The Effect of Spatial Correlation on the Capacity of Multi-Input Multi-Output Broadcast Channels with Partial Side Information

Principal Investigator: Dr. Tareq Y. Al-Naffouri

Department of Electrical Engineering
KFUPM, Dhahran 31261
Saudi Arabia
e-mail: naffouri@kfupm.edu.sa

University Project IN070342
Final Report

Contents

Abstract

This report 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 $\mathbf{R}$ (non-singular with $Tr(\mathbf{R}) = M$), we prove that the sum rate of dirty paper coding is $M \log \log n + M \log \frac{P}{M} + \log \det(\mathbf{R}) + o(1)$. 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 [26] and more generally, for random beamforming with precoding in which beams are pre-multiplied by a fixed matrix. In the second part of the report, we introduce various precoding matrices to reduce the hit that results from correlation. We obtain the optimum precoding matrix and various suboptimal precoding matrices. Our theoretical results are confirmed in both parts by simulations.

Chapter 1

Introduction to Broadcast Channels and Random Beamforming

1.1 Introduction

Future breakthroughs in wireless communications will be mostly driven by applications that require high data rates [6]. While increasing the link budget and/or bandwidth can accommodate this increase in data rate, such a solution would not be economical. A more cost effective solution is to exploit the space dimension by employing multiple antennas at the transmitter and receiver. Multiple input multiple output (MIMO) communication has thus 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 (an excellent overview of the research on this problem can be found in [28]).

1.1.1 Multiuser Communications, Broadcast Channels, and Dirty Paper Coding

Research focus has shifted recently to the role of multiple antennas in multiuser systems, especially broadcast scenarios (i.e., *point to multipoint* communication) as downlink scheduling is the major bottleneck for future broadband wireless networks. The broadcast channel resembles downlink communication in a cellular system, where the base station is to transmit to a group of users (other

applications with a similar broadcast scenario include the downlink of a DSL link and WLAN where the access point is to transmit to several laptops).

In these scenarios, one is usually interested in 1) quantifying the maximum possible rate of a transmitter when multiple users are present and in 2) devising techniques for achieving these rates [37]. These questions have been settled recently by using a technique similar to writing on dirty paper and hence known as dirty paper coding (DPC). While the dirty paper coding (DPC) solves the broadcast problem optimally, it is too computationally expensive and requires too much feedback as the transmitter needs the channel state information for all intended users.

Researchers have thus attempted to find simpler techniques to achieve performance close to (DPC) capacity. Of particular importance are techniques that exploit multiuser diversity to increase spectral efficiency in wireless networks. Specifically, in a network with a large number of users whose channels fade independently, the transmitter can choose the subset of users with the highest fade margin, thus maximizing the sum-rate capacity. Several schemes that make use of this phenomenon have been suggested in literature, e.g., [25], [27].

One (multiuser diversity) technique that has attracted a lot of attention recently is based on random beamforming [27]. Specifically, a transmitter equipped with M antennas sends M random beams. Each user would calculate the M SINR's (one for each beam) and feedback the maximum SINR along with its index. The transmitter would in turn rank the users according to their SINR's and transmits to the M best ones. Not only does this method require much less feedback than the DPC approach, but it also asymptotically (i.e., in the presence of large number of users) achieves the same performance [27].

The purpose of this project is to study the effect of transmit antenna (spatial) correlation on various multiuser schemes for multiuser broadcast access. Specifically, we investigate the effect of spatial correlation on the asymptotic capacity[1] of DPC and random beamforming, as well as other multiaccess techniques such as deterministic beamforming and channel inversion. The project also investigates the use of various precoding techniques to maximize the system's sum-rate.

---

[1]By asymptotic capacity we mean the capacity as we increase the number of users.

This report is organized as follows. This chapter introduces the system model and also performs a literature review of various methods for multiuser broadcast channels, including DPC and random beamforming. Chapter 2 studies the effect of spatial correlation on the sum-rate of DPC and random beamforming and its variants. Chapter 3 studies the effect of various precoding techniques on the sum-rate of RBF and determines the precoding that maximizes the sum-rate. We conclude the report in Chapter 4 which provides our recommendations and future work.

1.2   System Model

In this project, we consider a multi-antenna Gaussian broadcast channel. Specifically, we have one transmitter (base station) with M antennas and n users (receivers) each equipped with one antenna. Since in a typical cellular system, the number of users is much larger than the number of transmit and receive antennas, we will assume that $n \gg M$ throughout the report.

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

$$Y_i(t) = \sqrt{\rho_i}\mathbf{H}_i S(t) + W_i, \qquad\qquad i = 1, \ldots, n, \tag{1.1}$$

where W_i is the additive noise which is complex Gaussian with zero mean and unit variance, $CN(0,1)$.

The channel $\mathbf{H}_i$ is a $1 \times M$ complex channel vector, known perfectly to the receiver, and distributed as $CN(0, \mathbf{R})$. The $M \times M$ covariance matrix $\mathbf{R}$ is a measure of the spatial correlation and is assumed to be non-singular with $Tr(\mathbf{R}) = M$ [2]

We also assume that $\mathbf{H}_i$ 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.

---

[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 $\mathbf{R}$ is rank deficient, the results of this report apply with M replaced by the rank of the autocorrelation matrix and with the SNR kept fixed at P/M.

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

Remarks

1. Limiting the study to the Gaussian channel case is a common practice in literature. See for example the overview [28] which limits the discussion almost exclusively to the Gaussian case. Nevertheless, the Gaussian assumption makes the problem much more tractable and allows us to get a hint of how systems behave in the non-Gaussian case.

2. We assume that the channel to exhibit some spatial correlation. One might argue that the given the high frequencies that are in use today, such correlation is negligible. While this might be true, one also needs to keep in mind that correlation results from the presence of local scatterers around the base station which introduces correlation regardless of the frequency range.

1.3 Literature Review

1.3.1 Dirty Paper Coding

The capacity of point to point multi-antenna systems has been investigated with different assumption for the channel state information (CSI) (i.e., whether the receiver/transmitter knows the channel or not). As it is shown in [2, 3] if the receiver knows the channel perfectly, the capacity scales like $M \log \rho$ no matter the transmitter knows the channel or not.

While the full CSI in the transmitter does not seem to be beneficial in the point to point communication, the knowledge of the channel is crucial in *point to mutlipoint* broadcast channels [6, 28]. For the case with full CSI available at both the transmitter and the receivers, it is shown that the sum rate capacity[3] of the Gaussian broadcast channel can be achieved by using dirty paper coding [8, 18, 19]. Intuitively, one can write on a dirty paper by choosing the link in accordance with the

---

[3]The sum rate capacity is the maximum achievable throughput of the system, which corresponds to the sum of all users information rates.

type of dirt present. Similarly, if the transmitter knows the channels of all users, it knows the inter-ference signal felt by each user and so can design the transmitted signal accordingly. More precisely, the sum rate capacity, $\mathbf{R}^{DP}$, can be written as,

$$\mathbf{R}^{DP} = E\left\{ \max_{\{P_1,\dots,P_n,\sum P_i = M\rho\}} \log\det\left(1 + \sum_{i=1}^{n} \mathbf{H}_i^* P_i H_i\right)\right\} \qquad (1.2)$$

where $\mathbf{H}_i$ is $1 \times M$ channel matrix of receiver i and $M\rho$ is the total average power.

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

$$\mathbf{R}_{DPC} = M \log\log n + M \log \frac{P}{M} + o(1), \qquad (1.3)$$

when there is no spatial correlation, i.e., $\mathbf{R} = I$ [26]. Scaling of the sum rate capacity has also been investigated for other regions of n, M, and P (see [13, 10, 12] for details).

1.3.2 Shortcomings of DPC

There are two major drawbacks to the DPC approach. First, it is too computationally complex. The second problem is that 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). Moreover, and while (suboptimal) computationally less intensive techniques exist (e.g. channel inversion and Tomlinson-Harashima precoding [22, 23, 24]), these methods require full channel knowledge at the transmitter, just like DPC.

Research has thus focused on devising algorithms for multiuser broadcast channels which attain performance similar to DPC while avoiding its pitfalls. Towards this end, and to gain a better understanding of what is optimally possible, Sharif and Hassibi investigated the scaling laws of DPC in the presence of large number of users. They showed in [26] that for large number of users n, the DPC capacity scales like $M \log\log n$. [4]

---

[4]The term $M \log\log n$ is never negative because we are dealing here with high values of n.

1.3.3 Multiuser Diversity as an Alternative to DPC

The most distinct difficulty associated with a wireless network is the fading or time-variant nature of the communication channel. Many diversity techniques have been proposed to maximize the reliability of a single user (point-to-point) channel [28], [18]. Multiuser diversity provides a new dimension to play with. It stems from the fact that users' channels are independent as each user is located at a random position in the cell [6]. Serving the best user at each time instant guarantees using channels at fading peak levels rather than average level [3]. Many multiuser-based diversity techniques have been proposed in literature including [25]–[29]. One of the most promising techniques is the random beamforming technique [26] which we introduce next.

1.4 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 [25] and zero forcing [9] (see also [16, 7]). A random beamforming scheme was proposed in [26] 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

$$S(t) = \sum_{m=1}^{M} \phi_m(t) s_m(t), \qquad t = 1, \ldots, T \qquad (1.4)$$

After T channel uses, the transmitter independently chooses another set of orthogonal vectors $\{\phi_m\}$ and constructs the signal vector (according to (1.4)) 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

$$Y_i = \sqrt{P}\mathbf{H}_i S + W_i \tag{1.5}$$

$$= \sqrt{P} \sum_{m=1}^{M} \mathbf{H}_i \phi_m s_m + W_i, \qquad i = 1, \ldots, n \tag{1.6}$$

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 $\mathbf{H}_i \phi_m$, something that can be arranged by training, to calculate M SINR's, one for each transmitted beam

$$\text{SINR}_{i,m} = \frac{|\mathbf{H}_i \phi_m|^2}{\frac{M}{P} + \sum_{k \neq m} |\mathbf{H}_i \phi_k|^2}, \qquad m = 1, \ldots, M. \tag{1.7}$$

Each receiver then feeds back its maximum SINR, i.e. $\max\limits_{1 \leq m \leq M} \text{SINR}_{i,m}$, along with the maximizing index m. Thereafter, the transmitter assigns s_m to the user with the highest corresponding SINR, i.e. $\max\limits_{1 \leq i \leq n} \text{SINR}_{i,m}$. If we do the above scheduling, the throughput for large n can be written as [33] [5],

$$\mathbf{R}_{RBF} = ME \log \left(1 + \max_{1 \leq i \leq n} \text{SINR}_{i,m} \right) + o(1) \tag{1.8}$$

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 [26].

To further quantify (1.8), [26] used the fact that the $\text{SINR}_{i,m}$'s are iid over i and employed extreme value theory [36] to argue that $\max\limits_{i \leq n} \text{SINR}_{i,m}$ behaves like $\frac{P}{M} \log n$ and hence concluded that the sum rate capacity scales as in (1.3), meaning that the sum-rate of random beamforming behaves the same as that of DPC for large number of users.

1.4.1 Other Beamforming Schemes

The scaling result (1.3) applies for iid channels. As such, we derive in Section 2.2 the scaling law of this scheme for correlated channels. Alternatively, given this correlation, we consider the following beamforming schemes.

---

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

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[6], we would use $\sqrt{\alpha}\mathbf{R}^{-1/2}\Phi$ 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 2.2.1).

Random beamforming with general precoding More generally, we can precode with a general matrix $\sqrt{\alpha}\mathbf{A}^{-1/2}$ before beamforming, i.e. we use $\sqrt{\alpha}\mathbf{A}^{-1/2}\Phi$ 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 2.3 and 2.3.4. We go one step further and show how to compute the sum-rate when the beamforming matrix is premultiplied by the full rank matrix $\mathbf{A}$.

Deterministic beamforming Finally, by fixing the beamforming matrix Φ, we obtain deterministic beamforming, a scheme analyzed by Park and Park [21] (for the two antenna case) and which we further analyze in Section 2.3.3.

1.4.2   Performance of RBF under Non-ideal Conditions

This very promising performance of random beam-forming has prompted several researchers to study the effect of various nonideal conditions on its performance. Thus, Fakherdeen [14] studied the effect of frequency correlation, Vikali [15] studied the effect of estimation error, Gesbert [16] considered the effect of time-variation, and [21] quantified the effect of spatial correlation on deterministic beam forming (using two transmit antennas only).

1.5   Project Objective

The objective of this project is to study the scaling of various broadcast techniques for large number of users in the presence of spatial correlation. Specifically, the project studies the effect of correlation

---

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

on the asymptotic sum-rate of DPC. It also studies the effect of correlation on the asymptotic rate of random beamforming and some of its variants. As an extension of the original proposal, the project designs optimum precoding for random beamforming that minimizes the hit on the sum-rate due to correlation.

