How Much Does Transmit Correlation Affect the Sum-Rate Scaling of MIMO Gaussian Broadcast Channels?

Tareq Y. Al-Naffouri, Masoud Sharif, *Senior Member, IEEE*, and Babak Hassibi

Abstract— This paper considers the effect of spatial correlation between transmit antennas on the sum-rate capacity of the MIMO Gaussian broadcast channel (i.e., downlink of a cellular system). Specifically, for a system with a large number of users n, we analyze the scaling laws of the sum-rate for the dirty paper coding and for different types of beamforming transmission schemes. When the channel is i.i.d., it has been shown that for large n, the sum rate is equal to $M \log \log n + M \log \frac{P}{M} + o(1)$ where M is the number of transmit antennas, P is the average signal to noise ratio, and $o(1)$ refers to terms that go to zero as $n \to \infty$. When the channel exhibits some spatial correlation with a covariance matrix R (non-singular with $tr(R) = M$), we prove that the sum rate of dirty paper coding is $M \log \log n + M \log \frac{P}{M} + \log \det(R) + o(1)$. We further show that the sum-rate of various beamforming schemes achieves $M \log \log n + M \log \frac{P}{M} + M \log c + o(1)$ where $c \leq 1$ depends on the type of beamforming. We can in fact compute c for random beamforming proposed in [1] and more generally, for random beamforming with precoding in which beams are pre-multiplied by a fixed matrix. Simulation results are presented at the end of the paper.

Index Terms— Broadcast channel, channel state information, multi-user diversity, transmit correlation, wireless communications.

I. INTRODUCTION

MULTIPLE input multiple output (MIMO) communication has been the focus of a lot of research which basically demonstrated that the capacity of a point to point MIMO link increases linearly with the number of transmit and receive antennas. Research focus has shifted recently to the role of multiple antennas in multiuser systems, especially broadcast scenarios (i.e., *one to many* communication) as downlink scheduling is the major bottleneck for future broad-

band wireless networks. An overview of the research on this problem can be found in [3], [6], [16].

In these scenarios, when multiple users are present, one is usually interested in 1) quantifying the maximum possible sum rate to all users and 2) devising computationally efficient algorithms for capturing most of this rate [3]. The first question was settled recently by using a technique similar to writing on dirty paper and hence known as dirty paper coding (DPC). While DPC solves the broadcast problem optimally, it is computationally expensive and requires a great deal of feedback as the transmitter needs perfect channel state information for all users [3].

There has been increased interest recently to devise simple techniques that utilize multiuser diversity and achieve a sum-rate close to the sum-rate capacity of the MIMO broadcast channel (see, e.g., [1], [5]–[8], [12], [15]). The scheme proposed in [1], known as opportunistic multiple random beamforming (or concisely random beamforming), has been proved to asymptotically maximize the sum-rate (or throughput) of the downlink of single antenna cellular systems by transmitting to the users with the best channel conditions for a given set of random beams. The gain of this and other beamforming schemes can be attributed to multiuser diversity– each user experiences a different channel and therefore the transmitter can exploit this variation and choose the users that have the best channel conditions. Clearly, the multiuser gain would be specially magnified when the channels between the transmitter and the users are changing independently.

In this paper we focus on a multi-antenna downlink channel in the presence of correlation between transmit antennas. This correlation is caused by local scatterers around the base station or the fact that the transmit antennas in the base station are not spaced far enough to create independent channels. The overriding question then is to analyze the effect of this correlation on the sum-rate of DPC and various beamforming scheduling techniques.

Specifically, we consider three variations of random beamforming, namely, random beamforming with channel whitening, beamforming with general precoding, and deterministic beamforming. In the first, the transmitter spatially whitens the channel and then uses random beamforming. In random beamforming with precoding, the transmitter employs a more general precoding matrix. In both of these transmission schemes, the transmitted signal needs to be scaled properly to maintain the average power constraint. Finally, in deter-

ministic beamforming, as its names suggests, we use a fixed beamformer for all channel uses in place of the randomly varying one.

When the number of users is large and there is no correlation, the sum rate for DPC and random beamforming asymptotically coincide (see Lemma 1 and [1], [2])

$$R = M \log \log n + M \log \frac{P}{M} + o(1) \qquad (1)$$

where n is the number of users, M is the number of transmit antennas, and P is the average signal to noise ratio, and $o(1)$ represents terms that go to zero as $n \to \infty$. It turns out that this is not case for the channel with transmit correlation. In this case, the sum-rate can be written as

$$M \log \log n + M \log \frac{P}{M} + M \log c + o(1) \qquad (2)$$

where the constant $c \leq 1$ (which refers to the sum-rate loss due to correlation) depends on the scheduling scheme and the eigenvalues of the covariance matrix R.

The results of this paper strongly depend on assuming that the users have a common correlation matrix R.[1] This is essential to make the users's channels (and hence SINR's) iid. Otherwise, it would be very difficult to use extreme value theory to perform the scaling analysis. Fortunately, our analysis extends to the case when the users experience different path loss (i.e. when the users correlation matrices are the same modulo some multiplicative constant).

The paper is organized as follows. After introducing the channel model in the next section, we review in Section III the different scheduling schemes studied in this paper. We obtain the scaling law of the sum-rate for DPC and random beamforming schemes in Section IV and V, respectively. Section V-B, which is the heart of the paper, is devoted to deriving the scaling law of random beamforming in a spatially correlated environment. We use this result to derive the scaling laws for random beamforming with precoding and for deterministic beamforming. We conclude the paper with simulations and conclusions.

II. CHANNEL MODEL AND PROBLEM FORMULATION

In this paper we consider a multi-antenna Gaussian broadcast channel with n receivers equipped with one antenna and a transmitter (base station) with M antennas. Let $S(t)$ be the $M \times 1$ vector of the transmit symbols at time slot t, and let $Y_i(t)$ be the received signal at the i'th receiver. We can then write the received signal at the i'th user as

$$Y_i(t) = \sqrt{P} H_i S(t) + W_i, \qquad\qquad i = 1, \ldots, n, \quad (3)$$

where W_i is the additive noise which is complex Gaussian with zero mean and unit variance, $CN(0,1)$. Moreover, $S(t)$ is the transmit symbol satisfying the power constraint $E\{S^*S\} = 1$. Here P denotes the average transmit power (or equivalently the average SNR considering the normalization of the variances for channel and noise).

The channel H_i is a $1 \times M$ complex channel vector, known perfectly to the receiver, and distributed as $CN(0, R)$. The $M \times M$ covariance matrix R is a measure of the spatial correlation and is assumed to be non-singular with $tr(R) = M$ [2]. We also assume that H_i follows a block fading model, i.e., it remains constant during a coherence interval T and varies independently from one such interval to the next. We finally note that the channel is identically distributed across users but is independent from one user to another.

Denoting the average rate of the i'the user by R_i over all the channel realizations, we are interested in analyzing the behavior of the sum-rate, i.e., $\sum_{i=1}^n R_i$, of downlink for large n.

In the following section, we review the scheduling schemes that will be considered in this paper.

III. REVIEW OF TRANSMISSION SCHEMES IN THE DOWNLINK

A. Dirty Paper Coding (DPC)

The capacity region of the multi-antenna broadcast channel is achieved by dirty paper coding when full channel state information (CSI) is available to the transmitter and users. Intuitively, if the transmitter knows the channels of all users, it can use DPC to pre-subtract the interference for each user while preserving the average power constraint [3]. More precisely, the sum rate capacity, R_{DPC}, can be written as (see [12] and the references therein),

$$R_{DPC} = E \left\{ \max_{\{P_1, \ldots, P_n, \sum P_i \leq P\}} \log \det \left(1 + \sum_{i=1}^n H_i^* P_i H_i \right) \right\} \tag{4}$$

In a system with a large number of users n, and for fixed M and P, it has been shown that the sum-rate of DPC behaves as in (1),

$$R_{DPC} = M \log \log n + M \log \frac{P}{M} + o(1), \qquad (5)$$

when there is no spatial correlation, i.e., $R = I$ [1], [2]. Scaling of the sum rate capacity has also been investigated for other regions of n, M, and P (see [8]–[10] for details).

There are two major drawbacks of this scheme. First, it is very computationally complex, both at the receivers and transmitter. Moreover, it requires full CSI feedback from all active users to the transmitter of the base station (this feedback requirement increases with the number of antennas and users and with the decrease of the coherence time of the system).

