

University of  
**Waterloo**



## **On the Capacity of MIMO Rician Broadcast Channels**

Alireza Bayesteh, Kamyar Moshksar, and Amir K. Khandani

Technical Report UW-E&CE#2008-3

Jan. 7 2008

©A. Bayesteh, K. Moshksar, and A. K. Khandani 2008

# On the Capacity of MIMO Rician Broadcast Channels

Alireza Bayesteh, Kamyar Moshksar, and Amir K. Khandani

Dept. of Electrical Engineering

University of Waterloo

Waterloo, ON, N2L 3G1

alireza, kmoshksa, khandani@shannon2.uwaterloo.ca

## Abstract

In this paper, a downlink communication system, in which a Base Station (BS) equipped with  $M$  antennas communicates with  $N$  ( $N \gg 1$ ) single-antenna users, in a Rician fading environment is considered. The asymptotic (in terms of the number of users) sum-rate capacity of the system, as well as the capacity-achieving strategies, are derived. The main results of the paper are as follows: i) in the region of  $\mathcal{K} = o(\log N)$ , where  $\mathcal{K}$  denotes the *Rician factor*, the sum-rate capacity scales as  $M \log(1 + \frac{P}{M}\eta)$ , where  $P$  denotes the SNR and  $\eta \triangleq \frac{\log N}{1+\mathcal{K}}$ , which is achieved by Zero-Forcing Beam-Forming (ZFBF) along with a low-complexity user selection algorithm that considers only the scattered component of the users' channels, ii) in the region  $\mathcal{K} = \omega(\log N)$ , in the case of co-located transmit antennas, the capacity scales as  $\log(1 + MP)$ , which is achieved by transmitting to any arbitrary user, iii) in the region  $\mathcal{K} = \omega(\log N)$ , in the case of isotropically-distributed specular components, the sum-rate capacity behaves as  $M \log(1 + P)$ , which is achieved by ZFBF, along with a user selection algorithm that considers only the specular component of the users' channels. Simulation results are provided to examine the analytical results in practical networks. Simulation results are provided to examine the analytical results in practical networks.

## I. INTRODUCTION

Multiple-input multiple-output (MIMO) systems have proved their ability to achieve high bit rates on a scattering wireless network [1], [2]. In a MIMO broadcast channel, the base station equipped with multiple antennas communicates with several users. Recently, there has been a lot of interest in characterizing the capacity region of this channel [4]–[7]. In these works, it has been demonstrated that the sum-rate capacity of MIMO broadcast channels can be achieved by applying dirty-paper coding (DPC) [8] at the transmitter.

Despite the fact that the sum-rate capacity of Gaussian MIMO-BC is known, it is still interesting to study the behavior of sum-rate capacity in various scenarios. [9] compares the achievable sum-rate of MIMO-BC for DPC to that achieved by using linear precoding schemes, and characterizes the gap

Financial supports provided by Nortel, and the corresponding matching funds by the Federal government: Natural Sciences and Engineering Research Council of Canada (NSERC) and Province of Ontario: Ontario Centres of Excellence (OCE) are gratefully acknowledged.

between the achievable sum-rates in the high SNR regime. [10] compares the achievable sum-rate of DPC to that of Time Division Multiple Access (TDMA) for a Gaussian MIMO-BC. [11] considers a MIMO-BC with a large number of users and shows that i) the sum-rate capacity of the system scales as  $M \log \log N$ , when  $N$  is the number of users in the network, and ii) a simple scheme of “Random Beam-Forming” asymptotically achieves the sum-rate capacity as  $N \rightarrow \infty$ . References [12]–[14] consider the same network set-up and prove that one can achieve the sum-rate capacity of the system by utilizing Zero-Forcing Beam-Forming at the transmitter, provided that the user selection is performed wisely. In [15] the scaling laws of the sum-rate for fading MIMO Gaussian broadcast channels using time-sharing to the strongest user, DPC and beamforming, is derived for the asymptotic case of  $N \rightarrow \infty$ . In all the mentioned papers ([9]- [15]), the channel model is assumed to be Rayleigh fading. Therefore, it is of interest to investigate the sum-rate capacity of MIMO-BC, assuming more general channel models.

One of the most widely-used models for the wireless channels is Rician fading. This model is suitable for wireless links when there is a line of sight (LOS) link between the transmitter and receiver. Several papers in the literature consider Rician fading in the context of point-to-point MIMO communications. In [16], the authors derive the exact capacity of MIMO Rician channel, when perfect Channel State Information (CSI) is available at the receiver, but the transmitter has neither instantaneous nor statistical CSI. Reference [17] studies the capacity of MIMO Rician channel in the high and low SNR regimes, for both coherent and non-coherent communications. It is shown in [17] that in the low SNR regime, the specular component of the channel completely determines the form of the optimum signal whereas in the high SNR regime it has no effect on the optimum signal structure. In [18], the authors consider the min-capacity of a MIMO Rician channel with an unknown deterministic specular component. [19] analyzes the capacity of a MIMO Rician channel with isotropically random rank-one specular component, when the channel is unknown at both the transmitter and receiver sides.

In this paper, we consider a Rician MIMO-BC, in which a transmitter equipped with  $M$  antennas communicates with  $N$  ( $N \gg 1$ ) single-antenna users. The channels are assumed to be perfectly known at both the transmitter and receiver sides. The asymptotic (in terms of the number of users) sum-rate capacity of the system, as well as the capacity-achieving strategies, are derived. The main results of the paper are as follows: i) in the region of  $\mathcal{K} = o(\log N)$ , where  $\mathcal{K}$  denotes the *Rician factor*, the sum-rate capacity scales as  $M \log(1 + \frac{P}{M} \eta)$ , where  $P$  denotes the SNR and  $\eta \triangleq \frac{\log N}{1 + \mathcal{K}}$ , which is achieved by Zero-Forcing Beam-Forming (ZFBF) along with a low-complexity user selection algorithm that considers only the scattered component of the users’ channels, ii) in the region  $\mathcal{K} = \omega(\log N)$ , in the case of co-located transmit antennas, the capacity scales as  $\log(1 + MP)$ , which is achieved by TDMA, iii) in the region  $\mathcal{K} = \omega(\log N)$ , in the case of isotropically-distributed specular components, the sum-rate capacity behaves

as  $M \log(1+P)$ , which is achieved by ZFBF, along with a user selection algorithm that considers only the specular component of the users' channels. Simulation results confirm the validity of analytical results.

The rest of the paper is organized as follows. In II, we introduce the system model. Section III is devoted to analyzing the asymptotic sum-rate capacity of the underlying system. Some simulation results are presented in section IV, and finally, section V concludes the paper.

Throughout this paper, the norm of the vectors are denoted by  $\|\cdot\|$ , the Hermitian operation is denoted by  $(\cdot)^H$ , and the determinant and the trace operations are denoted by  $|\cdot|$  and  $\text{Tr}(\cdot)$ , respectively.  $\mathbb{E}\{\cdot\}$  represents the expectation, notation "log" is used for the natural logarithm, and the rates are expressed in *nats*. For any given functions  $f(N)$  and  $g(N)$ ,  $f(N) = O(g(N))$  is equivalent to  $\lim_{N \rightarrow \infty} \left| \frac{f(N)}{g(N)} \right| < \infty$ ,  $f(N) = o(g(N))$  is equivalent to  $\lim_{N \rightarrow \infty} \left| \frac{f(N)}{g(N)} \right| = 0$ ,  $f(N) = \Omega(g(N))$  is equivalent to  $\lim_{N \rightarrow \infty} \frac{f(N)}{g(N)} > 0$ ,  $f(N) = \omega(g(N))$  is equivalent to  $\lim_{N \rightarrow \infty} \frac{f(N)}{g(N)} = \infty$ , and  $f(N) = \Theta(g(N))$  is equivalent to  $\lim_{N \rightarrow \infty} \frac{f(N)}{g(N)} = c$ , where  $0 < c < \infty$ , and  $f(N) \sim g(N)$  is equivalent to  $\lim_{N \rightarrow \infty} \frac{f(N)}{g(N)} = 1$ .

## II. SYSTEM MODEL

In this work, a MIMO-BC in which a base station equipped with  $M$  antennas communicates with  $N$  users, each equipped with single antennas, is considered. The received signal by user  $k$  can be written as

$$y_k = \mathbf{H}_k \mathbf{x} + n_k, \quad (1)$$

where  $\mathbf{x} \in \mathbb{C}^{M \times 1}$  is the transmitted signal,  $\mathbf{H}_k \in \mathbb{C}^{1 \times M}$  is the channel vector from the transmitter to the  $k$ th user, which is assumed to be perfectly known at the receiver side and provided to the BS via a noiseless feedback channel<sup>1</sup>, and  $n_k \sim \mathcal{CN}(0, 1)$  is the AWGN at this receiver.

Under Rician channel model,  $\mathbf{H}_k$  can be written as

$$\mathbf{H}_k = \sqrt{1 - r_k} \mathbf{G}_k + \sqrt{r_k M} \mathbf{b}_k, \quad (2)$$

where  $\mathbf{G}_k$  is a circularly symmetric zero mean unit variance Gaussian vector, reflecting the scattered component and  $\mathbf{b}_k$  is a unit-norm vector representing the specular component of the channel, and  $r_k$  is a constant related to the Rician factor  $\mathcal{K}_k$ <sup>2</sup> via  $r_k = \frac{\mathcal{K}_k}{\mathcal{K}_k + 1}$ . We consider two scenarios for  $\mathbf{b}_k$ : (i) The entries of  $\mathbf{H}_k$  are i.i.d Gaussian with mean  $b_k$  and variance  $1 - |b_k|^2$ , where  $b_k$  is a complex number satisfying  $|b_k|^2 = r_k$ . In this case, it is easy to observe that  $\mathbf{b}_k = \frac{e^{j\theta_k}}{\sqrt{M}} \mathbf{1}$ , where  $\mathbf{1}$  is the vector of all ones. This model corresponds to the case that the transmit antennas are co-located, and consequently, the specular components from all transmit antennas to each of the users are equal<sup>3</sup>. (ii) The vector  $\mathbf{b}_k$  is isotropically

<sup>1</sup>In fact, the BS does not need to have the perfect CSI about all the users' channels. However, the partial CSI that the BS receives through feedback is based on the perfect CSI that the receivers have.

<sup>2</sup>Rician factor is defined as the ratio of the power of the specular component to the power of the scattered component.

<sup>3</sup>Note that however, the specular components from each transmit antenna to different users are not necessarily equal.

distributed in the unit sphere. This model has been used in [19]. It is assumed that  $r_k$  is fixed for all the users during the whole transmission period and is equal to a constant  $r$ , i.e.,  $r_1 = r_2 = \dots = r_N = r$ .

We assume that the transmitter has an average power constraint  $P$ , i.e.,

$$\mathbb{E} \{ \text{Tr}(\mathbf{x}\mathbf{x}^*) \} \leq P.$$

The power constraint is assumed to be *per frame*. In other words, the power constraint is independent of the channel realization. The channels are assumed to be quasi-static block fading, in which each channel  $\mathbf{H}_k$  is drawn randomly at the start of each transmission frame and remains constant for the whole transmission frame, and changes independently to another realization in the start of the next frame. The frame itself is assumed to be long enough to allow communication at rates close to the capacity. Defining the sum-rate capacity of the system in the channel realization  $\mathcal{H} \triangleq \{\mathbf{H}_k\}_{k=1}^N$ , when the transmitter has perfect CSI about all users' channels, as  $\mathcal{C}_{\text{sum}}(\mathcal{H})$ , the average sum-rate capacity, denoted as  $\mathcal{C}_{\text{sum}}$ , is defined as the average over time of  $\mathcal{C}_{\text{sum}}(\mathcal{H})$ , which is by the ergodicity of the channel, equal to  $\mathbb{E}_{\mathcal{H}} \{ \mathcal{C}_{\text{sum}}(\mathcal{H}) \}$ .  $\mathcal{C}_{\text{sum}}$  is shown in [4] to be equal to

$$\mathcal{C}_{\text{sum}} = \mathbb{E}_{\mathcal{H}} \left\{ \max_{\substack{P_k \\ \sum P_k = P}} \log \det \left( \mathbf{I}_M + \sum_{k=1}^N \mathbf{H}_k^* P_k \mathbf{H}_k \right) \right\}, \quad (3)$$

where  $P_k$  is the transmit power allocated to the  $k$ th user.

### III. ASYMPTOTIC ANALYSIS; CAPACITY COMPUTATION

In this section, we compute the capacity of MIMO-BC under Rician fading, in the asymptotic scenario of  $N \rightarrow \infty$ . To this end, we consider two cases; (i)  $\mathcal{K} = o(\log N)$  and (ii)  $\mathcal{K} = \omega(\log N)$ . For each case, we provide a lower-bound and upper-bound for the capacity and prove that as  $N \rightarrow \infty$ , these bounds converge to each other.

