k}{2}} \overset{(b)}{\le} \frac{1}{M\tilde{\theta}_k}$$

where (a) follows from $|\sin \frac{\pi \tilde{\theta}_k M}{2}| \le 1$ and $\sin \frac{\pi \tilde{\theta}_k}{2} > 0$ for $\tilde{\theta}_k \in (0, 1]$, and (b) follows from

$$\frac{1}{\tilde{\theta}_k} - \frac{1}{\sin \frac{\pi \tilde{\theta}_k}{2}} \ge 0 \iff f(\tilde{\theta}_k) := \sin \frac{\pi \tilde{\theta}_k}{2} - \tilde{\theta}_k \ge 0.$$

The RHS is true because $f(0) = f(1) = 0$ with $f''(\tilde{\theta}_k) = -\frac{\pi^2}{4} \sin \frac{\pi \tilde{\theta}_k}{2} < 0$ for $\tilde{\theta}_k \in (0, 1]$. ■

Now the following theorem shows the asymptotic result on the RBF per-user rate (55) when the total power P_t is fixed regardless of S.

Theorem 4: For $K = M^q$, $S = M^\ell$ with $q \in (0, 1)$ and $\ell \in (0, q - \frac{\epsilon}{2})$, asymptotic upper and lower bounds on the per-user rate $\mathcal{R}_{\kappa_b}$ of selected user κ_b for fixed total transmit power $P_t = 1$ are given by

$$\log\left(1 + M^{2q-1-\ell-\epsilon}\right) \lesssim_M \mathcal{R}_{\kappa_b} \lesssim_M \log\left(1 + M^{2q-1-\ell+\epsilon}\right) \tag{56}$$

for any sufficiently small $\epsilon > 0$.

Proof: The flow of proof is to first find lower and upper bounds on $\mathcal{R}_{\kappa_b}$, denoted by L and U, respectively, and then to show that the bounds L and U are asymptotically bounded as

$$\begin{aligned}\log\left(1 + M^{2q-1-\ell-\epsilon}\right) \lesssim_M L &\le \mathcal{R}_{\kappa_b} \\ &\le U \lesssim_M \log\left(1 + M^{2q-1-\ell+\epsilon}\right).\end{aligned} \tag{57}$$

To find L and U, we consider a virtual user selection method based on maximizing signal power not SINR for each beam $\mathbf{a}(\vartheta_b)$, i.e.,

$$\tilde{\kappa}_b = \arg\max_{1 \le k \le K} M|\mathbf{a}(\theta_k)^H \mathbf{a}(\vartheta_b)| \text{ for } b = 1, \cdots, S.$$

Since the user $\tilde{\kappa}_b$ is chosen based on maximizing signal power only, we have $\text{SINR}_{\tilde{\kappa}_b,b} \le \text{SINR}_{\kappa_b,b}$. Therefore, a lower bound on $\mathcal{R}_{\kappa_b}$ can be obtained as

$$\mathcal{R}_{\kappa_b} \ge \mathcal{R}_{\tilde{\kappa}_b} = \mathbb{E}\left[\log\left(1 + \frac{\rho Z_{bb}}{1 + \rho \sum_{b' \neq b} Z_{bb'}}\right)\right] =: L \tag{58}$$

where $Z_{bb'} := M|\mathbf{a}(\theta_{\tilde{\kappa}_b})^H \mathbf{a}(\vartheta_{b'}))|^2$ for $b' = 1, \cdots, S$. Furthermore, an upper bound on $\mathcal{R}_{\kappa_b}$ can be obtained by simply ignoring the inter-beam interference as

$$\begin{aligned}\mathcal{R}_{\kappa_b} &\le \mathbb{E}\left[\log\left(1 + \rho M |\mathbf{a}\left(\theta_{\kappa_b}\right)^H \mathbf{a}\left(\vartheta_b\right)|^2\right)\right] \\ &\le \mathbb{E}\left[\log\left(1 + \rho Z_{bb}\right)\right] =: U.\end{aligned} \tag{59}$$

By modifying Theorem 1 to include $\rho = 1/S = M^{-\ell}$ in front of Z_{bb} and applying the modified theorem to U in (59), we obtain

$$U \lesssim_M \log\left(1 + M^{2q-1-\ell+\epsilon}\right). \tag{60}$$

Hence, the claim on the upper bound follows.

Now consider the case of lower bound L. From the fact that $\mathbb{E}[f(X)] = \Pr\{A\}\mathbb{E}[f(X)|A] + \Pr\{A^c\}\mathbb{E}[f(X)|A^c] \geq \Pr\{A\}\mathbb{E}[f(X)|A]$ for a non-negative function $f(X)$, L in (58) with $\rho = 1/S = M^{-\ell}$ can be bounded as

$$\mathbb{E}\left[\log\left(1 + \frac{M^{-\ell}Z_{bb}}{1 + M^{-\ell}\sum_{b' \neq b} Z_{bb'}}\right)\right] \geq \Pr\{Z_{bb} \geq M^p\} \times \mathbb{E}\left[\log\left(1 + \frac{M^{-\ell}Z_{bb}}{1 + M^{-\ell}\sum_{b' \neq b} Z_{bb'}}\right)\middle| Z_{bb} \geq M^p\right]. \tag{61}$$

Under the condition that $\{Z_{bb} \geq M^p\}$, we have

$$|\theta_{\tilde{\kappa}_b} - \vartheta_b| \leq \frac{1}{\frac{\pi}{4}M^{(1+p)/2}}$$

by (13) of Lemma 1. Therefore, $|\theta_{\tilde{\kappa}_b} - \vartheta_{b'}| > \left|\frac{2}{S} - \frac{1}{\frac{\pi}{4}M^{(1+p)/2}}\right|$, $\forall b' \neq b$. Furthermore, we can re-arrange the indices of $\{\vartheta_{b'}\}_{b' \neq b}$ in the order of closeness to ϑ_b with the new indices $\{j\}$. Then, we have

$$|\theta_{\tilde{\kappa}_b} - \vartheta_{2j-1}|, |\theta_{\tilde{\kappa}_b} - \vartheta_{2j}| > \left|\frac{2j}{M^\ell} - \frac{1}{\frac{\pi}{4}M^{(1+p)/2}}\right|, \tag{62}$$

since $\frac{2}{S} = \frac{2}{M^\ell}$ is the angular spacing between two adjacent beams. We now have a lower bound on L, which is given by the expression (63) shown at the bottom of the page, where (a) holds by (61) and Lemma 2 with $Z_{bj} = MF_M^2(\theta_{\tilde{\kappa}_b} - \vartheta_j)$; (b) holds by (62); (c) holds because $\left|\frac{2j}{M^\ell} - \frac{1}{\frac{\pi}{4}M^{(1+p)/2}}\right|^2 \gtrsim_M \left|\frac{j}{M^\ell}\right|^2$ for large M provided that $\ell < (1+p)/2$; (d) follows from $\sum_{j=1}^{\infty}\frac{1}{j^2} = \frac{\pi^2}{6}$; and the last step holds because $\Pr\{Z_{bb} \geq M^p\} \to 1$ for $p < 2q-1$ by (30) and $\frac{\pi^2 M^{\ell-1}}{3} \to 0$ for $\ell < 1$. Hence the claim on the lower bound follows.