B. Random Beamforming

Given these drawbacks of DPC, research has focused on devising algorithms for multiuser broadcast channels that have less computational complexity and/or less feedback and still achieve most of the sum-rate promised by DPC such as random beamforming [15] and zero forcing [7] (see also [5], [11]). A random beamforming scheme was proposed in [1]

---

[1]One could argue that when users have different correlations matrices, we will continue to observe multiuser richness similar to that we observe in the white case.

[2]We assume that the spatial correlation is invariant across users. This assumption is realistic because this is effectively the transmit correlation among antennas at the base station. In the case when R is rank deficient, the results of this paper apply with M replaced by the rank of the autocorrelation matrix and with the SNR kept fixed at P/M.

where the transmitter sends multiple (in fact M) random orthonormal beams chosen to users with the best signal to interference ratio (SINR). In this scheme the only feedback required from each user is the SINR of the best beam and the corresponding index.

Specifically, the transmitter chooses M random orthonormal beam vectors ϕ_m (of size $M \times 1$) generated according to an isotropic distribution. Now these beams are used to transmit the symbols $s_1(t), s_2(t), \ldots, s_M(t)$ by constructing the transmitted vector

$$S(t) = \sum_{m=1}^{M} \phi_m(t) s_m(t), \qquad t = 1, \ldots, T \quad (6)$$

After T channel uses, the transmitter independently chooses another set of orthogonal vectors $\{\phi_m\}$ and constructs the signal vector (according to (6)) and so on. From now on and for simplicity, we will drop the time index t. The signal Y_i at the i'th receiver is given by

$$
\begin{aligned}
Y_i &= \sqrt{P} H_i S + W_i &&(7)\\
&= \sqrt{P} \sum_{m=1}^{M} H_i \phi_m s_m + W_i, &&i = 1, \ldots, n \,(8)
\end{aligned}
$$

where $E(SS^*) = \frac{1}{M} I$ since the s_i's are assumed to be identical and independently assigned to different users. The i'th receiver uses its knowledge of the effective channel gain $H_i \phi_m$, something that can be arranged by training, to calculate M SINR's, one for each transmitted beam

$$\text{SINR}_{i,m} = \frac{|H_i \phi_m|^2}{\frac{M}{P} + \sum_{k \neq m} |H_i \phi_k|^2}, \qquad m = 1, \ldots, M. \quad (9)$$

Each receiver then feeds back its maximum SINR, i.e. $\max_{1 \leq m \leq M} \text{SINR}_{i,m}$, along with the maximizing index m. Thereafter, the transmitter assigns s_m to the user with the highest corresponding SINR, i.e. $\max_{1 \leq i \leq n} \text{SINR}_{i,m}$. If we do the above scheduling, the throughput for large n can be written as [2] [3],

$$R_{RBF} = M E \log\left(1 + \max_{1 \leq i \leq n} \text{SINR}_{i,m}\right) + o(1) \quad (10)$$

where the term $o(1)$ accounts for the small probability that user i may be the strongest user for more than one signal s_m [1].

To further quantify (10), [1] used the fact that the $\text{SINR}_{i,m}$'s are iid over i and employed extreme value theory [20] to argue that $\max_{i \leq n} \text{SINR}_{i,m}$ behaves like $\frac{P}{M} \log n$ and hence concluded that the sum rate capacity scales as in (1), meaning that the sum-rate of random beamforming behaves the same as that of DPC for large number of users.

C. Other Beamforming Schemes

The scaling result (1) applies for iid channels. As such, we derive in Section V the scaling law of this scheme for correlated channels. Alternatively, given this correlation, we consider the following beamforming schemes.

Random beamforming with channel whitening

In the presence of correlation, one can first whiten the channel and then use random beamforming scheduling. In this case, and instead of using Φ as the beamforming matrix[4], we would use $\sqrt{\alpha} R^{-1/2} \Phi$ where α is a constant to make sure that the transmit symbol has an average power of 1. The scaling of this scheme would follow directly from the scaling of random beamforming over iid channels (see Section V-A).

Random beamforming with general precoding

More generally, we can precode with a general matrix $\sqrt{\alpha} A^{-1/2}$ before beamforming, i.e. we use $\sqrt{\alpha} A^{-1/2} \Phi$ to transmit the information symbols. The scaling of this scheme follows directly from the scaling of random beamforming over correlated channels and so is considered in Sections V-B and V-D. We go one step further and show how to compute the sum-rate when the beamforming matrix is premultiplied by the full rank matrix A.

Deterministic beamforming

Finally, by fixing the beamforming matrix Φ, we obtain deterministic beamforming, a scheme analyzed by Park and Park [14] (for the two antenna case) and which we further analyze in Section V-C.

As we mentioned above, and as we shall soon see, all these schemes have scaling similar to the iid case (1) with a penalty term $M \log c$ where $c \leq 1$ is a constant that depends only on the scheduling scheme and the correlation matrix R.

IV. EFFECT OF TRANSMIT CORRELATION ON THE SUM-RATE OF DPC

In this section, we derive the scaling laws of DPC for correlated channels. As mentioned earlier, dirty paper coding achieves the sum-rate capacity of the multi-antenna broadcast channel. The sum-rate capacity is given by (4) and its behavior when n is large is given by (5) for iid channels. It turns out that when the number of users is large, the sum-rate capacity will be decreased by a constant which depends on the covariance matrix of the channel. It should be mentioned that throughout the paper, we assume R is fixed and non-singular with $tr(R) = M$.

The next theorem proves this statement. The proof is along the same line as the proof for the i.i.d. case (as shown in [1]) with the only difference that the lower bound rather than being achieved with random beamforming is achieved with a special type of deterministic beamforming [5]. We first give the lower bound in the following lemma.

Lemma 1: Consider a Gaussian broadcast channel with a channel covariance matrix R which is non-singular with $tr(R) = M$. Let there be one transmitter with M antennas and n users with single antennas that have access to the CSI and the transmitter knows the CSI perfectly. We assume the transmitter uses the deterministic beamforming matrix

---

[3]The proof follows from the fact the when n is large the maximum SINR and the M'th maximum SINR behave quite similarly.

[4]Note that Φ is an orthonormal matrix composed of the beam (column) vectors $\phi_1, \ldots, \phi_M$.

[5]It should be stressed that the optimality here is in the asymptotic of large number of users.

$\Phi = U^*$ where U is the unitary matrix consisting of the eigenvectors of R. Then for large n, the sum-rate of this scheduling is

$$R_{BF-D} = M \log \log n + M \log \frac{P}{M} + M \log \sqrt[M]{\det(R)} + o(1). \tag{11}$$

Proof: See Section V-C for the proof. ∎

Clearly (11) is a lower bound for the sum-rate capacity. In the next theorem we show that (11) is indeed an upper bound for the sum-rate as well.

Theorem 1: Consider a Gaussian broadcast channel with an autocorrelation matrix R defined in Lemma 1. Let there be one transmitter with M antennas and n users with single antennas that have access to the CSI. Assume further that the transmitter knows the CSI perfectly. The sum-rate capacity (which is achieved by DPC) scales like

$$R_{DPC} = M \log \log n + M \log \frac{P}{M} + M \log \sqrt[M]{\det R} + o(1), \tag{12}$$

for large n.