Chapter 2

Effect of Correlation on DPC and Random Beamforming

This chapter quantifies the effect of correlation on DPC and random beamforming and its variants. The chapter shows that correlation results in a hit on the sum-rate that is a function of the eigenvalues of the correlation matrix. The dependence of the hit on the eigenvalues is in turn dependent on the broadcast technique used. As a by product of our development, we will come up with a new technique for evaluating the CDF of ratios of sum of squares of correlated Gaussian variables.

2.1 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 (1.2) and its behavior when n is large is given by (1.3) 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 $\mathbf{R}$ is fixed and non-singular with $Tr(\mathbf{R}) = M$.

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 [26]) with the only difference that the lower bound rather than being achieved

with random beamforming is achieved with a special type of deterministic beamforming. We first give the lower bound in the following lemma.

Lemma 2.1: Consider a Gaussian broadcast channel with a channel covariance matrix $\mathbf{R}$ which is non-singular with $Tr(\mathbf{R}) = M$. 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 $\Phi = U^*$ where U is the unitary matrix consisting of the eigenvectors of $\mathbf{R}$. Then for large n, the sum-rate of this scheduling is

$$\mathbf{R}_{BF-D} = M \log \log n + M \log \frac{P}{M} + M \log \sqrt[M]{\det(\mathbf{R})} + o(1). \tag{2.1}$$

Proof: See Section 2.3.3 for the proof.

Clearly (2.1) is a lower bound for the sum-rate capacity. In the next theorem we show that (2.1) is indeed an upper bound for the sum-rate as well.

Theorem 2.1: Consider a Gaussian broadcast channel with an autocorrelation matrix $\mathbf{R}$ defined in Lemma 2.1. 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

$$\mathbf{R}_{DPC} = M \log \log n + M \log \frac{P}{M} + M \log \sqrt[M]{\det \mathbf{R}} + o(1), \tag{2.2}$$

for large n.

Proof: Lemma 2.1 implies that the right hand side of (2.2) is achievable. All we need to prove the theorem is to show that the sum-rate of DPC can not be larger than (2.2). We use the sum rate capacity expression given in (1.2) to obtain an upper bound for the sum-rate. To this end, define $\mathbf{H}_i = \mathbf{H}_{w_i} \mathbf{R}^{\frac{1}{2}}$, where $\mathbf{H}_{w_i}$ is the gaussian channel with zero mean and variance 1 $(N(0, I))$. With this decomposition, the sum-rate capacity can be written as

$$\mathbf{R}_{DPC} = E \left\{ \max_{\{P_1,...,P_n, \sum P_i = P\}} \log \det \left(\mathbf{R}^{-1} + \sum_{i=1}^{n} \mathbf{H}_{w_i}^* P_i H_{w_i} \right) \det(\mathbf{R}) \right\} \tag{2.3}$$

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

$$\sum_{i=1}^{n} \text{Tr}\left(\mathbf{H}_{w_i}^* P_i H_{w_i}\right) \leq \max_i \text{Tr}(\mathbf{H}_{w_i}^* \mathbf{H}_{w_i}) \sum_{i=1}^{n} P_i$$
$$= \max_i \|\mathbf{H}_{w_i}\|^2 P$$

to replace the log det with an upper bound

$$\log \det \left(\mathbf{R}^{-1} + \sum_{i=1}^{n} \mathbf{H}_{w_i}^* P_i H_{w_i}\right) \leq M \log \left(\frac{1}{M}\text{Tr}(\mathbf{R}^{-1}) + \frac{1}{M}\sum_{i=1}^{n} \text{Tr}(\mathbf{H}_{w_i}^* P_i H_{w_i})\right)$$
$$\leq M \log \left(\frac{1}{M}\text{Tr}(\mathbf{R}^{-1}) + \max_i \|\mathbf{H}_{w_i}\|^2 \frac{P}{M}\right)$$

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

$$\mathbf{R}_{DPC} \leq M \log \left(\frac{Tr(\mathbf{R}^{-1})}{M} + \frac{P}{M}\log n\right) + \log \det \mathbf{R} + o(1) \tag{2.4}$$

For large n, the term $\frac{Tr(\mathbf{R}^{-1})}{M}$ is negligible compared to $\frac{P}{M}\log n$ and 2.4 simplifies to

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

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

2.2   Effect of Transmit Correlation on Random Beamforming

The deterministic beamforming scheme of Lemma 2.1 asymptotically achieves the DPC sum-rate. However it has the drawback that, unless the $\mathbf{H}_i$'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 [25, 26] 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 $\mathbf{R}^{-1/2}$ 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.

2.2.1 Random Beamforming with Channel Whitening

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

$$S(t) = \sum_{m=1}^{M} \sqrt{\alpha}\mathbf{R}^{-1/2}\phi_m(t)s_m(t) \tag{2.5}$$

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

$$
\begin{aligned}
E\{\alpha S^*\mathbf{R}^{-1}S\} &= \alpha E\{Tr(SR^{-1}S^*)\} \\
&= \alpha E\{Tr(\mathbf{R}^{-1}S^*S)\} \\
&= \alpha Tr(\mathbf{R}^{-1}E(S^*S))\} \\
&= \alpha\frac{Tr(\mathbf{R}^{-1})}{M} \tag{2.6}
\end{aligned}
$$

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

$$\text{SINR}_{i,m} = \frac{|\mathbf{H}_i R^{-1/2}\phi_m|^2}{\frac{M}{P\alpha} + \sum_{k\neq m}|\mathbf{H}_i\mathbf{R}^{-1/2}\phi_k|^2} = \frac{|\mathbf{H}_i^w\phi_m|^2}{\frac{M}{P\alpha} + \sum_{k\neq m}|\mathbf{H}_i^w\phi_k|^2}, \qquad m = 1,\ldots,M \tag{2.7}$$

where $\mathbf{H}_i^w = \mathbf{H}_i R^{-1/2}$ 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 [26] to obtain the sum rate of random beamforming with channel whitening. This is summarized in the following Theorem.

Theorem 2.2: Consider a Gaussian broadcast channel with a channel covariance matrix $\mathbf{R}$ defined in Lemma 2.1. 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 $\mathbf{R}_{BF-W}$) is given by

$$\mathbf{R}_{BF-W} = M\log\log n + M\log\frac{P}{M} - M\log\frac{Tr(\mathbf{R}^{-1})}{M} + o(1) \tag{2.8}$$

for sufficiently large n. When the the channel is i.i.d, Theorem 2.2 reduces to the already known result of [26]. It is also worth mentioning that (2.8) is less than the sum-rate achieved by DPC in (2.2).

2.3 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 (1.7). The sum rate capacity of random beamforming is given by (1.8). Now consider the expectation in (1.8). The averaging here is done over $\mathbf{H}_i$ and Φ in the following order,

$$E \log \left(1 + \max_{1 \leq i \leq n} \mathrm{SINR}_{i,m} \right) = E_\Phi \left\{ E_{\mathbf{H}'_i s | \Phi} \log \left(1 + \max_{1 \leq i \leq n} \mathrm{SINR}_{i,m} \right) | \Phi \right\} \tag{2.9}$$

i.e., we evaluate the expectation by first conditioning on Φ and calculating the expectation over $\mathbf{H}_i$ 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, $\mathrm{SINR}_{1,m}, \ldots, \mathrm{SINR}_{n,m}$ remain iid. This in turn allows us to evaluate $\max_{1 \leq i \leq n} \mathrm{SINR}_{i,m}$ using extreme value theory provided we can evaluate the CDF (and pdf) of the SINR. Once the CDF is available, we appeal to results in extreme value theory to obtain the behavior of $\max_{1 \leq i \leq n} \mathrm{SINR}_{i,m}$ when n is large and proceed to calculate the expectation in (2.9). 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.

It turns out that the main challenge lies in calculating the CDF of the SINR. When the channel is iid, calculating the CDF is straightforward as the SINR numerator and denominator are independent [26]. This ceases to be the case in the presence of correlation.

2.3.1 Indefinite Quadratic Forms in Gaussian Random variables

One can look at the CDF calculation of the SINR as calculating the CDF of an indefinite quadratic form in Gaussian random variables. To see this, note that

$$\Pr\{\mathrm{SINR} \leq x\} = \Pr\{\mathbf{H}_i((1+x)\phi_m\phi_m^* - xI)\mathbf{H}_i^* < \frac{x}{\rho}\} \tag{2.10}$$

Now the term $\mathbf{H}_i((1+x)\phi_m\phi_m^* - xI)\mathbf{H}_i^*$ is an indefinite sum of squares of correlated Gaussian random variables and rho represents the user's average SNIR. Before we evaluate the distribution of this variable, let's perform a quick review of approaches used to evaluate the distribution of quadratic forms.

Quadratic forms in Gaussian variables appear in many applications in signal processing, communications, and statistics. Several articles have been devoted to such study. Thus, Tzitras in [38] considered the distribution of positive definite quadratic forms in real and complex Gaussian variables. He provided necessary and sufficient conditions on when the quadratic form can be written as a sum of independent Gamma variables. His approach was to invert the expression for the characteristic function and the expressions he arrived at were almost always in the form of infinite series (for both the central and noncentral cases)

In [39], Raphaeli considered the distribution of special indefenite quaratic forms and computed the resulting CDF as an infinite series of Luagerre polynomials. The series obtained however are difficult to manipulate to find the pdf or moments. Also, it is not clear how Raphaeli's method can be used to treat the real case and how it simplifies in the central case. Shah and Li used Raphaeli's result in [40] to evaluate the distribution of quadratic forms in Gaussian mixtures.

Biyari and Lindsey considered in [41] a specific indefinite quadratic form and used the characteristic function approach to obtain expressions for the pdf and CDF. The series expansions obtained are difficult to manipulate.

More recently, Simon and Alouini [42] considered the CDF of the difference of two independent chi-square random variables and obtained a closed form expression for the value of the CDF at zero. They used their derivation to evaluate the pdf of a ratio of two such variables. In a related extension [44], Holm and Alouini evaluated the sum and difference of two correlated Nakagami variates in terms of the McKay distribution and then used that to evaluate the CDF of the ratio of such variables.

There are several drawbacks for approaches above:

1. The approaches above are not unified in nature. Various techniques are used to treat various cases (complex Gaussian, real Gaussian, central variables, noncentral variables, ... etc).

2. These approaches almost always end up with series expansions whose coefficients are difficult to evaluate. The series are difficult to manipulate further to obtain the moments or CDF.

3. They focus on obtaining the pdf from the characteristic function when the CDF is a more useful

expression. The reason is that the CDF (just like the pdf) can be used to obtain the moments (through integration). Moreover, the CDF directly gives an expression for the probability (when the pdf needs to be integrated to obtain this information).

2.3.2 Using Contour Integration for Calculating the CDF

The problem with the above approaches is that they attempt to characterize the behavior of a variable by evaluating the characteristic function. They subsequently attempt to invert it to evaluate the pdf (which requires another integration to evaluate the CDF). In contrast, and as we shall soon see, we will evaluate the CDF directly. In evaluating the CDF, we use a contour integral representation of the unit step and find the CDF using the Gaussian integral.

Distribution of $\mathrm{SINR}_{i,1}$ Given Φ

We first obtain the complementary CDF of $\mathrm{SINR}_{i,m}$ defined in (1.7) by defining the auxiliary variable S as

$$S = -\frac{x}{\rho} + \mathbf{H}_i((1+x)\phi_m\phi_m^* - xI)\mathbf{H}_i^* \tag{2.11}$$

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

$$P(\mathrm{SINR}_{i,1} > x) = P(S > 0) \quad = \quad \int_{-\infty}^{\infty} P(\mathbf{H}_i)u(S)dH_i \tag{2.12}$$

$$= \quad \frac{1}{\pi^M \det(\mathbf{R})} \int_{-\infty}^{\infty} e^{-\mathbf{H}_i\mathbf{R}^{-1}\mathbf{H}_i^*}u(S)dH_i \tag{2.13}$$

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

$$u(S) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{(j\omega+\beta)S}}{j\omega + \beta}d\omega \tag{2.14}$$

16

which is valid for any $\beta > 0$. This frees (2.13) from the constraint on S and, as we shall see, allows us to compute (2.13) in closed form.

