Optimality of Beamforming for MIMO Multiple Access Channels via Virtual Representation

Hong Wan, Rong-Rong Chen*, and Yingbin Liang

Abstract—In this paper, we consider the optimality of beamforming for achieving the ergodic capacity of multiple-input multiple-output (MIMO) multiple access channel (MAC) via virtual representation (VR) model. We assume that the receiver knows the channel state information (CSI) perfectly but that the transmitter knows only partial CSI, i.e., the channel statistics. For the single-user case, we prove that the capacity-achieving beamforming angle (c.b.a.) is unique, and there exists a signal-to-noise ratio (SNR) threshold below which beamforming is optimal and above which beamforming is strictly suboptimal. For the multi-user case, we show that the c.b.a is not unique and we obtain explicit conditions that determine the beamforming angles for a special class of correlated MAC-VR models. Under mild conditions, we show that a large class of power allocation schemes can achieve the sum-capacity within a constant as the number of users in the system becomes large. The beamforming scheme, in particular, is shown to be asymptotically capacity-achieving only for certain MAC-VR models.

I. INTRODUCTION

The multiple-input multiple-output (MIMO) techniques provide powerful means to improve reliability and capacity of wireless channels. Significant amount of work has been done to study optimal input distributions and the channel capacity of single-user and multi-user MIMO channels (see, e.g., [1]–[7]). Several models have been adopted to capture the spatial correlation between the channel gains corresponding to different transmit-receive antenna pairs. These models include the i.i.d. model [1], the Kronecker model [2], [8]–[10], the virtual representation (VR) model [4], [11], and the unitary-independent-unitary (UIU) model [5]. The i.i.d. model assumes that the channel gains are independent and identically distributed (i.i.d.), and the Kronecker model assumes that the correlation between the channel gains

can be written in terms of the product of the transmit correlation and the receive correlation. These two models apply only to wireless environments with rich or locally rich scattering at either the transmitter or the receiver. The VR and UIU models are more general, and both transform the MIMO channel to a domain such that the channel gains can be justified to be approximately independent.

In this paper, we adopt the VR model [11], which represents the MIMO channel in a virtual *angular* domain with each channel gain corresponding to one virtual transmit and receive angle pair. The channel gains in the angular domain can be justified to be approximately independent of each other, although not necessarily identically distributed, because they include different signal paths (corresponding to different transmit and receive angle pairs) with independent random phases.

The single-user MIMO channel based on VR was studied in [4]. In this paper, we generalize this study to the MIMO multiple access channel (MAC) based on VR, denoted by MAC-VR. We first characterize the optimal input distribution that achieves the sum-capacity. Then we study the optimality of beamforming, which is a simple scalar coding strategy desirable in practice. We first strengthen the conditions for the optimality of beamforming for the single-user VR model in [4] by proving that there exists a signal-to-noise ratio (SNR) threshold below which beamforming is optimal and above which beamforming is strictly suboptimal. This result was illustrated in [4] only numerically. For the multi-user case, we present an example to show that the capacity-achieving beamforming angle (c.b.a) of a given user may vary with SNR and beamforming angles of other users. This is in contrast to the single-user case in which the c.b.a. is independent of SNR. We also derive explicit conditions to determine possible c.b.a. for certain MAC-VR channels. For systems with K users, we show that as K goes to infinity, the sum-rates achieved by a large class of power allocation schemes are within a constant of the sum-capacity, and they grow in the order of $n_r \log K$, where n_r is the number of receive antennas. Furthermore, we obtain conditions under which beamforming is asymptotically capacity-

achieving.

Our study for the single-user case generalizes that in [2], [6] for the Kronecker model, and is different from [12] for the double-scattering model [13]. Our study for the MAC-VR also differs from [7] which assumes perfect channel state information at the transmitter, and from [14], which assumes finite feedback. We also note that the results we derive for the MAC-VR are applicable to the MIMO-MAC Kronecker (MAC-Kr) model in [9]. However, certain results valid for the MAC-Kr may not hold for the MAC-VR as demonstrated in later sections.

II. CHANNEL MODEL AND VIRTUAL REPRESENTATION

We consider the K-user MIMO MAC, in which K users transmit to one base station (BS) with each user equipped with n_t antennas and the BS equipped with n_r antennas. The channel between each user k and the BS is assumed to be a frequency-flat, MIMO fading channel. The received signal at the BS is an n_r-dimensional vector $Y \in \mathcal{C}^{n_r}$ and is given by

$$Y = \sum_{k=1}^{K} \sqrt{\frac{p^k}{n_t}} H^k X^k + W, \qquad (1)$$

where $X^k \in \mathcal{C}^{n_t}$ is the input vector of user k that satisfies the power constraint $E[X^{k\dagger} X^k] \leq n_t$, $(\cdot)^\dagger$ denotes the Hermitian operator, p^k represents the effective SNR of user k at each receive antenna, $W \in \mathcal{C}^{n_r}$ is a proper complex Gaussian noise vector that consists of i.i.d. entries with zero-mean and unit-variance, and $H^k \in \mathcal{C}^{n_r \times n_t}$ is the channel matrix of user k. The entries of H^k are identically distributed with unit variance, i.e., $E[|H_{m,j}|^2] = 1$ for all $m = 1, \ldots, n_r$ and $j = 1, \ldots, n_t$. In general, these entries are correlated because each channel gain in the antenna domain (corresponding to one transmit and receive antenna pair) captures all of the signal paths. For each H^k, we follow [4] to consider its virtual representation $H^k = A_r \tilde{H}^k A_t^\dagger$, where A_r and A_t are two-dimensional spatial Fourier matrices. The matrix $\tilde{H}^k$ is referred to as a virtual representation of H^k. Each element of $\tilde{H}^k$, referred to as the virtual coefficient, represents the channel gain corresponding to one transmit and receive virtual angle pair. The virtual coefficients are independent and not identically distributed (i.n.d.) random variables and each is assumed to be a zero-mean proper complex random variable with a symmetric distribution around the origin. The independency among virtual coefficients can be justified because they correspond to different pairs of transmit and receive angles and thus capture different sets of signal paths with independent random phases. The correlation

of the channel gains in the antenna domain is implicitly determined by the i.n.d. channel gains in virtual domain and the Fourier transform between the two domains.[1]

A virtual representation of the MIMO MAC channel (1) is given by

$$\tilde{Y} = \sum_{k=1}^{K} \sqrt{\frac{p^k}{n_t}} \tilde{H}^k \tilde{X}^k + \tilde{W}, \qquad (2)$$

where $\tilde{X}^k = A_t^\dagger X^k$, $\tilde{Y} = A_r^\dagger Y$, and $\tilde{W} = A_r^\dagger W$. Due to the unitarity of A_t, the input power constraint in the virtual domain does not change, i.e., $E(\tilde{X}^{k\dagger} \tilde{X}^k) \leq n_t$. Given $\tilde{H}^k$, we define the (m, j)-th element of the variance matrix V^k as $V_{m,j}^k = \text{Var}(\tilde{H}_{m,j}^k)$, for $1 \leq m \leq n_r$ and $1 \leq j \leq n_t$, which characterizes the second order statistics of $\tilde{H}^k$.