A.  $\mathcal{K} = o(\log N)$

**Theorem 1** *The capacity of the underlying MIMO-BC in the case of  $\mathcal{K} = o(\log N)$  equals*

$$\mathcal{C}_{\text{sum}} = M \log \left( 1 + \frac{P \log N}{M(1 + \mathcal{K})} \right) + o(1), \quad (4)$$

*which is asymptotically achievable by ZFBF.*

**Proof** - The proof is based on the upper-bound and lower-bound given as follows:

1) *Upper-bound:* Using (2), the upper-bound for the sum-rate capacity can be derived as [11]

$$\begin{aligned}
\mathcal{C}_{\text{sum}} &\leq M\mathbb{E} \left\{ \log \left( 1 + \frac{P}{M} \|\mathbf{H}\|_{\max}^2 \right) \right\} \\
&\leq M\mathbb{E} \left\{ \log \left( 1 + \frac{P}{M} \left[ \sqrt{1-r} \|\mathbf{G}_k\| + \sqrt{rM} \|\mathbf{b}_k\| \right]_{\max}^2 \right) \right\} \\
&\stackrel{(a)}{=} M\mathbb{E} \left\{ \log \left( 1 + \frac{P}{M} \max_k \left[ \sqrt{\frac{1}{1+\mathcal{K}}} \|\mathbf{G}_k\| + \sqrt{\frac{\mathcal{K}M}{1+\mathcal{K}}} \right]^2 \right) \right\} \\
&= M\mathbb{E} \left\{ \log \left( 1 + \frac{P}{M} \frac{1}{1+\mathcal{K}} \|\mathbf{G}\|_{\max}^2 \right) + \right. \\
&\quad \left. \log \left( 1 + \frac{\frac{P}{M} \left( \frac{2\sqrt{\mathcal{K}M}}{1+\mathcal{K}} \|\mathbf{G}\|_{\max} + \frac{M\mathcal{K}}{1+\mathcal{K}} \right)}{1 + \frac{P}{M} \frac{1}{1+\mathcal{K}} \|\mathbf{G}\|_{\max}^2} \right) \right\} \\
&\stackrel{(b)}{\leq} M\mathbb{E} \left\{ \log \left( 1 + \frac{P}{M} \frac{1}{1+\mathcal{K}} \|\mathbf{G}\|_{\max}^2 \right) + \right. \\
&\quad \left. \frac{2\sqrt{\mathcal{K}M} \|\mathbf{G}\|_{\max}}{\frac{M}{P}(1+\mathcal{K}) + \|\mathbf{G}\|_{\max}^2} + \frac{\mathcal{K}M}{\frac{M}{P}(1+\mathcal{K}) + \|\mathbf{G}\|_{\max}^2} \right\} \\
&\stackrel{(c)}{\leq} M \log \left( 1 + \frac{P}{M} \frac{1}{1+\mathcal{K}} \mathbb{E} \{ \|\mathbf{G}\|_{\max}^2 \} \right) + M\mathbb{E} \left\{ \frac{2\sqrt{\mathcal{K}M} \|\mathbf{G}\|_{\max}}{\frac{M}{P}(1+\mathcal{K}) + \|\mathbf{G}\|_{\max}^2} \right\} + \\
&\quad M\mathbb{E} \left\{ \frac{\mathcal{K}M}{\frac{M}{P}(1+\mathcal{K}) + \|\mathbf{G}\|_{\max}^2} \right\}, \tag{5}
\end{aligned}$$

where (a) follows from the facts that  $r = \frac{\mathcal{K}}{1+\mathcal{K}}$  and  $\|\mathbf{b}_k\| = 1$ , (b) results from upper-bounding  $\log(1+x)$  by  $x$ , and (c) follows from the concavity of  $\log(\cdot)$  function which incurs that  $\mathbb{E} \{ \log(1 + \frac{P}{M} \frac{1}{1+\mathcal{K}} \|\mathbf{G}\|_{\max}^2) \} \leq \log(1 + \frac{P}{M} \frac{1}{1+\mathcal{K}} \mathbb{E} \{ \|\mathbf{G}\|_{\max}^2 \})$ . Defining  $A \triangleq \frac{M}{P}(1+\mathcal{K})$ ,  $t \triangleq \log N + (M-3) \log \log N$ , and  $\mathfrak{A}$  as the event that  $\|\mathbf{G}\|_{\max}^2 \leq t$ , we have

$$\begin{aligned}
\mathbb{E} \left\{ \frac{2\sqrt{\mathcal{K}M} \|\mathbf{G}\|_{\max}}{A + \|\mathbf{G}\|_{\max}^2} \right\} &= \mathbb{E} \left\{ \frac{2\sqrt{\mathcal{K}M} \|\mathbf{G}\|_{\max}}{A + \|\mathbf{G}\|_{\max}^2} \middle| \mathfrak{A} \right\} \Pr\{\mathfrak{A}\} + \\
&\quad \mathbb{E} \left\{ \frac{2\sqrt{\mathcal{K}M} \|\mathbf{G}\|_{\max}}{A + \|\mathbf{G}\|_{\max}^2} \middle| \mathfrak{A}^C \right\} \Pr\{\mathfrak{A}^C\} \\
&\leq \sqrt{\frac{\mathcal{K}M}{A}} \Pr\{\mathfrak{A}\} + \frac{2\sqrt{\mathcal{K}M}t}{A+t}, \tag{6}
\end{aligned}$$

where the second line results from the fact that  $\frac{2\|\mathbf{G}\|_{\max}}{A+\|\mathbf{G}\|_{\max}^2} \leq \frac{1}{\sqrt{A}}$  and also the function  $\frac{2\|\mathbf{G}\|_{\max}}{A+\|\mathbf{G}\|_{\max}^2}$  is decreasing for  $\|\mathbf{G}\|_{\max}^2 \geq A$ , noting that as  $A = o(\log N)$  (since  $\mathcal{K} = o(\log N)$ ), we have  $t > A$ . By a similar approach, the third term in RH(5) can be upper-bounded as

$$\mathbb{E} \left\{ \frac{\mathcal{K}M}{\frac{M}{P}(1+\mathcal{K}) + \|\mathbf{G}\|_{\max}^2} \right\} \leq \frac{\mathcal{K}M}{A} \Pr\{\mathfrak{A}\} + \frac{\mathcal{K}M}{A+t}. \tag{7}$$

In Appendix V, it has been shown that  $\Pr\{\mathfrak{A}\} = o\left(\frac{1}{N}\right)$ . Noting that  $\frac{\mathcal{K}M}{A} = O(1)$  and  $\mathcal{K} = o(\log N)$ , which incurs that  $\mathcal{K} = o(t)$ , we have

$$\mathbb{E} \left\{ \frac{2\sqrt{\mathcal{K}M} \|\mathbf{G}\|_{\max}}{A + \|\mathbf{G}\|_{\max}^2} \right\} = o(1), \quad (8)$$

and

$$\mathbb{E} \left\{ \frac{\mathcal{K}M}{\frac{M}{P}(1 + \mathcal{K}) + \|\mathbf{G}\|_{\max}^2} \right\} = o(1). \quad (9)$$

Substituting in (5), the upper-bound on the sum-rate capacity can be written as

$$\begin{aligned} \mathcal{C}_{\text{sum}} &\leq M \log \left( 1 + \frac{P}{M} \frac{1}{1 + \mathcal{K}} \mathbb{E} \{ \|\mathbf{G}\|_{\max}^2 \} \right) + o(1) \\ &= M \log \left( 1 + \frac{P \log N}{M(1 + \mathcal{K})} \right) + o(1), \end{aligned} \quad (10)$$

where the second line follows from the fact that  $\mathbb{E} \{ \|\mathbf{G}\|_{\max}^2 \} = \log N + O(\log \log N)$  [11].

2) *Achievability: Scheduling based the scattered component:* Consider the following algorithm:

*Algorithm 1*

- Set the threshold  $t = \log N + (M - 3) \log \log N$
- Among the users in the following set:

$$\mathcal{S} \triangleq \{k \mid \|\mathbf{G}_k\|^2 > t\}, \quad (11)$$

select one user at random. Call this user  $s_1$ , and define  $\mathcal{S}_1 \triangleq \mathcal{S} - \{s_1\}$ .

- For  $m = 2$  to  $M$ , repeat the following:
  - Denote the set of selected users up to the  $(m - 1)$ th step as  $\mathcal{A}_m \triangleq \{s_1, \dots, s_{m-1}\}$ . Define  $\mathcal{S}_m \triangleq \mathcal{S} - \mathcal{A}_m$ .
  - Define  $\mathcal{P}_m$  as the sub-space spanned by the scattered channel components of the users selected in the previous steps, i.e.,  $\{\mathbf{v}_{s_j}\}_{j=1}^{m-1}$ , where  $\mathbf{v}_k \triangleq \frac{\mathbf{G}_k}{\|\mathbf{G}_k\|}$ ,  $k = 1, \dots, N$ .
  - Let  $\{\Phi_j\}_{j=1}^{m-1}$  be  $m - 1$  orthonormal bases for  $\mathcal{P}_m$ . Then,

$$s_m = \arg \min_{k \in \mathcal{S}_m} \sum_{j=1}^{m-1} |\mathbf{v}_k \Phi_j^H|. \quad (12)$$

In the above algorithm, the user selection is solely performed based on the scattered component of the channel. First, the users with scattered channel gains above the threshold  $t$  are candidated. After that, the algorithm tries to find a set of semi-orthogonal channel vectors out of the candidate users. To this end, at each step of the algorithm, the user whose scattered channel vector is the most orthogonal to the sub-space spanned by the previously selected users' scattered channel vectors is selected. After selecting the users, the BS performs zero-forcing beam-forming on the (whole) channel vectors of the selected

users. Defining  $\mathbb{H} \triangleq [\mathbf{H}_{s_1}^T | \cdots | \mathbf{H}_{s_M}^T]^T$  and  $\mathbf{u} = [u_1, \cdots, u_M]^T$  as the information vector for the selected users, we have

$$\mathbf{x} = \mathbb{H}^{-1} \mathbf{u}. \quad (13)$$

Therefore, the achievable sum-rate of this scheme can be written as

$$\mathcal{R} = M \mathbb{E}_{\mathbb{H}} \left\{ \log \left( 1 + \frac{P}{\text{Tr} \{ [\mathbb{H}^H \mathbb{H}]^{-1} \}} \right) \right\}. \quad (14)$$

Defining  $\mathfrak{B}$  as the event that  $L \triangleq |\mathcal{S}| > \log N$ ,  $\mathfrak{C}$  as  $\delta(\mathbb{G}^H) > 1 + 2M (\log N)^{-\frac{1}{2(M-1)}}$ , and  $\mathfrak{D}$  as the event that  $\|\mathbf{G}\|_{\max}^2 \leq t^+$ , where  $\delta(\mathbf{A})$  denotes the orthogonality defect [20] of  $\mathbf{A}$ ,  $\|\mathbf{G}\|_{\max}^2 \triangleq \max_k \|\mathbf{G}_k\|^2$ , and  $t^+ \triangleq \log N + M \log \log N$ , we have

$$\begin{aligned} \mathcal{R} &= M \mathbb{E}_{\mathbb{H}|\mathfrak{B}, \mathfrak{C}, \mathfrak{D}} \left\{ \log \left( 1 + \frac{P}{\text{Tr} \{ [\mathbb{H}^H \mathbb{H}]^{-1} \}} \right) \middle| \mathfrak{B}, \mathfrak{C}, \mathfrak{D} \right\} \Pr\{\mathfrak{B}, \mathfrak{C}, \mathfrak{D}\} + \\ &\quad M \mathbb{E}_{\mathbb{H}|\mathfrak{B}^C \cup \mathfrak{C}^C \cup \mathfrak{D}^C} \left\{ \log \left( 1 + \frac{P}{\text{Tr} \{ [\mathbb{H}^H \mathbb{H}]^{-1} \}} \right) \middle| \mathfrak{B}^C \cup \mathfrak{C}^C \cup \mathfrak{D}^C \right\} \times \\ &\quad \Pr\{\mathfrak{B}^C \cup \mathfrak{C}^C \cup \mathfrak{D}^C\} \\ &\geq M \mathbb{E}_{\mathbb{H}|\mathfrak{B}, \mathfrak{C}, \mathfrak{D}} \left\{ \log \left( 1 + \frac{P}{\text{Tr} \{ [\mathbb{H}^H \mathbb{H}]^{-1} \}} \right) \middle| \mathfrak{B}, \mathfrak{C}, \mathfrak{D} \right\} \Pr\{\mathfrak{B}, \mathfrak{C}, \mathfrak{D}\} \\ &\geq \left( M \log P - M \mathbb{E}_{\mathbb{H}|\mathfrak{B}, \mathfrak{C}} \left\{ \log \left( \text{Tr} \{ [\mathbb{H}^H \mathbb{H}]^{-1} \} \right) \middle| \mathfrak{B}, \mathfrak{C}, \mathfrak{D} \right\} \right) \Pr\{\mathfrak{B}\} \Pr\{\mathfrak{C}|\mathfrak{B}\} \times \\ &\quad \Pr\{\mathfrak{D}|\mathfrak{B}, \mathfrak{C}\}. \end{aligned} \quad (15)$$

In Appendix V, it has been shown that  $\Pr\{\mathfrak{B}\} = 1 + o\left(\frac{1}{N}\right)$  and  $\Pr\{\mathfrak{C}|\mathfrak{B}\} = 1 + o\left(\frac{1}{\log N}\right)$ , and  $\Pr\{\mathfrak{D}|\mathfrak{B}, \mathfrak{C}\} = 1 + O\left(\frac{1}{\log^2 N}\right)$ .