Proof: Lemma 1 implies that the right hand side of (12) is achievable. All we need to prove the theorem is to show that the sum-rate of DPC can not be larger than (12). We use the sum rate capacity expression given in (4) to obtain an upper bound for the sum-rate. To this end, define $H_i = H_{w_i} R^{\frac{1}{2}}$, where H_{w_i} is $N(0, I)$. With this decomposition, the sum-rate capacity can be written as

$$R_{DPC} = E \max_{\{P_1, \dots, P_n, \sum P_i = P\}}$$

$$\left\{ \log \det \left(R^{-1} + \sum_{i=1}^{n} H_{w_i}^* P_i H_{w_i} \right) \det(R) \right\} \tag{13}$$

Now using the geometric-arithmetic mean inequality $\det(A) \le \left(\frac{\text{tr}(A)}{M} \right)^M$, we obtain

$$\sum_{i=1}^{n} \text{tr} \left(H_{w_i}^* P_i H_{w_i} \right) \quad \le \quad \max_i \text{tr}(H_{w_i}^* H_{w_i}) \sum_{i=1}^{n} P_i$$

$$= \quad \max_i \|H_{w_i}\|^2 P$$

to replace the $\log \det$ with an upper bound

$$\log \det \left(R^{-1} + \sum_{i=1}^{n} H_{w_i}^* P_i H_{w_i} \right) \quad \le$$

$$M \log \left(\frac{1}{M} \text{tr}(R^{-1}) + \frac{1}{M} \sum_{i=1}^{n} \text{tr}(H_{w_i}^* P_i H_{w_i}) \right) \le$$

$$M \log \left(\frac{1}{M} \text{tr}(R^{-1}) + \max_i \|H_{w_i}\|^2 \frac{P}{M} \right)$$

Since $\|H_{w_i}\|^2$ is $\chi^2(2M)$ distributed, with high probability, the maximum $\max_i \|H_{w_i}\|^2$ behaves like $\log n + O(\log \log n)$. Thus,

$$R_{DPC} \le M \log \left(\frac{tr(R^{-1})}{M} + \frac{P}{M} \log n + O(\log \log n) \right) +$$

$$\log \det R + o(1) \tag{14}$$

or using the fact that for large n

$$\log(\frac{tr(R^{-1})}{M} + \frac{P}{M} \log n + O(\log \log n)) \quad =$$

$$\log \frac{P}{M} + \log \log n + \log \left(\frac{tr(R^{-1})}{P} + O \left(\frac{\log \log n}{\log n} \right) \right) =$$

$$\frac{P}{M} + \log \log n + O \left(\frac{\log \log n}{\log n} \right), \tag{15}$$

we can further simplify (14) and obtain

$$R_{DPC} \le M \log \log n + M \log \frac{P}{M} + M \log \sqrt[M]{\det R} + o(1)$$

which is the desired upper bound. This completes the proof of the theorem. ∎

V. EFFECT OF TRANSMIT CORRELATION ON RANDOM BEAMFORMING

The deterministic beamforming scheme of Lemma 1 asymptotically achieves the DPC sum-rate. However it has the drawback that, unless the H_i's change very rapidly over different channel uses, it will often transmit to a fixed set of users. To make the scheduling more short-term fair, it is useful to further randomize the user selection by random beamforming (see [1], [15] for more details). In this section, we analyze the effect of correlation on the sum-rate of random beamforming. We start by the simplest case in which the beamforming matrix is multiplied by $R^{-1/2}$ in order to whiten the channel. We then turn our attention to the random beamforming scheme and finally use it to deduce the sum rates of deterministic beamforming and beamforming with general precoding.

A. Random Beamforming with Channel Whitening

To whiten the channel, we multiply all the beams with $\sqrt{\gamma} R^{-1/2}$ where γ is a normalization factor. The transmit symbol is therefore equal to

$$S(t) = \sum_{m=1}^{M} \sqrt{\gamma} R^{-1/2} \phi_m(t) s_m(t) \tag{16}$$

We choose γ to satisfy the power constraint– that the transmit symbol average power is bounded by unity,

$$E\{\gamma S^* R^{-1} S\} = \gamma E\{tr(SR^{-1}S^*)\}$$
$$= \gamma E\{tr(R^{-1}S^*S)\}$$
$$= \gamma tr(R^{-1}E(S^*S))\}$$
$$= \gamma \frac{tr(R^{-1})}{M} \tag{17}$$

Thus, the constraint $E\{\gamma S^* R^{-1} S\} \le 1$ implies that $\gamma \le \frac{M}{tr(R^{-1})}$. We can therefore write the SINR as

$$\text{SINR}_{i,m} = \frac{|H_i R^{-1/2} \phi_m|^2}{\frac{M}{P\gamma} + \sum_{k \ne m} |H_i R^{-1/2} \phi_k|^2} \tag{18}$$

$$= \frac{|H_i^w \phi_m|^2}{\frac{M}{P\gamma} + \sum_{k \ne m} |H_i^w \phi_k|^2}, m = 1, \dots, M \tag{19}$$

where $H_i^w = H_i R^{-1/2}$ has covariance of I and therefore has i.i.d. Gaussian entries with zero mean and unit variance. Therefore we can apply the random beamforming result of [1] to obtain the sum rate of random beamforming with channel whitening. This is summarized in the following Theorem.

Theorem 2: Consider a Gaussian broadcast channel with a channel covariance matrix R defined in Lemma 1. Let there be one transmitter with M antennas and n users with single antennas that have access to the CSI. If the transmitter knows the channel autocorrelation perfectly, then the sum rate capacity for random beam forming with channel whitening (denoted by R_{BF-W}) is given by

$$R_{BF-W} = M \log\log n + M \log \frac{P}{M} - M \log \frac{tr(R^{-1})}{M} + o(1) \tag{20}$$

for sufficiently large n.

When the the channel is i.i.d, Theorem 2 reduces to the already known result of [1]. It is also worth mentioning that (20) is less than the sum-rate achieved by DPC in (12).

B. Sum-Rate of Random Beamforming

In this section, we study the effect of transmit correlation on random beam-forming. To do this, we need to derive the CDF and pdf of the SINR defined in (9).

The sum rate capacity of random beamforming is given by (10). Now consider the expectation in (10). The averaging here is done over H_i and Φ in the following order,

$$E \log\left(1 + \max_{1 \leq i \leq n} \text{SINR}_{i,m}\right) =$$

$$E_\Phi \left\{ E_{H_i's|\Phi} \log\left(1 + \max_{1 \leq i \leq n} \text{SINR}_{i,m}\right) | \Phi \right\} \tag{21}$$

i.e., we evaluate the expectation by first conditioning on Φ and calculating the expectation over H_i and we subsequently average over Φ. The advantage of doing so is that Φ is common among all users and so, by conditioning over Φ, all the SINR's, $\text{SINR}_{1,m}, \ldots, \text{SINR}_{n,m}$ remain iid. This in turn allows us to evaluate $\max_{1 \leq i \leq n} \text{SINR}_{i,m}$ using extreme value theory provided we can evaluate the CDF (and pdf) of the SINR.

It turns out that the main challenge lies in calculating the CDF. When the channel is iid, calculating the CDF is straightforward as the SINR numerator and denominator are independent [1]. This ceases to be the case in the presence of correlation and in evaluating the CDF, we use a contour integral representation of the unit step and find the CDF using the Gaussian integral. Once the CDF is available, we appeal to results in extreme value theory to obtain the behavior of $\max_{1 \leq i \leq n} \text{SINR}_{i,m}$ when n is large and proceed to calculate the expectation in (21)

With the scaling law for random beamforming at hand, it becomes straightforward to obtain the scaling laws of random beamforming with precoding and of deterministic beamforming.

1) Distribution of $\text{SINR}_{i,1}$ *Given* Φ *:* We first obtain the complementary CDF of $\text{SINR}_{i,m}$ defined in (9) by defining the auxiliary variable S as

$$S = -\frac{x}{\rho} + H_i((1+x)\phi_m\phi_m^* - xI)H_i^* \tag{22}$$

Here $\rho = \frac{P}{M}$ just to simplify the notation and where the beamforming matrix Φ is given and H_i is an $1 \times M$ vector with Gaussian entries and with covariance matrix R. We can write the probability that $\text{SINR}_{i,m} > x$ as,

$$P(\text{SINR}_{i,1} > x) = P(S > 0) \quad = \quad \int_{-\infty}^{\infty} P(H_i)u(S)dH_i$$

$$= \frac{1}{\pi^M \det(R)} \int_{-\infty}^{\infty} e^{-H_i R^{-1} H_i^*} u(S) dH_i \tag{23}$$

where $u(S)$ is the unit-step function. To evaluate $P(S > 0)$, we can view S as a weighted sum of correlated Gaussian random variables and employ one of various techniques that have been suggested in the literature. Unfortunately, the expressions we get involve recursions and infinite sums and hence don't lend themselves to further mathematical manipulations. Instead, we use the following representation of the unit step function [4]

$$u(S) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{(j\omega+\beta)S}}{j\omega+\beta} d\omega \tag{24}$$

which is valid for any $\beta > 0$. This frees (23) from the constraint on S and, as we shall see, allows us to compute (23) in closed form.