Remark 1: It can be shown that the MAC-Kr also takes the form of (2) under unitary transformations. However, it imposes more constraints on $\tilde{H}^k$ than MAC-VR does. For MAC-Kr, each element of $\tilde{H}^k$ is assumed to be Gaussian distributed. This is not required for the MAC-VR. Furthermore, for MAC-VR, the elements of $\tilde{H}^k$ are i.n.d., and thus elements of V^k can take any non-negative value. For MAC-Kr, $V_{m,j}^k$ takes on a product form:

$$V_{m,j}^k = a_m \cdot b_j^k, \quad \text{for every } 1 \leq m \leq n_r \text{ and } 1 \leq j \leq n_t, \qquad (3)$$

where $\{a_m\}$ and $\{b_j^k\}$ are square-roots of the eigenvalues of the receive correlation and transmit correlation matrix, respectively.

III. PRELIMINARIES

In this section, we present some preliminary results that will be used in the proofs of the main results in Sections IV and V. Since these preliminary results can be easily derived following techniques in [4], [9], [10], the proofs are omitted for brevity.

First, we show that the sum-capacity of MAC-VR is achieved when the inputs of all users in the virtual domain are zero-mean proper complex Gaussian and are independent of each other. Let $\tilde{Q}^k = E(\tilde{X}^k \tilde{X}^{k\dagger})$ be the input covariance matrix of user k. The sum-capacity is given by

$$C = \max_{\substack{\text{Tr}(\tilde{Q}^k) \leq n_t \\ k=1,\cdots,K}} E\left[\log \det \left(I_{n_r} + \sum_{k=1}^{K} \frac{p^k}{n_t} \tilde{H}^k \tilde{Q}^k \tilde{H}^{k\dagger}\right)\right], \qquad (4)$$

---

[1] We note that a channel matrix with arbitrary correlation in the antenna domain may not necessarily have a meaningful virtual representation with an i.n.d. channel matrix in the angular domain [11].

where the expectation is over $\{\tilde{H}^k, k = 1, \cdots, K\}$, and the constraint on $\tilde{Q}^k$ is due to the power constraint for user k. Here we assume that $\{\tilde{H}^k\}$ changes in time and the receiver knows the perfect channel state information (CSI) $\{\tilde{H}^k\}$. The transmitter knows only partial CSI, i.e., the channel statistics $\{V^k\}$, which does not change over the time duration of interest. Following techniques of [4], we can prove that the capacity-achieving covariance matrices are diagonal [15]. Furthermore, a necessary and sufficient condition for the optimality of the input covariance matrices can be derived following similar approaches of [9], [10]. This leads to Theorem 1 below.

Theorem 1: The diagonal covariance matrices $\{\tilde{Q}^k, k = 1, \cdots, K\}$ achieve the sum-capacity if and only if for every $1 \le k \le K$ and $1 \le j \le n_t$, we have

$$
E \operatorname{Tr} \left[A^{-1} \left(\tilde{h}_j^k \tilde{h}_j^{k\dagger} - \frac{1}{n_t} \tilde{H}^k \tilde{Q}^k \tilde{H}^{k\dagger} \right) \right] \begin{cases} = 0, & \text{if } \lambda_j^k > 0, \\ \le 0, & \text{if } \lambda_j^k = 0, \end{cases}
$$

$$(5)$$

where $\tilde{Q}^k = \operatorname{diag}(\lambda_1^k, \cdots, \lambda_{n_t}^k)$, $A = I_{n_r} + \sum_{l=1}^{K} \frac{p^l}{n_t} \tilde{H}^l \tilde{Q}^l \tilde{H}^{l\dagger}$, and $\tilde{h}_j^k$ denotes the j-th column of $\tilde{H}^k$.

Remark 2: Since the diagonal element λ_j^k, $j = 1, \ldots, n_t$, of $\tilde{Q}^k$ can be interpreted as the amount of transmission power allocated for the j-th virtual transmit angle, computation of the optimal covariance matrices to maximize the ergodic capacity is equivalent to finding the optimal power allocation parameters $\{\lambda_j^k\}$. Although the iterative algorithm in [10] can be applied to compute the optimal power allocation, it has high complexity due to statistical averaging over fading distributions. A low-complexity algorithm that utilizes only the second order statistics of the fading distribution, i.e., the variance matrix V^k, can be developed for the MAC-VR following similar approaches of [16].

When all users perform beamforming, i.e., each user k allocates full transmission power to a virtual angle i_k (the beamforming angle), then from (5) we obtain Corollary 1 below that characterizes the optimality of beamforming. It is equivalent to [9, Theorem 2] for the MAC-Kr.

Corollary 1: Consider a K-user MAC-VR in which each user k performs beamforming along virtual angle i_k. Let $A = I_{n_r} + \sum_{l=1}^{K} p^l \tilde{h}_{i_l}^l \tilde{h}_{i_l}^{l\dagger}$. Then beamforming is optimal, in terms of achieving the sum-capacity (4), if and only if for every $j = 1, \cdots, n_t$ and $k = 1, \cdots, K$, the following condition is satisfied

$$
f_j^k(p^k) = E \left[\tilde{h}_j^{k\dagger} A^{-1} \tilde{h}_j^k - \tilde{h}_{i_k}^{k\dagger} A^{-1} \tilde{h}_{i_k}^k \right] \le 0. \quad (6)
$$

IV. OPTIMALITY OF BEAMFORMING

A. Single-User Case

It is observed in [4] numerically that there exists an SNR threshold below which beamforming is optimal and above which beamforming is suboptimal. Here we provide a mathematical proof of this threshold behavior. The key step of the proof follows from Lemma 1 below, which characterizes an important property of the beamforming condition $f_j(p)$ in (6). For $K = 1$, we assume that the user beamforms to virtual angle i, and thus we have $A = I_{n_r} + p\tilde{h}_i\tilde{h}_i^{\dagger}$, and $f_j(p) = E \left[\tilde{h}_j^{\dagger} A^{-1} \tilde{h}_j - \tilde{h}_i^{\dagger} A^{-1} \tilde{h}_i \right]$.