Defining  $\mathbb{G} \triangleq [\mathbf{G}_{s_1}^T | \cdots | \mathbf{G}_{s_M}^T]^T$ , and  $\mathbb{B} \triangleq [\mathbf{b}_{s_1}^T | \cdots | \mathbf{b}_{s_M}^T]^T$ , the term  $\text{Tr} \{ [\mathbb{H}^H \mathbb{H}]^{-1} \}$  can be written as

$$\begin{aligned} \text{Tr} \{ [\mathbb{H}^H \mathbb{H}]^{-1} \} &= \text{Tr} \left\{ \left[ \left( \sqrt{\frac{1}{\mathcal{K}+1}} \mathbb{G} + \sqrt{\frac{M\mathcal{K}}{1+\mathcal{K}}} \mathbb{B} \right)^H \left( \sqrt{\frac{1}{\mathcal{K}+1}} \mathbb{G} + \sqrt{\frac{M\mathcal{K}}{1+\mathcal{K}}} \mathbb{B} \right) \right]^{-1} \right\} \\ &= (\mathcal{K}+1) \text{Tr} \left\{ \left[ \mathbb{G}^H \mathbb{G} + \sqrt{M\mathcal{K}} (\mathbb{G}^H \mathbb{B} + \mathbb{B}^H \mathbb{G}) + M\mathcal{K} \mathbb{B}^H \mathbb{B} \right]^{-1} \right\} \\ &\stackrel{(a)}{\leq} (\mathcal{K}+1) \text{Tr} \left\{ \left[ \mathbb{G}^H \mathbb{G} - 2M \sqrt{\mathcal{K} \text{Tr} \{ \mathbb{G}^H \mathbb{G} \}} \mathbf{I} \right]^{-1} \right\} \\ &\stackrel{(b)}{\leq} (\mathcal{K}+1) \text{Tr} \left\{ \left[ \lambda_{\min} \{ \mathbb{G}^H \mathbb{G} \} \mathbf{I} - 2M \sqrt{\mathcal{K} \text{Tr} \{ \mathbb{G}^H \mathbb{G} \}} \mathbf{I} \right]^{-1} \right\} \\ &= M(\mathcal{K}+1) \left( \lambda_{\min} \{ \mathbb{G}^H \mathbb{G} \} - 2M \sqrt{\mathcal{K} \text{Tr} \{ \mathbb{G}^H \mathbb{G} \}} \right)^{-1}. \end{aligned} \quad (16)$$

In the above equation, (a) follows from the facts that for any two positive definite matrices  $\mathbf{A}$  and  $\mathbf{B}$ : i) if  $\mathbf{A} \preceq \mathbf{B}$ , then  $\text{Tr}\{\mathbf{A}\} \leq \text{Tr}\{\mathbf{B}\}$ , ii) if  $\mathbf{A} \preceq \mathbf{B}$ , then  $\mathbf{B}^{-1} \preceq \mathbf{A}^{-1}$ , iii)  $\mathbb{B}^H \mathbb{B} \succeq \mathbf{0}$ , and iv)  $\mathbb{G}^H \mathbb{B} + \mathbb{B}^H \mathbb{G} \succeq$



$-2\sqrt{M\text{Tr}\{\mathbb{G}^H\mathbb{G}\}}\mathbf{I}$ . The latter results from the fact that for any  $M \times M$  matrices  $\mathbf{A}$  and  $\mathbf{B}$ , and any  $M \times 1$  unit norm vector  $\mathbf{x}$ , we have

$$\begin{aligned} \mathbf{x}^H (\mathbf{A}^H \mathbf{B} + \mathbf{B}^H \mathbf{A}) \mathbf{x} &= 2\Re\{\mathbf{x}^H \mathbf{A}^H \mathbf{B} \mathbf{x}\} \\ &\geq -2|\mathbf{A}\mathbf{x}| |\mathbf{B}\mathbf{x}| \\ &\geq -2\sqrt{\lambda_{\max}(\mathbf{A})\lambda_{\max}(\mathbf{B})} \\ &\geq -2\sqrt{\text{Tr}\{\mathbf{A}^H \mathbf{A}\}\text{Tr}\{\mathbf{B}^H \mathbf{B}\}}, \end{aligned} \quad (17)$$

where  $\lambda_{\max}(\mathbf{A})$  denotes the maximum eigenvalue of  $\mathbf{A}^H \mathbf{A}$ . This implies that

$$|\lambda_i(\mathbf{C})| \leq 2\sqrt{\text{Tr}\{\mathbf{A}^H \mathbf{A}\}\text{Tr}\{\mathbf{B}^H \mathbf{B}\}},$$

$i = 1, \dots, M$ , where  $\lambda_i(\mathbf{C})$  denotes the  $i$ th singular value of  $\mathbf{C} = \mathbf{A}^H \mathbf{B} + \mathbf{B}^H \mathbf{A}$ . Hence,

$$\mathbf{A}^H \mathbf{B} + \mathbf{B}^H \mathbf{A} \succeq -2\sqrt{\text{Tr}\{\mathbf{A}^H \mathbf{A}\}\text{Tr}\{\mathbf{B}^H \mathbf{B}\}}\mathbf{I}. \quad (18)$$

Substituting  $\mathbf{A}$  by  $\mathbb{G}$  and  $\mathbf{B}$  by  $\mathbb{B}$ , noting that  $\text{Tr}\{\mathbf{B}^H \mathbf{B}\} = M$ , (a) follows. Also, (b) results from the fact that  $\mathbb{G}^H \mathbb{G} \succeq \lambda_{\min}(\mathbb{G}^H \mathbb{G})\mathbf{I}$ .

Conditioned on  $\mathfrak{B}$  and  $\mathfrak{D}$ ,  $\text{Tr}\{\mathbb{G}^H \mathbb{G}\}$  is upper-bounded by  $Mt^+$ . Defining  $\varepsilon \triangleq 2M(\log N)^{-\frac{1}{2(M-1)}}$ , conditioned on  $\mathfrak{C}$ , we have

$$\begin{aligned} \delta(\mathbb{G}^H) &= \frac{\prod_{i=1}^M \|\mathbf{G}_i\|^2}{|\mathbb{G}^H \mathbb{G}|} < 1 + \varepsilon \\ &\stackrel{(a)}{\longrightarrow} \frac{t^M}{\prod_{i=1}^M \lambda_i(\mathbb{G}^H \mathbb{G})} < 1 + \varepsilon \\ &\stackrel{(b)}{\longrightarrow} \frac{t^M}{\lambda_{\min}(\mathbb{G}^H \mathbb{G}) \left[ \frac{\text{Tr}(\mathbb{G}^H \mathbb{G}) - \lambda_{\min}(\mathbb{G}^H \mathbb{G})}{M-1} \right]^{M-1}} < 1 + \varepsilon \\ &\stackrel{(c)}{\longrightarrow} \frac{t^M}{\lambda_{\min}(\mathbb{G}^H \mathbb{G}) \left[ \frac{Mt^+ - \lambda_{\min}(\mathbb{G}^H \mathbb{G})}{M-1} \right]^{M-1}} < 1 + \varepsilon, \end{aligned} \quad (19)$$

where (a) follows from the fact that conditioned on  $\mathfrak{B}$ , we have  $\|\mathbf{G}_i\|^2 \geq t$ , (b) results from the fact that knowing  $\lambda_{\min}(\mathbb{G}^H \mathbb{G})$ , the product of the rest of the eigenvalues is maximized when they are equal, i.e.,

$$\prod_{i=1}^M \lambda_i(\mathbb{G}^H \mathbb{G}) \leq \lambda_{\min}(\mathbb{G}^H \mathbb{G}) \left[ \frac{Mt^+ - \lambda_{\min}(\mathbb{G}^H \mathbb{G})}{M-1} \right]^{M-1},$$

and (c) follows from the fact that conditioned on  $\mathfrak{D}$ ,  $\text{Tr}(\mathbb{G}^H \mathbb{G}) < Mt^+$ .

Defining  $\gamma \triangleq \frac{\lambda_{\min}(\mathbb{G}^H \mathbb{G})}{t^+}$ , from the above equation, we can write

$$\frac{\gamma(M-\gamma)^{M-1}}{(M-1)^{M-1}} > \frac{(t/t^+)^M}{1+\varepsilon}. \quad (20)$$

Since  $t = \log N + (M - 3) \log \log N$  and  $t^+ = \log N + M \log \log N$ , it follows that  $\frac{t}{t^+} > 1 - \frac{3 \log \log N}{\log N}$ . Hence, using the inequality  $(1 - x)^n \geq 1 - nx$ , for  $0 \leq x \leq 1$ , we have  $\left(\frac{t}{t^+}\right)^M > 1 - \frac{3M \log \log N}{\log N}$ . Moreover, using the fact that  $\frac{1}{1+\varepsilon} > 1 - \varepsilon$ , the above equation can be rewritten as

$$\frac{\gamma(M - \gamma)^{M-1}}{(M - 1)^{M-1}} > 1 - \psi, \quad (21)$$

where  $\psi \triangleq \frac{3M \log \log N}{\log N} + \varepsilon$ . Since the function  $f(\gamma) = \frac{\gamma(M-\gamma)^{M-1}}{(M-1)^{M-1}}$  is an increasing function of  $\gamma$  over the interval  $[0, 1]$ , writing the Taylor series of  $f(\gamma)$  about 1, noting that  $f(1) = 1$ ,  $f'(1) = 0$ , and  $f''(1) = -\frac{M}{M-1}$ , we have

$$\begin{aligned} \frac{\gamma(M - \gamma)^{M-1}}{(M - 1)^{M-1}} &> 1 - \psi \\ \implies \frac{M(1 - \gamma)^2}{2(M - 1)} &< \psi \\ \implies \gamma &> 1 - \sqrt{\frac{2(M - 1)\psi}{M}}. \end{aligned} \quad (22)$$

Having the fact that  $\psi = O\left((\log N)^{-\frac{1}{2(M-1)}}\right)$ , the above equation yields that conditioned on  $\mathfrak{B}$ ,  $\mathfrak{C}$  and  $\mathfrak{D}$ ,

$$\begin{aligned} \lambda_{\min}(\mathbb{G}^H \mathbb{G}) &= t^+ \left[ 1 + O\left((\log N)^{-\frac{1}{4(M-1)}}\right) \right] \\ &= \log N \left[ 1 + O\left((\log N)^{-\frac{1}{4(M-1)}}\right) \right], \end{aligned} \quad (23)$$

where the second line follows from the fact that  $t^+ = \log N + M \log \log N = \log N \left[ 1 + O\left(\frac{\log \log N}{\log N}\right) \right] = \log N \left[ 1 + O\left((\log N)^{-\frac{1}{4(M-1)}}\right) \right]$ . Substituting in (16) yields that conditioned on  $\mathfrak{B}$ ,  $\mathfrak{C}$ , and  $\mathfrak{D}$ ,

$$\begin{aligned} \text{Tr} \left\{ [\mathbb{H}^H \mathbb{H}]^{-1} \right\} &\leq M(\mathcal{K} + 1) \left( t^+ \left[ 1 + O\left((\log N)^{-\frac{1}{4(M-1)}}\right) \right] - 2M\sqrt{\mathcal{K}Mt^+} \right)^{-1} \\ &= \frac{M(\mathcal{K} + 1)}{t^+} \left[ 1 + O\left((\log N)^{-\frac{1}{4(M-1)}}\right) + O\left(\frac{1}{\sqrt{t^+}}\right) \right] \\ &\stackrel{(a)}{=} \frac{M(\mathcal{K} + 1)}{t^+} \left[ 1 + O\left((\log N)^{-\frac{1}{4(M-1)}}\right) \right] \\ &= \frac{M(\mathcal{K} + 1)}{\log N} \left[ 1 + O\left((\log N)^{-\frac{1}{4(M-1)}}\right) \right], \end{aligned} \quad (24)$$

where (a) follows from the fact that  $\frac{1}{\sqrt{t^+}} = O\left(\frac{1}{\sqrt{\log N}}\right) = o\left((\log N)^{-\frac{1}{4(M-1)}}\right)$ . Substituting in (15) yields

$$\begin{aligned} \mathcal{R} &\geq M \log \left( \frac{P \log N}{M(1 + \mathcal{K})} \left[ 1 + O\left((\log N)^{-\frac{1}{4(M-1)}}\right) \right] \right) \Pr\{\mathfrak{B}, \mathfrak{C}, \mathfrak{D}\} \\ &\stackrel{(b)}{=} \left[ M \log \left( \frac{P \log N}{M(1 + \mathcal{K})} \right) + O\left((\log N)^{-\frac{1}{4(M-1)}}\right) \right] \left[ 1 + O\left(\frac{1}{\log N}\right) \right] \\ &= M \log \left( \frac{P \log N}{M(1 + \mathcal{K})} \right) + O\left((\log N)^{-\frac{1}{4(M-1)}}\right). \end{aligned} \quad (25)$$

Since  $\mathcal{K} = o(\log K)$ , it follows that  $\log\left(\frac{P \log N}{M(1+\mathcal{K})}\right) = \log\left(1 + \frac{P \log N}{M(1+\mathcal{K})}\right) + o(1)$ . Noting this fact and comparing the above lower-bound with the upper-bound derived in (10) completes the proof of Theorem 1. ■

B.  $\mathcal{K} = \omega(\log N)$

1) *Co-located transmit antennas:* In this scenario, the specular components from all transmit antennas to each receiver are equal. In other words,  $\mathbf{b}_k = \frac{e^{i\theta_k}}{\sqrt{M}} \mathbf{1}_M$ , where  $\mathbf{1}_M$  the all-one vector with size  $M$ . However, the scattered component of all users' channels follow the circularly symmetric complex Gaussian distribution. The following theorem gives the capacity of MIMO-BC in this scenario:

**Theorem 2** *The capacity of MIMO-BC in the case of  $\mathcal{K} = \omega(\log N)$  and co-located transmit antennas scales as*

$$\mathcal{C}_{\text{sum}} = \log(1 + MP) + o(1), \quad (26)$$

*which is achievable by TDMA.*

**Proof** - Like the proof of Theorem 1, we first give an upper-bound on the sum-rate capacity and then, by giving an achievable rate which is asymptotically equal to the upper-bound the theorem is proved.