Using (2.14), we can express (2.13) as

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

Using the definition of S in (2.11), we get

$$
\begin{aligned}
P(S > 0) &= \frac{1}{2\pi\det(\mathbf{R})}\int_{-\infty}^{\infty}d\omega\frac{e^{-(j\omega+\beta)\frac{x}{\rho}}}{j\omega+\beta}\int_{-\infty}^{\infty}dH_i e^{-\mathbf{H}_i\tilde{\mathbf{R}}\mathbf{H}_i^*} \\
&= \frac{1}{2\pi\det(\mathbf{R})}\int_{-\infty}^{\infty}d\omega\frac{e^{-(j\omega+\beta)\frac{x}{\rho}}}{j\omega+\beta}\frac{1}{\det(\tilde{\mathbf{R}})}
\end{aligned}
\tag{2.15}
$$

where

$$\tilde{\mathbf{R}} = \mathbf{R}^{-1} + x(j\omega+\beta)I - (1+x)(j\omega+\beta)\phi_m\phi_m^* \tag{2.16}$$

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

$$
\begin{aligned}
\det(\tilde{\mathbf{R}}) &= \det(U^*\Lambda^{-1}U + (j\omega+\beta)(xI - (1+x)\phi_m\phi_m^*)) & (2.17) \\
&= \det(\Lambda^{-1} + (j\omega+\beta)(xI - (1+x)\overline{\phi}_m\overline{\phi}_m^*)) & (2.18) \\
&= \det(\Lambda^{-1})\det(-\mathbf{A})\det((j\omega+\beta)I - \mathbf{A}^{-1}) & (2.19)
\end{aligned}
$$

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

$$\mathbf{A} = (1+x)\Lambda^{1/2}\overline{\phi}_m\overline{\phi}_m^*\Lambda^{1/2} - x\Lambda \tag{2.20}$$

Now

$$
\begin{aligned}
\det(\Lambda^{-1})\det(-\mathbf{A}) &= \det(xI - (1+x)\overline{\phi}_m\overline{\phi}_m^*) & (2.21) \\
&= x^{M-1}(x - (1+x)) & (2.22) \\
&= -x^{M-1} & (2.23)
\end{aligned}
$$

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

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

17

Now consider the equation

$$\det((j\omega + \beta)I - \mathbf{A}^{-1}) = 0 \tag{2.24}$$

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

$$\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2$$

where

$$\mathbf{A}_1 = (1 + x)\Lambda^{1/2}\overline{\phi}_m\overline{\phi}_m^*\Lambda^{1/2} \quad \text{and} \quad \mathbf{A}_2 = -x\Lambda$$

The matrix $\mathbf{A}_1$ has only one nonzero eigenvalue, $(1 + x)\overline{\phi}_m^*\Lambda\overline{\phi}_m$. The eigenvalues of $\mathbf{A}_2$ 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) [1]. The second largest eigenvalue of $\mathbf{A}$ thus satisfies [20]

$$\lambda_{M-1}(\mathbf{A}) \leq \begin{cases} \lambda_{M-1}(\mathbf{A}_1) + \lambda_M(\mathbf{A}_2) \\ \lambda_M(\mathbf{A}_1) + \lambda_{M-1}(\mathbf{A}_2) \end{cases} \tag{2.25}$$

$$= \begin{cases} 0 - x\lambda_1 \\ \overline{\phi}_m^*\Lambda\overline{\phi}_m - x\lambda_2 \end{cases} \tag{2.26}$$

This means that $\lambda_{M-1}(\mathbf{A}) \leq -x\lambda_1 < 0$. So the second largest eigenvalue is negative. The largest eigenvalue, however, is positive (otherwise $\mathbf{A}$ would be negative definite or singular, neither of which is the case). This means that (2.24) has exactly one positive root

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

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

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

[1]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.

Deriving the CDF of SINR With the above factorization of $\det(\tilde{\mathbf{R}})$, we can proceed to evaluate

the probability $P(\lambda > 0)$ in (2.15) and hence the CDF of the SINR can be written as,

$$P(S > 0) = -\frac{1}{x^{M-1}} \frac{1}{2\pi \det(\mathbf{R})} \int \frac{e^{-(j\omega+\beta)\frac{x}{\rho}}}{(j\omega + \beta)((j\omega + \beta) - \frac{1}{\lambda_M})\prod_{i=1}^{M-1}((j\omega + \beta) - \frac{1}{\lambda_i})} d\omega \qquad (2.27)$$

Using partial fraction expansion, we can write

$$\frac{1}{(j\omega + \beta)(j\omega + \beta - \frac{1}{\lambda_M})\prod_{i=1}^{M-1}(j\omega + \beta - \frac{1}{\lambda_i})} = \frac{\alpha_M}{j\omega + \beta - \frac{1}{\lambda_M}} + \sum_{i=1}^{M-1} \frac{\alpha_i}{j\omega + \beta - \frac{1}{\lambda_i}} + \frac{\alpha_0}{j\omega + \beta}$$

The term $\frac{\alpha_M}{j\omega+\beta-\frac{1}{\lambda_M}}$ is the only one that contributes to the integral in (2.27) (the other terms integrate

to zero since the poles are outside the contour of integration), and so we only need to calculate α_M

$$\alpha_M = \left. \frac{1}{(j\omega + \beta) \prod_{i=1}^{M-1}(j\omega + \beta - \frac{1}{\lambda_i})} \right|_{j\omega+\beta=\frac{1}{\lambda_M}} \qquad (2.28)$$

$$= \frac{1}{\frac{1}{\lambda_M} \prod_{i=1}^{M-1}(\frac{1}{\lambda_M} - \frac{1}{\lambda_i})} \qquad (2.29)$$

and

$$P(S > 0) = \frac{1}{2\pi \det(\mathbf{R})} \frac{1}{x^{M-1}} \int \frac{\alpha_M e^{-(j\omega+\beta)\frac{x}{\rho}}}{\frac{1}{\lambda_M} - (j\omega + \beta)} d\omega \qquad (2.30)$$

$$= \frac{1}{\det(\mathbf{R})} \frac{\alpha_M}{x^{M-1}} e^{-\frac{1}{\rho}\frac{x}{\lambda_M}} \qquad (2.31)$$

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

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

Or, upon replacing α_M by its value obtained in (2.29),

$$\boxed{F(x) = 1 - \frac{1}{\det(\mathbf{R})} \lambda_M \prod_{i=1}^{M-1} \frac{\lambda_i \lambda_M}{x(\lambda_i - \lambda_M)} e^{-\frac{1}{\rho}\frac{x}{\lambda_M}}} \qquad (2.32)$$

We would like to emphasize that the eigenvalues of $\mathbf{A}$, λ_i, are functions of x.

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

19

can write

$$\lambda_i = \|q_i\|_{\mathbf{A}}^2$$

$$= q_i^* \Lambda^{1/2} (\overline{\phi}_m \overline{\phi}_m^* - x \sum_{k \neq m} \overline{\phi}_k \overline{\phi}_k^*) \Lambda^{1/2} q_i$$

where we used the notation $\|q_i\|_{\mathbf{A}}^2 = q_i^* A q_i$. We can use this to show that

$$\frac{d\lambda_i}{dx} = \|q_i\|_{\mathbf{B}}^2 \qquad (2.33)$$

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

$$\frac{d}{dx} \left(\frac{\lambda_i \lambda_M}{x(\lambda_i - \lambda_M)} \right) = \frac{\lambda_M^2 \|q_i\|_{\mathbf{C}}^2 - \lambda_i^2 \|q_M\|_{\mathbf{C}}^2}{x^2(\lambda_i - \lambda_M)^2} \qquad (2.34)$$

where $\mathbf{C} = \Lambda^{1/2} \overline{\phi}_m \overline{\phi}_m^* \Lambda^{1/2}$. From (2.32)–(2.34), we can show that the SINR pdf is given by

$$\boxed{f(x) = \frac{1}{\det(\mathbf{R})} e^{-\frac{1}{\rho}\frac{x}{\lambda_M}} \prod_{i=1}^{M-1} \frac{\lambda_i \lambda_M}{x(\lambda_i - \lambda_M)} \left\{ \frac{1}{\rho} \frac{\|q_M\|_{\mathbf{C}}^2}{\lambda_M} - \|q_M\|_{\mathbf{B}}^2 - \sum_{i=1}^{M} \frac{1}{\lambda_i} \frac{\lambda_M^2 \|q_i\|_{\mathbf{C}}^2 - \lambda_i^2 \|q_M\|_{\mathbf{C}}^2}{x(\lambda_i - \lambda_M)} \right\}} \qquad (2.35)$$

Scaling Law of the Maximum SINR

Lemma 2.2: Let $F(x)$ denote the CDF of $\text{SINR}_{i,m}$ given by (2.32) and let $f(x)$ denote the associated pdf (given by (2.35)). 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 [26].

Using extreme value theory, and the lemma above, we know that $\max_{1 \leq i \leq n} \text{SINR}_{i,m}$ behaves like $\frac{\rho}{\|\overline{\phi}_m\|_{\Lambda^{-1}}^2} \log n$. Upon substituting this in (2.9) and noting that the $\overline{\phi}$'s are identically distributed, we can write

$$\begin{aligned}
\mathbf{R}_{RBF} &= \sum_{m=1}^{M} E_{\phi_m} \log \left(1 + \frac{P}{M\|\overline{\phi}_m\|_{\Lambda^{-1}}^2} \log n + o(\log \log n) \right) + o(1) \\
&= \sum_{m=1}^{M} E_{\phi_m} \log \left(\frac{P}{M\|\overline{\phi}_m\|_{\Lambda^{-1}}^2} \log n \right) + o(1) \\
&= M \log \log n + M \log \frac{P}{M} + M E_{\phi_m} \log \left(\frac{1}{\|\overline{\phi}_m\|_{\Lambda^{-1}}^2} \right) + o(1). \qquad (2.36)
\end{aligned}$$

It thus remains to calculate the expectation in (2.36) for which we need to derive the CDF of $\frac{1}{\|\bar{\phi}_m\|_{\Lambda^{-1}}^2}$.

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

Lemma 2.3: 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 $\eta_i = \frac{1}{\prod_{j\neq i}\left(\frac{1}{\lambda_j(\Lambda)} - \frac{1}{\lambda_i(\Lambda)}\right)}$.

Proof: See Appendix B for the proof.

Calculating the sum-rate

Now all we need to do to calculate the sum-rate in (2.36) is to compute $E\log\left(\frac{1}{\|\phi\|_{\Lambda^{-1}}^2}\right)$ where the distribution of $\frac{1}{\|\phi\|_{\Lambda^{-1}}^2}$ is given in the above Lemma. We employ integration by parts and use the CDF to calculate the expectation as follows

$$
\begin{aligned}
E\left(\log\left(\frac{1}{\|\bar{\phi}\|_{\Lambda^{-1}}^2}\right)\right) &= \left. G(y)\log(y)\right|_{\lambda_1(\Lambda)}^{\lambda_M(\Lambda)} - \int_{\lambda_1(\Lambda)}^{\lambda_M(\Lambda)} G(y)\frac{1}{y}dy \\
&= G(\lambda_M(\Lambda))\log(\lambda_M(\Lambda)) - \int_{\lambda_1(\Lambda)}^{\lambda_M(\Lambda)} G(y)\frac{1}{y}dy \\
&= \log(\lambda_M(\Lambda)) - \int_{\lambda_1(\Lambda)}^{\lambda_M(\Lambda)} \frac{1}{y} + \sum_{i=1}^{M}\eta_i\int_{\lambda_1(\Lambda)}^{\lambda_i(\Lambda)}\left(\frac{1}{y}-\frac{1}{\lambda_i}\right)^{M-1}\frac{1}{y}dy \\
&= \log(\lambda_1(\Lambda)) + \sum_{i=1}^{M}\eta_i\int_{\lambda_1(\Lambda)}^{\lambda_i(\Lambda)}\left(\frac{1}{y}-\frac{1}{\lambda_i}\right)^{M-1}\frac{1}{y}dy \\
&= \log(\lambda_1(\Lambda)) + \sum_{i=1}^{M}\eta_i\sum_{k=0}^{M-1}\int_{\lambda_1(\Lambda)}^{\lambda_i(\Lambda)}\frac{1}{y^{k+1}}\left(\frac{-1}{\lambda_i}\right)^{M-1-k}dy \\
&= \log(\lambda_1(\Lambda)) + \sum_{i=1}^{M}\eta_i\log\left(\frac{\lambda_i}{\lambda_1}\right)\sum_{k=1}^{M-1}\int_{\lambda_1(\Lambda)}^{\lambda_i(\Lambda)}\frac{1}{y^{k+1}}\left(\frac{-1}{\lambda_i}\right)^{M-1-k}dy \\
&= \log(\lambda_1(\Lambda)) + \sum_{i=1}^{M}\eta_i\log\left(\frac{\lambda_i}{\lambda_1}\right)\sum_{k=1}^{M-1}\frac{1}{k+2}\left(\frac{-1}{\lambda_i}\right)^{M-1-k}\left.\frac{1}{y^{k+2}}\right|_{\lambda_1(\Lambda)}^{\lambda_i(\Lambda)}
\end{aligned}
$$

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

$$
\begin{aligned}
\mathbf{R}_{RBF} &= M\log\log n + M\log\frac{P}{M} + M\log\lambda_1(\Lambda) + \\
&\quad M\sum_{i=1}^{M}\eta_i\log\left(\frac{\lambda_i}{\lambda_1}\right)\sum_{k=1}^{M-1}\frac{1}{k+2}\left(\frac{-1}{\lambda_i}\right)^{M-1-k}\left\{\frac{1}{(\lambda_i(\Lambda))^{k+2}} - \frac{1}{(\lambda_1(\Lambda))^{k+2}}\right\} + o(1) \quad (2.37)
\end{aligned}
$$

21

2.3.3 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,

$$\mathbf{R}_{BF-D} = M \log \log n + M \log \frac{P}{M} + \sum_{i=1}^{M} \log \left(\frac{1}{\phi_i^* U^* \Lambda^{-1} U \phi_i} \right) + o(1) \qquad (2.38)$$

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

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 reduces to 2.1.