Note that the conditions used to derive (63) and (60) are $p < 2q-1$, $\ell < (1+p)/2$, and $\ell < 1$, and q and ℓ are given. Set $p = 2q - 1 - \epsilon$ for $\epsilon > 0$. Then, $\ell < q - \epsilon/2 < 1$ since $q \in (0, 1)$. This concludes proof. ■

In Theorem 4, the condition $\ell < q - \epsilon/2 < 1$ guarantees that the S beams are asymptotically orthonormal by (42) and there exist more users than the number of beams in the cell by the difference in the fractional orders q and ℓ.

Now consider the sum rate (54) of the RBF scheme with multi-beam multi-user selection. The result is summarized in the following corollary to Theorem 4.

Corollary 2: Under the UR-SP channel model with constant path gain and fixed total power at the BS, RBF achieves linear sum rate scaling w.r.t. the number M of antennas when K and S approach M, i.e. $K, S \uparrow M$. Furthermore, RBF yields optimal sum rate performance for large M, when $K, S \uparrow M$.

Proof: The sum rate $\mathcal{R}_M$ corresponding to Theorem 4 scales as $M^\ell \log(1 + M^{2q-1-\ell})$ when $2q - 1 - \ell > 0$. By sending $q \uparrow 1$ and $\ell \uparrow 1$, i.e., $K, S \uparrow M$, we have $\mathcal{R}_M \to M\log 2$. Thus, $\mathcal{R}_M$ achieves linear scaling w.r.t. M as $K, S \uparrow M$. Furthermore, when $K \geq M$, the capacity of the overall downlink channel from the BS with M transmit antennas to K single-antenna users is at most $\min(M, K)\log 2 = M\log 2$. This is because from (55) the best per-user rate $\mathcal{R}_{\kappa_b}$ is $\log 2$ for $P_t = 1$ when there is no interference, since $\rho = 1/M$ for $P_t = 1$. Thus, for large M, RBF achieves optimal sum rate under the UR-SP channel model. ■

The result in Corollary 2 is a significant difference from the sum rate behavior (7) of the RBF method [8] in large-scale MIMO, i.e., $\lim_{K\to\infty}\frac{\log K}{M} = 0$, under the i.i.d. Rayleigh fading channel model (5) representing rich scattering environments. The major performance difference results from

$$\begin{aligned}
L &\overset{(a)}{\geq} \Pr\{Z_{bb} \geq M^p\} \cdot \mathbb{E}\left[\log\left(1 + \frac{M^{-\ell}Z_{bb}}{1 + M^{-\ell}\sum_{j\neq b}\frac{1}{M|\theta_{\tilde{\kappa}_b} - \vartheta_j|^2}}\right)\middle| Z_{bb} \geq M^p\right] \\
&\overset{(b)}{\geq} \Pr\{Z_{bb} \geq M^p\} \cdot \mathbb{E}\left[\log\left(1 + \frac{M^{-\ell}Z_{bb}}{1 + M^{-\ell}\sum_{j=1}^{\frac{S}{2}}\frac{2}{M\left|\frac{2j}{S} - \frac{1}{\frac{\pi}{4}M^{(1+p)/2}}\right|^2}}\right)\middle| Z_{bb} \geq M^p\right] \\
&\overset{(c)}{\gtrsim_M} \Pr\{Z_{bb} \geq M^p\} \cdot \mathbb{E}\left[\log\left(1 + \frac{M^{-\ell}Z_{bb}}{1 + 2M^{\ell-1}\sum_{j=1}^{\frac{S}{2}}\frac{1}{j^2}}\right)\middle| Z_{bb} \geq M^p\right] \\
&\overset{(d)}{\geq} \Pr\{Z_{bb} \geq M^p\} \cdot \log\left(1 + \frac{M^{-\ell}M^p}{1 + \frac{\pi^2 M^{\ell-1}}{3}}\right) \\
&\sim_M \log\left(1 + M^{p-\ell}\right),
\end{aligned} \tag{63}$$

the difference in degrees-of-freedom in the two channels, the UR-SP channel (4) and the i.i.d. Rayleigh fading channel (5). In the i.i.d. Rayleigh fading channel case, we have M independent parameters and the channel vector is randomly located within a ball in the M-dimensional space. Consider a cone around each axis in the M-dimensional space so that channel vectors each of which is contained in each of the cones are roughly orthogonal, as shown in Fig. 3 of [15]. Then, the probability that a channel vector generated randomly according to (5) falls into such a cone is exponentially decreasing as M increases (See Appendix C). Hence, if the number K of users randomly distributed within the ball does not increases exponentially fast w.r.t. the dimension M (i.e., $\lim_{K\to\infty}\frac{\log K}{M}=0$), it is difficult to find M users whose channel vectors are contained in the M roughly-orthogonal cones (one for each) [8], [10], [12] (the goal of SUS [9], RBF [8] or ReDOS-PBR [15] scheduling is to find such M users), and linear sum rate scaling by RBF w.r.t. the dimension M (i.e., the number of antennas) is not attainable. In the considered mm-wave MIMO with the UR-SP channel model, however, the situation is quite different. Theorem 4 and Corollary 2 state that linear sum rate scaling w.r.t. M is possible only with $K \uparrow M$ in this case. This is because the degree-of-freedom in the UR-SP channel model (4) with $\alpha_k = 1$ is *one* regardless of the value of M. The orthogonality of the multiple transmit beams is attained by simply dividing the line of the normalized angle θ with length 2 by line segments each with length $2/S = 2/M^{\ell}$. Thus, if $K = M^q$ with $q > \ell$, there exists sufficient users in each line segment one of which is well matched to the transmit beam direction associated with each line segment if $2q - 1 - \ell > 0$. In fact, the channel matrix composed of the channel vectors of the users scheduled in such a way satisfies the asymptotically favorable propagation condition in [20].