**Upper-bound:** Writing the sum-rate capacity of MIMO-BC from (3), we have

$$\begin{aligned}
\mathcal{C}_{\text{sum}} &= \mathbb{E}_{\mathcal{H}} \left\{ \max_{\substack{P_k \\ \sum P_k = P}} \log \left| \mathbf{I}_M + \sum_{k=1}^N \mathbf{H}_k^H P_k \mathbf{H}_k \right| \right\} \\
&= \mathbb{E}_{\mathfrak{G}} \left\{ \max_{\substack{P_k \\ \sum P_k = P}} \log \left| \mathbf{I}_M + \sum_{k=1}^N \left[ \sqrt{1-r_k} \mathbf{G}_k + \sqrt{r_k M} \mathbf{b}_k \right]^H P_k \left[ \sqrt{1-r_k} \mathbf{G}_k + \sqrt{r_k M} \mathbf{b}_k \right] \right| \right\} \\
&= \mathbb{E}_{\mathfrak{G}} \left\{ \max_{\substack{P_k \\ \sum P_k = P}} \log \left| \mathbf{I}_M + rM \sum_{k=1}^N \mathbf{b}_k^H P_k \mathbf{b}_k \right| + \right. \\
&\quad \left. \log \left| \mathbf{I}_M + \left( \sqrt{r(1-r)M} \sum_{k=1}^N [\mathbf{G}_k^H P_k \mathbf{b}_k + \mathbf{b}_k^H P_k \mathbf{G}_k] + (1-r) \sum_{k=1}^N \mathbf{b}_k^H P_k \mathbf{b}_k \right) \mathbf{P} \right| \right\} \\
&= \mathbb{E}_{\mathfrak{G}} \left\{ \max_{\substack{P_k \\ \sum P_k = P}} \log \left| \mathbf{I}_M + r \mathbf{1}_M^H \left( \sum_{k=1}^N P_k \right) \mathbf{1}_M \right| + \right. \\
&\quad \left. \log \left| \mathbf{I}_M + \left( \sqrt{r(1-r)M} \sum_{k=1}^N [\mathbf{G}_k^H P_k \mathbf{b}_k + \mathbf{b}_k^H P_k \mathbf{G}_k] + (1-r) \sum_{k=1}^N \mathbf{b}_k^H P_k \mathbf{b}_k \right) \mathbf{P} \right| \right\} \\
&\stackrel{(a)}{\leq} \log(1+rMP) + \\
&\quad \mathbb{E}_{\mathfrak{G}} \left\{ \max_{\substack{P_k \\ \sum P_k = P}} \log \left| \mathbf{I}_M + \sqrt{r(1-r)M} \sum_{k=1}^N [\mathbf{G}_k^H P_k \mathbf{b}_k + \mathbf{b}_k^H P_k \mathbf{G}_k] + (1-r) \sum_{k=1}^N \mathbf{G}_k^H P_k \mathbf{G}_k \right| \right\} \\
&\stackrel{(b)}{\leq} \log(1+rP) + \\
&\quad M \mathbb{E}_{\mathfrak{G}} \left\{ \max_{\substack{P_k \\ \sum P_k = P}} \log \left( 1 + \frac{\sum_{k=1}^N 2P_k \left( \sqrt{r(1-r)M} \text{Tr} \{ \mathbf{G}_k^H \mathbf{b}_k \} + (1-r) \text{Tr} \{ \mathbf{G}_k^H \mathbf{G}_k \} \right)}{M} \right) \right\}, \tag{27}
\end{aligned}$$

where  $\mathfrak{G} \triangleq \{\mathbf{G}_k\}_{k=1}^N$  and  $\mathbf{P} \triangleq (\mathbf{I}_M + rP\mathbf{1}_M^H\mathbf{1}_M)^{-1}$ . In the above equation, (a) follows from i)  $|\mathbf{I} + \mathbf{A}\mathbf{B}| = |\mathbf{I} + \mathbf{B}\mathbf{A}|$ , and hence,  $|\mathbf{I}_M + rP\mathbf{1}_M^H\mathbf{1}_M| = 1 + rP\mathbf{1}_M^H\mathbf{1}_M$ , noting that  $\mathbf{1}_M^H\mathbf{1}_M = M$ , and ii) as  $\mathbf{P} \preceq \mathbf{I}$ , we have

$$\begin{aligned}
&\log \left| \mathbf{I}_M + \left( \sqrt{r(1-r)M} \sum_{k=1}^N [\mathbf{G}_k^H P_k \mathbf{b}_k + \mathbf{b}_k^H P_k \mathbf{G}_k] + (1-r) \sum_{k=1}^N \mathbf{b}_k^H P_k \mathbf{b}_k \right) \mathbf{P} \right| \leq \\
&\quad \log \left| \mathbf{I}_M + \sqrt{r(1-r)M} \sum_{k=1}^N [\mathbf{G}_k^H P_k \mathbf{b}_k + \mathbf{b}_k^H P_k \mathbf{G}_k] + (1-r) \sum_{k=1}^N \mathbf{b}_k^H P_k \mathbf{b}_k \right|.
\end{aligned}$$

Moreover, (b) results from the fact that for any  $\mathbf{A} \succeq 0$ ,  $|\mathbf{A}| \leq \left( \frac{\text{Tr}\{\mathbf{A}\}}{M} \right)^M$ . Noting that  $\text{Tr}\{\mathbf{G}_k^H \mathbf{b}_k\} = \text{Tr}\{\mathbf{b}_k^H \mathbf{G}_k\} \leq \|\mathbf{G}_k\| \|\mathbf{b}_k\| = \frac{1}{\sqrt{M}} \|\mathbf{G}_k\|$ , and  $\text{Tr}\{\mathbf{G}_k^H \mathbf{G}_k\} = \|\mathbf{G}_k\|^2$ , the second term in the right hand side

of the above equation, denoted by  $R_2$ , can be further upper-bounded as follows:

$$\begin{aligned}
R_2 &\leq M \mathbb{E}_{\mathfrak{G}} \left\{ \max_{\substack{P_k \\ \sum P_k = P}} \log \left( 1 + \frac{\sum_{k=1}^N 2P_k \left( \sqrt{r(1-r)} \|\mathbf{G}_k\| + (1-r) \|\mathbf{G}_k\|^2 \right)}{M} \right) \right\} \\
&\stackrel{(a)}{=} M \mathbb{E} \left\{ \log \left( 1 + \frac{P \left( \sqrt{r(1-r)} \|\mathbf{G}\|_{\max} + (1-r) \|\mathbf{G}\|_{\max}^2 \right)}{M} \right) \right\} \\
&\stackrel{(b)}{\leq} M \log \left( 1 + \frac{P \left( \frac{\mathbb{E}\{\|\mathbf{G}\|_{\max}\}}{\sqrt{1+\mathcal{K}}} + \frac{\mathbb{E}\{\|\mathbf{G}\|_{\max}^2\}}{1+\mathcal{K}} \right)}{M} \right) \\
&\stackrel{(c)}{=} M \log \left( 1 + \frac{P \left( \frac{O(\sqrt{\log N})}{\sqrt{1+\mathcal{K}}} + \frac{O(\log N)}{1+\mathcal{K}} \right)}{M} \right) \\
&\stackrel{(d)}{=} o(1), \tag{28}
\end{aligned}$$

where  $\|\mathbf{G}\|_{\max} = \max_k \|\mathbf{G}_k\|$ . In the above equation, (a) results from the fact that the solution to the optimization problem in (28) is to allocate all the transmit power to the user with the highest scattered gain. (b) follows from i) the concavity of log function along with the Jensen's inequality which enables us to move the expectation inside the log, and ii) the fact that  $r = \frac{\mathcal{K}}{1+\mathcal{K}}$ , which incurs that  $r \leq 1$ , and  $1-r = \frac{1}{1+\mathcal{K}}$ . (c) results from [11], in which it is shown that  $\|\mathbf{G}\|_{\max}^2 = \log N + O(\log \log N)$  with probability one, which incurs that  $\mathbb{E}\{\|\mathbf{G}\|_{\max}^2\} = O(\log N)$  and  $\mathbb{E}\{\|\mathbf{G}\|_{\max}\} = O(\sqrt{\log N})$ , and finally, (d) follows from the assumption of  $\mathcal{K} = \omega(\log N)$  and the fact that  $\log(1+o(1)) = o(1)$ . Substituting  $R_2$  in (27) yields

$$\begin{aligned}
\mathcal{C}_{\text{sum}} &\leq \log(1+rPM) + o(1) \\
&\leq \log(1+PM) + o(1), \tag{29}
\end{aligned}$$

where the last line comes from the fact that  $r \leq 1$ .

*Achievability* - In order to show that the sum-rate given in (26) is achievable, we propose a random selection scheme, in which the transmitter selects a user at random and communicates with that user. Therefore, the maximum achievable rate is equal to the capacity of a MISO Rician channel, expressed as bellow:

$$\begin{aligned}
\mathcal{R} &= \mathbb{E}_{\mathbf{H}_k} \left\{ \max_{\substack{\mathbf{Q} \\ \text{Tr}\{\mathbf{Q}\} \leq P}} \log \left( 1 + \mathbf{H}_k \mathbf{Q} \mathbf{H}_k^H \right) \right\} \\
&= \mathbb{E} \left\{ \log \left( 1 + P \|\mathbf{H}_k\|^2 \right) \right\} \\
&\geq \mathbb{E} \left\{ \log \left( 1 + P \left| \sqrt{rM} - \sqrt{1-r} \|\mathbf{G}_k\| \right|^2 \right) \right\}. \tag{30}
\end{aligned}$$

Let us define  $\mathfrak{E}$  as the event that  $\|\mathbf{G}_k\|^2 < \log N$ .  $\mathcal{R}$  can be lower-bounded as

$$\begin{aligned} \mathcal{R} &\geq \mathbb{E} \left\{ \log \left( 1 + P \left| \sqrt{rM} - \sqrt{1-r} \|\mathbf{G}_k\| \right|^2 \right) \middle| \mathfrak{E} \right\} \Pr \{ \mathfrak{E} \} \\ &\stackrel{(a)}{=} \log(1 + PrM + o(1)) \Pr \{ \mathfrak{E} \} \\ &\stackrel{(b)}{=} \log(1 + PM) + o(1). \end{aligned} \quad (31)$$

In the above equation, (a) follows from the assumption of  $\mathcal{K} = \omega(\log N)$ , which implies that conditioned on  $\mathfrak{E}$ ,  $\sqrt{1-r} \|\mathbf{G}_k\| = \frac{\|\mathbf{G}_k\|}{\sqrt{\mathcal{K}+1}} = o(1)$ . (b) follows from i) as  $\|\mathbf{G}_k\|^2$  has Chi-Square distribution with  $2M$  degrees of freedom,  $\Pr \{ \mathfrak{E} \} \sim \frac{\log^M N}{(M-1)!N} = o(1)$  and ii) as  $r = \frac{\mathcal{K}}{\mathcal{K}+1}$  and  $\mathcal{K} = \omega(\log N)$ , we have  $r = 1 + o\left(\frac{1}{\log N}\right)$ . This completes the proof of achievability and hence, the proof of Theorem 2. ■

2) *The case of isotropic specular components:* In this case, it is assumed that the specular component of all users' channels, i.e.,  $\mathbf{b}_k$ ,  $k = 1, \dots, N$ , are isotropically distributed in the unit sphere. The difference between this case and the previous case is that in the case of co-located transmit antennas, there is only one available coordinate in the system (the coordinate of  $\mathbf{1}_M$ ) for transmission, and as a result, we don't have the  $M$ -fold capacity increase, as we expect in Gaussian MIMO-BC. However, in this case, by wisely selecting the users one can achieve the  $M$ -fold capacity increase. The following theorem gives the capacity in this case:

**Theorem 3** *The capacity of Rician MIMO-BC in the case of  $\mathcal{K} = \omega(\log N)$  and isotropic specular components is equal to*

$$\mathcal{C}_{\text{sum}} = M \log(1 + P) + o(1). \quad (32)$$

**Proof - Upper-bound:** In [11], Appendix B, an upper-bound on the capacity of MIMO-BC is given as

$$\begin{aligned} \mathcal{C}_{\text{sum}} &\leq M \mathbb{E} \left\{ \log \left( 1 + \frac{P}{M} \|\mathbf{H}\|_{\max}^2 \right) \right\} \\ &\leq M \mathbb{E} \left\{ \log \left( 1 + \frac{P}{M} \left| \sqrt{rM} + \sqrt{1-r} \|\mathbf{G}\|_{\max} \right|^2 \right) \right\} \\ &\stackrel{(a)}{\leq} M \log \left( 1 + \frac{P}{M} \mathbb{E} \left\{ \left| \sqrt{rM} + \sqrt{1-r} \|\mathbf{G}\|_{\max} \right|^2 \right\} \right) \\ &\stackrel{(b)}{=} M \log \left( 1 + \frac{P}{M} \left| \sqrt{rM} + o(1) \right|^2 \right) \\ &= M \log(1 + rP) + o(1) \\ &\stackrel{(c)}{=} M \log(1 + P) + o(1). \end{aligned} \quad (33)$$

In the above equation, (a) follows from the concavity of log function along with the Jensen's inequality, (b) results from the fact that  $\|\mathbf{G}\|_{\max} = O(\log N)$  and since  $1 - r = \frac{1}{1+\mathcal{K}} = o\left(\frac{1}{\log N}\right)$ , we have  $\sqrt{1-r} \|\mathbf{G}\|_{\max} = o(1)$ , and (c) results from  $r = 1 + o(1)$ .