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. (??) in fact proves Lemma 2.1. It should be also mentioned that this result is obtained in [21] for $M = 2$.

2.3.4 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 $\sqrt{\alpha}\mathbf{A}^{-1/2}\Phi$ where $\mathbf{A}$ is a positive definite matrix and α is just a normalization factor to adjust the transmit power. Again similar to Section 2.3, we can state that α has to be less than $\frac{M}{Tr(\mathbf{A}^{-1})}$.

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 $\mathbf{R} = \mathbf{A}^{*-1/2}\mathbf{R}\mathbf{A}^{-1/2}$. Therefore the same result holds for this case with the new covariance matrix $\tilde{\mathbf{R}}$. Here is the main result.

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

$$\mathbf{R}_{BF-Prec} = M \log \log n + M \log \frac{P}{M} + \sum_{i=1}^{M} E \log \left(1 + \frac{Tr(\Lambda^{-1})}{M} \frac{P}{\phi_i^* U^* \Lambda^{-1} U \phi_i} \right) + o(1). \qquad (2.39)$$

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

2.4 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

$$\mathbf{R}_2 = \begin{bmatrix} 1 & \gamma \\ \gamma & 1 \end{bmatrix} \tag{2.40}$$

where γ is the correlation. Fig. 1 shows the sum-rate loss (relative to the case of no correlation) versus the correlation coefficient γ 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

$$\mathbf{R}_3 = \begin{bmatrix} 1 & \gamma & \gamma^2 \\ \gamma & 1 & \gamma \\ \gamma^2 & \gamma & 1 \end{bmatrix} \tag{2.41}$$

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

2.5 Conclusion

This chapter 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.

23
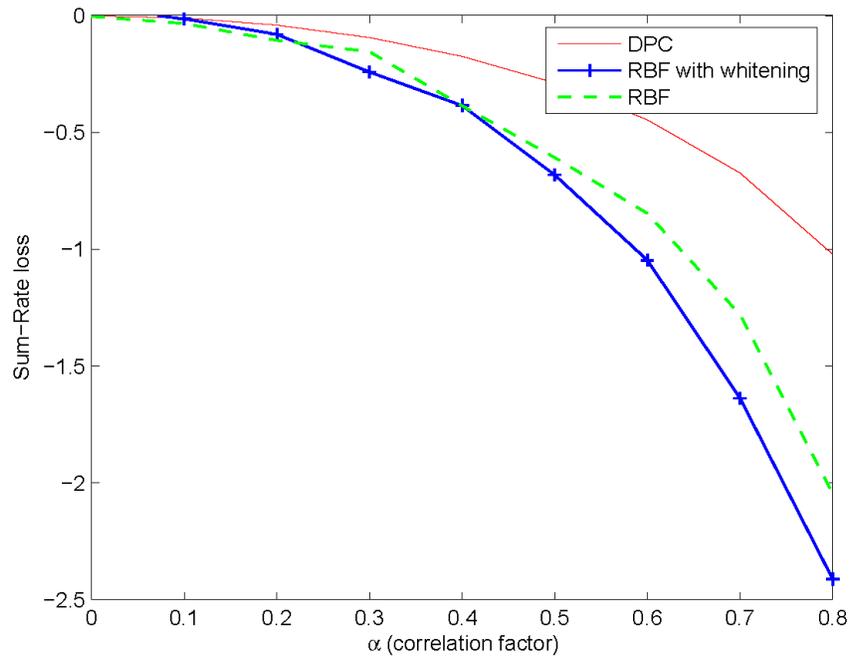
Figure 2.1: Sum-rate loss versus the correlation factor γ for a system with $M = 2$ and $n = 100$.
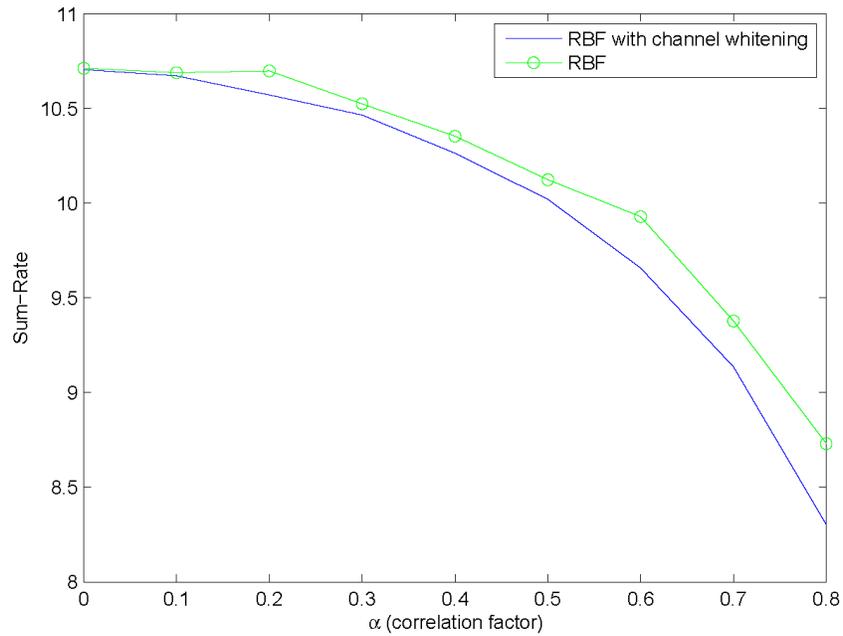
Figure 2.2: Sum-rate versus the correlation factor γ for a system with $M = 2$, $P = 10$, and $n = 100$.
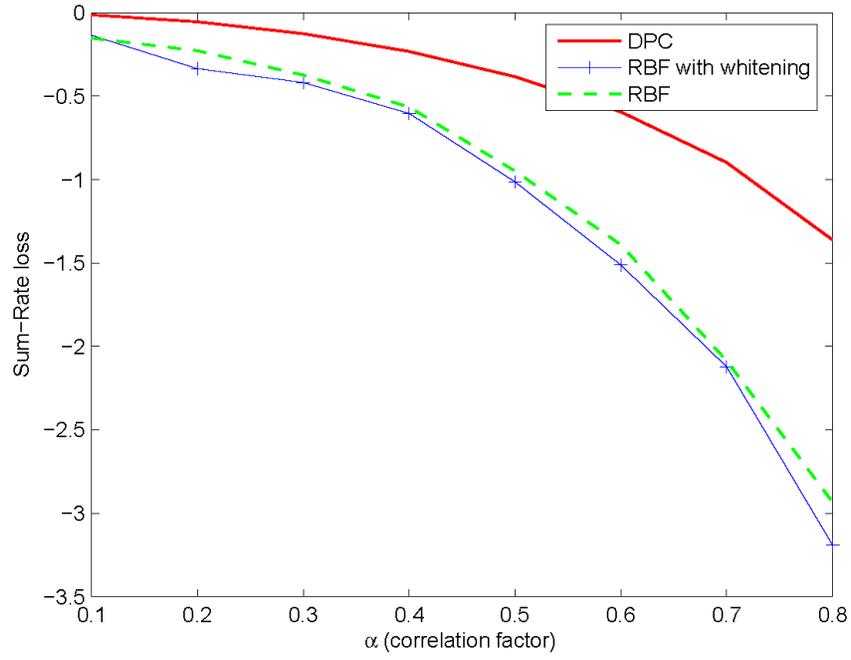
Figure 2.3: Sum-rate loss versus the correlation factor γ for a system with $M = 3$ and $n = 100$.
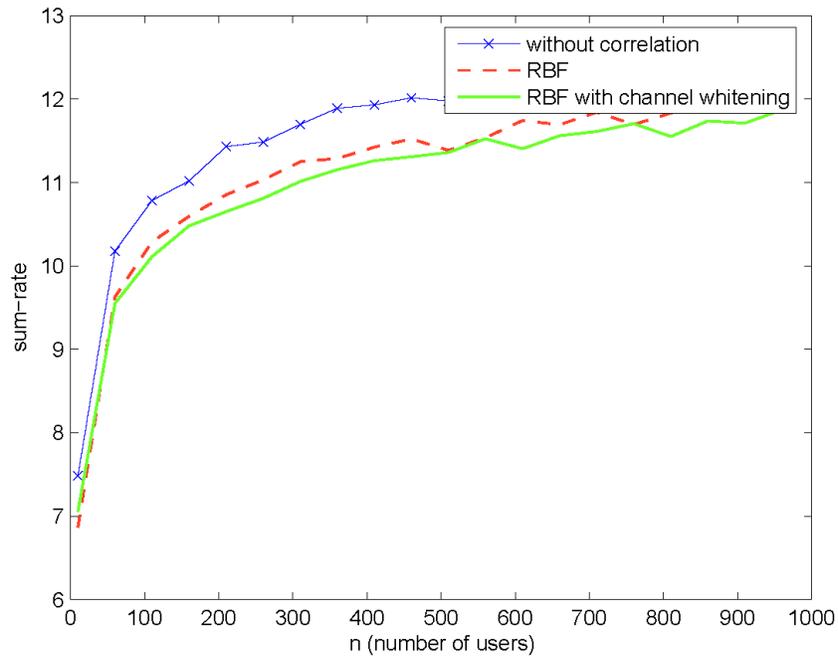
Figure 2.4: Sum-rate versus the number of users in a system with $M = 2$, $P = 10$, and $\gamma = 0.5$

In the presence of a correlation between transmit antennas, the channel matrix has a covariance matrix $\mathbf{R}$ which is assumed to be non-singular and $Tr(\mathbf{R}) = M$. In this case, the sum-rate of DPC and beamforming schemes will be 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 $\mathbf{R}$. For DPC, c is just the geometric mean of the eigenvalues of $\mathbf{R}$. 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 $\mathbf{R}$. It is worth mentioning, numerical results suggest that sum-rate of random beamforming outperforms that of the random beamforming with channel whitening [2].

2.6 Appendix A: Proof of Lemma 2.2

From (2.32) and (2.35), we can write

$$\frac{1 - F(x)}{f(x)} = \frac{\lambda_M}{\frac{1}{\rho} \frac{\|q_M\|_{\mathbf{C}}^2}{\lambda_M} - \|q_M\|_{\mathbf{B}}^2 - \sum_{i=1}^{M-1} \frac{1}{\lambda_i} \frac{\lambda_M^2 \|q_i\|_{\mathbf{C}}^2 - \lambda_i^2 \|q_M\|_{\mathbf{C}}^2}{x(\lambda_i - \lambda_M)}} \tag{2.42}$$

To evaluate the limit of this expression, we need to investigate the behavior of the eigenvalues and eigenvectors of $\mathbf{A}$ as $x \to \infty$. Now from the bound (2.26), 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,

$$\lambda_M = \max_{\|u\|_2 = 1} u^* A u = \max_{\|u\|_2 = 1} u^* (\Lambda^{1/2} \overline{\phi}_m \overline{\phi}_m^* \Lambda^{1/2} - x \Lambda^{1/2} \sum_{m \neq i} \overline{\phi}_i \overline{\phi}_i^*) u \tag{2.43}$$

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

$$\|Au\|_2 = u^* A u = u^* \left(\alpha_m \Lambda^{1/2} \overline{\phi}_m - x \sum_{i \neq m} \alpha_i \Lambda^{1/2} \overline{\phi}_i \right) = \alpha_m^2 - x \sum_{i \neq m} \alpha_i^2 \tag{2.44}$$

---

[2]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.