C. Performance Comparison: The Sparse User Regime and the Fractional Rate Order

In this subsection, we compare the asymptotic performance of the three schemes considered in the previous sections. Previously, it is shown that RBF is optimal in sum rate for large M when K, S approaching M under the UR-SP channel model. When M is order of hundred as in massive MIMO, the number K of active users in the cell may be smaller than M and the network may be operated in the sparse user regime [29], [30]. In order to compare the relative performance in the sparse user regime, where the sparse number of active users in the cell is captured as the fractional power function $K = M^q$, $0 \le q \le 1$, we define the *fractional rate order (FRO)* γ as

$$\gamma := \lim_{M\to\infty}\frac{\log \mathcal{R}}{\log M}. \tag{64}$$

Here, we assume $\alpha_k = 1$, $\forall\, k$ and $P_t = 1$. Note that $\mathcal{R} = \Theta(M^{\gamma})$ for $\gamma \neq 0$. For $\gamma > 0$, $\mathcal{R}$ increases to infinity as $M \to \infty$, whereas for $\gamma < 0$, $\mathcal{R}$ decreases to zero as $M \to \infty$. Now

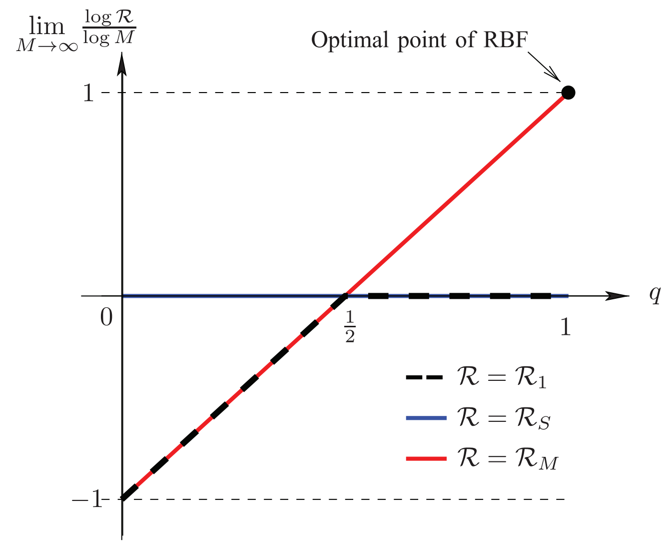

Fig. 2. Fractional rate order versus q

consider the three rates $\mathcal{R} = \mathcal{R}_1, \mathcal{R}_S$, and $\mathcal{R}_M$. First, for the single beam RDB rate $\mathcal{R}_1$ we have by Theorem 1 that

$$\gamma_1 = \lim_{M\to\infty}\frac{\log \mathcal{R}_1}{\log M} = \begin{cases} 0, & \text{for } q \in (\frac{1}{2}, 1) \\ 2q-1 & \text{for } q \in (0, \frac{1}{2}), \end{cases} \tag{65}$$

where we used $\log(1+x) = x$ for small x for the second part. Next, for the multi-beam RDB scheme with single-user selection, we have

$$\gamma_S = \lim_{M\to\infty}\frac{\log \mathcal{R}_S}{\log M} = 0, \quad \text{for} \quad q \in (0, 1) \tag{66}$$

by Theorem 3 with setting ℓ such that $1/2 < \ell + q < 1$. Here, $\gamma_S = 0$ is achieved even for $q \in (0, 1/2)$ because of added ℓ. Finally, we consider the multi-beam multi-user selection RDB or RBF. In this case, $\mathcal{R}_M = \Theta(M^{\ell}\log(1 + M^{2q-\ell-1}))$ from Theorem 4 and (54). Using $M^{\ell}\log(1 + M^{2q-\ell-1}) = \log(1 + 1/M^{-2q+\ell+1})^{M^{\ell}M^{-2q+\ell+1-(-2q+\ell+1)}} = \log(1 + \frac{1}{M^{-2q+\ell+1}})^{M^{-2q+\ell+1}M^{2q-1}} \sim_M M^{2q-1}$ by setting ℓ such that $2q - 1 < \ell < q$, we obtain

$$\lim_{M\to\infty}\frac{\log \mathcal{R}_M}{\log M} = 2q - 1, \quad \text{for} \quad q \in (0, 1). \tag{67}$$

Fig. 2 shows (65), (66) and (67) versus $q \in (0, 1)$, and shows which strategy is better for different q determining the number of users in the cell relative to the number of transmit antennas. $\mathcal{R}_M$ has the largest FRO for $q \in (\frac{1}{2}, 1)$, whereas $\mathcal{R}_S$ has the largest FRO for $q \in (0, \frac{1}{2})$. γ_1 is a lower bound on both γ_S and γ_M for all $q \in (0, 1)$. Note that $\gamma_M \uparrow 1$ as $q \uparrow 1$, which shows the optimality of RDB as mentioned already.

VI. Numerical Results

In this section, we provide some numerical results to validate our asymptotic analysis in the previous sections. All the expectations in the below are average over 5000 channel realizations and we set $P_t = 1$.