Using (24), we can express (23) as

$$P(S > 0) = \frac{1}{2\pi^{M+1}\det(R)} \int_{-\infty}^{\infty} d\omega \frac{1}{j\omega+\beta}$$
$$\int_{-\infty}^{\infty} dH_i e^{(j\omega+\beta)S - H_i R^{-1} H_i^*}$$

Using the definition of S in (22), we get

$$P(S > 0) = \frac{1}{2\pi\det(R)} \int_{-\infty}^{\infty} d\omega \frac{e^{-(j\omega+\beta)\frac{x}{\rho}}}{j\omega+\beta}$$
$$\int_{-\infty}^{\infty} dH_i e^{-H_i \tilde{R} H_i^*} \tag{25}$$

$$= \frac{1}{2\pi\det(R)} \int_{-\infty}^{\infty} d\omega \frac{e^{-(j\omega+\beta)\frac{x}{\rho}}}{j\omega+\beta} \frac{1}{\det(\tilde{R})} \tag{26}$$

where

$$\tilde{R} = R^{-1} + x(j\omega+\beta)I - (1+x)(j\omega+\beta)\phi_m\phi_m^* \tag{27}$$

a) Evaluating the roots of $\tilde{R}$*:* Now to evaluate the integral with respect to ω, we need to find the roots of $\det(\tilde{R})$ with respect to ω. To this end, note that

$$\det(\tilde{R}) = \det(U^*\Lambda^{-1}U + (j\omega+\beta)(xI - (1+x)\phi_m\phi_m^*))$$
$$= \det(\Lambda^{-1} + (j\omega+\beta)(xI - (1+x)\overline{\phi}_m\overline{\phi}_m^*)) \tag{28}$$
$$= \det(\Lambda^{-1})\det(-A)\det((j\omega+\beta)I - A^{-1}) \tag{29}$$

where $U^*\Lambda^{-1}U$ represents the eigenvalue decomposition of R^{-1}, $\overline{\phi}_m \triangleq U\phi_m$, and

$$A = (1+x)\Lambda^{1/2}\overline{\phi}_m\overline{\phi}_m^*\Lambda^{1/2} - x\Lambda \tag{30}$$

Now

$$\begin{aligned}
\det(\Lambda^{-1})\det(-A) &= \det(xI - (1+x)\overline{\phi}_m\overline{\phi}_m^*) \quad (31)\\
&= x^{M-1}(x-(1+x)) \quad (32)\\
&= -x^{M-1} \quad (33)
\end{aligned}$$

because $xI - (1+x)\overline{\phi}_m\overline{\phi}_m^*$ has x as an eigenvalue with multiplicity $M-1$ and an eigenvalue at $x - (1+x)\|\overline{\phi}_m\|^2 = -1$. We can thus write

$$\det(\tilde{R}) = -x^{M-1}\det((j\omega+\beta)I - A^{-1})$$

Now consider the equation

$$\det((j\omega+\beta)I - A^{-1}) = 0 \quad (34)$$

The roots of this equation, with respect to $j\omega+\beta$, are $1/\lambda_i(A)$ where $\lambda_i(A)$ is an eigenvalue of the matrix A. Since A is Hermitian and nonsingular, these eigenvalues are real and nonzero. To find these eigenvalues, decompose A as

$$A = A_1 + A_2$$

where

$$A_1 = (1+x)\Lambda^{1/2}\overline{\phi}_m\overline{\phi}_m^*\Lambda^{1/2} \quad \text{and} \quad A_2 = -x\Lambda$$

The matrix A_1 has only one nonzero eigenvalue, $(1+x)\overline{\phi}_m^*\Lambda\overline{\phi}_m$. The eigenvalues of A_2 are

$$-x\lambda_M(\Lambda) \leq -x\lambda_{M-1}(\Lambda) \leq \cdots \leq -x\lambda_1(\Lambda)$$

where $\lambda_1(\Lambda) \leq \lambda_2(\Lambda) \leq \cdots \leq \lambda_M(\Lambda)$ are the diagonal elements of Λ (ordered) [6]. The second largest eigenvalue of A thus satisfies [13]

$$\begin{aligned}
\lambda_{M-1}(A) &\leq \begin{cases} \lambda_{M-1}(A_1) + \lambda_M(A_2) \\ \lambda_M(A_1) + \lambda_{M-1}(A_2) \end{cases} \quad (35)\\
&= \begin{cases} 0 - x\lambda_1 \\ \overline{\phi}_m^*\Lambda\overline{\phi}_m - x\lambda_2 \end{cases} \quad (36)
\end{aligned}$$

This means that $\lambda_{M-1}(A) \leq -x\lambda_1 < 0$. So the second largest eigenvalue is negative. The largest eigenvalue, however, is positive (otherwise A would be negative definite or singular, neither of which is the case). This means that (34) has exactly one positive root

$$\lambda = \frac{1}{\lambda_M(A)}$$

Henceforth, we drop the dependence upon the matrix A as it is understood. From above, we can express $\tilde{R}$ as

$$\boxed{\det(\tilde{R}) = -x^{M-1}((j\omega+\beta) - \tfrac{1}{\lambda_M})\prod_{i=1}^{M-1}((j\omega+\beta) - \tfrac{1}{\lambda_i})}$$

[6]In general, the M eigenvalues of a size M matrix K are written as $\lambda_1(K) \leq \lambda_2(K) \leq \cdots \leq \lambda_M(K)$. We will drop the dependence on K for notational convenience whenever it is understood.

b) Deriving the CDF of SINR: With the above factorization of $\det(\tilde{R})$, we can proceed to evaluate the probability $P(\lambda > 0)$ in (26) and hence the CDF of the SINR can be written as,

$$P(S > 0) = -\frac{1}{x^{M-1}}\frac{1}{2\pi\det(R)}$$

$$\int \frac{e^{-(j\omega+\beta)\frac{x}{\rho}}}{(j\omega+\beta)((j\omega+\beta) - \frac{1}{\lambda_M})\prod_{i=1}^{M-1}((j\omega+\beta) - \frac{1}{\lambda_i})}d\omega \quad (37)$$

Using partial fraction expansion, we can write

$$\frac{1}{(j\omega+\beta)(j\omega+\beta - \frac{1}{\lambda_M})\prod_{i=1}^{M-1}(j\omega+\beta - \frac{1}{\lambda_i})} =$$
$$\frac{\zeta_M}{j\omega+\beta - \frac{1}{\lambda_M}} + \sum_{i=1}^{M-1}\frac{\zeta_i}{j\omega+\beta - \frac{1}{\lambda_i}} + \frac{\zeta_0}{j\omega+\beta} \quad (38)$$

The term $\frac{\zeta_M}{j\omega+\beta - \frac{1}{\lambda_M}}$ is the only one that contributes to the integral in (37) (the other terms integrate to zero since the poles are outside the contour of integration), and so we only need to calculate ζ_M

$$\begin{aligned}
\zeta_M &= \left.\frac{1}{(j\omega+\beta)\prod_{i=1}^{M-1}(j\omega+\beta - \frac{1}{\lambda_i})}\right|_{j\omega+\beta = \frac{1}{\lambda_M}} \quad (39)\\
&= \frac{1}{\frac{1}{\lambda_M}\prod_{i=1}^{M-1}(\frac{1}{\lambda_M} - \frac{1}{\lambda_i})} \quad (40)
\end{aligned}$$

and

$$\begin{aligned}
P(S > 0) &= \frac{1}{2\pi\det(R)}\frac{1}{x^{M-1}}\int \frac{\zeta_M e^{-(j\omega+\beta)\frac{x}{\rho}}}{\frac{1}{\lambda_M} - (j\omega+\beta)}d\omega\\
&= \frac{1}{\det(R)}\frac{\zeta_M}{x^{M-1}}e^{-\frac{1}{\rho}\frac{x}{\lambda_M}} \quad (41)
\end{aligned}$$

This represents the probability $P(\mathrm{SINR}_{i,m} > x)$. Thus, the CDF of the SINR is given by

$$F(x) = 1 - \frac{1}{\det(R)}\frac{\zeta_M}{x^{M-1}}e^{-\frac{1}{\rho}\frac{x}{\lambda_M}}$$

Or, upon replacing ζ_M by its value obtained in (40),

$$\boxed{F(x) = 1 - \frac{1}{\det(R)}\lambda_M\prod_{i=1}^{M-1}\frac{\lambda_i\lambda_M}{x(\lambda_i - \lambda_M)}e^{-\frac{1}{\rho}\frac{x}{\lambda_M}}} \quad (42)$$

We would like to emphasize that the eigenvalues of A, λ_i, are functions of x.