Lemma 1: If $f_j(p)$ satisfies $f_j(p) \le 0$, then we must have $f_j'(p) > 0$.

Proof: Let $f_j'(p)$ denote the derivative of $f_j(p)$ with respect to p. We will prove that for every $j \ne i$,

$$
f_j'(p) > \frac{1}{p} \left(E \left[\frac{1}{1 + p\|\tilde{h}_i\|^2} \right] - 1 \right) f_j(p). \quad (7)
$$

Once (7) is proved, Lemma 1 immediately follows because $\left(E \left[\frac{1}{1+p\|\tilde{h}_i\|^2} \right] - 1 \right) \le 0$. To prove (7), we apply the matrix inversion lemma to write $f_j(p) = u_j(p) + v_j(p)$, where $u_j(p) = E \left[\frac{\|\tilde{h}_j\|^2 - \|\tilde{h}_i\|^2\|}{1 + p\|\tilde{h}_i\|^2} \right]$ and $v_j(p) = E \left[\frac{\|\tilde{h}_j\|^2\|\tilde{h}_i\|^2 - \|\tilde{h}_j^{\dagger}\tilde{h}_i\|^2}{p^{-1} + \|\tilde{h}_i\|^2} \right]$. Due to the Cauchy-Schwartz inequality, we have $\|\tilde{h}_j\|^2\|\tilde{h}_i\|^2 - \|\tilde{h}_j^{\dagger}\tilde{h}_i\|^2 > 0$ and thus $v_j(p)$ is a positive, strictly increasing function of p with $v_j'(p) > 0$. It follows that

$$
\begin{aligned}
f_j'(p) > u_j'(p) &= E \left[\frac{\|\tilde{h}_i\|^4 - \|\tilde{h}_j\|^2\|\tilde{h}_i\|^2}{(1 + p\|\tilde{h}_i\|^2)^2} \right] \\
&= \frac{1}{p} E \left[\frac{\|\tilde{h}_i\|^2 - \|\tilde{h}_j\|^2}{1 + p\|\tilde{h}_i\|^2} \right] + \frac{1}{p} E \left[\frac{\|\tilde{h}_j\|^2}{(1 + p\|\tilde{h}_i\|^2)^2} \right] \\
&\quad - \frac{1}{p} E \left[\frac{\|\tilde{h}_i\|^2}{(1 + p\|\tilde{h}_i\|^2)^2} \right]. \quad (8)
\end{aligned}
$$

For the second term of (8), since $\tilde{h}_j$ is independent of $\tilde{h}_i$, we apply the inequality $E(1/x^2) > [E(1/x)]^2$ with $x = 1 + p\|\tilde{h}_i\|^2$ to obtain

$$
\begin{aligned}
E \left[\frac{\|\tilde{h}_j\|^2}{(1 + p\|\tilde{h}_i\|^2)^2} \right] &> \left(E \left[\frac{1}{1 + p\|\tilde{h}_i\|^2} \right] \right)^2 E \left[\|\tilde{h}_j\|^2 \right] \\
&= E \left[\frac{1}{1 + p\|\tilde{h}_i\|^2} \right] E \left[\frac{\|\tilde{h}_j\|^2}{1 + p\|\tilde{h}_i\|^2} \right]. \quad (9)
\end{aligned}
$$

For the third term of (8), we apply the same inequality to obtain

$$
\begin{aligned}
-E\left[\frac{\|\tilde{h}_i\|^2}{\left(1+p\|\tilde{h}_i\|^2\right)^2}\right] &= \frac{1}{p}E\left[\frac{1}{\left(1+p\|\tilde{h}_i\|^2\right)^2}\right] \\
&\quad - \frac{1}{p}E\left[\frac{1}{\left(1+p\|\tilde{h}_i\|^2\right)}\right] \\
&> E\left[\frac{1}{1+p\|\tilde{h}_i\|^2}\right]E\left[\frac{-\|\tilde{h}_i\|^2}{1+p\|\tilde{h}_i\|^2}\right].
\end{aligned}
\tag{10}
$$

We then substitute (9) and (10) into (8) to obtain (7). ∎

Next, we characterize the threshold behavior of beamforming in Theorem 2.

Theorem 2: For a single-user VR channel, the i-th virtual angle is the c.b.a. if and only if

(a) The i-th virtual angle has a sum-variance, defined by $\sum_{j=1}^{n_r} V_{j,i}$, that is strictly larger than the sum-variance of any other virtual angles. (This condition implies that the c.b.a. is unique.)

(b) The SNR is below a threshold, i.e., $p < p_s$, where p_s is a fixed constant.

Proof: It follows from Lemma 1 that $f_j(p)$ is strictly increasing at any p such that $f_j(p) \leq 0$. This implies that (i) If $f_j(0) \geq 0$, then $f_j(p)$ is nonzero over $p \in (0,\infty)$, and thus we have $f_j(p) > 0$ for all $p > 0$. (ii) If $f_j(0) < 0$, then $f_j(p)$ has a unique zero point p_j such that $f_j(p) < 0$ if and only if $p < p_j$. Hence, if virtual angle i is the c.b.a, i.e., $f_j(p) \leq 0$ for some $p > 0$ and $j \neq i$, then it follows from (i) that we must have $f_j(0) < 0$. Since $f_j(0) = E(\|\tilde{h}_j\|^2 - \|\tilde{h}_i\|^2)$, angle i must have the largest sum-variance. This proves Theorem 2 (a). Theorem 2 (b) follows from (ii) by letting $p_s = \min_{j \neq i} p_j$. ∎

B. Multi-User Case

For the multi-user case, we first present an example to show that for the c.b.a. of a particular user is not unique, and it may vary with the SNRs and beamforming angles of other users in the system.