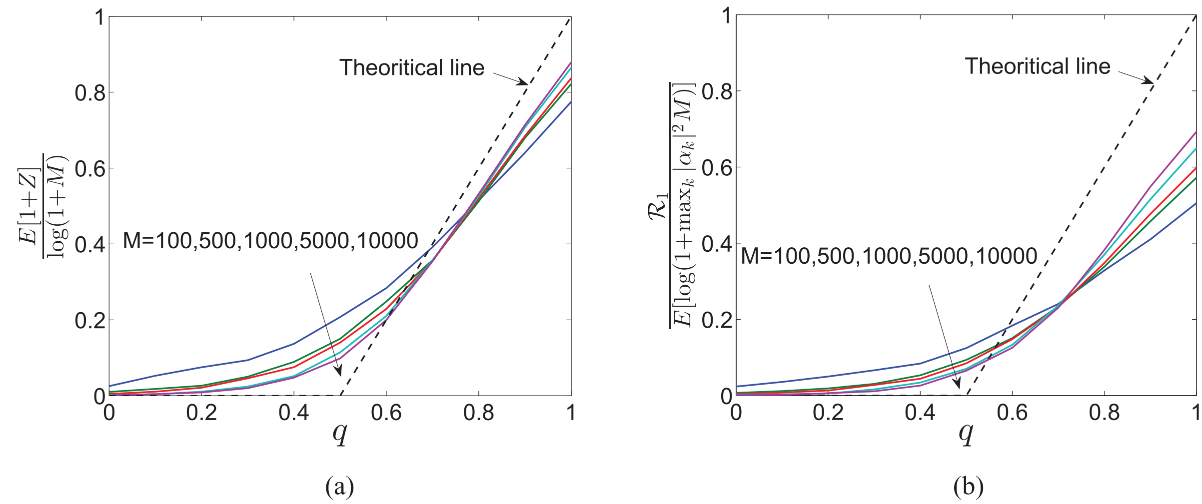

Fig. 3. The ratio of the RDB rate $\mathcal{R}_1$ to the rate with perfect CSI $\mathbb{E}[\log(1 + \max_k |\alpha_k|^2 M)]$ versus q for different M: (a) $\alpha_k = 1$ and (b) $\alpha_k \sim \mathcal{CN}(0, 1)$.

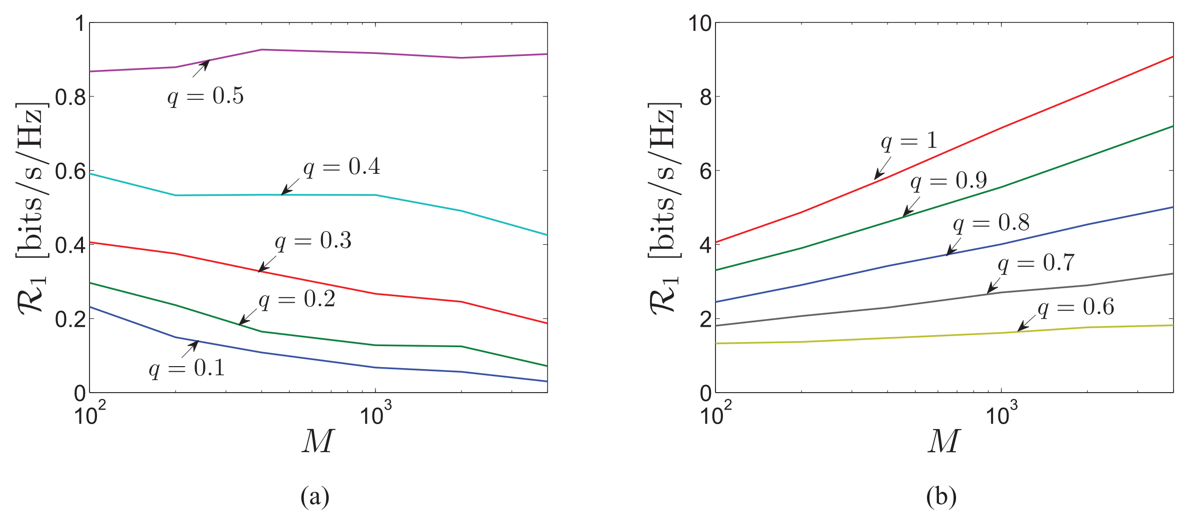

Fig. 4. The RDB rate $\mathcal{R}_1$ versus M with $\alpha_k \sim \mathcal{CN}(0, 1)$ for different q (log scale on x-axis): (a) $q = 0.1, 0.2, \cdots, 0.5$ and (b) $q = 0.6, 0.7, \cdots, 1$.

A. The Single Beam Case

To verify the asymptotic analysis in Section IV, we considered a mm-wave MU-MISO downlink system with the UR-SP channel model. Fig. 3(a) and (b) shows the value of $\frac{\mathbb{E}[1+Z]}{\log(1+M)}$ versus q for $M = 100,\ 500,\ 1000,\ 5000,\ 10000$ for $\alpha_k = 1$ and $\alpha_k \sim \mathcal{CN}(0, 1)$, respectively. It is seen that the curve of $\frac{\mathbb{E}[1+Z]}{\log(1+M)}$ versus q gradually converges to the theoretical line of $2q - 1$ for $q > \frac{1}{2}$ and 0 for $q \leq \frac{1}{2}$ as M increases. Note that there exist some gap between the theoretical asymptotic line and the finite-sample results. This results from the slow rate of convergence. Fig. 4(a) and (b) show the actual RDB rate w.r.t. M for $q = 0.1$ to 0.5 and $q = 0.6$ to 1, respectively, in the case of $\alpha_k \sim \mathcal{CN}(0, 1)$. It is seen in Fig. 4(a) that the RDB rate for q below 0.5 decreases as M increases, but it almost remains the same when $q = 0.5$. On the other hand, it is seen in Fig. 4(b) that the RDB rate for q above 0.5 increases as M increases. (Since x-axis is in log scale, the rate curve is linear as expected by Theorem 2 when $q > 0.5$.) The results in Figs. 3 and 4 coincide with Theorems 1 and 2.

B. The Multiple Beam Case

We first considered the multiple beam RDB with single user selection. Fig. 5(a) and (b) show the ratio of the multiple beam RDB rate $\mathcal{R}_S$ with single-user selection to the rate with perfect CSI versus q for different ℓ in the cases of $\alpha_k = 1$ and $\alpha_k \sim \mathcal{CN}(0, 1)$, respectively, when $M = 1000$. It is seen that the simulation curves roughly match the theoretical lines. We then verified the rate $\mathcal{R}_S$ for $q = 0.3$ with different ℓ. It is seen in Fig. 6(a) that $\mathcal{R}_S$ increases as M increases for the cases of $\ell > 0.2$ (i.e., $q + \ell > 0.5$), as predicted by Theorem 3. On the other hand, the rate decreases for the case of $\ell < 0.2$ as M increases. Finally, we verified the multi-beam multi-user selection RDB. We set to $q = 0.7$ and used $\alpha_k \sim \mathcal{CN}(0, 1), \forall\, k$. Fig. 6(b) shows the per-user rate $\mathcal{R}_{\kappa_b}$ in Theorem 4 versus M for different ℓ.