2) Probability Density Function of SINR: To find the pdf of the SINR, we simply evaluate the derivative $\frac{dF(x)}{dx}$. To do this, we first need to find the derivative of the eigenvalues $\frac{d\lambda_i}{dx}$. So let q_i be the eigenvector associated with λ_i. Then, we can write

$$\begin{aligned}
\lambda_i &= \|q_i\|_A^2\\
&= q_i^*\Lambda^{1/2}(\overline{\phi}_m\overline{\phi}_m^* - x\sum_{k\neq m}\overline{\phi}_k\overline{\phi}_k^*)\Lambda^{1/2}q_i
\end{aligned}$$

where we used the notation $\|q_i\|_A^2 = q_i^*Aq_i$. We can use this to show that

$$\frac{d\lambda_i}{dx} = \|q_i\|_B^2 \quad (43)$$

where $B = \Lambda^{1/2}(\overline{\phi}_m\overline{\phi}_m^* - I)\Lambda^{1/2}$. We can in turn use this result to show that

$$\frac{d}{dx}\left(\frac{\lambda_i\lambda_M}{x(\lambda_i - \lambda_M)}\right) = \frac{\lambda_M^2\|q_i\|_C^2 - \lambda_i^2\|q_M\|_C^2}{x^2(\lambda_i - \lambda_M)^2} \quad (44)$$

where $C = \Lambda^{1/2}\overline{\phi}_m\overline{\phi}_m^*\Lambda^{1/2}$. From (42)–(44), we can show that the SINR pdf is given by

$$f(x) = \frac{1}{\det(R)}e^{-\frac{1}{\rho}\frac{x}{\lambda_M}}\prod_{i=1}^{M-1}\frac{\lambda_i\lambda_M}{x(\lambda_i - \lambda_M)}$$

$$\left\{\frac{1}{\rho}\frac{\|q_M\|_C^2}{\lambda_M} - \|q_M\|_B^2 - \sum_{i=1}^{M}\frac{1}{\lambda_i}\frac{\lambda_M^2\|q_i\|_C^2 - \lambda_i^2\|q_M\|_C^2}{x(\lambda_i - \lambda_M)}\right\}$$
$$(45)$$

3) Scaling Law of the Maximum SINR:

Lemma 2: Let $F(x)$ denote the CDF of $\text{SINR}_{i,m}$ given by (42) and let $f(x)$ denote the associated pdf (given by (45)). Then

$$\lim_{x\to\infty}\frac{1 - F(x)}{f(x)} = \frac{\rho}{\|\overline{\phi}_m\|_{\Lambda^{-1}}^2}$$

Proof: See Appendix A for the proof. ∎

Note that in the absence of spatial correlation, $\Lambda = I$, and the above limit reduces to

$$\lim_{x\to\infty}\frac{1 - F(x)}{f(x)} = \frac{\rho}{\|\overline{\phi}_m\|^2} = \rho$$

which is the scaling obtained in [1].

Using extreme value theory [17]–[19], and the lemma above, we know that $\max_{1\le i\le n}\text{SINR}_{i,m}$ behaves like $\frac{\rho}{\|\overline{\phi}_m\|_{\Lambda^{-1}}^2}\log n$. Upon substituting this in (21) and noting that the $\overline{\phi}$'s are identically distributed, we can write

$$R_{RBF} =$$

$$\sum_{m=1}^{M}E_{\phi_m}\log\left(1 + \frac{P}{M\|\overline{\phi}_m\|_{\Lambda^{-1}}^2}\log n + o(\log\log n)\right) + o(1)$$

$$= \sum_{m=1}^{M}E_{\phi_m}\log\left(\frac{P}{M\|\overline{\phi}_m\|_{\Lambda^{-1}}^2}\log n\right) + o(1) \quad (46)$$

$$= M\log\log n + M\log\frac{P}{M} + ME_{\phi_m}\log\left(\frac{1}{\|\overline{\phi}_m\|_{\Lambda^{-1}}^2}\right)$$
$$+ o(1). \quad (47)$$

It thus remains to calculate the expectation in (47) for which we need to derive the CDF of $\frac{1}{\|\overline{\phi}_m\|_{\Lambda^{-1}}^2}$.

4) Calculating the CDF of $\frac{1}{\|\phi\|_{\Lambda^{-1}}^2}$:

Lemma 3: The CDF of $y = \frac{1}{\|\phi\|_{\Lambda^{-1}}^2}$ is given by

$$G(x) = Pr\left(\frac{1}{\|\phi\|_{\Lambda^{-1}}^2} < x\right) = 1 - \sum_i \eta_i\left(\frac{1}{x} - \frac{1}{\lambda_i(\Lambda)}\right)^{M-1}u\left(1 - \frac{x}{\lambda_i(\Lambda)}\right)$$

where $\eta_i = \frac{1}{\prod_{j\ne i}\left(\frac{1}{\lambda_j(\Lambda)} - \frac{1}{\lambda_i(\Lambda)}\right)}$.

Proof: See Appendix B for the proof. ∎

5) Calculating the sum-rate: Now all we need to do to calculate the sum-rate in (47) is to compute $E\log(\frac{1}{\|\phi\|_{\Lambda^{-1}}^2})$ where the distribution of $\frac{1}{\|\phi\|_{\Lambda^{-1}}^2}$ is given in Lemma 3. We employ integration by parts and use the CDF to calculate the expectation as follows

$$E\left(\log(\frac{1}{\|\overline{\phi}\|_{\Lambda^{-1}}^2})\right) = G(y)\log(y)|_{\lambda_1(\Lambda)}^{\lambda_M(\Lambda)}$$

$$- \int_{\lambda_1(\Lambda)}^{\lambda_M(\Lambda)}G(y)\frac{1}{y}dy$$

$$= G(\lambda_M(\Lambda))\log(\lambda_M(\Lambda)) - \int_{\lambda_1(\Lambda)}^{\lambda_M(\Lambda)}G(y)\frac{1}{y}dy$$

$$= \log(\lambda_M(\Lambda)) - \int_{\lambda_1(\Lambda)}^{\lambda_M(\Lambda)}\frac{1}{y} + \sum_{i=1}^{M}\eta_i$$

$$\int_{\lambda_1(\Lambda)}^{\lambda_i(\Lambda)}(\frac{1}{y} - \frac{1}{\lambda_i})^{M-1}\frac{1}{y}dy$$

$$= \log(\lambda_1(\Lambda)) + \sum_{i=1}^{M}\eta_i\int_{\lambda_1(\Lambda)}^{\lambda_i(\Lambda)}(\frac{1}{y} - \frac{1}{\lambda_i})^{M-1}\frac{1}{y}dy$$

$$= \log(\lambda_1(\Lambda)) + \sum_{i=1}^{M}\eta_i\sum_{k=0}^{M-1}\int_{\lambda_1(\Lambda)}^{\lambda_i(\Lambda)}\frac{1}{y^{k+1}}$$

$$(\frac{-1}{\lambda_i})^{M-1-k}dy$$

$$= \log(\lambda_1(\Lambda)) + \sum_{i=1}^{M}\eta_i\log(\frac{\lambda_i}{\lambda_1})\sum_{k=1}^{M-1}$$

$$\int_{\lambda_1(\Lambda)}^{\lambda_i(\Lambda)}\frac{1}{y^{k+1}}(\frac{-1}{\lambda_i})^{M-1-k}dy$$

$$= \log(\lambda_1(\Lambda)) + \sum_{i=1}^{M}\eta_i\log(\frac{\lambda_i}{\lambda_1})\sum_{k=1}^{M-1}\frac{1}{k+2}(\frac{-1}{\lambda_i})^{M-1-k}$$

$$\frac{1}{y^{k+2}}\Big|_{\lambda_1(\Lambda)}^{\lambda_i(\Lambda)}$$

Therefore the sum-rate of beamforming can be written as,

$$R_{RBF} = M\log\log n + M\log\frac{P}{M} + M\log\lambda_1(\Lambda) +$$

$$M\sum_{i=1}^{M}\eta_i\log\left(\frac{\lambda_i}{\lambda_1}\right)\sum_{k=1}^{M-1}\frac{1}{k+2}(\frac{-1}{\lambda_i})^{M-1-k}$$

$$\left\{\frac{1}{(\lambda_i(\Lambda))^{k+2}} - \frac{1}{(\lambda_1(\Lambda))^{k+2}}\right\} + o(1). \quad (48)$$

C. Sum-Rate of Deterministic Beamforming

Here we consider the case where the beamforming matrix Φ is fixed over all channel uses. In this case, we can use the same analysis as we done in the case of random beamforming with the only exception that we do not need to take expectation over the beamforming matrix. Therefore, we may write the sum-rate for the deterministic beamforming matrix Φ as,

$$R_{BF-D} =$$

$$M \log \log n + M \log \frac{P}{M} + \sum_{i=1}^{M} \log \left(\frac{1}{\phi_i^* U^* \Lambda^{-1} U \phi_i} \right) + o(1) \tag{49}$$

where $U^* \Lambda^{-1} U$ is the eigenvalue decomposition of the correlation matrix R^{-1}.