Example 1: Consider a two-user MAC-VR with $n_t = n_r = 2$. The variance matrices are $V^1 = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$, $V^2 = \begin{pmatrix} 2.5 & 0 \\ 0 & 1.5 \end{pmatrix}$. From condition (6) we find that, if $(p^1, p^2) = (-15 \text{ dB}, -10 \text{ dB})$, the first virtual angle is the c.b.a. for both users. If $(p^1, p^2) = (-3 \text{ dB}, -10 \text{ dB})$, however, the second virtual angle becomes the c.b.a. for user 2, while the first angle is still the c.b.a. for user 1.

Next, we show in Theorem 3 that when the variance matrix satisfies certain properties, some of the n_t virtual angles cannot be the c.b.a. of a given user. We first present a useful Lemma.

Lemma 2: Assume that there exists virtual angels i, j for a user k in the MAC-VR such that

$$
V_{m,i}^k \leq V_{m,j}^k \quad \text{for every } m = 1, \cdots, n_r. \tag{11}
$$

Then we have $E\left[\frac{c_j - c_i}{1 + p^k c_i}\right] > 0$, where $c_j = \tilde{h}_j^{k\dagger} D \tilde{h}_j^k, c_i = \tilde{h}_i^{k\dagger} D \tilde{h}_i^k$, and $D = \left(I_{n_r} + \sum_{l \neq k}^{K} p^l \tilde{h}_{i_l}^l \tilde{h}_{i_l}^{l\dagger}\right)^{-1}$.

Proof: It suffices to show that for any fixed D, we have $\bar{E}\left[\frac{c_j - c_i}{1 + p^k c_i}\right] > 0$, here $\bar{E}(\cdot)$ denotes the conditional expectation over $\tilde{h}_j^k$ and $\tilde{h}_i^k$, given D. Since c_j and c_i are independent given D, we obtain

$$
\begin{aligned}
\bar{E}\left[\frac{c_j - c_i}{1 + p^k c_i}\right] &= \frac{1}{p^k}\left(\bar{E}\left[\frac{1 + p^k c_j}{1 + p^k c_i}\right] - 1\right) \\
&= \frac{1}{p^k}\left(\bar{E}[1 + p^k c_j] \cdot \bar{E}\left[\frac{1}{1 + p^k c_i}\right] - 1\right) \\
&> \frac{1}{p^k}\left(\frac{\bar{E}[1 + p^k c_j]}{\bar{E}[1 + p^k c_i]} - 1\right) \\
&= \frac{1}{p^k}\left(\frac{1 + p^k \bar{E}[c_j]}{1 + p^k \bar{E}[c_i]} - 1\right),
\end{aligned}
\tag{12}
$$

where the inequality follows from the Jensen's inequality $E(1/x) > 1/(Ex)$. Next, we show that $\bar{E}[c_j] \geq \bar{E}[c_i]$, and thus the right-hand side of (12) is non-negative. We can write $\bar{E}[c_j] - \bar{E}[c_i] = \bar{E}[\tilde{h}_j^{k\dagger} D \tilde{h}_j^k] - \bar{E}[\tilde{h}_i^{k\dagger} D \tilde{h}_i^k] = \text{Tr}\left[D \cdot \bar{E}\left(\tilde{h}_j^k \tilde{h}_j^{k\dagger} - \tilde{h}_i^k \tilde{h}_i^{k\dagger}\right)\right]$. It follows from (11) that $\bar{E}\left(\tilde{h}_j^k \tilde{h}_j^{k\dagger} - \tilde{h}_i^k \tilde{h}_i^{k\dagger}\right) = \text{diag}(V_{1,j}^k - V_{1,i}^k, \cdots, V_{n_r,j}^k - V_{n_r,i}^k)$ is a diagonal matrix with non-negative entries. This, combined with D being a positive definite Hermitian matrix with positive diagonal entries, imply that $\bar{E}[c_j] \geq \bar{E}[c_i]$. ∎

Theorem 3: Consider user k in a K-user MAC-VR. The i-th virtual angle cannot be the c.b.a. of user k if there exists another virtual angle j such that (11) holds.

Proof: It is sufficient to prove that the beamforming condition (6), with $i_k = i$, is violated for any j that satisfies (11), i.e., $f_j^k(p^k) > 0$. To prove this, we apply the matrix inversion lemma to write

$$
\begin{aligned}
f_j^k(p^k) &= E\left[\frac{c_j - c_i}{1 + p^k c_i}\right] + E\left[\frac{c_j c_i - (\tilde{h}_j^{k\dagger} D \tilde{h}_i^k)(\tilde{h}_i^{k\dagger} D \tilde{h}_j^k)}{(p^k)^{-1} + c_i}\right] \\
&> E\left[\frac{c_j - c_i}{1 + p^k c_i}\right].
\end{aligned}
\tag{13}
$$

The inequality above is due to the Cauchy-Schwartz inequality that $c_j c_i - (\tilde{h}_j^{k\dagger} D \tilde{h}_i^k)(\tilde{h}_i^{k\dagger} D \tilde{h}_j^k) > 0$. It then follows from Lemma 2 and (13) that $f_j^k(p^k) > 0$. Thus, angle i cannot be the c.b.a. of user k. ∎

Example 2: Assume that $V^k = \begin{pmatrix} 1.5 & 1.6 & 0.3 \\ 0.6 & 0.8 & 1.2 \end{pmatrix}$. Since (11) is satisfied for $i = 1$ and $j = 2$, it follows from Theorem 3 that the first virtual angle cannot be the c.b.a for user k. This result holds independent of the SNR and other users' beamforming angles.

An immediate corollary of Theorem 3 is as follows:

Corollary 2: If there exists a virtual angle i such that $V^k_{m,i} > V^k_{m,j}$, for every $1 \leq m \leq n_r$ and $j \neq i$, then angle i is the only possible c.b.a.

Corollary 2 is applicable to both MAC-VR and MAC-Kr. For MAC-Kr, we see that the c.b.a. of user k is the i-th angle that maximizes $\{b^k_j, j = 1, \cdots, n_t\}$ (defined in (3)), and thus the c.b.a is unique. In comparison, the c.b.a. is not unique for MAC-VR. Such differences arise because the variance matrix for MAC-VR takes a general form, while that of MAC-Kr is restricted to the product form (3).