*Achievability; Scheduling based on specular component* Consider the following algorithm:

*Algorithm 2*

- select one user at random. Call this user  $s_1$ , and define  $\mathcal{S}_1 \triangleq \mathcal{S} - \{s_1\}$ .
- For  $m = 2$  to  $M$ , repeat the following:
  - Denote the set of selected users up to the  $(m - 1)$ th step as  $\mathcal{A}_m \triangleq \{s_1, \dots, s_{m-1}\}$ . Define  $\mathcal{S}_m \triangleq \mathcal{S} - \mathcal{A}_m$ .
  - Define  $\mathcal{P}_m$  as the sub-space spanned by the specular channel components of the users selected in the previous steps, i.e.,  $\{\mathbf{b}_{s_j}\}_{j=1}^{m-1}$ .
  - Let  $\{\Phi_j\}_{j=1}^{m-1}$  be  $m - 1$  orthonormal bases for  $\mathcal{P}_m$ . Then,

$$s_m = \arg \min_{k \in \mathcal{S}_m} \sum_{j=1}^{m-1} |\mathbf{b}_k \Phi_j^H|. \quad (34)$$

- After selecting the users, the BS performs zero-forcing beam-forming on the (whole) channel vectors of the selected users. Defining  $\mathbb{H} \triangleq [\mathbf{H}_{s_1}^T | \dots | \mathbf{H}_{s_M}^T]^T$  and  $\mathbf{u} = [u_1, \dots, u_M]^T$  as the information vector for the selected users, we have

$$\mathbf{x} = \mathbb{H}^{-1} \mathbf{u}. \quad (35)$$

Defining the event  $\mathfrak{F} \triangleq \{\delta(\mathbb{B}^H) < 1 + \epsilon\}$  and  $\Omega \triangleq \{\text{Tr}\{\mathbb{G}^H \mathbb{G}\} < \log N\}$ , where  $\mathbb{B} = [\mathbf{b}_{s_1}^T | \dots | \mathbf{b}_{s_M}^T]^T$ ,  $\mathbb{G} = [\mathbf{G}_{s_1}^T | \dots | \mathbf{G}_{s_M}^T]^T$ , and  $\epsilon \triangleq 2MN^{-\frac{1}{2(M-1)}}$ , similar to (15), we have

$$\mathcal{R} \geq M \mathbb{E}_{\mathbb{H}|\mathfrak{F}, \Omega} \left\{ \log \left( 1 + \frac{P}{\text{Tr}\{[\mathbb{H}^H \mathbb{H}]^{-1}\}} \right) \middle| \mathfrak{F}, \Omega \right\} \Pr\{\mathfrak{F}, \Omega\}. \quad (36)$$

Since  $\mathbf{b}_k$ 's are isotropic unit vectors,  $\Pr\{\mathfrak{F}\}$  can be computed similar to  $\Pr\{\mathcal{C}|\mathfrak{B}\}$ , which is performed in Appendix V, and shown to be  $1 + o(\frac{1}{N})$ <sup>4</sup>. Moreover, since the scattered component is not considered in the selection, it follows that  $\mathbb{G}$  can be considered as an  $M \times M$  circularly symmetric complex Gaussian matrix, and as a result,  $\text{Tr}\{\mathbb{G}^H \mathbb{G}\}$  has Chi-Square distribution with  $2M^2$  degrees of freedom which implies that  $\Pr\{\text{Tr}\{\mathbb{G}^H \mathbb{G}\} > \log N\} = \frac{[\log N]^{M^2-1} e^{-\log N}}{(M-1)!} [1 + o(1)] = O\left(\frac{[\log N]^{M^2-1}}{N}\right)$ . Therefore,  $\Pr\{\Omega\} = 1 + O\left(\frac{[\log N]^{M^2-1}}{N}\right) = o\left(\frac{1}{\sqrt{N}}\right)$ . Having computed  $\Pr\{\mathfrak{F}\}$  and  $\Pr\{\Omega\}$ , noting that as the specular and scattered components of the channels are independent,  $\mathfrak{F}$  and  $\Omega$  are also independent, we have

$$\Pr\{\mathfrak{F}, \Omega\} = 1 + o\left(\frac{1}{\sqrt{N}}\right). \quad (37)$$

<sup>4</sup>To this end, it is sufficient to substitute  $\log N$  by  $N$  in the steps of proof.

Similar to (16),  $\text{Tr} \left\{ [\mathbb{H}^H \mathbb{H}]^{-1} \right\}$  can be upper-bounded as

$$\begin{aligned} \text{Tr} \left\{ [\mathbb{H}^H \mathbb{H}]^{-1} \right\} &\leq (\mathcal{K} + 1) \text{Tr} \left\{ \left( M\mathcal{K}\mathbb{B}^H \mathbb{B} - 2M\sqrt{\mathcal{K}\text{Tr}\{\mathbb{G}^H \mathbb{G}\}}\mathbf{I} \right)^{-1} \right\} \\ &\leq (\mathcal{K} + 1) \text{Tr} \left\{ \left( M\mathcal{K}\lambda_{\min}(\mathbb{B}^H \mathbb{B})\mathbf{I} - 2M\sqrt{\mathcal{K}\text{Tr}\{\mathbb{G}^H \mathbb{G}\}}\mathbf{I} \right)^{-1} \right\} \\ &= \left( 1 + \frac{1}{\mathcal{K}} \right) \left( \lambda_{\min}(\mathbb{B}^H \mathbb{B}) - 2\sqrt{\frac{\text{Tr}\{\mathbb{G}^H \mathbb{G}\}}{\mathcal{K}}} \right)^{-1}. \end{aligned} \quad (38)$$

Conditioned on  $\Omega$ , we have  $\text{Tr}\{\mathbb{G}^H \mathbb{G}\} < \log N$ , and since  $\mathcal{K} = \omega(\log N)$ , it follows that  $2\sqrt{\frac{\text{Tr}\{\mathbb{G}^H \mathbb{G}\}}{\mathcal{K}}} = o(1)$ . Moreover, conditioned on  $\mathfrak{F}$ , i.e.,  $\delta(\mathbb{B}^H) < 1 + \epsilon$ , and following the equations (19)-(23) with  $t = t^+ = 1$ , we have

$$\lambda_{\min}(\mathbb{B}^H \mathbb{B}) = 1 + O\left(N^{-\frac{1}{4(M-1)}}\right). \quad (39)$$

Combining the above equation with (38) yields

$$\text{Tr} \left\{ [\mathbb{H}^H \mathbb{H}]^{-1} \right\} \leq 1 + o(1). \quad (40)$$

Substituting in (36), noting (37), we have

$$\mathcal{R} \geq M \log(1 + P) + o(1), \quad (41)$$

which completes the proof of Theorem 3. ■

*Remark* - Comparing the sum-rate capacity of the system in the two cases of co-located transmit antennas and isotropic specular components when  $\mathcal{K} = \omega(\log N)$ , it follows that in the first case, the capacity grows logarithmically with  $M$ , while in the second case it scales linearly with  $M$ . Moreover, since  $(1 + x)^M > 1 + Mx$ ,  $\forall x, M$ , it follows that

$$\mathcal{C}_{\text{sum}}^{\text{isotropic}} \geq \mathcal{C}_{\text{sum}}^{\text{co-located}}. \quad (42)$$

### C. $\mathcal{K} = \Theta(\log N)$ , Isotropic specular components

The following theorem gives the asymptotic sum-rate in this case:

**Theorem 4** *The sum-rate capacity of the system in the case of  $\mathcal{K} = \Theta(\log N)$  and isotropic specular components can be obtained as*

$$\mathcal{C}_{\text{sum}} = M \log \left( 1 + P \left[ 1 + \sqrt{\frac{\eta}{M}} \right]^2 \right) + o(1), \quad (43)$$

where  $\eta \triangleq \lim_{N \rightarrow \infty} \frac{\log N}{\mathcal{K}}$ .



**Proof - Upper-bound:** Similar to (33), we can write

$$\begin{aligned}
\mathcal{C}_{\text{sum}} &\leq M \log \left( 1 + \frac{P}{M} \mathbb{E} \left\{ \left| \sqrt{rM} + \sqrt{1-r} \|\mathbf{G}\|_{\max} \right|^2 \right\} \right) \\
&= M \log \left( 1 + \frac{P}{M} \mathbb{E} \left\{ \left| \sqrt{rM} + \sqrt{\frac{\|\mathbf{G}\|_{\max}^2}{1+\mathcal{K}}} \right|^2 \right\} \right) \\
&\stackrel{(a)}{=} M \log \left( 1 + \frac{P}{M} \left[ \sqrt{rM} + \sqrt{\eta[1+o(1)]} \right]^2 \right) \\
&\stackrel{(b)}{=} M \log \left( 1 + P \left[ 1 + \sqrt{\frac{\eta}{M}} \right]^2 \right) + o(1), \tag{44}
\end{aligned}$$

where (a) follows from the facts that i)  $\|\mathbf{G}\|_{\max}^2 = \log N + o(\log N)$ , with probability one, and ii)  $\eta \sim \frac{\log N}{\mathcal{K}+1}$ , and (b) results from the fact that as  $\mathcal{K} = \Theta(\log N)$ , we have  $r \sim 1$ .

*Achievability; Scheduling based on both specular and scattered components:* Consider the following algorithm:

*Algorithm 3:*

- Select the thresholds  $t = \log N - 2.5 \log \log N$  and  $\gamma = \frac{2}{\log N}$ .
- Construct the following set:

$$\mathcal{S}_0 \triangleq \{k \mid \Re(\mathbf{v}_k, \mathbf{b}_k) \geq 1 - \gamma\}, \tag{45}$$

where  $\Re(x)$  denotes the real part of  $x$ , and  $\mathbf{v}_k \triangleq \frac{\mathbf{G}_k}{\|\mathbf{G}_k\|}$ ,  $k = 1, \dots, N$ .

- Among the users in the following set:

$$\mathcal{S} \triangleq \{k \in \mathcal{S}_0 \mid \|\mathbf{G}_k\|^2 > t\}, \tag{46}$$

select one user at random. Call this user  $s_1$ , and define  $\mathcal{S}_1 \triangleq \mathcal{S} - \{s_1\}$ .

- For  $m = 2$  to  $M$ , repeat the following:
  - Denote the set of selected users up to the  $(m-1)$ th step as  $\mathcal{A}_m \triangleq \{s_1, \dots, s_{m-1}\}$ . Define  $\mathcal{S}_m \triangleq \mathcal{S} - \mathcal{A}_m$ .
  - Define  $\mathcal{P}_m$  as the sub-space spanned by the scattered channel components of the users selected in the previous steps, i.e.,  $\{\mathbf{v}_{s_j}\}_{j=1}^{m-1}$ .
  - Let  $\{\Phi_j\}_{j=1}^{m-1}$  be  $m-1$  orthonormal bases for  $\mathcal{P}_m$ . Then,

$$s_m = \arg \min_{k \in \mathcal{S}_m} \sum_{j=1}^{m-1} |\mathbf{v}_k \Phi_j^H|. \tag{47}$$

- After selecting the users, the BS performs zero-forcing beam-forming on the (whole) channel vectors of the selected users, i.e.,

$$\mathbf{x} = \mathbb{H}^{-1} \mathbf{u}. \tag{48}$$

As can be observed, the above algorithm is very similar to Algorithm 1, with the difference in putting an extra constraint for the user selection, which is, the scattered and specular components of the selected users must be almost in the same direction.