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

$$\lim_{x \to \infty} q_M = \lim_{x \to \infty} u = \frac{\Lambda^{-1/2}\overline{\phi}_m}{\sqrt{\|\overline{\phi}_m\|_{\Lambda^{-1}}^2}} \tag{2.45}$$

and

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

Using the above, it is easy to verify that

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

and

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

From (2.45) and the defining expression of $\mathbf{B}$, we also deduce that

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

Thus, the only nonzero limit in the denominator of (2.42) is $\frac{1}{\rho} \frac{\|q_m\|_{\mathbf{C}}^2}{\lambda_M}$ and

$$\lim_{x \to \infty} \frac{1 - F(x)}{f(x)} = \frac{\lambda_M^2}{\frac{1}{\rho}\|q_m\|_{\mathbf{C}}^2} = \frac{\rho}{\|\overline{\phi}_m\|_{\Lambda^{-1}}^2} \tag{2.46}$$

2.7   Appendix B: Proof of Lemma 2.3

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

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

27

Alternatively, following the approach of [5], 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

$$
\begin{aligned}
p(\frac{1}{\|\phi\|_{\Lambda^{-1}}^2} > x) &= \frac{\Gamma(M)}{4\pi^{M+2}} \int d\omega_1 \int d\omega_2 \int d\phi \frac{e^{(j\omega_1+\beta_1)(1-x\|\phi\|_{\Lambda^{-1}}^2)} e^{j\omega_2(\|\phi\|^2-1)}}{j\omega_1 + \beta_1} \\
&= \frac{\Gamma(M)}{4\pi^{M+2}} \int d\omega_1 \frac{e^{(j\omega_1+\beta_1)}}{j\omega_1 + \beta_1} \int d\omega_2 e^{-j\omega_2} \int d\phi e^{-\phi^*\left(x(j\omega_1+1)\Lambda^{-1}-j\omega_2 I\right)\phi} \\
&= \frac{\Gamma(M)}{4\pi^{M+2}} \int d\omega_1 \frac{e^{(j\omega_1+\beta_1)}}{j\omega_1 + \beta_1} \int d\omega_2 e^{-j\omega_2} \frac{1}{\det\left(x(j\omega_1+\beta_1)\Lambda^{-1}-j\omega_2 I\right)}
\end{aligned}
$$

Now use partial fraction expansion to show that

$$
\frac{1}{\det\left(x(j\omega_1+\beta_1)\Lambda^{-1}-j\omega_2 I\right)} = \frac{1}{\prod_{i=1}^M \left(\frac{x}{\lambda_i(\Lambda)}(j\omega_1+\beta_1)-j\omega_2\right)} \tag{2.47}
$$

$$
= \frac{1}{x^{M-1}} \frac{1}{(j\omega_1+\beta_1)^{M-1}} \sum_{i=1}^M \frac{\eta_i}{\frac{x}{\lambda_i(\Lambda)}(j\omega_1+\beta_1)-j\omega_2} \tag{2.48}
$$

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

$$
\begin{aligned}
p(\frac{1}{\|\phi\|_{\Lambda^{-1}}^2} > x) &= \frac{\Gamma(M)}{4\pi^2} \frac{1}{x^{M-1}} \int d\omega_1 \frac{e^{j\omega_1+\beta_1}}{(j\omega_1+\beta_1)^M} \sum_i \int d\omega_2 \frac{\eta_i}{\frac{x}{\lambda_i(\Lambda)}(j\omega_1+\beta_1)-j\omega_2} e^{-j\omega_2} \tag{2.49} \\
&= \frac{\Gamma(M)}{2\pi} \frac{1}{x^{M-1}} \sum_i \eta_i \int d\omega_1 \frac{e^{(j\omega_1+\beta_1)(1-\frac{x}{\lambda_i(\Lambda)})}}{(j\omega_1+\beta_1)^M} \tag{2.50}
\end{aligned}
$$

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.

Chapter 3

Opportunistic Beamforming with Precoding

In the previous chapter, we showed that correlation always results in a hit on the various broadcast techniques. However, the hit varies from one technique to another and the sum-rate of random beamforming does not anymore match that of DPC for large number of users (in direct contrast to the white channel case). The chapter also showed that a technique like zero-forcing which attempts to counter the effect of correlation performs even worse than random beamforming.

In this chapter, we explore other precoding techniques that in contrast to zero forcing, can actually improve the performance of random beamforming. Thus, after revisiting random beamforming with percoding, we derive the optimum precoding that minimizes the hit on the sum rate of random beamforming. The precoding matrix arrived is obtained by solving M nonlinear equations in M unknowns. Thus, we also introduce 3 suboptimum precoding techniques, two of which are intuitively justified (namely zero forcing and MMSE precoding), and a third that minimizes an upper bound on the hit and gives the precoding matrix in closed form. Simulations are presented at the end of the chapter.

3.1 Beamforming with Precoding

3.2 Random Beamforming with Optimum Precoding

To counter the effect of correlation, we introduce beamforming with precoding. Specifically, instead of transmitting the signal (1.4), we transmit

$$\alpha \mathbf{A} S(t) = \alpha \mathbf{A} \sum_{m=1}^{M} \phi_m(t) s_m(t), \qquad t = 1, \ldots, T \tag{3.1}$$

where the constant α is chosen to maintain a power constraint of P and $\mathbf{A}$ is the desired precoding matrix. Specifically we should have

$$\alpha^2 E\{S^* \mathbf{A}^* \mathbf{A} S\} \leq P$$

i.e. $\alpha^2 Tr(\mathbf{A}^* \mathbf{A} E[SS^*])\} \leq P$, or as $E[SS^*] = \frac{P}{M} I$, it follows that

$$\alpha \leq \sqrt{\frac{M}{\mathrm{Tr}(\mathbf{A}^* \mathbf{A})}}$$

Lets consider the input/output equation for the new choice of $S(t)$,

$$Y_i = \alpha \mathbf{H}_i \mathbf{A} S(t) + W_i$$

Random beamforming over this channel is nothing but the familiar random beamforming over the effective channel

$$\tilde{\mathbf{H}}_i = \alpha \mathbf{H}_i \mathbf{A}$$

which exhibits correlation of $\alpha^2 \tilde{\mathbf{R}} = \alpha^2 \mathbf{A}^* \mathbf{R} \mathbf{A}$. From (2.36), we know that the sum-rate capacity scales as

$$
\begin{aligned}
\mathbf{R}_{\mathrm{PC}} &= M \log \log n + M \log \frac{P}{M} + ME \log \frac{1}{\|\phi_m\|^2_{\frac{1}{\alpha^2}\tilde{\mathbf{R}}^{-1}}} \\
&= M \log \log n + M \log \frac{P}{M} + M \log \frac{M}{Tr(\mathbf{A}^* \mathbf{A})} + E \log \frac{1}{\|\phi_m\|^2_{\tilde{\mathbf{R}}^{-1}}}
\end{aligned} \tag{3.2}
$$

where the last equality follows from the fact that the choice

$$\alpha^2 = \frac{M}{Tr(\mathbf{A}^* \mathbf{A})}$$

will maximize the sum-rate. Alternatively, we can write

$$\mathbf{R}_{\mathrm{PC}} = M \log \log n + M \log \frac{P}{M} - h(\mathbf{A}) \tag{3.3}$$

where $h(\mathbf{A})$ is the hit incurred by using a general precoding matrix $\mathbf{A}$

$$\boxed{h(\mathbf{A}) = M \log \frac{\mathrm{Tr}(\mathbf{A}^*\mathbf{A})}{M} + ME \log \|\phi_m\|_{\tilde{\mathbf{R}}^{-1}}^2} \tag{3.4}$$

Finding the general precoding matrix $\mathbf{A}$ that minimizes the hit is difficult. The following lemma shows that the optimum $\mathbf{A}$ has a special structure.

Lemma: The optimum precoding matrix $\mathbf{A}_{opt}$ can be written as

$$\mathbf{A}_{opt} = \mathbf{Q}_{Aopt}\mathbf{D}_{Aopt}$$

where $\mathbf{Q}_{Aopt}$ is an orthonormal matrix and $\mathbf{D}_{Aopt}$ is a diagonal matrix with positive entries.

Proof: Consider the general precoding hit $h(\mathbf{A})$ in (3.4) and consider the eigenvalue decomposition of $\tilde{\mathbf{R}} = \tilde{\mathbf{Q}}\tilde{\Lambda}\tilde{\mathbf{Q}}^*$. It is easy to see that

$$
\begin{aligned}
\|\phi_m\|_{\tilde{\mathbf{R}}^{-1}}^2 &= \|\phi_m\|_{\tilde{\mathbf{Q}}\tilde{\Lambda}^{-1}\tilde{\mathbf{Q}}^*}^2 \\
&= \|\tilde{\mathbf{Q}}\phi_m\|_{\tilde{\Lambda}^{-1}}^2 = \|\phi_m\|_{\tilde{\Lambda}^{-1}}^2
\end{aligned} \tag{3.5}
$$

where the last equality follows from the fact that ϕ_m is an isotropic vector and hence is invariant under multiplication by an orthonormal matrix $\mathbf{Q}$. With this in mind, the hit can be equivalently written as

$$h(\mathbf{A}) = M \log \frac{\mathrm{Tr}(\mathbf{A}^*\mathbf{A})}{M} + ME \log \|\phi_m\|_{\tilde{\Lambda}^{-1}}^2$$

Now the first term of the hit depends on $\mathrm{Tr}(\mathbf{A}^*\mathbf{A})$ and hence $\mathrm{Tr}(AA^*)$. The second term depends on the eigenvalues of $\tilde{\mathbf{R}}$, i.e. of $\mathbf{A}^*\mathbf{R}\mathbf{A}$, or equivalently the eigenvalues of RAA^* as it coincides with the eigenvalues of $\mathbf{A}^*\mathbf{R}\mathbf{A}$. Thus both terms of the hit are determined by AA^*. One choice of the optimum matrix $\mathbf{A}_{opt}$ is thus

$$\boxed{\mathbf{A}_{opt} = \mathbf{Q}_{Aopt}\mathbf{D}_{Aopt}}$$

where $\mathbf{Q}_{Aopt}$ is orthonormal and $\mathbf{D}_{Aopt}$ is diagonal with positive entries. This proves the lemma.

3.2.1 Determining $\mathbf{Q}_{opt}$

An intuitive choice however is to set $\mathbf{Q}_{Aopt}$ to $\mathbf{Q_R}$ of the channel autocorrelation matrix (as this will diagonalize this matrix). In the following, we show that this choice is actually optimum. To this end, let Π_l be a diagonal matrix with all $1's$ on the diagonal except for a -1 at the lth entry and define $\hat{\mathbf{A}} = \mathbf{Q}\Pi_l\mathbf{D}$. This induces the effective correlation $\tilde{\mathbf{R}}_l$. The hit that results by using either of the precoding matrices $\mathbf{A}$ or $\hat{\mathbf{A}}_l$ is the same. To see this, note that

$$\mathrm{Tr}(\mathbf{A}^*\mathbf{A}) = \mathrm{Tr}(\hat{\mathbf{A}}_l^*\hat{\mathbf{A}}_l) = \mathrm{Tr}(\mathbf{D})$$

Moreover,

$$\|\phi\|^2_{\hat{\mathbf{R}}_l^{-1}} = \|\phi\|^2_{\mathbf{D}^{-\frac{1}{2}}\Pi_l\mathbf{Q}^*\mathbf{R}\mathbf{Q}\Pi_l\mathbf{D}^{-\frac{1}{2}}} = \|\Pi_l\mathbf{D}^{-\frac{1}{2}}\phi\|^2_{\mathbf{Q}^*\mathbf{R}\mathbf{Q}} \tag{3.6}$$

Now the distribution of ϕ is unchanged by the changing the sign of the lth entry. Hence,

$$E\log\|\phi\|^2_{\hat{\mathbf{R}}_l^{-1}} = E\log\|\mathbf{D}^{-\frac{1}{2}}\phi\|^2_{\mathbf{Q}\mathbf{R}\mathbf{Q}^*} = E\log\|\phi\|^2_{\tilde{\mathbf{R}}^{-1}} \tag{3.7}$$

Thus, both terms of the hits are the same and

$$h(\tilde{\mathbf{A}}) = h(\mathbf{A})$$

Now note that

$$\begin{aligned}
\frac{h(\mathbf{A}) + h(\tilde{\mathbf{A}})}{2} &= M\log\frac{\mathrm{Tr}(\mathbf{D})}{M} + \frac{M}{2}E\log\|\phi\|^2_{\tilde{\mathbf{R}}^{-1}} + \frac{M}{2}E\log\|\phi\|^2_{\tilde{\mathbf{R}}_l^{-1}} \\
&\geq M\log\frac{\mathrm{Tr}(\mathbf{D})}{M} + ME\log\|\phi\|^2_{(\frac{1}{2}\tilde{\mathbf{R}}+\frac{1}{2}\tilde{\mathbf{R}}_l)^{-1}} \\
&= M\log\frac{\mathrm{Tr}(\mathbf{D})}{M} + ME\log\|\mathbf{D}^{-\frac{1}{2}}\phi\|^2_{(\frac{1}{2}\mathbf{Q}^*\mathbf{R}\mathbf{Q}+\frac{1}{2}\Pi_l\mathbf{Q}^*\mathbf{R}^{-1}\mathbf{Q}\Pi_l)^{-1}}
\end{aligned}$$

Or as $h(\tilde{\mathbf{A}}) = h(\mathbf{A})$,

$$h(\mathbf{A}) = M\log\frac{\mathrm{Tr}(\mathbf{D})}{M} + \frac{M}{2}E\log\|\mathbf{D}^{-\frac{1}{2}}\phi\|^2_{(\frac{1}{2}\mathbf{Q}^*\mathbf{R}\mathbf{Q}+\frac{1}{2}\Pi_l Q^*\mathbf{R}^{-1}\mathbf{Q}\Pi_l)^{-1}}$$

Note that the weight matrix $\frac{1}{2}\mathbf{Q}^*\mathbf{R}\mathbf{Q} + \frac{1}{2}\Pi_l\mathbf{Q}^*\mathbf{R}\mathbf{Q}\Pi_l$ has entries equal to those of $\mathbf{Q}^*\mathbf{R}\mathbf{Q}$ except the off diagonals lying on the lth column or lth row which are zero. What the last equation says is that nulling the off diagonal elements of $\mathbf{Q}^*\mathbf{R}\mathbf{Q}$ at the lth row and lth column can only reduce the hit. This argument can be repeated for $l = 1, \ldots, M$. Hence, nulling the off diagonal entries of $\mathbf{Q}\mathbf{R}\mathbf{Q}^*$ can only reduce the hit. Thus, $\mathbf{Q}\mathbf{R}\mathbf{Q}^*$ should be diagonal, i.e. $\mathbf{Q}_{opt} = \mathbf{Q_R}$.