V. POWER ALLOCATION FOR LARGE SYSTEMS

In this section, we show that under mild conditions, the sum-capacity of a K-user system, denoted by $C(K)$, grows in the order of $n_r \log K$. Furthermore, we present conditions under which the sum-rate achieved by a power allocation scheme $\boldsymbol{\lambda}$ is within a constant of $C(K)$ as K goes to infinity.

Given a power allocation $\boldsymbol{\lambda} = \{\lambda^k_j, k = 1, \cdots, K, j = 1, \cdots, n_t\}$, let $A_K = I_{n_r} + \sum_{k=1}^{K} \frac{p^k}{n_t} \tilde{H}^k \tilde{Q}^k \tilde{H}^{k\dagger} = I_{n_r} + \sum_{k=1}^{K} \sum_{j=1}^{n_t} \frac{p^k}{n_t} \lambda^k_j \tilde{h}^k_j \tilde{h}^{k\dagger}_j$. The sum-rate achieved by $\boldsymbol{\lambda}$ is given by $I(\boldsymbol{\lambda}, K) = E[\log \det A_K]$. We apply Jensen's inequality to obtain an upper bound $\bar{I}(\boldsymbol{\lambda}, K)$ such that

$$
\begin{aligned}
I(\boldsymbol{\lambda}, K) &= E[\log \det A_K] \leq \log \det E(A_K) \\
&= \sum_{m=1}^{n_r} \log \left(1 + \sum_{k=1}^{K} \sum_{j=1}^{n_t} \frac{p^k}{n_t} \lambda^k_j V^k_{m,j}\right) \\
&= \bar{I}(\boldsymbol{\lambda}, K). \quad (14)
\end{aligned}
$$

Proposition 1 below shows that under mild conditions, $\bar{I}(\boldsymbol{\lambda}, K)$ is asymptotically tight as $K \to \infty$.

Proposition 1: Assume that $M_4 = \sup_{m,j,k}(p^k)^2 E(|\tilde{h}^k_{m,j}|^4) < \infty$. If there exists a constant $c > 0$ such that $\boldsymbol{\lambda}$ satisfies

$$
\min_{1 \leq m \leq n_r} \liminf_{K \to \infty} \frac{1}{Kn_t} \sum_{k=1}^{K} \sum_{j=1}^{n_t} \frac{p^k}{n_t} \lambda^k_j V^k_{m,j} \geq c > 0, \quad (15)
$$

then we have $I(\boldsymbol{\lambda}, K) = \bar{I}(\boldsymbol{\lambda}, K) + o(1)$, where $o(1)$ converges to zero as $K \to \infty$.

Proof. First, we write

$$
\begin{aligned}
I(\boldsymbol{\lambda}, K) - \bar{I}(\boldsymbol{\lambda}, K) &= E \log \frac{\det A_K}{\det E(A_K)} \\
&= E \log \frac{\det A_K}{\prod_{m=1}^{n_r} \left[1 + \sum_{k=1}^{K} \sum_{j=1}^{n_t} \frac{p^k}{n_t} \lambda^k_j V^k_{m,j}\right]} \\
&= E \log \det \tilde{A}_K,
\end{aligned}
$$

where the (m, i)-th element of $\tilde{A}_K$, denoted by $(\tilde{A}_K)_{m,i}$, is given by

$$
\begin{aligned}
(\tilde{A}_K)_{m,i} &= \frac{\left(1 + \sum_{k=1}^{K} \sum_{j=1}^{n_t} \frac{p^k}{n_t} \lambda^k_j \tilde{h}^k_{m,j} \tilde{h}^{k\dagger}_{i,j}\right)}{\left(1 + \sum_{k=1}^{K} \sum_{j=1}^{n_t} \frac{p^k}{n_t} \lambda^k_j V^k_{m,j}\right)} \\
&= \frac{\left(1 + \sum_{k=1}^{K} \sum_{j=1}^{n_t} \frac{p^k}{n_t} \lambda^k_j \tilde{h}^k_{m,j} \tilde{h}^{k\dagger}_{i,j}\right)/(1 + Kn_t)}{\left(1 + \sum_{k=1}^{K} \sum_{j=1}^{n_t} \frac{p^k}{n_t} \lambda^k_j V^k_{m,j}\right)/(1 + Kn_t)}. \quad (16)
\end{aligned}
$$

We let $\bar{S}_K$ denote the numerator in the second fraction of (16), which equals the average of the summation of independent random variables. It follows from the Strong Law of Large Number (SLLN) and (15) that

$$
\begin{aligned}
0 &= \lim_{K \to \infty} |\bar{S}_K - E(\bar{S}_K)| = \limsup_{K \to \infty} \left(E(\bar{S}_K) \left| \frac{\bar{S}_K}{E(\bar{S}_K)} - 1\right|\right) \\
&\geq \left(\liminf_{K \to \infty} E(\bar{S}_K)\right) \left(\limsup_{K \to \infty} \left| \frac{\bar{S}_K}{E(\bar{S}_K)} - 1\right|\right) \\
&\geq c \cdot \limsup_{K \to \infty} \left| \frac{\bar{S}_K}{E(\bar{S}_K)} - 1\right|.
\end{aligned}
$$

Thus, for $m = i$ we obtain $\lim_{K \to \infty}(\tilde{A}_K)_{i,i} = \lim_{K \to \infty} \bar{S}_K/E(\bar{S}_K) = 1$. For $m \neq i$, since $E(\bar{S}_K) = 0$, it follows from (15) and the SLLN that $\lim_{K \to \infty}(\tilde{A}_K)_{m,i} \leq \frac{1}{c} \lim_{K \to \infty} \bar{S}_K = \frac{1}{c} \lim_{K \to \infty} (\bar{S}_K - E(\bar{S}_K)) = 0$. This proves that $\tilde{A}_K$ converges to the identity matrix and thus $E \lim_{K \to \infty} [\log \det \tilde{A}_K] = 0$. Assume that the distributions of the random vectors $\{\tilde{h}^k_j\}$ are sufficiently smooth to facilitate exchange of the limit and the expectation operator, we obtain $\lim_{K \to \infty} [I(\boldsymbol{\lambda}, K) - \bar{I}(\boldsymbol{\lambda}, K)] = \lim_{K \to \infty} [E \log \det \tilde{A}_K] = E[\lim_{K \to \infty} \log \det \tilde{A}_K] = 0$. ∎

Note that although both proofs apply the SLLN, the proof of Proposition 1 differs from that of [9, Lemma 2] for the MAC-Kr in that we provide a sufficient condition (15), which guarantees that (17) holds. From Proposition 1 we obtain Corollary 3 below.