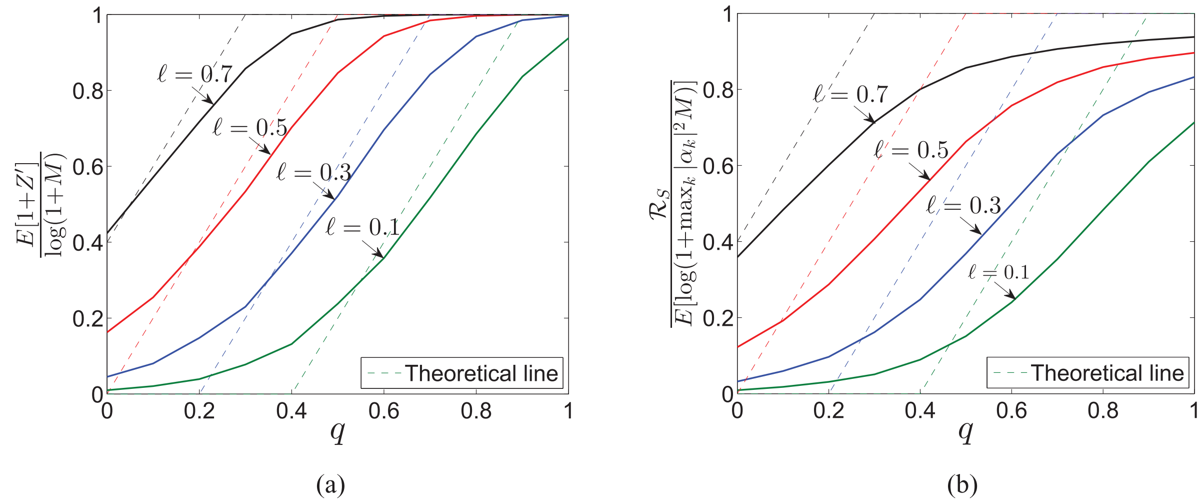

Fig. 5. The ratio of $\mathcal{R}_S$ to the rate with perfect CSI $\mathbb{E}[\log(1+\max_k|\alpha_k|^2M)]$ versus q for different ℓ: (a) $\alpha_k = 1, \forall k$ and (b) $\alpha_k \sim \mathcal{CN}(0,1), \forall k$.

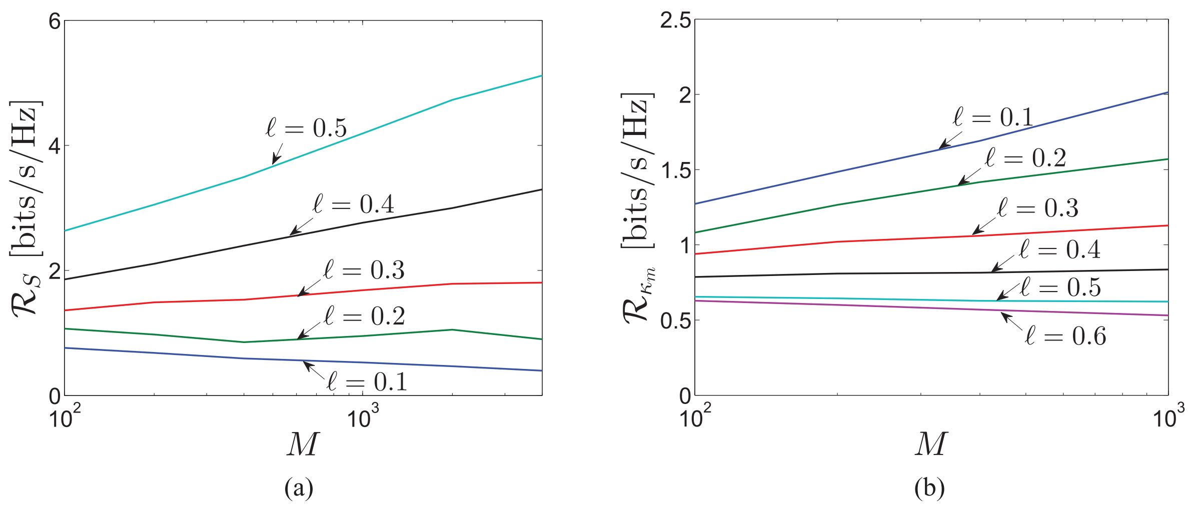

Fig. 6. (a) $\mathcal{R}_S$ versus M ($q = 0.3$ and $\alpha_k \sim \mathcal{CN}(0,1)$) (log scale on x-axis) and (b) the per-user rate of the multi-beam multi-user selection RDB versus M ($q = 0.7$ and $\alpha_k \sim \mathcal{CN}(0,1)$).

It is seen that the per-user rate $\mathcal{R}_{\kappa_b}$ increases when $\ell < 0.4$, whereas it decreases when $\ell > 0.4$, as M increases, as predicted by Theorem 4 (i.e., $2q-1-\ell > 0$ or $2q-1-\ell < 0$).

VII. Conclusion

We have considered RDB for millimeter-wave MU-MISO and examined the associated MU gain, using asymptotic performance analysis based on the UR-SP channel model which well captures radio propagation channels with few multipaths in the mm-wave band. We have shown that there exists a transition point on the number of users relative to the number of transmit antennas for non-trivial performance of the RDB scheme and have identified the case in which downlink training and channel estimation are important for good performance. We have also shown that linear sum rate scaling w.r.t. the number of transmit antennas can be achieved under the UR-SP channel model by RBF based on multiple beams equi-spaced in the angle domain and proper user scheduling, if the number of users in the cell increases linearly w.r.t. the number of transmit antennas. We have compared three RDB schemes composed of beamforming and user scheduling based on the newly defined fractional rate order, yielding insights into the most effective beamforming and scheduling choices for mm-wave MU-MISO in various operating conditions. Simulation results validate our theoretical analysis. The results here are based on the simplified UR-SP channel model capturing high propagation directivity, and thus extension to a general multipath channel model is left as future work.