3.2.2   Determining $\mathbf{D}_{opt}$

We have so far established that $\mathbf{A}_{\text{opt}} = \mathbf{Q_R}\mathbf{D}_{opt}^{\frac{1}{2}}$ where $\mathbf{D}_{\text{opt}}$ is a diagonal matrix to be determined. The hit in this case is given by

$$h(\mathbf{A}_{\text{opt}}) = M \log \frac{\text{Tr}(\mathbf{D}_{\text{opt}})}{M} + E \log \|\phi\|^2_{\mathbf{D}_{\text{opt}}^{-1}\Lambda^{-1}}$$

Now, taking the derivative with respect to ith diagonal d_i element of $\mathbf{D}_{\text{opt}}$ and setting it to zero, we obtain

$$\frac{1}{d_i} E \left[\frac{\frac{1}{d_i\lambda_i}|\phi(i)|^2}{\|\phi\|^2_{\mathbf{D}_{\text{opt}}^{-1}\Lambda^{-1}}} \right] = \frac{1}{Tr(\mathbf{D}_{\text{opt}})} \tag{3.8}$$

where in arriving at (3.8), we exchanged the differentiation and expectation operations. Thus, we have a set of M implicit equations for $d_1, d_2, \ldots, d_M$. We can solve these equations numerically provided we first obtain the expectation of the random variable Z_1 that appears in (3.8). In the Appendix, we evaluate the CDF of the more general random variable Z for diagonal matrices $\mathbf{B}$ and $\mathbf{C}$. By setting $\mathbf{B} = \text{diag}(0, \ldots, \frac{1}{d_i\lambda_i}, \ldots, 0)$ and $\mathbf{C} = \mathbf{D}^{-1}\Lambda^{-1}$, we obtain the CDF of Z_1. Since the support of Z_1 is over the interval $(0, 1)$, it is expectation is given by

$$E[Z_1] = \int_0^1 (1 - F_{Z_1}(z_1))dz_1$$

3.3   Approximate Precoding Matrcies

In the previous section, we obtained the optimum precoding matrix which we showed to be of the form

$$\mathbf{A}_{\text{opt}} = \mathbf{Q_R}\mathbf{D}_{opt}^{\frac{1}{2}}$$

To obtain $\mathbf{D}_{\text{opt}}$, we need to solve M nonlinear equations in M unknowns. In the following, we derive 3 approximate precoding matrices, two of which are intuitively justified and the third is obtained by minimizing an upper bound on the hit.

3.3.1 Random Beamforming with Zero Forcing

A natural choice of the precoding matrix is one which whitens the channel, i.e.

$$\mathbf{A}_{\text{ZF}} = \mathbf{Q_R} \Lambda_{\mathbf{R}}^{-\frac{1}{2}}$$

From (3.4), this zero-forcing choice of the precoding matrix results in the hit

$$h_{ZF} = M \log \frac{Tr(\mathbf{R}^{-1})}{M}$$

Comparing this with the sumrate of DPC, we note that the hit in the DPC case is equal to geometric mean of the eigenvalues, while it is equal to the harmonic mean in the channel whitening case, making the latter inferior to DPC.

3.3.2 Random Beamforming with MMSE Precoding

The zero-focing solution invests most of the input power taking care of the minimum eigenvalue which explains its inferior behavior. So we consider here the MMSE solution. Specifically, consider the choice

$$\mathbf{A}_{\text{MMSE}} = \mathbf{Q_R}(\Lambda + \beta I)^{-\frac{1}{2}}$$

for some constant β which we now determine. For this choice of $\mathbf{A}$, the sum-rate hit is given by

$$h_{\text{MMSE}} = M \log \frac{Tr(\Lambda + \beta I)^{-1}}{M} + M E \log \left(1 + \beta \|\Phi_m\|_{\Lambda^{-1}}^2\right)$$

Upon setting the first derivative of the hit to zero, we obtain the following implicit equation for β

$$\frac{Tr(\Lambda + \beta^* I)^{-2}}{Tr(\Lambda + \beta^* I)^{-1}} = E \left(\frac{1}{\beta + \frac{1}{\|\Phi_m\|_{\Lambda^{-1}}^2}} \right) \tag{3.9}$$

where in arriving at this equation, we interchanged the expectation and differentiation operations. To solve this implicit equation, we need to evaluate the expectation and hence the CDF of $\frac{1}{\|\phi_m\|_{\Lambda^{-1}}^2}$.

34

In the Appendix, we evaluate the CDF of the more general random variable

$$Z = \frac{\|\phi\|_{\mathbf{B}}^2}{\|\phi\|_{\mathbf{C}}^2} \tag{3.10}$$

for diagonal matrices $\mathbf{B}$ and $\mathbf{C}$ which for the special case of $\mathbf{B} = I$ and $\mathbf{C} = \Lambda^{-1}$ reads

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

where $\eta_i = \frac{1}{\prod_{j \neq i} \frac{1}{\lambda_j} - \frac{1}{\lambda_i}}$ and $\lambda_M \geq \cdots \geq \lambda_1 > 0$ are the eigenvalues of $\mathbf{R}$ and where $u(x)$ is the unit step function. Thus, the moment in (3.9) is given by

$$E\left(\frac{1}{\beta + \frac{1}{\|\Phi_m\|_{\Lambda^{-1}}^2}}\right) = \frac{1}{1 + \lambda_M} + \int_{\lambda_1}^{\lambda_M} \frac{1}{(\beta + x)^2} G(x) dx$$

3.3.3 An approximate Precoding Matrix

In the previous section, we obtained an approximate precoding matrix (up to an orthogonal transformation). The problem with the solution obtained is that we need to simultaneously solve M nonlinear equations in the M diagonal unknowns. We derive in this section an approximate precoding that is 1) explicit and 2) does not assume that $\mathbf{Q_{A_{opt}}} = \mathbf{Q_R}$ but actually proves it.

The difficult part in minimizing the hit is the term that depends on ϕ_m. So we rewrite this hit as

$$
\begin{aligned}
h(\mathbf{A}) &= M \log \operatorname{Tr}(\mathbf{A}^*\mathbf{A}) + ME \log \|\phi\|_{(\mathbf{A}^*\mathbf{R}\mathbf{A})^{-1}}^2 & (3.11) \\
&= M \log \operatorname{Tr}(\mathbf{A}^*\mathbf{A}) + M \log \operatorname{Tr}((\mathbf{A}^*\mathbf{R}\mathbf{A})^{-1}) + ME \log \|\phi\|_{\frac{(\mathbf{A}^*\mathbf{R}\mathbf{A})^{-1}}{\operatorname{Tr}(\mathbf{A}^*\mathbf{R}\mathbf{A})^{-1}}}^2 & (3.12)
\end{aligned}
$$

We now minimize the sum of the first two terms of the hit and ignore the 3rd term. There are two justifications for doing so

1. The first two terms constitute an upper bound on the hit. To see this, note that

$$
\begin{aligned}
\log \|\phi\|_{\frac{(\mathbf{A}^*\mathbf{R}\mathbf{A})^{-1}}{\operatorname{Tr}(\mathbf{A}^*\mathbf{R}\mathbf{A})^{-1}}}^2 &= E \log \|\phi\|_{\frac{(\tilde{\Lambda})^{-1}}{\operatorname{Tr}(\tilde{\Lambda}^{-1})}}^2 & (3.13) \\
&\leq \log \|\phi\|^2 \frac{\operatorname{Tr}(\tilde{\Lambda}^{-1})}{\operatorname{Tr}(\tilde{\Lambda}^{-1})} = 0 & (3.14)
\end{aligned}
$$

where $\tilde{\Lambda}$ is the diagonal matrix of eigenvalues of $\mathbf{A}^*\mathbf{R}\mathbf{A}$.
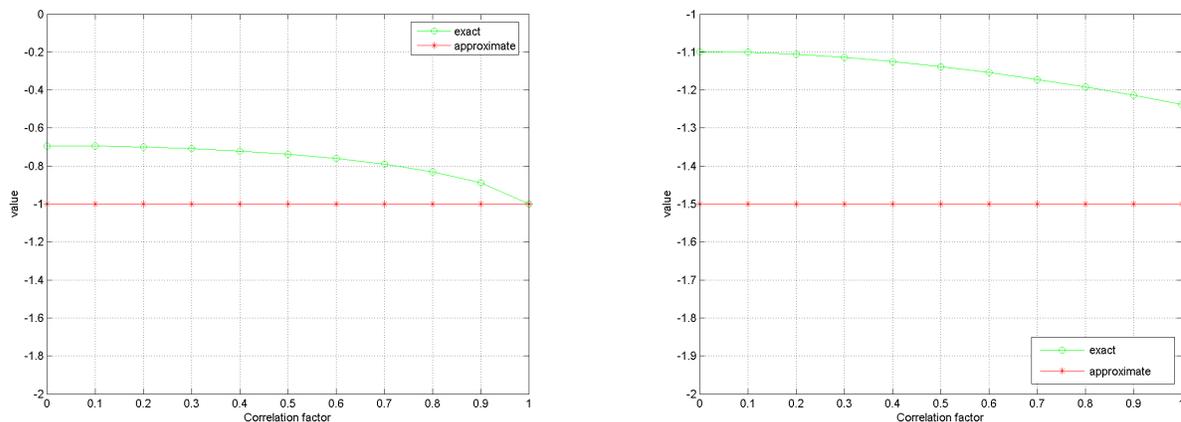
Figure 3.1: Comparison between the exact and approximation for M=2 (left) and M=3 (right).

2. One can consider the term $\|\phi\|^2_{\frac{(\tilde{\Lambda})^{-1}}{\mathrm{Tr}(\tilde{\Lambda}^{-1})}}$ as the squared dot product of two unit norm vectors ϕ

   and

$$c = \frac{\mathrm{diag}(\tilde{\Lambda})^{-\frac{1}{2}}}{\sqrt{\mathrm{Tr}(\tilde{\Lambda}^{-1})}}$$

This squared dot product can be approximated as the squared dot product of two *uniformly*

distributed unit norm vectors υ which has a CDF [32]

$$F(\upsilon) = 1 - (1 - \upsilon)^{M-1} \qquad \upsilon \in [0, 1]$$

Hence, we can approximate the expectation in (3.12) as

$$E \log \|\phi\|^2_{\frac{(\tilde{\Lambda})^{-1}}{\mathrm{Tr}(\tilde{\Lambda}^{-1})}} \quad \simeq \quad E[\log \upsilon] \tag{3.15}$$

$$= \quad -\sum_{m=1}^{M-1} \frac{1}{m} \tag{3.16}$$

Figure 3.1 plots the two sides of (3.15) for various values of the correlation coefficient α and

shows that they are almost the same.