Corollary 3: For any $\boldsymbol{\lambda}$ that satisfies (15), we have $I(\boldsymbol{\lambda}, K) = C(K) + O(1) = n_r \log K + O(1)$, where $O(1)$

denotes a bounded quantity as $K \to \infty$. Hence, $I(\boldsymbol{\lambda}, K)$ grows in the order of $n_r \log K$, and asymptotically it differs from the sum-capacity $C(K)$ by only a constant. *Sketch of proof.* It is sufficient to show that there exists constants u_1 and u_c such that

$$
\begin{aligned}
n_r \log K + u_c + o(1) \leq I(\boldsymbol{\lambda}, K) \leq C(K) \leq \\
\max_{\boldsymbol{\lambda}} \bar{I}(\boldsymbol{\lambda}, K) \leq n_r \log K + u_1 + o(1).
\end{aligned} \quad (17)
$$

From (15) we can find u_c such that $\bar{I}(\boldsymbol{\lambda}, K) \geq n_r \log K + u_c$. This, combined with Proposition 1, leads to the first inequality of (17). The last inequality of (17) utilizes $M_2 = \sup_{m,j,k} p^k V_{m,j}^k < \infty$. ∎

Next, we consider a simple example for which we can characterize the term $O(1)$ in Corollary 3 for various $\boldsymbol{\lambda}$. The accuracy of these computations will be verified in Section VI.

Example 3: Assume that all users have the same $p^k = 2$ and the same variance matrix $V = \begin{bmatrix} 2 & 0.5 \\ 0.5 & 1 \end{bmatrix}$. The virtual elements in $\tilde{H}^k$ are complex Gaussian distributed. Assume that each user adopts the same power allocation $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ such that $\lambda_1 + \lambda_2 = 2$. For each $\boldsymbol{\lambda}$ that satisfies the assumptions of Proposition 1, we have

$$
\begin{aligned}
I(\boldsymbol{\lambda}, K) &= \bar{I}(\boldsymbol{\lambda}, K) + o(1) \\
&= 2\log(K) + \log(2\lambda_1 + 0.5 \cdot \lambda_2) + \\
&\quad \log(0.5 \cdot \lambda_1 + \lambda_2) + o(1).
\end{aligned} \quad (18)
$$

Consider the following three power allocations for which Proposition 1 and Corollary 3 are applicable.

(1) The beamforming scheme $\boldsymbol{\lambda}_{\mathrm{BF}}$ where each user beamforms to the first virtual angle which has the largest sum-variance. Since $\lambda_1 = 2, \lambda_2 = 0$, it follows from (18) that

$$
\begin{aligned}
I(\boldsymbol{\lambda}_{\mathrm{BF}}, K) &= \bar{I}(\boldsymbol{\lambda}_{\mathrm{BF}}, K) + o(1) \\
&= 2\log K + \log(4) + o(1).
\end{aligned} \quad (19)
$$

(2) The equal power allocation $\boldsymbol{\lambda}_{\mathrm{Eq}}$ such that $\lambda_1 = \lambda_2 = 1$. From (18) we have

$$
\begin{aligned}
I(\boldsymbol{\lambda}_{\mathrm{Eq}}, K) &= \bar{I}(\boldsymbol{\lambda}_{\mathrm{Eq}}, K) + o(1) \\
&= 2\log K + \log(15/4) + o(1) \\
&= 2\log K + 1.9069 + o(1).
\end{aligned} \quad (20)
$$

(3) We can choose (λ_1, λ_2) to maximize the summation of the two constant terms in (18). This yields the optimized solution $\boldsymbol{\lambda}^* = (\lambda_1^*, \lambda_2^*) = (5/3, 1/3)$. Hence, from (18) we have

$$
\begin{aligned}
I(\boldsymbol{\lambda}^*, K) &= \bar{I}(\boldsymbol{\lambda}^*, K) + o(1) \\
&= 2\log K + \log(49/12) + o(1) \\
&= 2\log K + 2.0297 + o(1).
\end{aligned} \quad (21)
$$

The constant term in (21) is slightly greater than that of $\boldsymbol{\lambda}_{\mathrm{BF}}$ in (19) and that of $\boldsymbol{\lambda}_{\mathrm{Eq}}$ in (20). This example demonstrates that beamforming may not be asymptotically optimal for the MAC-VR, even though it is asymptotically optimal for MAC-Kr [9, Theorem 7]. Corollary 4 below provides a sufficient condition under which beamforming is asymptotically optimal for MAC-VR.

Corollary 4: Beamforming is asymptotically optimal for the MAC-VR: $I(\boldsymbol{\lambda}_{\mathrm{BF}}, K) = C(K) + o(1)$, as $K \to \infty$, if each user k beamforms to a virtual angle i_k that satisfies $V_{m,i_k}^k \geq V_{m,j}^k$ for every $1 \leq m \leq n_r$ and $1 \leq j \leq n_t$, and there exists a constant $c > 0$ such that

$$
\min_{1 \leq m \leq n_r} \liminf_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} p^k V_{m,i_k}^k \geq c > 0. \quad (22)
$$

Sketch of proof. Given $V_{m,i_k}^k \geq V_{m,j}^k$, one can show that $C(K) \leq \bar{I}(\boldsymbol{\lambda}_{\mathrm{BF}}, K)$. Condition (22) ensures that $I(\boldsymbol{\lambda}_{\mathrm{BF}}, K) = \bar{I}(\boldsymbol{\lambda}_{\mathrm{BF}}, K) + o(1)$. Thus Corollary 4 follows.

Considering a special case of Corollary 4 in which we let i_k be the virtual angle that maximizes $\{b_j^k, j = 1, \cdots, n_t\}$, then we obtain the same result as [9, Theorem 7] that beamforming is always asymptotically optimal for MAC-Kr. In comparison, as shown in Example 3, there exists MAC-VR such that beamforming is not asymptotically optimal. This difference, again, is due to the general structure of the variance matrix for MAC-VR.