Appendix A
Distribution of $\vartheta - \theta_k$

Since $\vartheta, \theta_k \overset{\text{i.i.d.}}{\sim} \text{Unif}[-1,1]$, the difference random variable $\tilde{\theta}_k$ has the distribution, given by

$$p(\tilde{\theta}) = \begin{cases} \frac{1}{4}\tilde{\theta} + \frac{1}{2}, & -2 \le \tilde{\theta} \le 0 \\ -\frac{1}{4}\tilde{\theta} + \frac{1}{2}, & 0 \le \tilde{\theta} \le 2. \end{cases} \tag{68}$$

For any function $f(\tilde{\theta})$ with the periodicity of period two, we have $f(\tilde{\theta}) = f(\tilde{\theta}+2)$ for $\tilde{\theta} \in [-1,0]$ and $f(\tilde{\theta}) = f(\tilde{\theta}-2)$

for $\tilde{\theta} \in [0, 1]$. Therefore, we can regard $p(\tilde{\theta})$ on the function $f(\tilde{\theta})$ as

$$p(\tilde{\theta}) = \begin{cases} \frac{1}{4}\tilde{\theta} + \frac{1}{2} - \frac{1}{4}\tilde{\theta}, & -1 \le \tilde{\theta} \le 0 \\ -\frac{1}{4}\tilde{\theta} + \frac{1}{2} + \frac{1}{4}\tilde{\theta}, & 0 \le \tilde{\theta} \le 1 \end{cases} \tag{69}$$

i.e., $\tilde{\theta} \sim \text{Unif}[-1, 1]$.

Appendix B
Proof of Theorem 2

Before proving Theorem 2, we prove another interesting lemma of which proof is partly used in proof of Theorem 2.

Lemma 3: For $K = M^q$, $q \in (\frac{1}{2}, 1)$ and $\alpha_k \overset{\text{i.i.d.}}{\sim} \mathcal{CN}(0, 1)$, we have

$$\lim_{M\to\infty} \frac{\mathcal{R}_1}{\mathbb{E}\left[\log\left(1 + |\alpha_{k'}|^2 Z_{k'}\right)\right]} = 1, \tag{70}$$

where $k' = \arg\max_k Z_k$, and $\mathcal{R}_1$ is the optimal RDB rate in (9) considering the random path gain.

Proof: $\mathcal{R}_1$ is bounded as

$$\begin{aligned} \mathbb{E}[\log(1 + |\alpha_{k'}|^2 Z_{k'})] &\le \mathcal{R}_1 \\ &\le \mathbb{E}\left[\log\left(1 + \left(\max_k |\alpha_k|^2\right) Z_{k'}\right)\right]. \end{aligned} \tag{71}$$

Eq. (22) in Theorem 1 can easily be modified to

$$\begin{aligned} \log\left(1 + \beta M^{2q-1-\epsilon}\right) &\lesssim_M \mathbb{E}\left[\log\left(1 + \beta Z\right)\right] \\ &\lesssim_M \log\left(1 + \beta M^{2q-1+\epsilon}\right) \end{aligned} \tag{72}$$

for $q \in (\frac{1}{2}, 1)$ and $\beta > 0$. Note that $\mathbb{E}[\log(1 + |\alpha_{k'}|^2 Z_{k'})] = \mathbb{E}[\mathbb{E}[\log(1 + |\alpha_{k'}|^2 Z_{k'}) \mid |\alpha_{k'}|^2]]$ by the law of iterated expectations. Applying the lower bound in (72) to $\mathbb{E}[\log(1 + |\alpha_{k'}|^2 Z_{k'}) \mid |\alpha_{k'}|^2]$, we have

$$\mathbb{E}\left[\log\left(1 + |\alpha_{k'}|^2 M^{2q-1-\epsilon}\right)\right] \lesssim_M \mathbb{E}\left[\log\left(1 + |\alpha_{k'}|^2 Z_{k'}\right)\right]. \tag{73}$$

For $q \in (\frac{1}{2}, 1)$, we have

$$\begin{aligned} &\mathbb{E}\left[\log\left(1 + |\alpha_{k'}|^2 M^{2q-1-\epsilon}\right)\right] \\ &\sim_M \mathbb{E}\left[\log\left(|\alpha_{k'}|^2 M^{2q-1-\epsilon}\right)\right] \\ &= \mathbb{E}\left[\log|\alpha_{k'}|^2\right] + (2q-1-\epsilon)\log M \\ &\sim_M (2q-1-\epsilon)\log M \end{aligned} \tag{74}$$

for any sufficiently small $\epsilon > 0$ such that $2q - 1 - \epsilon > 0$. Since $|\alpha_{k'}|^2 \sim \chi^2(2)$, $\mathbb{E}[\log|\alpha_{k'}|^2]$ is a constant.

Now consider the upper bound in (71). Again applying the law of iterated expectations and the upper bound in (72), we have $\mathbb{E}\left[\log\left(1 + \left(\max_k |\alpha_k|^2\right) Z_{k'}\right)\right] \le \mathbb{E}\left[\log\left(1 + \left(\max_k |\alpha_k|^2\right) M^{2q-1+\epsilon}\right)\right]$. From the fact that $\mathbb{E}[\log(1 + \max_k |\alpha_k|^2)] \sim_M \log(\log K)$ [8], the above bound can further be simplified as

$$\begin{aligned} &\mathbb{E}\left[\log\left(1 + \left(\max_k |\alpha_k|^2\right) Z_{k'}\right)\right] \\ &\lesssim_M \log\left(M^{2q-1+\epsilon}\log K\right) \\ &\sim_M (2q-1+\epsilon)\log M + \log(\log M). \end{aligned} \tag{75}$$

Dividing (71) by $\mathbb{E}[\log(1 + |\alpha_{k'}|^2 Z_{k'})]$, we have

$$\begin{aligned} 1 &\le \frac{\mathcal{R}_1}{\mathbb{E}[\log(1 + |\alpha_{k'}|^2 Z_{k'})]} \\ &\le \frac{\mathbb{E}\left[\log\left(1 + \left(\max_k |\alpha_k|^2\right) Z_{k'}\right)\right]}{\mathbb{E}[\log(1 + |\alpha_{k'}|^2 Z_{k'})]} \\ &\overset{(a)}{\lesssim_M} \frac{(2q-1+\epsilon)\log M + \log\log M}{(2q-1-\epsilon)\log M} \\ &\sim_M \frac{2q-1+\epsilon}{2q-1-\epsilon} \end{aligned} \tag{76}$$

where step (a) follows from (74) and (75). Since (76) holds for any small $\epsilon > 0$, the claim follows. ■