Defining the events  $\mathfrak{B}$ ,  $\mathfrak{C}$ , and  $\mathfrak{D}$  as in the proof of the achievability part of Theorem 1, similar to (15), we have

$$\mathcal{R} \geq \frac{M \mathbb{E}_{\mathbb{H}|\mathfrak{B}, \mathfrak{C}, \mathfrak{D}} \left\{ \log \left( 1 + \frac{P}{\text{Tr} \{ [\mathbb{H}^H \mathbb{H}]^{-1} \}} \right) \right\}}{\Pr\{\mathfrak{B}\} \Pr\{\mathfrak{C}|\mathfrak{B}\} \Pr\{\mathfrak{D}|\mathfrak{B}, \mathfrak{C}\}} \times \quad (49)$$

$\Pr\{\mathfrak{C}|\mathfrak{B}\}$  and  $\Pr\{\mathfrak{D}|\mathfrak{B}, \mathfrak{C}\}$  can be computed from Appendix V as  $1 + o\left(\frac{1}{\log N}\right)$  and  $1 + O\left(\frac{1}{\log N}\right)$ , respectively. For computing  $\Pr\{\mathfrak{B}\}$ , we first compute  $\xi \triangleq \Pr\{k \in \mathcal{S}\}$  as follows:

$$\begin{aligned} \xi &= \Pr\{\Re(\mathbf{v}_k, \mathbf{b}_k) > 1 - \gamma, \|\mathbf{G}_k\|^2 > t\} \\ &\stackrel{(a)}{=} \Pr\{\Re(\mathbf{v}_k, \mathbf{b}_k) > 1 - \gamma\} \Pr\{\|\mathbf{G}_k\|^2 > t\} \\ &\stackrel{(b)}{\geq} \Pr\{z(\mathbf{v}_k, \mathbf{b}_k) > 1 - 0.5\gamma\} \Pr\{\cos[\Theta(\mathbf{v}_k \mathbf{b}_k^H)] > 1 - 0.5\gamma\} \Pr\{\|\mathbf{G}_k\|^2 > t\} \\ &\stackrel{(c)}{=} (0.5\gamma)^{M-1} \frac{\sqrt{\gamma} t^{M-1} e^{-t}}{\pi (M-1)!} [1 + O(1/t)] \end{aligned} \quad (50)$$

$$\stackrel{(d)}{=} \frac{\sqrt{2} \log^2 N}{\pi (M-1)! N} \left[ 1 + O\left(\frac{\log \log N}{\log N}\right) \right], \quad (51)$$

where  $\Theta(x)$  denotes the phase of a complex number  $x$ , and for any  $1 \times M$  vectors  $\mathbf{u}$  and  $\mathbf{v}$ ,  $z(\mathbf{u}, \mathbf{v})$  is defined as  $\frac{|\mathbf{u}\mathbf{v}^H|^2}{\|\mathbf{u}\|^2 \|\mathbf{v}\|^2}$ . In the above equation, (a) follows from the facts that i)  $z(\mathbf{v}_k, \mathbf{b}_k)$  is a function of only the direction of  $\mathbf{G}_k$  and for Gaussian vectors, norm and direction are independent, and ii)  $\mathbf{b}_k$  and  $\mathbf{v}_k$  are independent of each other. (b) comes from the fact that since  $\Re(\mathbf{v}_k, \mathbf{b}_k) = z(\mathbf{v}_k, \mathbf{b}_k) \cos[\Theta(\mathbf{v}_k \mathbf{b}_k^H)]$ , having the events  $z(\mathbf{v}_k, \mathbf{b}_k) > 1 - 0.5\gamma$  and  $\cos[\Theta(\mathbf{v}_k \mathbf{b}_k^H)] > 1 - 0.5\gamma$  yields  $\Re(\mathbf{v}_k, \mathbf{b}_k) > 1 - \gamma$ , and also the fact that the norm and phase of  $\mathbf{v}_k \mathbf{b}_k^H$  are independent of each other. (c) results from i) as  $\mathbf{b}_k$  and  $\mathbf{v}_k$  are two independent isotropic unit vectors, the pdf of  $Z \triangleq z(\mathbf{v}_k, \mathbf{b}_k)$  is computed in Lemma 3 of [12], as

$$p_Z(z) = (M-1)(1-z)^{M-2} \implies \Pr\{Z > 1 - \gamma\} = \gamma^{M-1}, \quad (52)$$

ii) for small enough  $x$ ,  $\cos(x) \approx 1 - \frac{x^2}{2}$ , and hence, the event  $\cos[\Theta(\mathbf{v}_k \mathbf{b}_k^H)] > 1 - 0.5\gamma$  is equivalent to  $|\Theta(\mathbf{v}_k, \mathbf{b}_k)| < \sqrt{\gamma}$ , and since  $\Theta(\mathbf{v}_k, \mathbf{b}_k)$  is uniformly distributed between 0 and  $2\pi$ , we have  $\Pr\{\cos[\Theta(\mathbf{G}_k \mathbf{b}_k^H)] > 1 - 0.5\gamma\} \approx \frac{\sqrt{\gamma}}{\pi}$ , and iii) Since  $\|\mathbf{G}_k\|^2$  has Chi-square distribution with  $2M$  degrees of freedom [21], it can be shown that

$$\Pr\{\|\mathbf{G}_k\|^2 > t\} = \frac{t^{M-1} e^{-t}}{(M-1)!} [1 + O(1/t)]. \quad (53)$$

Finally, (d) follows from substitution of  $t = \log N - 2.5 \log \log N$  and  $\gamma = \frac{2}{\log N}$  in (50).

Similar to (16),  $\text{Tr} \left\{ [\mathbb{H}^H \mathbb{H}]^{-1} \right\}$  can be written as

$$\text{Tr} \left\{ [\mathbb{H}^H \mathbb{H}]^{-1} \right\} = \text{Tr} \left\{ \left[ \frac{1}{\mathcal{K}+1} \mathbb{G}^H \mathbb{G} + \frac{\sqrt{M\mathcal{K}}}{\mathcal{K}+1} (\mathbb{G}^H \mathbb{B} + \mathbb{B}^H \mathbb{G}) + \frac{M\mathcal{K}}{\mathcal{K}+1} \mathbb{B}^H \mathbb{B} \right]^{-1} \right\}. \quad (54)$$

In Appendix V, it has been shown that  $z(\mathbf{G}_{s_i}, \mathbf{G}_{s_j}) \leq \epsilon$ , for  $i \neq j$ , where  $\epsilon = 1 - \frac{1}{1+2M(\log N)^{-\frac{1}{2(M-1)}}} \approx 2M(\log N)^{-\frac{1}{2(M-1)}}$ . As a result, conditioned on  $\mathfrak{B}$ ,  $\mathfrak{C}$ , and  $\mathfrak{D}$ , we have

$$\left| \mathbf{G}_{s_i} \mathbf{G}_{s_j}^H \right|^2 \leq t^{+2} \epsilon. \quad (55)$$

Moreover, since conditioned on  $\mathfrak{B}$ ,  $\|\mathbf{G}_{s_i}\|^2 > t$ , we have

$$\mathbb{G}^H \mathbb{G} \succeq \mathbb{D}, \quad (56)$$

where  $\mathbb{D}$  is an  $M \times M$  matrix with  $\mathbb{D}_{ii} = t$  and  $\mathbb{D}_{ij} = \mathbf{G}_{s_i} \mathbf{G}_{s_j}^H$ . Since  $\frac{t^+}{t} = 1 + O\left(\frac{\log \log N}{\log N}\right)$ , from (55) it follows that  $\frac{\mathbb{D}}{t} = \mathbf{I} + \epsilon O(\mathbf{I})$ , where  $O(\mathbf{I})$  denotes a matrix whose eigenvalues are  $O(1)$ . Moreover, since  $\log N \sim \eta \mathcal{K}$  and  $t = \log N - 2.5 \log \log N$ , we have

$$\frac{1}{\mathcal{K}+1} \mathbb{G}^H \mathbb{G} \succeq \eta \mathbf{I} + o(\mathbf{I}). \quad (57)$$

For computing  $\frac{\sqrt{M\mathcal{K}}}{\mathcal{K}+1} \mathbb{F}$ , where  $\mathbb{F} \triangleq \mathbb{G}^H \mathbb{B} + \mathbb{B}^H \mathbb{G}$ , we need to find  $\mathbb{F}_{ij} = \mathbf{G}_{s_i} \mathbf{b}_{s_j}^H + \mathbf{b}_{s_i} \mathbf{G}_{s_j}^H = 2\Re\left(\mathbf{G}_{s_i} \mathbf{b}_{s_j}^H\right) = 2\|\mathbf{G}_{s_i}\| \Re\left(\mathbf{v}_{s_i} \mathbf{b}_{s_j}^H\right)$ ,  $\forall i, j$ . For  $i = j$ , due to the algorithm, we have  $\mathbb{F}_{ii} \geq 2\sqrt{t}(1-\gamma)$ . Also, for  $i \neq j$ ,  $\mathbb{F}_{ij}$  can be upper-bounded as  $2\left|\mathbf{G}_{s_i} \mathbf{b}_{s_j}^H\right|$ . Writing  $\mathbf{b}_{s_j}$  as  $\alpha_j^{\parallel} \mathbf{v}_{s_j} + \alpha_j^{\perp} \mathbf{v}_{s_j}^{\perp}$ , where  $\mathbf{v}_{s_j}^{\perp}$  is perpendicular to  $\mathbf{v}_{s_j}$ ,  $\alpha_j^{\parallel} = \mathbf{b}_{s_j} \mathbf{v}_{s_j}^H$ , and  $\|\mathbf{v}_{s_j}^{\perp}\| = 1$ . Hence,

$$\begin{aligned} \mathbb{F}_{ij} &\leq 2 \left| \mathbf{G}_{s_i} \mathbf{b}_{s_j}^H \right| \\ &= 2 \left| \mathbf{G}_{s_i} \left( \alpha_j^{\parallel} \mathbf{v}_{s_j} + \alpha_j^{\perp} \mathbf{v}_{s_j}^{\perp} \right)^H \right| \\ &\stackrel{(a)}{\leq} 2 \left| \mathbf{G}_{s_i} \mathbf{v}_{s_j}^H \right| + |\alpha_j^{\perp}| \|\mathbf{G}_{s_i}\| \\ &\stackrel{(b)}{\leq} 2 \left[ \sqrt{t^+ \epsilon} + \sqrt{2\gamma t^+} \right] \\ &= \sqrt{t^+} O(\sqrt{\epsilon}). \end{aligned} \quad (58)$$

where (a) follows from i)  $|a+b| \leq |a| + |b|$ , ii)  $|\alpha_j^{\parallel}| \leq 1$ , and iii)  $\left| \mathbf{G}_{s_i} \mathbf{v}_{s_j}^{\perp H} \right| \leq \|\mathbf{G}_{s_i}\|$ , and (b) results from i)  $\left| \mathbf{G}_{s_i} \mathbf{v}_{s_j}^H \right| = \|\mathbf{G}_{s_i}\| \sqrt{z(\mathbf{G}_{s_i}, \mathbf{G}_{s_j})}$ , which is conditioned on  $\mathfrak{C}$  and  $\mathfrak{D}$  upper-bounded by  $\sqrt{t^+ \epsilon}$ , and ii)  $|\alpha_j^{\perp}| = \sqrt{1 - |\alpha_j^{\parallel}|^2} = \sqrt{1 - |\mathbf{b}_{s_j} \mathbf{v}_{s_j}^H|^2} \leq \sqrt{1 - [\Re(\mathbf{b}_{s_j} \mathbf{v}_{s_j}^H)]^2} \leq \sqrt{1 - (1-\gamma)^2} \leq \sqrt{2\gamma}$ . This implies that  $\mathbb{F} = 2\sqrt{t} [\mathbf{I} + \sqrt{\epsilon} O(\mathbf{I})]$ . Consequently, noting that  $\log N \sim \eta \mathcal{K}$  and  $t = \log N - 2.5 \log \log N$ , we have

$$\frac{\sqrt{M\mathcal{K}}}{\mathcal{K}+1} \mathbb{F} = 2\sqrt{\eta M} [\mathbf{I} + o(\mathbf{I})]. \quad (59)$$