VI. NUMERICAL RESULTS

In this section, we present numerical examples to illustrate the theoretical results given in previous sections. Four power allocation schemes are considered: the equal power allocation ($\boldsymbol{\lambda}_{\mathrm{Eq}}$), the beamforming scheme ($\boldsymbol{\lambda}_{\mathrm{BF}}$), the optimal power allocation ($\boldsymbol{\lambda}_{\mathrm{Opt}}$) found by the algorithm of [10], and a low-complexity power allocation algorithm derived based on [16] ($\boldsymbol{\lambda}_{\mathrm{Low}}$). Let $I(\boldsymbol{\lambda})$ denote the sum-rate achieved by $\boldsymbol{\lambda}$. We first consider a single-user system with $n_r = n_t = 5$. The virtual coefficients in $\tilde{H}$ are assumed to be complex Gaussian distributed with the same variance matrix as the one in [4], given by

$$
V = \frac{25}{5.7} \begin{bmatrix} 0.1 & 0 & 1 & 0 & 0 \\ 0 & 0.1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0.25 & 0 \\ 0 & 0 & 1 & 0 & 0.25 \end{bmatrix}. \quad (23)
$$

Please refer to [4] for the physical meaning of such a variance matrix. The third virtual angle is the beamforming angle because it has the largest sum-variance. Fig.1
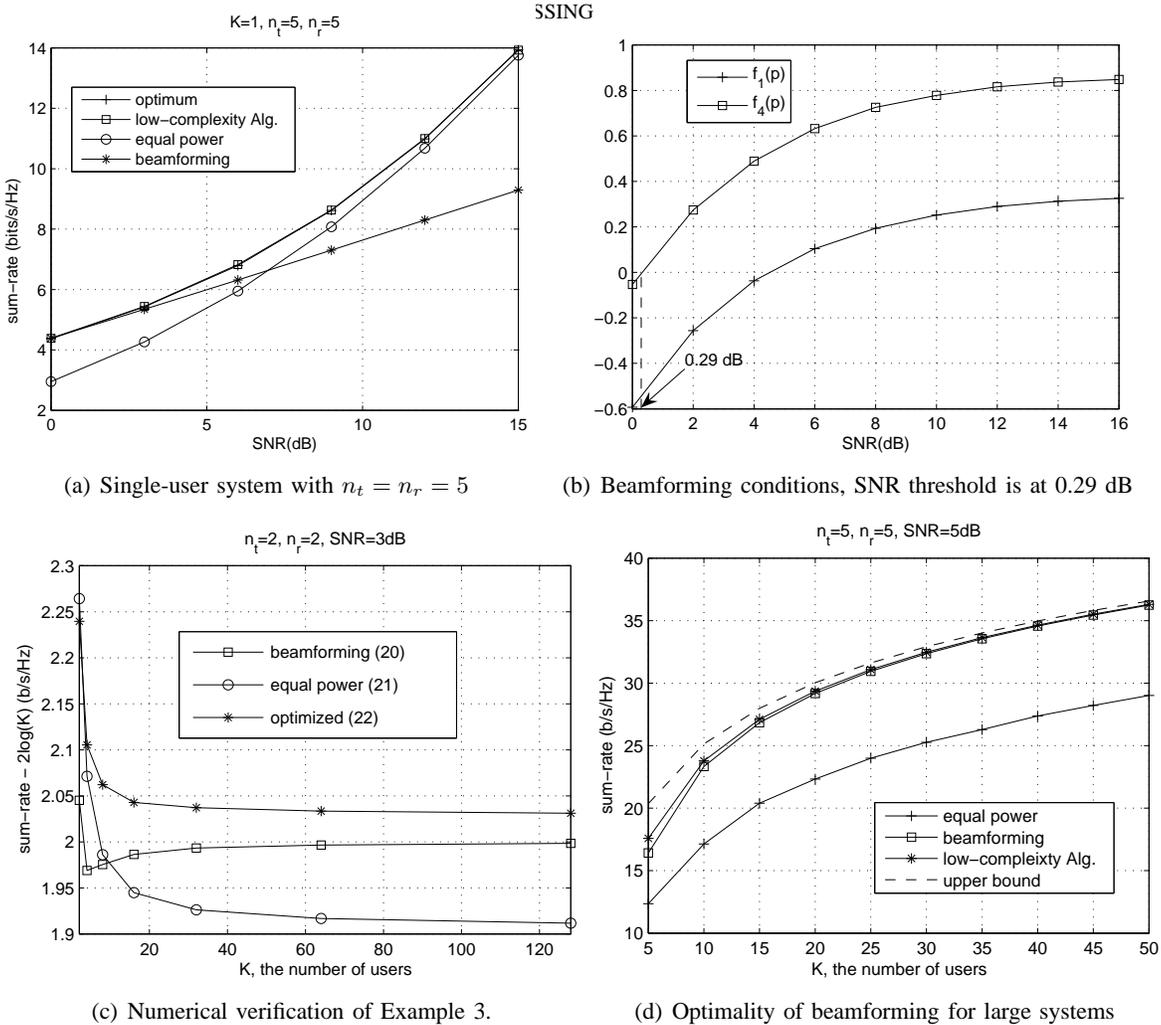
(a) Single-user system with $n_t = n_r = 5$

(b) Beamforming conditions, SNR threshold is at 0.29 dB

(c) Numerical verification of Example 3.

(d) Optimality of beamforming for large systems

Fig. 1. Optimality of beamforming for MAC-VR.

(a) shows that $I(\boldsymbol{\lambda}_{\mathrm{Low}})$ is very close to $I(\boldsymbol{\lambda}_{\mathrm{Opt}})$ for the entire range of SNRs considered. $I(\boldsymbol{\lambda}_{\mathrm{Eq}})$ is near optimal only at high SNR and $I(\boldsymbol{\lambda}_{\mathrm{BF}})$ is optimal only when SNR is below the threshold of 0.29 dB. This is consistent with the threshold behavior proved in Theorem 2. In Fig. 1 (b), we plot the beamforming conditions defined in Section IV-A for the third virtual angle $i = 3$. Since $f_2(p) = f_1(p)$ and $f_5(p) = f_4(p)$, Fig. 1 (b) plots $f_1(p)$ and $f_4(p)$ and shows that when SNR is below 0.29 dB, both functions are negative and thus beamforming to the third virtual angle is optimal.