Proof of Theorem 2: By (71), (74), and (75), we have

$$\begin{aligned} (2q-1-\epsilon)\log M \lesssim_M \mathcal{R}_1 \lesssim_M (2q-1+\epsilon)\log M \\ + \log\log M. \end{aligned}$$

Dividing the above equation by $\mathbb{E}[\log(1 + M\max_k |\alpha_k|^2)]$ and using the fact that $\mathbb{E}[\log(1 + M\max_k |\alpha_k|^2)] \sim_M \log M + \log\log M$ [8], we have

$$\begin{aligned} \frac{(2q-1-\epsilon)\log M}{\log M + \log\log M} &\lesssim_M \frac{\mathcal{R}_1}{\mathbb{E}\left[\log\left(1 + M\max_k |\alpha_k|^2\right)\right]} \\ &\lesssim_M \frac{(2q-1+\epsilon)\log M + \log\log M}{\log M + \log\log M} \end{aligned}$$

for arbitrarily and sufficiently small $\epsilon > 0$. Hence, we have

$$2q - 1 - \epsilon \lesssim_M \frac{\mathcal{R}_1}{\mathbb{E}\left[\log\left(1 + M\max_k |\alpha_k|^2\right)\right]} \lesssim_M 2q - 1 + \epsilon. \tag{77}$$

Now consider the case of $q \in (0, \frac{1}{2})$. In this case, Eq. (22) in Theorem 1 can be modified to

$$\beta M^{2q-1-\epsilon} \lesssim_M \mathbb{E}\left[\log\left(1 + \beta Z\right)\right] \lesssim_M \beta M^{2q-1+\epsilon} \tag{78}$$

for $\beta > 0$ and sufficiently small $\epsilon > 0$ such that $2q - 1 + \epsilon < 0$, since $\log(1 + \beta M^{2q-1\pm\epsilon}) \sim_M \beta M^{2q-1\pm\epsilon}$ from $\log(1 + x) \to x$ as $x \to 0$. Again applying the law of iterated expectations and the upper bound in (78), $\mathcal{R}_1$ is upper bounded as

$$\mathcal{R}_1 \le \mathbb{E}\left[\log\left(1 + \left(\max_k |\alpha_k|^2\right) Z_{k'}\right)\right] \tag{79}$$

$$\lesssim_M \mathbb{E}\left(\max_k |\alpha_k^2|\right) M^{2q-1+\epsilon} \tag{80}$$

$$\sim_M (q\log M) M^{2q-1+\epsilon} \to 0 \tag{81}$$

as $M \to \infty$. This concludes the proof. ■

APPENDIX C
PROBABILITY OF FINDING QUASI-ORTHOGONAL USERS IN THE RAYLEIGH FADING CASE

A double cone (or cone) $\mathcal{C}_j$ around each axis j in the M-dimensional space is defined as

$$\mathcal{C}_j(\eta) = \left\{ \mathbf{h} : \frac{|\mathbf{h}^H \mathbf{e}_j|}{\|\mathbf{h}\|} \geq \eta \right\}, \tag{82}$$

where $\mathbf{e}_j$ is the j-th column of the $M \times M$ identity matrix, and $\eta \in (0, 1)$. The probability that the channel vector $\mathbf{h}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ is contained in the cone $\mathcal{C}_j$ is given by

$$\begin{aligned} \Pr\{\mathbf{h}_k \in \mathcal{C}_j(\eta)\} &= \Pr\{|h_{k,j}| \geq \eta\|\mathbf{h}_k\|\} \\ &\overset{(a)}{\approx} \Pr\{|h_{k,j}|^2 \geq \eta^2 M\} \\ &\overset{(b)}{=} e^{-\eta^2 M} \end{aligned} \tag{83}$$

where (a) becomes tight for large M due to $\|\mathbf{h}_k\|^2/M \to 1$, and (b) holds by $|h_{k,j}|^2 \sim \chi^2(2)$. Therefore, the probability that the cone $\mathcal{C}_j$ contains at least one out of the K channel vectors is given by

$$\begin{aligned} \Pr\{\mathcal{C}_j \neq \emptyset\} &= 1 - \Pr\{\mathcal{C}_j = \emptyset\} \\ &= 1 - \Pr\{\mathbf{h}_k \notin \mathcal{C}_j\}^K \\ &\approx 1 - \left(1 - \frac{1}{e^{\eta^2 M}}\right)^K \\ &\to \begin{cases} 1, & \text{for} \quad \lim_{M,K\to\infty} \frac{\log K}{M} = \infty \\ c_1, & \text{for} \quad K = \Theta(\exp(\eta^2 M)), \\ & \quad\ \text{or} \quad M = \Theta(\log K) \\ 0, & \text{for} \quad \lim_{M,K\to\infty} \frac{\log K}{M} = 0 \end{cases} \end{aligned} \tag{84}$$

as $M, K \to \infty$, where $c_1 \in (0, 1)$ is a constant. This is the physical intuition behind the results in [8].

ACKNOWLEDGMENT

The authors wish to thank the anonymous reviewers for their valuable comments and discussion, especially for pointing out some issues in the proofs, the speed of convergence, the issue of optimality of random beamforming under the UR-SP channel model beyond linear sum rate scaling w.r.t. the number M of transmit antennas, and the issue of setting the number K of users as $K = M^q$, $0 \leq q \leq 1$.

REFERENCES

[1] G. Lee, Y. Sung, and M. Kountouris, "On the performance of randomly directional beamforming between line-of-sight and rich scattering channels," in *Proc. IEEE Int. Workshop Signal Process. Adv. Wireless Commun*. Stockholm, Sweden, Jun. 2015, pp. 141–145.

[2] G. Lee, Y. Sung, and M. Kountouris, "On the performance of random beamforming in sparse millimeter wave channels," *IEEE J. Sel. Topics Signal Process.*, Jun. 2015, submitted for publication.

[3] G. Lee, Y. Sung, and J. Seo, "How many users are needed for non-trivial performance of random beamforming in highly-directional mm-wave MIMO downlink?," in *Proc. IEEE Inf. Theory Workshop*, Jeju, Korea, Oct. 2015.

[4] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath Jr., "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[5] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proc. IEEE*, vol. 98, no. 6, pp. 1058–1076, Jun. 2010.

[6] J. Seo, Y. Sung, G. Lee, and D. Kim, "Pilot beam sequence design for channel estimation in millimeter-wave MIMO systems: A POMDP framework," in *Proc. IEEE Signal Process. Adv. Wireless Commun.*, Stockholm, Sweden, Jun. 2015, pp. 236–240.