One interesting special case would be the case where the $U\phi_i$'s are the columns of the identity matrix. In this case, the beamforming matrix is in fact equal to U^* and the argument in the logarithm would therefore reduce to λ_m. Thus, when n is large, the sum-rate is given by

$$R_{BF-D} = M \log \log n + M \log \frac{P}{M} + M \log \sqrt[M]{\det R} + o(1). \tag{50}$$

Keeping in mind that the eigenvalues of Λ are such that $\sum_{i=1}^{M} \lambda_i(\Lambda) = M$, it is clear that the geometric mean of λ_i's would be less than 1. Eq. (50) in fact proves Lemma 1. It should be also mentioned that this result is obtained in [14] for $M = 2$.

D. Sum-Rate of Random Beamforming with Precoding

We can consider a generalization of the random beamforming by using precoding. In this scheme the new beamforming matrix is $\sqrt{\gamma} A^{-1/2} \Phi$ where A is a positive definite matrix and γ is just a normalization factor to adjust the transmit power. Again similar to Section V-B, we can state that γ has to be less than $\frac{M}{tr(A^{-1})}$.

In order to analyze the sum-rate, we can follow along the same line as what we did for the analysis of the random beamforming with the only exception that the covariance matrix of the channel is replaced with $\mathcal{R} = A^{-*/2} R A^{-1/2}$. Therefore the same result holds for this case with the new covariance matrix $\tilde{R}$. Here is the main result.

Corollary 1: Considering the random beamforming scheduling with beamforming matrix $\sqrt{\gamma} A^{-1/2} \Phi$ where Φ is a random unitary matrix, the sum-rate of this scheme can be written as

$$R_{BF-Prec} = M \log \log n + M \log \frac{P}{M} +$$

$$\sum_{i=1}^{M} E \log \left(1 + \frac{tr(\Lambda^{-1})}{M} \frac{P}{\phi_i^* U^* \Lambda^{-1} U \phi_i} \right) + o(1). \tag{51}$$

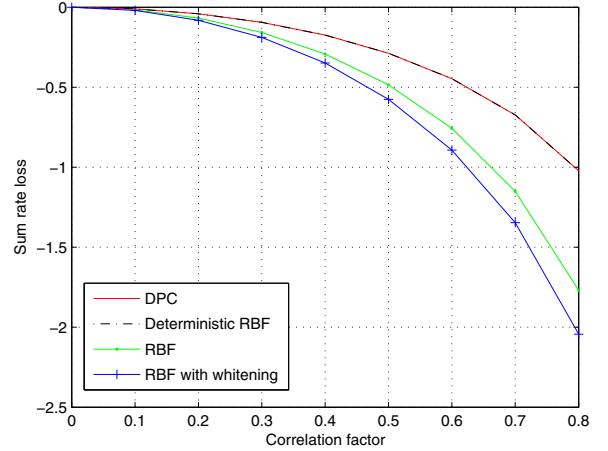for large n, where $U^* \Lambda^{-1} U$ represents the eigenvalue decomposition of $\mathcal{R}^{-1}$.

Fig. 1.   Sum-rate loss versus the correlation factor α for a system with $M = 2$ and $n = 100$.

VI. SIMULATION RESULTS

In this section we present the simulation results for the sum-rate of beamforming schemes and DPC. In the first example, we consider a system with two transmit antennas, i.e., $M = 2$, and 100 users. The covariance matrix is assumed to be like

$$F = \begin{bmatrix} 1 & \alpha \\ \alpha & 1 \end{bmatrix} \tag{52}$$

where α is the correlation. Fig. 1 shows the sum-rate loss (relative to the case of no correlation) versus the correlation coefficient α for DPC, RBF and RBF with whitening. It is clear that RBF outperforms the one with channel whitening. Fig. 2 also shows the actual sum-rate for such a setting for RBF and RBF with whitening. Fig. 3 shows the sum-rate loss for the there antenna case $M = 3$ where the covariance matrix is now given by

$$F = \begin{bmatrix} 1 & \alpha & \alpha^2 \\ \alpha & 1 & \alpha \\ \alpha^2 & \alpha & 1 \end{bmatrix} \tag{53}$$

where α is changing from 0 to 0.8. In Fig. 4, we show the sum-rate versus the number of users in system with $M = 2$, $\alpha = 0.5$, $P = 10$ for beamforming scheme and it is compared to the case of having no correlation.

VII. CONCLUSION

This paper considers the effect of spatial correlation on various multiuser scheduling schemes for MIMO broadcast channels. Specifically, we considered dirty paper coding and various (random, deterministic, and channel whitening) beamforming schemes. When the channel is i.i.d. and for large number of users, the sum rate of all these techniques exhibits the same scaling, namely, as $M \log \log n + M \log \frac{P}{M} + o(1)$ where n is the number of users, M is the number of transmit antennas and P is the average SNR.

In the presence of a correlation between transmit antennas, the channel matrix has a covariance matrix R which is assumed to be non-singular and $tr(R) = M$. In this case, the sum-rate of DPC and beamforming schemes will be
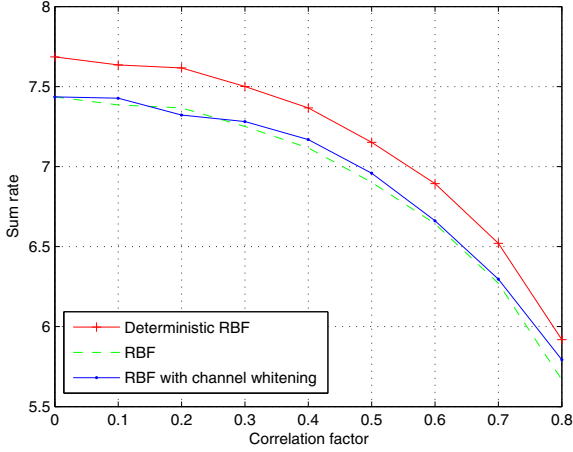
Fig. 2. Sum-rate versus the correlation factor α for a system with $M = 2$, $P = 10$, and $n = 100$.
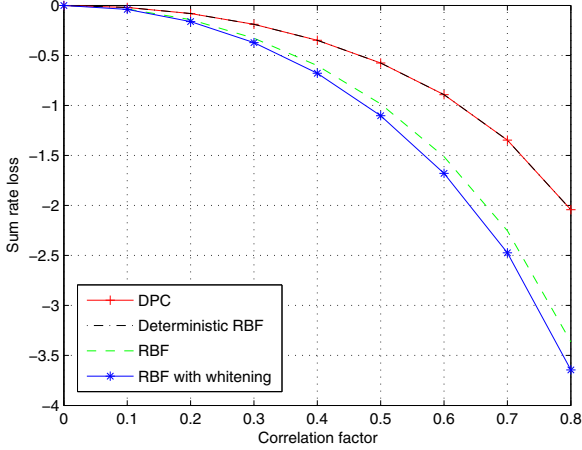

Fig. 3. Sum-rate loss versus the correlation factor α for a system with $M = 3$ and $n = 100$.
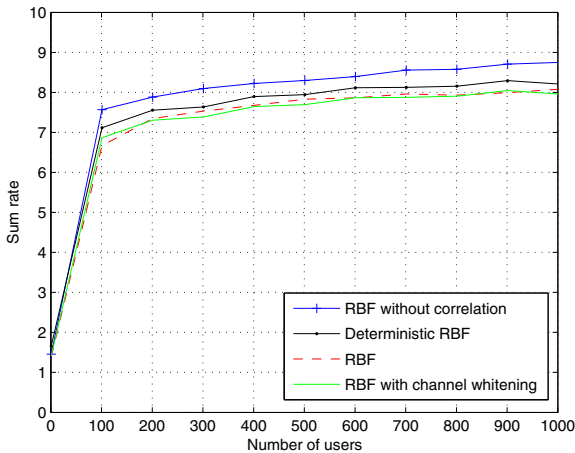

Fig. 4. Sum-rate versus the number of users in a system with $M = 2$, $P = 10$, and $\alpha = 0.5$

different. It turns out that in these case, the sum-rate can be written as $M \log \log n + M \log \frac{P}{M} + M \log c + o(1)$ where $c < 1$ is a constant that only depends on the scheduling scheme and the covariance matrix R. For DPC, c is just the geometric mean of the eigenvalues of R. We further obtain c for different beamforming schemes; For example, for the case of beamforming with channel whitening, c will be equal to the harmonic mean of the eigenvalues of R. It is worth mentioning, numerical results suggest that sum-rate of random beamforming outperforms that of the random beamforming with channel whitening [7].