In Fig. 1 (c), we examine the accuracy of Proposition 1 and Corollary 3 by comparing the asymptotic expressions (19)-(21) in Example 3 with numerical values of $I(\boldsymbol{\lambda}, K)$ obtained through Monte Carlo integration. Three functions $I(\boldsymbol{\lambda}_{\mathrm{BF}}, K) - 2\log K$, $I(\boldsymbol{\lambda}_{\mathrm{Eq}}, K) - 2\log K$, and $I(\boldsymbol{\lambda}^*, K) - 2\log K$ are plotted to confirm that as K increases, they indeed converge to the predicted constants 2, 1.9069, and 2.0297, respectively.

Fig. 1 (d) considers a multi-user system in which each V^k is generated independently, taking a form similar

to (23). The beamforming angle of user k is chosen to be the virtual angle with the largest sum-variance. Because $\{V^k\}$ satisfy conditions of Corollary 4, $\boldsymbol{\lambda}_{\mathrm{BF}}$ is asymptotically optimal. This is confirmed in Fig. 1 (d). The curve for $I(\boldsymbol{\lambda}_{\mathrm{Opt}}, K)$ is not shown due to high complexity for computing $\boldsymbol{\lambda}_{\mathrm{Opt}}$. Instead, we provide a simple sum-capacity upper bound $C(K) \leq n_r \log(1 + KM_2)$ as a performance benchmark for large K. Hence, the gap between $I(\boldsymbol{\lambda}_{\mathrm{BF}}, K)$ and $C(K)$ is less than the small gap shown in Fig. 1 (d) between $I(\boldsymbol{\lambda}_{\mathrm{BF}}, K)$ and the upper bound. The gap becomes negligible as K increases, confirming the optimality of beamforming. For small K, $I(\boldsymbol{\lambda}_{\mathrm{Low}}, K)$ closely approximates $I(\boldsymbol{\lambda}_{\mathrm{Opt}}, K)$ (not shown). For large K, $I(\boldsymbol{\lambda}_{\mathrm{Low}}, K)$ and $I(\boldsymbol{\lambda}_{\mathrm{BF}}, K)$ merge quickly and become indistinguishable after $K \geq 30$. We note that $I(\boldsymbol{\lambda}_{\mathrm{Eq}}, K)$ is inferior to $I(\boldsymbol{\lambda}_{\mathrm{BF}}, K)$ by roughly a constant, even though it achieves the same asymptote of $n_r \log K = 5 \log K$. This is consistent with Corollary 3.

VII. CONCLUSION

In this paper, we study the optimality of beamforming for the MAC-VR. For the single-user case, we provide a mathematical proof of the threshold behavior for the optimality of beamforming which is applicable to both Kronecker model and VR model. We present useful criteria in determining the capacity-achieving beamforming angles for a class of MAC-VR models. Due to the generality of the VR model, we demonstrate by examples that existing results for the MAC-Kr may not be valid for the MAC-VR. These include the uniqueness of the capacity-achieving beamforming angle, and the optimality of beamforming for systems with large number of users.

REFERENCES

[1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecom*, vol. 10, pp. 585–595, Nov. 1999.

[2] S. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1165–1175, July 2004.

[3] Z. Wang and G. B. Giannakis, "Outage mutual information rate of space-time MIMO channels," *IEEE Trans. on Information Theory*, vol. 50, no. 4, pp. 657–662, Sep. 2001.

[4] V. V. Veeravalli, Y. Liang, and A. M. Sayeed, "Correlated MIMO wireless channels: Capacity, optimal signaling, and asymptotics," *IEEE Transactions on information theory*, vol. 51, pp. 2058–2072, 2005.

[5] A. M.Tulino, A. Lozano, and S. Verdu, "Impact of antenna correlation on the capacity of multiantenna channels," *IEEE Trans. on Information Theory*, vol. 51, pp. 2491–2509, July 2005.

[6] E. Jorswieck and H. Boche, "Channel capacity and capacity range of beamforming in MIMO wireless systems under correlated fading with covariance feedback," *IEEE Trans. on Wireless Commu.*, vol. 3, pp. 1543–1553, Sep. 2004.

[7] W. Rhee, W. Yu, and J. M. Cioffi, "The optimality of beamforming in uplink multiuser wireless systems," *IEEE Trans. on Wireless Commu.*, vol. 3, no. 1, pp. 86–96, Jan. 2004.

[8] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering, and keyhole," *IEEE Trans. on Info. Theory*, vol. 49, pp. 2636–2647, Oct. 2003.

[9] A. Soysal and S. Ulukus, "Optimality of beamforming in fading multiple access channels," *IEEE Trans. on Communications*, pp. 1171–1183, April, 2009.

[10] ——, "Optimum power allocation for single-user MIMO and multi-user MIMO-MAC with partial CSI," *IEEE Trans. on Selected Areas in Communication*, no. 7, pp. 1402–1412, Sep. 2007.

[11] A. M. Sayeed, "Deconstructing multi-antenna fading channels," *IEEE Trans. Signal Processing*, pp. 2563–2579, Oct, 2002.

[12] X. Li, S. Jin, X. Gao, M. Mckay, and K. Wong, "Transmitter optimization and beamforming optimality conditions for double-scattering MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 9, pp. 3647–3654, Sep. 2008.

[13] D. Gesbert, H. Bolcskei, D. A. Gore, and A. J. Paulraj, "Ourdoor MIMO wireless channels: Models and performance prediction," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 1886–1888, Dec. 2002.

[14] W. Dai, B. Rider, and Y. Liu, "Joint beamforming for multiaccess MIMO systems with finite rate feedback," *IEEE Transactions on Wireless Communications*, vol. 8, no. 5, pp. 2618–2628, May. 2009.

[15] H. Wan, R.-R. Chen, and Y. Liang, "Optimality of beamforming in MIMO multi-access channels via virtual representation," *in Proc. IEEE International Symposium on Information Theory (ISIT'08)*, pp. 2573–2577, July, 2008.

[16] S. Lasaulce, A. Suarez, M. Debbah, and L. Cottatellucci, "Power allocation game for fading MIMO multiple access channels with antenna correlation," *Conference on GameComm*, Oct. 2007.