Finally, having the facts that  $[\mathbb{B}^H \mathbb{B}]_{ii} = \|\mathbf{b}_i\|^2 = 1$ , and for  $i \neq j$ ,

$$\begin{aligned}
\left| [\mathbb{B}^H \mathbb{B}]_{ij} \right| &= \left| \mathbf{b}_{s_i} \mathbf{b}_{s_j}^H \right| \\
&= \left| \left( \alpha_i^\parallel \mathbf{v}_{s_i} + \alpha_i^\perp \mathbf{v}_{s_i}^\perp \right) \left( \alpha_j^\parallel \mathbf{v}_{s_j} + \alpha_j^\perp \mathbf{v}_{s_j}^\perp \right)^H \right| \\
&\stackrel{(a)}{\leq} \left| \mathbf{v}_{s_i} \mathbf{v}_{s_j}^H \right| + |\alpha_i^\perp| + |\alpha_j^\perp| + |\alpha_i^\perp| |\alpha_j^\perp| \\
&\stackrel{(b)}{\leq} \sqrt{\epsilon} + 2\sqrt{2\gamma} + 4\gamma, \\
&= O(\sqrt{\epsilon}),
\end{aligned} \tag{60}$$

in which (a) follows from the facts that i)  $|\alpha_i^\parallel| \leq 1, |\alpha_j^\parallel| \leq 1, |\alpha_i^\perp| \leq 1, |\alpha_j^\perp| \leq 1$ , and ii)  $|\mathbf{v}_{s_i} \mathbf{v}_{s_j}^H| \leq 1, |\mathbf{v}_{s_i}^\perp \mathbf{v}_{s_j}^H| \leq 1, |\mathbf{v}_{s_i}^\perp \mathbf{v}_{s_j}^\perp{}^H| \leq 1$ , and (b) results from the facts that i)  $|\alpha_i^\perp| \leq \sqrt{2\gamma}$  and conditioned on  $\mathfrak{C}$ ,  $|\mathbf{v}_{s_i} \mathbf{v}_{s_j}^H| \leq \sqrt{\epsilon}$ , we have  $\mathbb{B}^H \mathbb{B} = \mathbf{I} + \sqrt{\epsilon} O(\mathbf{I})$ , and consequently,

$$\frac{MK}{\mathcal{K} + 1} \mathbb{B}^H \mathbb{B} = M\mathbf{I} + \sqrt{\epsilon} O(\mathbf{I}). \tag{61}$$

Combining (54), (57), (59), and (61) yields

$$\begin{aligned}
\text{Tr} \left\{ [\mathbb{H}^H \mathbb{H}]^{-1} \right\} &= \text{Tr} \left\{ \left[ \eta \mathbf{I} + o(\mathbf{I}) + 2\sqrt{\eta M} \mathbf{I} + o(\mathbf{I}) + M\mathbf{I} + \sqrt{\epsilon} O(\mathbf{I}) \right]^{-1} \right\} \\
&= \text{Tr} \left\{ \left[ (\sqrt{\eta} + \sqrt{M})^2 \mathbf{I} + o(\mathbf{I}) \right]^{-1} \right\} \\
&\stackrel{(a)}{=} \frac{M}{(\sqrt{\eta} + \sqrt{M})^2} + o(1),
\end{aligned} \tag{62}$$

where (a) follows from the fact that  $[\mathbf{I} + o(\mathbf{I})]^{-1} = \mathbf{I} + o(\mathbf{I})$ . Substituting in (49) yields

$$\begin{aligned}
\mathcal{R} &\geq M \log \left( 1 + P \left[ 1 + \sqrt{\frac{\eta}{M}} \right]^2 + o(1) \right) \Pr\{\mathfrak{B}\} \Pr\{\mathfrak{C}|\mathfrak{B}\} \Pr\{\mathfrak{D}|\mathfrak{B}, \mathfrak{C}\} \\
&= \left[ M \log \left( 1 + P \left[ 1 + \sqrt{\frac{\eta}{M}} \right]^2 \right) + o(1) \right] \left[ 1 + O\left(\frac{1}{\log N}\right) \right] \\
&= M \log \left( 1 + P \left[ 1 + \sqrt{\frac{\eta}{M}} \right]^2 \right) + o(1).
\end{aligned} \tag{63}$$

This completes the proof of Theorem 4. ■

#### IV. SIMULATION RESULTS

In this section, we examine the analytical results in the previous section by simulation. Figures 1-3 present the plots of the sum-rate capacity versus the number of users, for different values of Rician factor  $\mathcal{K} = 1, \mathcal{K} = 10$ , and  $\mathcal{K} = 100$ , respectively. The SNR ( $P$ ) is assumed to be 10 dB in these figures and the number of transmit antennas  $M$  is set to 2. Also, the plots of the achievable sum-rate for ZFBF and

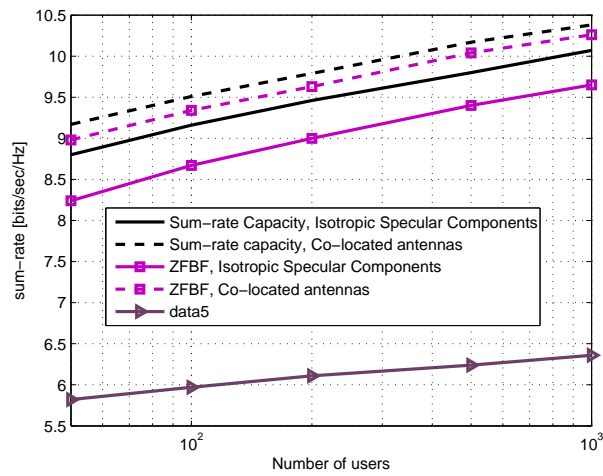


Fig. 1. Sum-rate capacity versus the number of users;  $\mathcal{K} = 1$ .

TDMA are given for comparison. The user selection algorithm used for ZFBF is the same as Algorithm 1 in [12]. As can be observed in the figures the following observations can be made: i) The sum-rate capacity of the system in the case of isotropic specular components is larger than the sum-rate capacity in the case of co-located transmit antennas. ii) In the case of isotropic specular components, ZFBF performs well for all values of  $\mathcal{K}$ , while in the case of co-located transmit antennas the performance of ZFBF is degraded significantly by increasing  $\mathcal{K}$ . iii) in the case of co-located transmit antennas and  $K = 100$ , the sum-rate of TDMA is almost close to the sum-rate capacity, which is compatible with the result of Theorem 2.

Figure 4 presents the plots of sum-rate capacity versus SNR for different values of Rician factor and two cases of isotropic specular components and co-located transmit antennas. It is assumed that  $N = 100$  and  $M = 2$  in this figure. As can be observed, by increasing the value of the Rician factor, the difference between the sum-rate capacity of the system in the two cases of isotropic specular components and co-located transmit antennas increases. Moreover, the slope of the curves in the case of isotropic specular components is equal to 2, regardless of the value of  $\mathcal{K}$ , while the slope of the curves in the case of co-located transmit antennas decreases with  $\mathcal{K}$ , but increases with SNR. However, for high values of SNR, the slope of all curves approaches 2, implying that the multiplexing gain of the system is 2, regardless of the distribution of the specular components and the value of the Rician factor.

## V. CONCLUSION

In this paper, we have derived the asymptotic sum-rate capacity of MIMO-BC with large number of users in a Rician fading environment. It is observed that in the region  $\mathcal{K} = o(\log N)$ , the capacity achieving strategy is exactly the same as the Rayleigh fading case. In the region  $\mathcal{K} = \omega(\log N)$ , the sum-rate capacity depends on the distribution of the specular component; in the case of co-located transmit antennas, it is

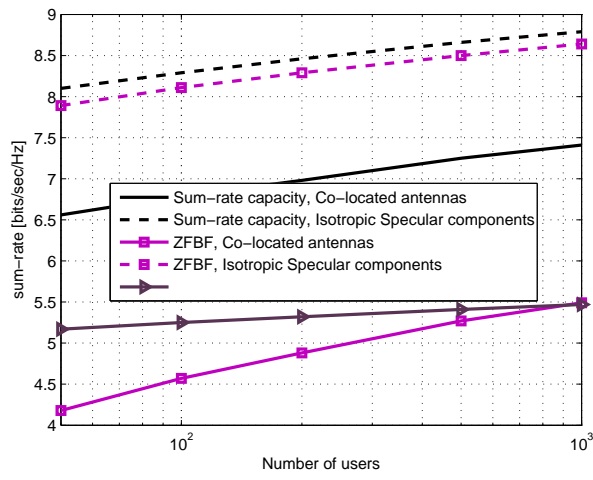


Fig. 2. Sum-rate capacity versus the number of users;  $\mathcal{K} = 10$ .

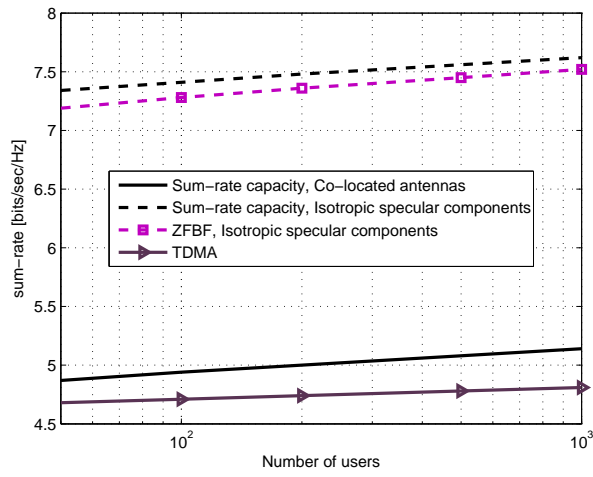


Fig. 3. Sum-rate capacity versus the number of users;  $\mathcal{K} = 100$ .

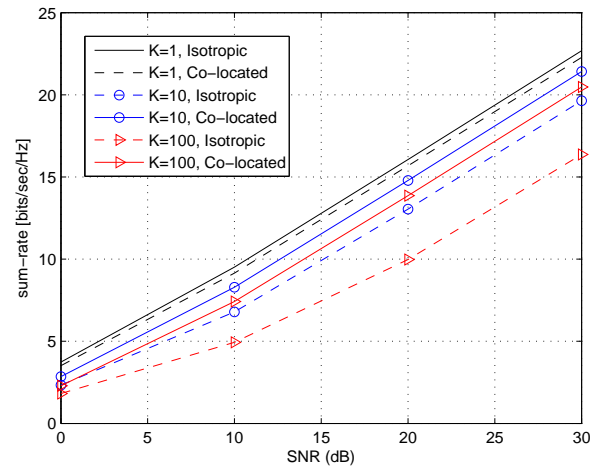


Fig. 4. Sum-rate capacity versus SNR;  $N = 100$ .

demonstrated that TDMA achieves the sum-rate capacity and the capacity grows logarithmically with the number of transmit antennas. In the case of isotropically distributed specular components, ZFBF along with a user selection strategy which selects  $M$  users with semi-orthogonal specular components is shown to be optimum. Moreover, the sum-rate capacity grows linearly with the number of transmit antennas.

#### APPENDIX A

We have

$$\begin{aligned}\Pr\{\mathfrak{A}\} &= \Pr\{\|\mathbf{G}\|_{\max}^2 \leq t\} \\ &= (\Pr\{\|\mathbf{G}\|_k^2 \leq t\})^N.\end{aligned}\quad (64)$$

Using the fact that  $\|\mathbf{G}_k\|^2$  has Chi-square distribution with  $2M$  degrees of freedom [21], we have

$$\Pr\{\|\mathbf{G}\|_k^2 \leq t\} = 1 - \sum_{m=0}^{M-1} \frac{t^m}{m!} e^{-t}.\quad (65)$$

Substituting  $t = \log N + (M - 3) \log \log N$ , the above equation can be rewritten as

$$\Pr\{\|\mathbf{G}\|_k^2 \leq t\} = 1 - \frac{\log^2 N}{N} \left[ 1 + O\left(\frac{\log \log N}{\log N}\right) \right].\quad (66)$$

Substituting in (64), it is concluded that

$$\begin{aligned}\Pr\{\mathfrak{A}\} &= \left( 1 - \frac{\log^2 N}{N} \left[ 1 + O\left(\frac{\log \log N}{\log N}\right) \right] \right)^N \\ &= o\left(\frac{1}{N}\right).\end{aligned}\quad (67)$$

#### APPENDIX B

*Calculation of  $\Pr\{\mathfrak{B}\}$ :*

Consider a randomly selected user  $k$ . Using (65) and (66), we have

$$\Pr\{k \in \mathcal{S}\} = \frac{\log^2 N}{N} \left[ 1 + O\left(\frac{\log \log N}{\log N}\right) \right].\quad (68)$$

Therefore,  $L = |\mathcal{S}|$  is a Binomial random variable with parameters  $(N, p)$ , where  $p \triangleq \Pr\{k \in \mathcal{S}\}$ . Using the Gaussian approximation for Binomial distribution, we have

$$\begin{aligned}
\Pr\{\mathfrak{B}\} &= \sum_{l=\lfloor \log N \rfloor + 1}^N \binom{N}{l} p^l (1-p)^{N-l} \\
&\approx Q\left(\frac{\log N - Np}{\sqrt{Np(1-p)}}\right) \\
&= 1 - Q\left(\frac{Np - \log N}{\sqrt{Np(1-p)}}\right) \\
&\stackrel{(a)}{\geq} 1 - e^{-\frac{(Np - \log N)^2}{2Np(1-p)}} \\
&\geq 1 - Ne^{-\frac{Np}{2}} \\
&\stackrel{(b)}{=} 1 - o(1/N),
\end{aligned} \tag{69}$$

where (a) results from the fact that  $Q(x) \leq e^{-x^2/2}$  and (b) follows from the fact that  $Np \sim \frac{\log^2 N}{(M-1)!}$ , which incurs that  $e^{-Np/2} = o(N^{-2})$ .