Thus up to an almost constant term, the hit is given by

$$h(\mathbf{A}_{\mathrm{Appx}}) = M \log \frac{\mathrm{Tr}(\mathbf{A}^*\mathbf{A})}{M} + \log Tr(\mathbf{A}^*\mathbf{R}\mathbf{A})^{-1}$$

Let's try to minimize the hit by designing A properly. Taking the first derivative with respect to $\mathbf{A}$ and setting the result to zero yields

$$\frac{\partial Tr(\mathbf{A}^*\mathbf{A})}{\partial \mathbf{A}} = 2\mathbf{A} - 2\mathbf{R}\mathbf{A}(\mathbf{A}^*\mathbf{R}\mathbf{A})^{-2} \tag{3.17}$$

$$\frac{2}{\text{Tr}(\mathbf{A}^*\mathbf{A})}\mathbf{A} = \frac{2}{\text{Tr}(\mathbf{A}^*\mathbf{R}\mathbf{A})^{-2}}\mathbf{R}\mathbf{A}(\mathbf{A}^*\mathbf{R}\mathbf{A})^{-1}$$

Alternatively, we can write

$$AA^*RAA^* = \frac{\text{Tr}(\mathbf{A}^*\mathbf{A})}{\text{Tr}(\mathbf{A}^*\mathbf{R}\mathbf{A})^{-1}}I \tag{3.18}$$

This shows that AA^* is a left and a right inverse of $\mathbf{R}$. So, using the eigenvalue decomposition, $\mathbf{R} = \mathbf{Q_R}\Lambda_\mathbf{R}\mathbf{Q_R^*}$, we can show that the choice

$$\boxed{\mathbf{A}_{\text{Appx}} = \mathbf{Q_R}\Lambda_\mathbf{R}^{-1/4}}$$

satisfies (3.18). The resulting hit for this choice of precoding is given by

$$h_{\text{Appx}} = M \log \frac{\text{Tr}(\Lambda^{-\frac{1}{2}})}{M} + M \log \|\phi\|^2_{\Lambda^{-\frac{1}{2}}}$$

3.4 Simulations

We consider a broadcast scenario with a base station having $M = 2$ and $M = 3$ antennas. The channels exhibit the following correlations respectively $0 \leq \gamma < 1$

$$\mathbf{R}_2 = \begin{bmatrix} 1 & \gamma \\ \gamma & 1 \end{bmatrix} \quad \mathbf{R}_3 = \begin{bmatrix} 1 & \gamma & \gamma^2 \\ \gamma & 1 & \gamma \\ \gamma^2 & \gamma & 1 \end{bmatrix}$$

In what follows, we evaluate the scaling of RBF, zero-forcing RBF, MMSE, diagonal and approximate RBF precoding for different channel correlation strength. Figures (3.2)-(3.6) clearly show that RBF incurs a hit in the presence of correlation. They compare the sum-rate capacity against the number of users in the presence of spatial correlation between the transmit antennas for a system with $M = 2$ and $M = 3$ transmit antennas, a correlation factor (γ) of 0.5 and 0.7 respectively and a power (P)
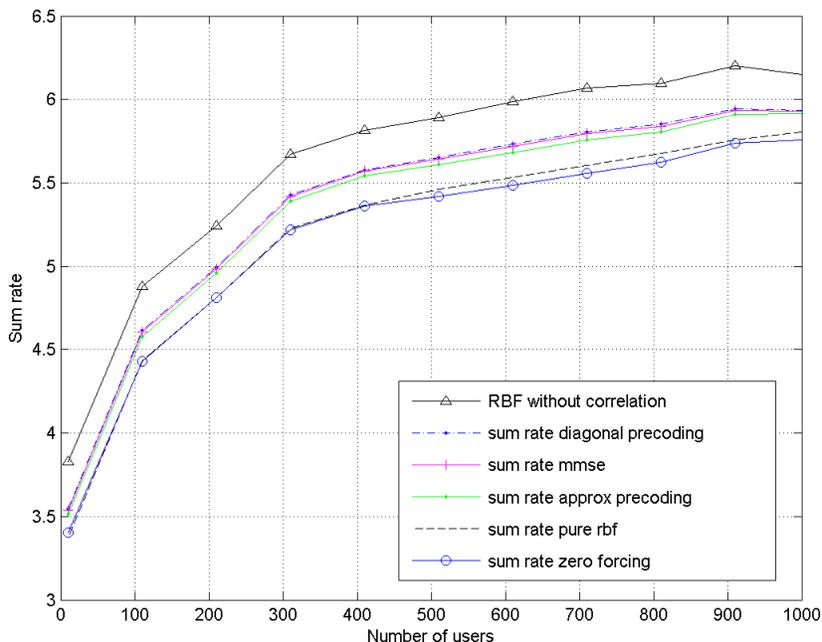
Figure 3.2: Sum-rate versus the number of users in a system with M = 2, P = 10 and $\gamma = 0.5$

of 10. Both theoretical and simulation results show that RBF with precoding (diagonal, MMSE and approximate) performs better than RFB with zero forcing and thus optimizes the sum-rate capacity. Figures (3.7)-(3.11) compares the sum-rate loss against the channel correlation γ for $M = 2$ and $M = 3$ transmit antennas with $n = 400$ and $n = 100$ users respectively. Both simulation and theoretical results show that RBF with diagonal, MMSE and approximate precoding outperforms RBF especially for highly correlated channels, while zero-forcing is inferior to RBF. Figure (3.12) shows the CDF of Z for general diagonal matrices **A** and **B**.