APPENDIX A: PROOF OF LEMMA 2

From (42) and (45), we can write

$$\frac{1 - F(x)}{f(x)} =$$

$$\frac{\lambda_M}{\frac{1}{\rho} \frac{\|q_M\|_C^2}{\lambda_M} - \|q_M\|_B^2 - \sum_{i=1}^{M-1} \frac{1}{\lambda_i} \frac{\lambda_M^2 \|q_i\|_C^2 - \lambda_i^2 \|q_M\|_C^2}{x(\lambda_i - \lambda_M)}} \quad (54)$$

To evaluate the limit of this expression, we need to investigate the behavior of the eigenvalues and eigenvectors of A as $x \to \infty$. Now from the bound (36), we deduce that

$$\lim_{x \to \infty} \lambda_i = -\infty \quad \text{for all} \quad i \neq m$$

We now have to evaluate the behavior of the maximum of eigenvalue as x tends to infinity. This is done by using the Rayleigh quotient for the maximum eigenvalue as,

$$\lambda_M = \max_{\|u\|_2 = 1} u^* A u =$$

$$\max_{\|u\|_2 = 1} u^*(\Lambda^{1/2} \overline{\phi}_m \overline{\phi}_m^* \Lambda^{1/2} - x \Lambda^{1/2} \sum_{m \neq i} \overline{\phi}_i \overline{\phi}_i^*) u \quad (55)$$

The vector u that maximizes λ_M is the associated eigenvector. Since any vector u of dimension M can be written as $u = \sum_{i=1}^{M} \alpha_i \Lambda^{-1/2} \overline{\phi}_i$, we can write $\|Au\|_2$ as

$$\|Au\|_2 = u^* A u = u^* \left(\alpha_m \Lambda^{1/2} \overline{\phi}_m - x \sum_{i \neq m} \alpha_i \Lambda^{1/2} \overline{\phi}_i \right)$$

$$= \alpha_m^2 - x \sum_{i \neq m} \alpha_i^2 \quad (56)$$

where we used the fact that the $\overline{\phi}_i$'s are orthonormal vectors. Now as x tends to infinity, $\|Au\|_2$ could go to $-\infty$ and is maximized when $\sum_{i \neq m} \alpha_i^2$ is equal to zero (i.e., $\alpha_i = 0$ for $i \neq m$ and as a result $\alpha_m = \frac{1}{\sqrt{\|\overline{\phi}_m\|_{\Lambda^{-1}}^2}}$). We have thus proved that

$$\lim_{x \to \infty} q_M = \lim_{x \to \infty} u = \frac{\Lambda^{-1/2} \overline{\phi}_m}{\sqrt{\|\overline{\phi}_m\|_{\Lambda^{-1}}^2}} \quad (57)$$

and

$$\lim_{x \to \infty} \lambda_M = \frac{1}{\|\overline{\phi}_m\|_{\Lambda^{-1}}^2}$$

---

[7]Channel whitening is like zero forcing in that it takes care of the worst eigenvalue and thus would result in a big waste of power.

Using the above, it is easy to verify that

$$\lim_{x \to \infty} \frac{1}{\lambda_i} \frac{\lambda_M^2 \|q_i\|_C^2}{x(\lambda_i - \lambda_M)} = 0$$

and

$$\lim_{x \to \infty} -\frac{1}{\lambda_i} \frac{\lambda_i^2 \|q_M\|_C^2}{x(\lambda_i - \lambda_M)} = \lim_{x \to \infty} -\frac{\lambda_i \|q_M\|_C^2}{x(\lambda_i - \lambda_M)} = 0$$

From (57) and the defining expression of B, we also deduce that

$$\lim_{x \to \infty} \|q_M\|_B^2 = \frac{1}{\|\overline{\phi}_m\|_{\Lambda^{-1}}^2} \overline{\phi}_m^*(\overline{\phi}_m \overline{\phi}_m^* - I)\overline{\phi}_m = 0$$

Thus, the only nonzero limit in the denominator of (54) is $\frac{1}{\rho} \frac{\|q_m\|_C^2}{\lambda_M}$ and

$$\lim_{x \to \infty} \frac{1 - F(x)}{f(x)} = \frac{\lambda_M^2}{\frac{1}{\rho} \|q_m\|_C^2} = \frac{\rho}{\|\overline{\phi}_m\|_{\Lambda^{-1}}^2} \qquad (58)$$

APPENDIX B: PROOF OF LEMMA 3

Consider the inequality

$$y = \frac{1}{\|\phi\|_{\Lambda^{-1}}^2} > x$$

which can be equivalently written as $1 - x\|\phi\|_{\Lambda^{-1}}^2 > 0$. As we did to derive the SINR CDF above, we use the unit-step representation

$$u(1 - x\|\phi\|_{\Lambda^{-1}}^2) = \frac{1}{2\pi} \int \frac{e^{(1 - x\|\phi\|_{\Lambda^{-1}}^2)(j\omega_1 + \beta_1)}}{j\omega_1 + \beta_1} d\omega_1$$

Now the pdf of ϕ is

$$p(\phi) = \frac{\Gamma(M)}{\pi^M} \delta(\|\phi\|^2 - 1)$$

Alternatively, following the approach of [4], we can use an integral representation for the Dirac delta

$$p(\phi) = \frac{\Gamma(M)}{\pi^M} \frac{1}{2\pi} \int d\omega_2 e^{j\omega_2(\|\phi\|^2 - 1)}$$

So the probability $p(\frac{1}{\|\phi\|_{\Lambda^{-1}}^2} > x) = p(1 - x\|\phi\|_{\Lambda^{-1}}^2 > 0)$ is given by

$$p(\frac{1}{\|\phi\|_{\Lambda^{-1}}^2} > x) =$$

$$\frac{\Gamma(M)}{4\pi^{M+2}} \int d\omega_1 \int d\omega_2$$

$$\int d\phi \frac{e^{(j\omega_1 + \beta_1)(1 - x\|\phi\|_{\Lambda^{-1}}^2)} e^{j\omega_2(\|\phi\|^2 - 1)}}{j\omega_1 + \beta_1}$$

$$= \frac{\Gamma(M)}{4\pi^{M+2}} \int d\omega_1 \frac{e^{(j\omega_1 + \beta_1)}}{j\omega_1 + \beta_1} \int d\omega_2 e^{-j\omega_2}$$

$$\int d\phi e^{-\phi^*(x(j\omega_1 + 1)\Lambda^{-1} - j\omega_2 I)\phi}$$

$$= \frac{\Gamma(M)}{4\pi^{M+2}} \int d\omega_1 \frac{e^{(j\omega_1 + \beta_1)}}{j\omega_1 + \beta_1} \int d\omega_2 e^{-j\omega_2}$$
$$\frac{1}{\det\left(x(j\omega_1 + \beta_1)\Lambda^{-1} - j\omega_2 I\right)} \qquad (59)$$