*Calculation of  $\Pr\{\mathfrak{C}|\mathfrak{B}\}$ :*

Using equation (144) in Appendix E of [12], we have

$$\delta(\mathbb{G}^H) = \delta(\Psi) = \frac{1}{\prod_{i=1}^{M-1} \beta_i}, \tag{70}$$

where  $\Psi \triangleq [\mathbf{v}_{s_1}^H | \dots | \mathbf{v}_{s_M}^H]$ , and  $\beta_i$  denotes the projection of  $\mathbf{v}_{s_{i+1}}$  over  $\mathcal{P}_{i+1}^\perp$ , which denotes the null space of  $\mathcal{P}_{i+1}$ , the subspace spanned by  $\{\mathbf{v}_{s_j}\}_{j=1}^i$ . Defining  $\epsilon \triangleq 1 - \frac{1}{1+2M(\log N)^{-\frac{1}{2(M-1)}}}$ , from (70),  $\Pr\{\mathfrak{C}|\mathfrak{B}\}$  can be written as

$$\begin{aligned}
\Pr\{\mathfrak{C}|\mathfrak{B}\} &= \Pr\left\{\prod_{i=1}^{M-1} \beta_i < 1 - \epsilon \mid \mathfrak{B}\right\} \\
&\stackrel{(a)}{\leq} \sum_{i=1}^{M-1} \Pr\left\{\beta_i < 1 - \frac{\epsilon}{M} \mid \mathfrak{B}\right\} \\
&\stackrel{(b)}{\leq} \sum_{i=1}^{M-1} \Pr\left\{\beta_i < 1 - (\log N)^{-\frac{1}{2(M-1)}} \mid \mathfrak{B}\right\},
\end{aligned} \tag{71}$$

where (a) follows from the fact that the event  $\prod_{i=1}^{M-1} \beta_i < 1 - \epsilon$  is a subset of the event  $\bigcup_{i=1}^{M-1} \{\beta_i < 1 - \frac{\epsilon}{M}\}$ . To show this, we observe that if none of the events  $\{\beta_i < 1 - \frac{\epsilon}{M}\}_{i=1}^{M-1}$  occur, it means that  $\prod_{i=1}^{M-1} \beta_i > (1 - \frac{\epsilon}{M})^M > 1 - \epsilon$ . Also, (b) results from the fact that as  $N \rightarrow \infty$ ,  $2M(\log N)^{-\frac{1}{2(M-1)}} < 1$ .



From the algorithm,  $\beta_i$  can be written as  $1 - \min_{k \in \mathcal{S}_{i+1}} z_{k,i}$ , where  $z_{k,i}$  denotes the projection of  $\mathbf{v}_k$  over  $\mathcal{P}_{i+1}$ . The probability density function (pdf) of  $z_{k,i}$  is given in the equation (146) in [12] as <sup>5</sup>

$$p_{z_{k,i}}(z) = \frac{\Gamma(M)}{\Gamma(M-i)\Gamma(i)} z^{i-1} (1-z)^{M-i-1}. \quad (72)$$

Since  $\mathbf{V}_k$ 's are i.i.d. random variables (since the channel vector of users are independent of each other), it follows that  $z_{k,i}$ 's are also i.i.d.. Hence, defining  $\theta \triangleq 1 - (\log N)^{-\frac{1}{2(M-1)}}$ , we have

$$\begin{aligned} \Pr \{ \beta_i < \theta | \mathfrak{B} \} &\stackrel{(a)}{=} (\Pr \{ z_{k,i} > 1 - \theta \})^{L-i} \\ &= (1 - I_{i, M-i}(1 - \theta))^{L-i} \\ &\stackrel{(b)}{\leq} (1 - (1 - \theta)^{M-1})^{L-i} \\ &= \left( 1 - \frac{1}{\sqrt{\log N}} \right)^{L-i} \\ &\stackrel{(c)}{\leq} \left( 1 - \frac{1}{\sqrt{\log N}} \right)^{\log N - i} \\ &\sim e^{-\sqrt{\log N}} \\ &= o\left(\frac{1}{\log N}\right). \end{aligned} \quad (73)$$

In the above equation, (a) results from the fact that  $|\mathcal{S}_{i+1}| = L - i$ , and (b) follows from the the fact that  $I_{i, M-i}(\theta) \geq I_{M-1, 1}(\theta) = \theta^{M-1}$ . (c) comes from the fact that conditioned on  $\mathfrak{B}$ ,  $L > \log N$ . Combining the above equation with (71) yields  $\Pr\{\mathfrak{C}^C | \mathfrak{B}\} = o\left(\frac{1}{\log N}\right)$ , which implies that  $\Pr\{\mathfrak{C} | \mathfrak{B}\} = 1 + o\left(\frac{1}{\log N}\right)$ .

#### Computation of $\Pr\{\mathfrak{D} | \mathfrak{B}, \mathfrak{C}\}$

To compute  $\Pr\{\mathfrak{D} | \mathfrak{B}, \mathfrak{C}\}$ , we first note that since the norm and direction of circularly symmetric complex Gaussian vectors are independent of each other and having the facts that  $\mathfrak{B}$  and  $\mathfrak{D}$  depend solely on the norm of  $\{\mathbf{G}_k\}_{k=1}^N$  and  $\mathfrak{C}$  depends only on the direction of these vectors, it follows that  $\mathfrak{B}$  and  $\mathfrak{D}$  are independent of  $\mathfrak{C}$ . Therefore, we can ignore  $\mathfrak{C}$  in the condition and write

$$\begin{aligned} \Pr\{\mathfrak{D} | \mathfrak{B}, \mathfrak{C}\} &= \Pr\{\mathfrak{D} | \mathfrak{B}\} \\ &= \frac{\Pr\{\mathfrak{D}, \mathfrak{B}\}}{\Pr\{\mathfrak{B}\}} \\ &\stackrel{(a)}{\geq} \frac{\Pr\{\mathfrak{D}\} - \Pr\{\mathfrak{B}^C\}}{\Pr\{\mathfrak{B}\}}. \end{aligned} \quad (74)$$

<sup>5</sup>Note that since the norm and direction of circularly symmetric complex Gaussian vectors are independent of each other, the distribution of  $z_{k,i}$  is independent of the condition  $\mathfrak{B}$ .

Since we have already computed  $\Pr\{\mathfrak{B}\}$  in this appendix, it suffices to compute  $\Pr\{\mathfrak{D}\}$ .  $\Pr\{\mathfrak{D}^C\}$  is computed in Lemma 3 of [12], and shown to scale as  $O\left(\frac{1}{\log N}\right)$ . Hence,

$$\begin{aligned}\Pr\{\mathfrak{D}|\mathfrak{B}, \mathfrak{C}\} &= \frac{1 + O\left(\frac{1}{\log N}\right) + o\left(\frac{1}{N}\right)}{1 + o\left(\frac{1}{N}\right)} \\ &= 1 + O\left(\frac{1}{\log N}\right).\end{aligned}\tag{75}$$

### APPENDIX C

Conditioned on  $\mathfrak{C}$ , we have  $\prod_{i=1}^{M-1} \beta_i > \frac{1}{1+2M(\log N)^{-\frac{1}{2(M-1)}}} = 1 - \epsilon$ . Since  $\beta_i \leq 1$ , this incurs that  $\beta_i \geq 1 - \epsilon, \forall i$ . In other words,  $\gamma_i \leq \epsilon$ , where  $\gamma_i$  denotes the projection of  $\mathbf{v}_{s_{i+1}}$  over  $\mathcal{P}_{i+1}$ . Now, consider  $\{\Phi_j\}_{j=1}^i$  as  $j$  orthonormal bases for  $\mathcal{P}_{i+1}$ . Since  $\mathbf{v}_{s_j} \in \mathcal{P}_{i+1}, \forall j \leq i$ , we can write

$$\mathbf{v}_{s_j} = \sum_{l=1}^i a_l \Phi_l,\tag{76}$$

where  $\sum_{l=1}^i |a_l|^2 = 1$ . Therefore, for all  $i, j, j \leq i$ , we have

$$\begin{aligned}z(\mathbf{G}_{s_{i+1}}, \mathbf{G}_{s_j}) &= z(\mathbf{v}_{s_{i+1}}, \mathbf{v}_{s_j}) \\ &= \left| \mathbf{v}_{s_{i+1}} \mathbf{v}_{s_j}^H \right|^2 \\ &= \left| \sum_{l=1}^i a_l^H (\mathbf{v}_{s_{i+1}} \Phi_l^H) \right|^2 \\ &\stackrel{(a)}{\leq} \sum_{l=1}^i \left| \mathbf{v}_{s_{i+1}} \Phi_l^H \right|^2 \\ &= \gamma_i \\ &\leq \epsilon,\end{aligned}\tag{77}$$

where (a) follows from the fact that  $\left| \sum_{l=1}^i a_l b_l \right|^2 \leq \left( \sum_{l=1}^i |a_l|^2 \right) \left( \sum_{l=1}^i |b_l|^2 \right)$ , noting that  $\sum_{l=1}^i |a_l|^2 = 1$ .

### REFERENCES

- [1] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecommunications*, vol. 10, no. 6, pp. 585–595, November 1999.
- [2] G. J. Foschini and M.J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, pp. 311–335, 1998.
- [3] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. on Inform. Theory*, vol. 49, pp. 1691–1706, July 2003.
- [4] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. on Inform. Theory*, vol. 49, no. 10, pp. 2658–2668, October 2003.
- [5] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. on Inform. Theory*, vol. 49, pp. 1912–1921, Aug. 2003.

- [6] Wei Yu and John Cioffi, "Sum capacity of vector Gaussian broadcast channels," *IEEE Trans. on Inform. Theory*, vol. 50, pp. 1875–1892, Sep. 2004.
- [7] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian MIMO broadcast channel," in *Conference on Information Sciences and Systems*, 2004.
- [8] M. Costa, "Writing on dirty paper," *IEEE Trans. on Inform. Theory*, vol. 29, pp. 439–441, May 1983.
- [9] J. Lee and N. Jindal, "High SNR Analysis for MIMO Broadcast Channels: Dirty Paper Coding vs. Linear Precoding," *IEEE Trans. on Inform. Theory*, Vol. 53, No. 12, pp. 4787–4792, Dec. 2007.
- [10] N. Jindal and A. Goldsmith, "Dirty paper coding vs. TDMA for mimo broadcast channels," in *IEEE Int. Conf. Commun.*, June 2004, vol. 2, pp. 682–686.
- [11] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channel with partial side information," *IEEE Trans. on Inform. Theory*, vol. 51, pp. 506–522, Feb. 2005.
- [12] Alireza Bayesteh and Amir K. Khandani, "On the user selection for MIMO broadcast channels," *To Appear in IEEE Trans. Inform. Theory*, March 2008, Available online at <http://cst.uwaterloo.ca/>, Technical Report UW-ECE 2005-16.
- [13] Taesang Yoo and Andrea Goldsmith, "On the optimality of Multi-Antenna Broadcast Scheduling Using Zero-Forcing Beamforming," *IEEE JSAC Special Issue on 4G Wireless Systems*, Vol. 24, No. 3, March 2006.
- [14] C. Swannack, E. Uysal-Biyikoglu, and G. W. Wornell, "Low Complexity Multiuser Scheduling for Maximizing Throughput in the MIMO Broadcast Channel," in *Proc. Allerton Conf. Commun., Contr., and Computing*, Oct. 2004.
- [15] M. Sharif and B. Hassibi, "A comparison of time-sharing, DPC, and beamforming for MIMO broadcast channels with many users," *IEEE Trans. on Communications*, Vol. 55, pp. 11–15, Jan. 2007.
- [16] Ming Kang and Mohamed-Slim Alouini, "Capacity of MIMO Rician Channels," *Trans. on Wireless Communi.*, Vol. 50, pp. 112–122, Jan. 2006.
- [17] M. Godavarti and A. O. Hero III, "Multiple antenna capacity in a deterministic Rician fading channel," *Submitted to IEEE Trans. on Information Theory*
- [18] Mahesh Godavarti, Alfred O. Hero, III, and Thomas L. Marzetta, "Min-Capacity of a Multiple-Antenna Wireless Channel in a Static Rician Fading Environment," *IEEE Trans. on Communi.*, Vol. 4, pp. 1715–1723, July 2005.
- [19] Mahesh Godavarti, Thomas L. Marzetta, and Shlomo Shamai (Shitz), "Capacity of a Mobile Multiple-Antenna Wireless Link With Isotropically Random Rician Fading," *IEEE Transactions on Information Theory*, Vol. 49, pp. 3330–3334, Dec. 2003.
- [20] Mahmoud Taherzadeh, Amin Mobasher, and Amir Keyvan Khandani, "Communication over mimo broadcast channels using lattice-basis reduction," *IEEE Transactions on Information Theory*, Accepted to be published., 2006.
- [21] Sheldon Ross, *A First Course in Probability*, Prentice-Hall Inc., sixth edition, 2002.