[7] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath Jr., "Hybrid precoding for millimeter wave cellular systems with partial channel knowledge," in *Proc. Inf. Theory Appl. Workshop*, San Diego, CA, USA, 2013, pp. 1–5.

[8] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channels with partial side information," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 506–522, Feb. 2005.

[9] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.

[10] A. Tomasoni, G. Caire, M. Ferrari, and S. Bellini, "On the selection of semi-orthogonal users for zero-forcing beamforming," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2009, pp. 1100–1104.

[11] T. Al-Naffouri, M. Sharif, and B. Hassibi, "How much does transmit correlation affect the sum-rate scaling of MIMO Gaussian broadcast channels?," *IEEE Trans. Commun.*, vol. 57, no. 2, pp. 562–572, Feb. 2009.

[12] H. Hur, A. M. Tulino, and G. Caire, "Network MIMO with linear zero-forcing beamforming: Large system analysis, impact of channel estimation, and reduced-complexity scheduling," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2911–2934, May 2012.

[13] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[14] J. Nam, A. Adhikary, J. Ahn, and G. Caire, "Joint spatial division and multiplexing: Opportunistic beamforming, user grouping and simplified downlink scheduling," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 876–890, Oct. 2014.

[15] G. Lee and Y. Sung, "A new approach to user scheduling in massive multi-user MIMO broadcast channels," *IEEE Trans. Inf. Theory.*, Mar. 2014 [Online]. Available: http://arxiv.org/pdf/1403.6931.pdf

[16] J. Chung, C. Hwang, K. Kim, and Y. K. Kim, "A random beamforming technique in MIMO systems exploiting multiuser diversity," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 848–855, Jun. 2003.

[17] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dump antennas," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1277–1294, Jun. 2002.

[18] T. S. Rappaport, E. Ben-Dor, J. N. Murdock, and Y. Qiao, "38 GHz and 60 GHz angle-dependent propagation for cellular & peer-to-peer wireless communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2012, pp. 4568–4573.

[19] A. Sayeed and J. Brady, "Beamspace MIMO for high-dimensional multiuser communication at millimeter-wave frequencies," in *Proc. IEEE Global Telecommun. Conf. (Globecom)*, Dec. 2013, pp. 3679–3684.

[20] H. Q. Ngo, E. G. Larsson, and T. L Marzetta, "Aspects of favorable propagation in massive MIMO," in *Proc. IEEE 22nd Eur. Signal Process. Conf.*, Sep. 2014, pp. 76–80.

[21] A. Alkhateeb, G. Leus, and R. W. Heath Jr., "Limited feedback hybrid precoding for multi-user millimeter wave systems," *arXiv preprint arXiv:1409.5162v2*, Mar. 2015.

[22] A. Alkhateeb, R. W. Heath Jr., and G. Leus, "Achievable rates of multi-user millimeter wave systems with hybrid precoding," *arXiv preprint arXiv:1505.00303v1*, May 2015.

[23] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4391–4403, Oct. 2013.

[24] C. N. Barati, S. A. Hosseini, S. Rangan, P. Liu, T. Korakis, and S. S. Panwar, "Directional cell search for millimeter wave cellular systems," in *Proc. IEEE Int. Workshop Signal Process. Adv. Wireless Commun.*, Toronto, ON, Canada, Jun. 2014, pp. 120–124.

[25] A. Alkhateeb, G. Leus, and R. W. Heath Jr., "Compressed sensing based multi-user millimeter wave systems: How many measurements are needed?," *arXiv preprint arXiv:1505.00299v1*, May 2015.

[26] T. Bai, V. Desai, and R. W. Heath, "Millimeter wave cellular channel models for system evaluation," in *Proc. IEEE Int. Comput. Netw. Commun.*, Feb. 2014, pp. 178–182.

[27] R. S. Strichartz, *The Way of Analysis*. Boston, MA, USA: Jones & Bartlett, 2000.

[28] H. L. V. Trees, *Optimum Array Processing*. New York, NY, USA: Wiley, 2002.

[29] M. Kountouris, D. Gesbert, and T. Salzer, "Enhanced multiuser random beamforming: Dealing with the not so large number of users case," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 28, pp. 1536–1545, Oct. 2008.

[30] J. L. Vicario, B. Bosisio, C. Anton-Haro, and U. Spagnolini, "Beam selection strategies for orthogonal random beamforming in sparse networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 9, pp. 3385–3396, Sep. 2008.

Gilwon Lee (S'10) received the B.S. and M.S. degrees in electrical engineering from KAIST, Daejeon, South Korea, in 2010 and 2012, respectively. He is currently pursuing the Ph.D. degree at the Wireless Information Systems Research Group, KAIST. His research interests include the design and analysis of large-scale MIMO systems and signal processing for next wireless communications.

Youngchul Sung (S'92–M'93–SM'09) received the B.S. and M.S. degrees in electronics engineering from Seoul National University, Seoul, South Korea, and the Ph.D. degree in electrical and computer engineering from Cornell University, Ithaca, NY, USA, in 1993, 1995, and 2005, respectively. From 1995 to 2000, he worked at LG Electronics, Ltd., Seoul, South Korea. From 2005 to 2007, he was a Senior Engineer with the Corporate R&D Center of Qualcomm, Inc., San Diego, CA, USA, and participated in design of WCDMA base station modem. Since 2007, he has been on the Faculty with the Department of Electrical Engineering, KAIST, Daejeon, South Korea. His research interests include signal processing for communications, statistical signal processing, asymptotic statistics, and information geometry. He is currently a Member of the UNESCO/Netexplo University Advisory Board, Signal and Information Processing Theory and Methods (SIPTM) Technical Committee of Asia-Pacific Signal and Information Processing Association (APSIPA), Vice-Chair of the IEEE ComSoc Asia-Pacific Board ISC, and was an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS from 2012 to 2014.

Junyeong Seo (S'11) received the B.S. and M.S. degrees in electrical engineering from KAIST, Daejeon, South Korea, in 2011 and 2013, respectively. He is currently pursuing the Ph.D. degree at the Wireless Information Systems Research Group, KAIST. His research interests include the design and analysis of large-scale MIMO systems and signal processing for next wireless communications.