Now use partial fraction expansion to show that

$$\frac{1}{\det\left(x(j\omega_1 + \beta_1)\Lambda^{-1} - j\omega_2 I\right)} =$$

$$\frac{1}{\prod_{i=1}^M \left(\frac{x}{\lambda_i(\Lambda)}(j\omega_1 + \beta_1) - j\omega_2\right)}$$

$$= \frac{1}{x^{M-1}} \frac{1}{(j\omega_1 + \beta_1)^{M-1}} \sum_{i=1}^M \frac{\eta_i}{\frac{x}{\lambda_i(\Lambda)}(j\omega_1 + \beta_1) - j\omega_2} \qquad (60)$$

where $\eta_i = \frac{1}{\prod_{j \neq i}\left(\frac{1}{\lambda_j(\Lambda)} - \frac{1}{\lambda_i(\Lambda)}\right)}$. We thus have

$$p(\frac{1}{\|\phi\|_{\Lambda^{-1}}^2} > x) =$$

$$\frac{\Gamma(M)}{4\pi^2} \frac{1}{x^{M-1}} \int d\omega_1 \frac{e^{j\omega_1 + \beta_1}}{(j\omega_1 + \beta_1)^M} \sum_i$$
$$\int d\omega_2 \frac{\eta_i}{\frac{x}{\lambda_i(\Lambda)}(j\omega_1 + \beta_1) - j\omega_2} e^{-j\omega_2} \qquad (61)$$

$$= \frac{\Gamma(M)}{2\pi} \frac{1}{x^{M-1}} \sum_i \eta_i \int d\omega_1 \frac{e^{(j\omega_1 + \beta_1)(1 - \frac{x}{\lambda_i(\Lambda)})}}{(j\omega_1 + \beta_1)^M} \qquad (62)$$

or after some straight-forward calculations,

$$p(\frac{1}{\|\phi\|_{\Lambda^{-1}}^2} > x) = \sum_i \eta_i \left(\frac{1}{x} - \frac{1}{\lambda_i(\Lambda)}\right)^{M-1} u\left(1 - \frac{x}{\lambda_i(\Lambda)}\right)$$

Alternatively, the CDF, $G(x) = p(\frac{1}{\|\phi\|_{\Lambda^{-1}}^2} < x)$ is given by

$$G(x) = 1 - \sum_i \eta_i \left(\frac{1}{x} - \frac{1}{\lambda_i(\Lambda)}\right)^{M-1} u\left(1 - \frac{x}{\lambda_i(\Lambda)}\right)$$

which completes the proof of the Lemma.

REFERENCES

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

[2] A. F. Dana, M. Sharif, A. Vakili, and B. Hassibi, "Differentiated rate scheduling for the down-link of cellular systems," to be published.

[3] Q. H. Spencer, C. B. Peel, A. L. Swindlehurst, and M. Haardt, "An introduction to the multi-user MIMO downlink," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 60–67, Oct. 2004.

[4] B. Hassibi and T. L. Marzetta, "Multiple-antennas and isotropically random unitary inputs: The received signal density in closed form," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1473–1484, June 2002.

[5] B. Hochwald, B. Peel, and A. L. Swindlehurst, "A Vector perturbation technique for near capacity multiantenna multiuser communication–Part II: Perturbation," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 537–544, Mar. 2005.

[6] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1691–1706, July 2003.

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

[8] H. Viswanathan and S. Venkatesan, "Asymptotics of sum rate for dirty paper coding and beamforming in multiple antenna broadcast channels," in *Proc. 41'st Annual Allerton Conf.*, 2003.

[9] B. Hochwald and S. Viswanath, "Space time multiple access: Linear growth in the sum rate," in *Proc. 40'th Annual Allerton Conf.*, 2002.

[10] N. Jindal and A. Goldsmith, "Dirty-paper coding versus TDMA for MIMO Broadcast channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 5, pp. 1783–1794, May 2005.

[11] M. Kountouris and D. Gesbert, "Robust multi-user opportunistic beamforming for sparse networks," in *Proc. IEEE Workshop Signal Processing Advances Wireless Commun., (SPAWC)*, 2005, pp. 975–979.

[12] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[13] H. Lütkepohl, *Handbook of Matrices*. New York: John Wiley & Sons, 1996.

[14] D. Park and S. Y. Park, "Effect of transmit antenna correlation on multiuser diversity," in *Proc. IEEE ISIT*, 2005, pp. 1421–1425.

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

[16] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas. Commun.*, vol. 21, no. 5, pp. 684–702, June 2003.

[17] M. R. Leadbetter, "Extreme value theory under weak mixing conditions," *Studies in Probability Theory, MAA Studies in Mathematics*, pp. 46–110, 1978.

[18] N. T. Uzgoren, "The asymptotic development of the distribution of the extreme values of a sample," *Studies in Mathematics and Mechanics Presented to Richard von Mises*. New York: Academic Press, pp. 346–353, 1954.

[19] H. A. David, *Order Statistics*. New York: Wiley, 1970.

[20] M. R. Leadbetter and H. Rootzen, "Extremal theory for stochastic processes," *The Annals of Probability*, vol. 16, pp. 431–478, 1988.

Tareq Al-Naffouri obtained his M.S. degree in electrical engineering from Georgia Institute of Technology, Atlanta, in 1998, and his Ph.D. in Electrical Engineering from Stanford University, CA in 2004. He was a visiting scholar at the California Institute of Technology, Pasadena, CA Jan–Aug, 2005 and during the summer of 2006. He is currently a Fulbright scholar at the University of Southern California. He has held internship positions at NEC Research Labs, Tokyo, Japan in 1998; Adaptive Systems Lab, University of California at Los Angeles in 1999; National Semiconductor, Santa Clara, CA, in 2001 and 2002; and Beceem Communications, Santa Clara, CA, in 2004. Dr Al- Naffouri is currently an Assistant Professor at the Electrical Engineering Department, King Fahd University of Petroleum and Minerals, Saudi Arabia. His research interests lie in the areas of adaptive and statistical signal processing and their applications to wireless communications and in multiuser information theory. He has over 50 publications in journal and conference Proceedings. Dr Al-Naffouri is the recipient of a 2001 best student paper award at the IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing (NSIP) 2001 for his work on adaptive filtering analysis.

Masoud Sharif (S'99) received his Ph.D. in Electrical Engineering (2005) from the California Institute of Technology. In 2005, he was a post-doctoral scholar in the EE department at Caltech. Since January 2006, he has been an assistant Professor at Boston University. Dr. Sharif was awarded the C. H. Wilts Prize in 2006 for best doctoral thesis in Electrical Engineering at Caltech. He is a member of the Center for Information and Systems Engineering at Boston University. His research interests include ad-hoc and sensor networks, multiple-user multiple-antenna communication channels, crosslayer design for wireless networks, and multi-user information theory. His recent research has focused on collaborative communication schemes in ad-hoc and sensor networks and the capacity of multiple antenna broadcast channels.

Babak Hassibi was born in Tehran, Iran, in 1967. He received the B.S. degree from the University of Tehran in 1989, and the M.S. and Ph.D. degrees from Stanford University in 1993 and 1996, respectively, all in electrical engineering. From October 1996 to October 1998 he was a research associate at the Information Systems Laboratory, Stanford University, and from November 1998 to December 2000 he was a Member of the Technical Staff in the Mathematical Sciences Research Center at Bell Laboratories, Murray Hill, NJ. Since January 2001 he has been with the department of electrical engineering at the California Institute of Technology, Pasadena, CA, where he is currently an associate professor. He has also held short-term appointments at Ricoh California Research Center, the Indian Institute of Science, and Linkoping University, Sweden. His research interests include wireless communications, robust estimation and control, adaptive signal processing and linear algebra.

He is the coauthor of the books *Indefinite Quadratic Estimation and Control: A Unified 1 Approach to H_2 and H_1 Theories* (New York: SIAM, 1999) and *Linear Estimation* (Englewood Cliffs, NJ: Prentice Hall, 2000). He is a recipient of an Alborz Foundation Fellowship, the 1999 O. Hugo Schuck best paper award of the American Automatic Control Council, the 2002 National Science Foundation Career Award, the 2002 Okawa Foundation Research Grant for Information and Telecommunications, the 2003 David and Lucille Packard Fellowship for Science and Engineering, and the 2003 Presidential Early Career Award for Scientists and Engineers (PECASE). He has been a Guest Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY special issue on space-time transmission, reception, coding and signal processing, was an Associate Editor for Communications of the IEEE TRANSACTIONS ON INFORMATION THEORY during 2004-2006, and is currently an Editor for the *Journal of Foundations and Trends in Information and Communication*.