3.5   Conclusion

In this chapter, we considered random beamforming in a spatially correlated regime. While RFB matches DPC for uncorrelated channels (in the large number of users regime), it incurs an SNR hit in the presence of correlation. We suggested precoding techniques as a way to counter the effect of correlation. Specifically, it was shown that RBF with diagonal and MMSE precoding outperforms
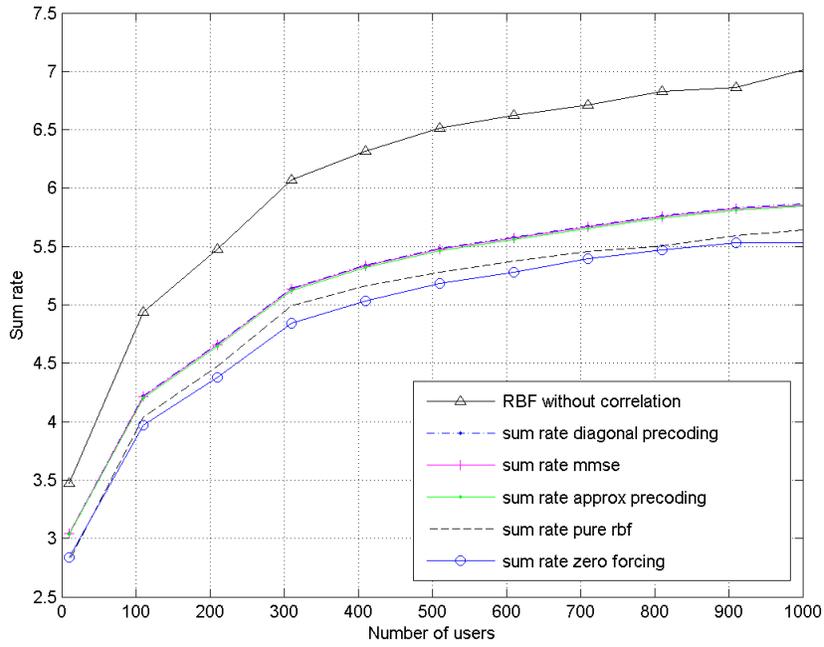
Figure 3.3: Sum-rate versus the number of users in a system with M = 3, P = 10 and $\gamma = 0.7$
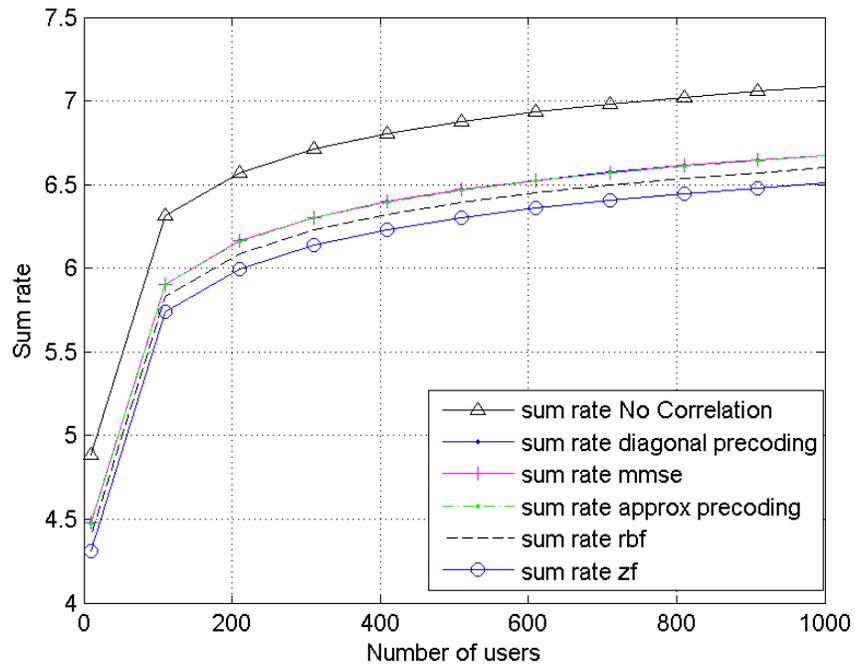
Figure 3.4: Theoretical sum-rate versus the number of users in a system with M = 2 and $\gamma = 0.5$
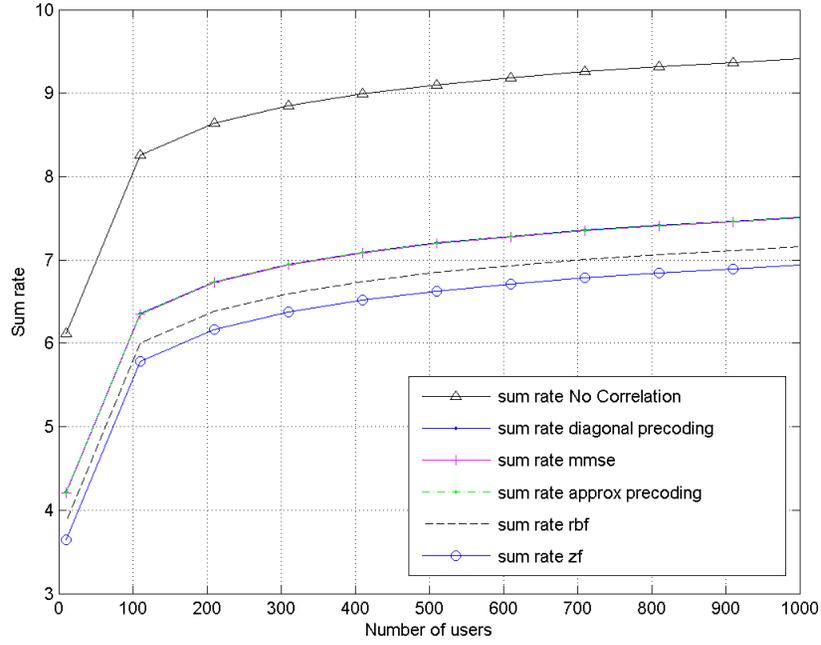
Figure 3.5: Theoretical sum-rate versus the number of users in a system with M = 3 and $\gamma = 0.7$
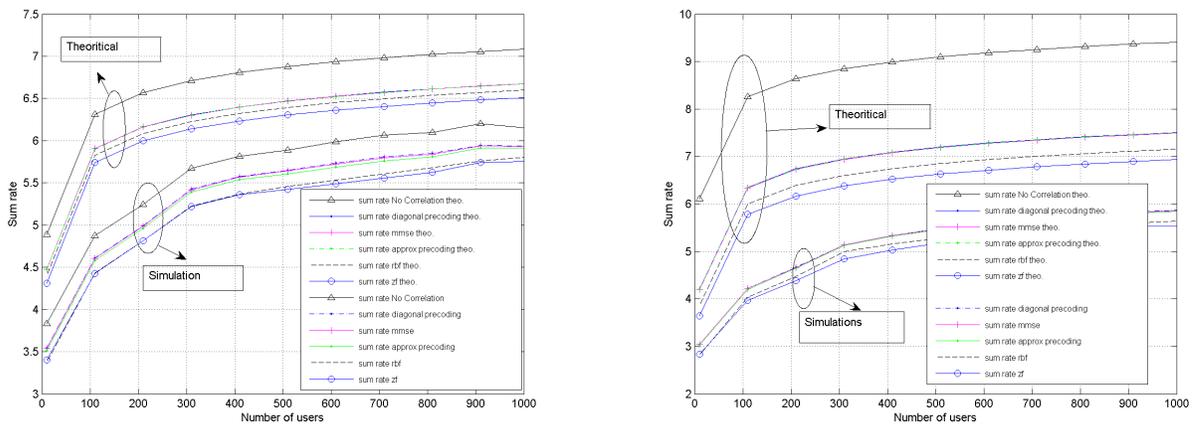
Figure 3.6: Theoretical and experimental sum-rate versus the number of users in a system with M = 2, P = 10, $\gamma = 0.5$ (left) and M = 3, P = 10, $\gamma = 0.7$ (right).
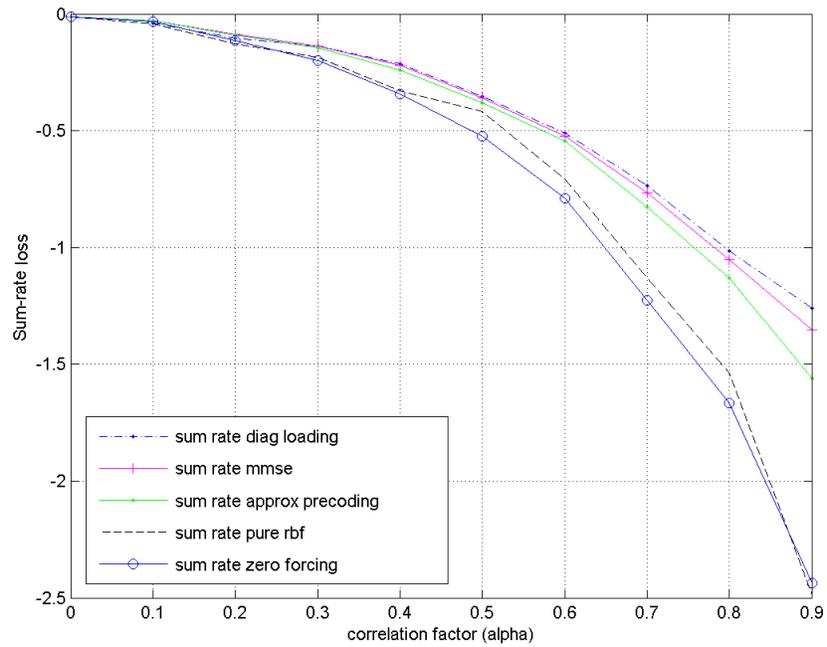
Figure 3.7: Sum-rate loss versus correlation factor γ in a system with M = 2, P=10 and n=400
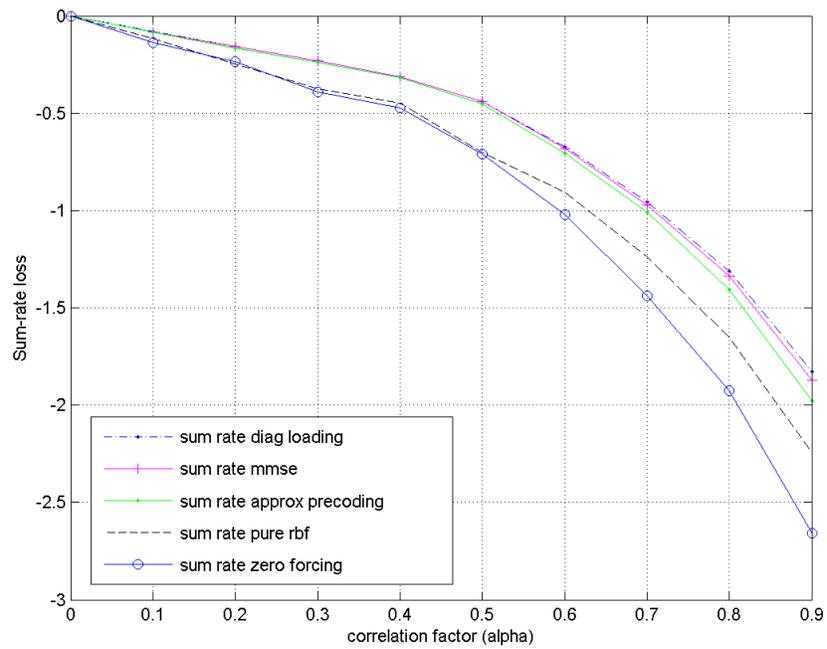
Figure 3.8: Sum-rate loss versus correlation factor γ in a system with M = 3, P=10 and n=100
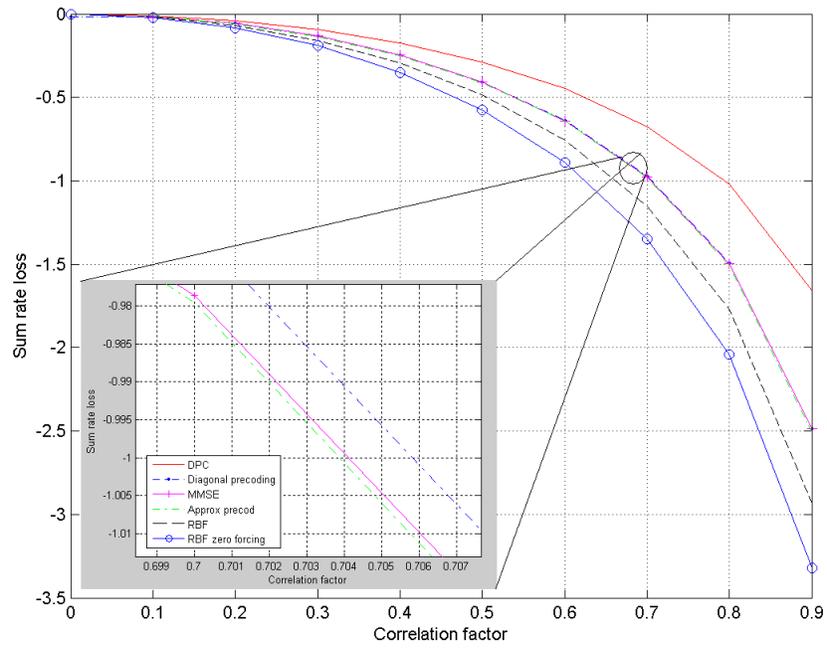
Figure 3.9: Theoretical sum-rate loss versus correlation factor γ in a system with M = 2
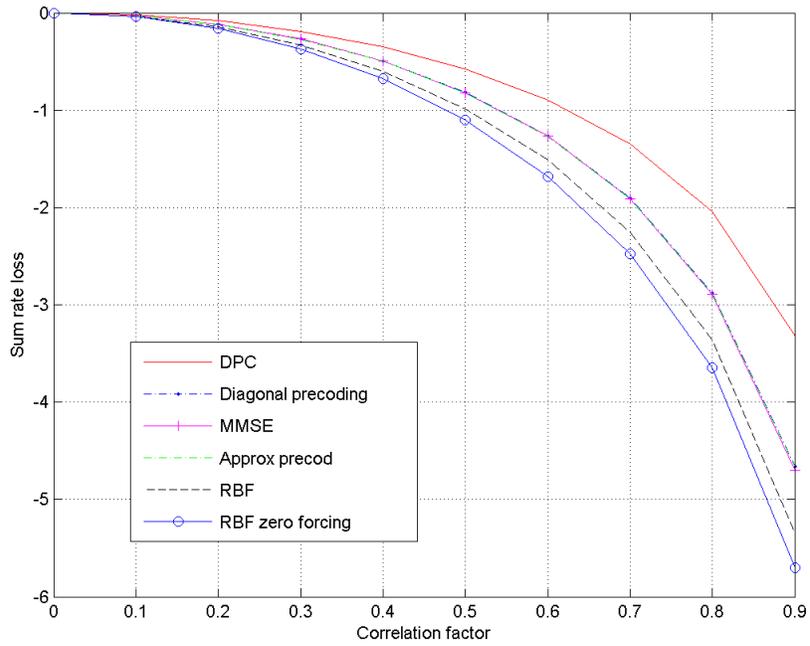
Figure 3.10: Theoretical sum-rate loss versus correlation factor γ in a system with M = 3
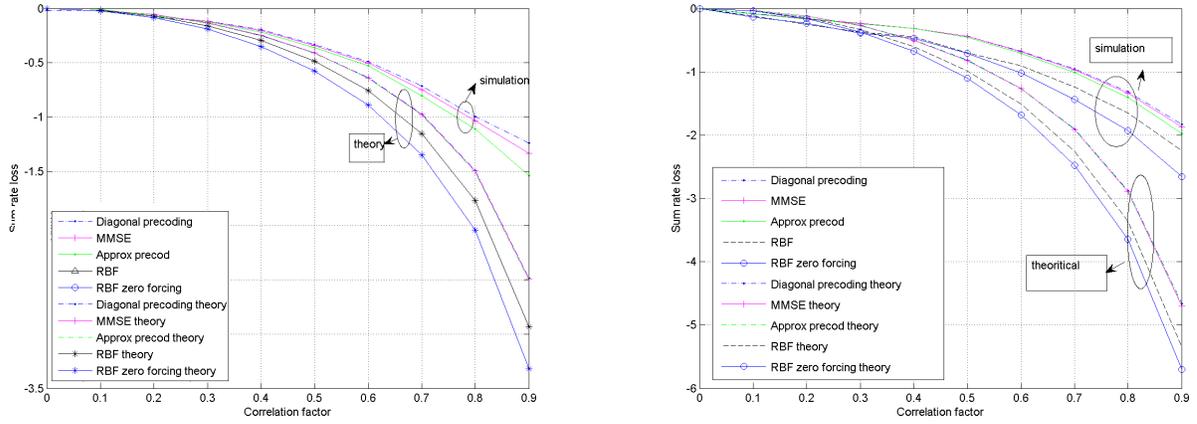
Figure 3.11: Theoretical and experimental sum-rate loss versus correlation factor γ in a system with M = 2, P=10, n=400 (left) and M = 3, P=10, n=100 (right).
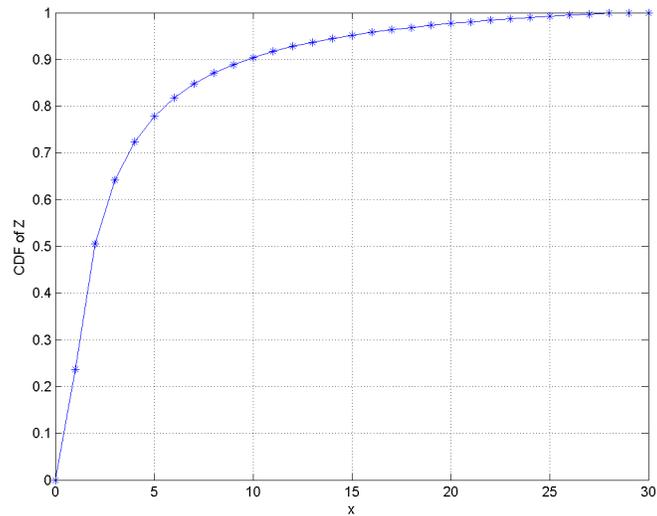
Figure 3.12: CDF of Z for general diagonal matrices **A** and **B**.

pure RBF and RBF with zero forcing. We also showed that diagonal precoding did not yield much optimization as compared with MMSE and approximate precoding. Although diagonal and MMSE precoding scaled well in the presence of channel correlation, we introduced a less computational and rather a direct and approximate technique (approximate precoding) that closely matches the performance of diagonal and MMSE precoding techniques. Simulation and theoretical results show that RBF with diagonal, MMSE and approximate precoding optimized the sum-rate capacity and reduced the SNR hit especially in a very highly correlated channel and hence were able to reduce gap between RBF and DPC.

3.6   Appendix: Calculating the CDF of Z

For convenience of presentation, we evaluate in this section the CDF of a general quantity

$$Z = \frac{\|\phi\|_{\mathbf{A}}^2}{\|\phi\|_{\mathbf{B}}^2}$$

where $\mathbf{A}$ and $\mathbf{B}$ are diagonal matrices. To this end, note that the inequality $Z \leq x$ can be written as

$$\|\phi_m\|_{x\mathbf{B}-\mathbf{A}}^2 \geq 0$$

The CDF is then given by

$$P\{Z \leq x\} \quad = \quad \int_{\|\phi_m\|_{\mathbf{xB}-\mathbf{A}}^2 \geq 0} p(\phi)d\phi = \int p(\phi)u(\|\phi_m\|_{\mathbf{xB}-\mathbf{A}}^2)d\phi \qquad (3.19)$$

where $p(\phi)$ is the pdf of ϕ defined in (3.9)and $u(x)$ is the step function. This integral is very difficult to calculate due the inequality constraint (the unit step function) and due to the delta function. To go around this, we use the following unit step representation [5]

$$u(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{x(j\omega_1+\beta_1)}}{j\omega_1 + \beta}$$

which is valid for any $\beta_1 > 0$. We can thus write

$$u(\|\phi_m\|_{\mathbf{xB}-\mathbf{A}}^2) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{(\|\phi_m\|_{\mathbf{xB}-\mathbf{A}}^2)(j\omega_1+\beta_1)}}{j\omega_1 + \beta_1} d\omega_1$$

We can also replace the delta function with a similar integral representation

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

We thus have the following integral representation of the CDF of Z

$$P\{r \leq x\} = \frac{\Gamma(M)}{4\pi^{M+2}} \int_{-\infty}^{\infty} d\omega_1 \frac{1}{j\omega_1 + \beta_1} \int_{-\infty}^{\infty} d\omega_2 e^{-j\omega_2}$$

$$\int d\phi e^{-\phi^*((\mathbf{A}-\mathbf{x}\mathbf{B})(j\omega_1+\beta_1)-j\omega_2 I)\phi}$$

By inspecting the inner integral, we note that it is similar to the Gaussian density integral. Specifically, we have

$$\frac{1}{\pi^M} \int d\phi e^{-\phi^*((\mathbf{A}-\mathbf{x}\mathbf{B})(j\omega_1+\beta_1)-j\omega_2 I)\phi} =$$

$$\frac{1}{\det((\mathbf{A} - \mathbf{x}\mathbf{B})(j\omega_1 + 1) - j\omega_2 I)}$$

This allows us to write

$$P\{Z \leq x\} = \frac{\Gamma(M)}{4\pi^{M+2}} \int d\omega_1 \frac{1}{j\omega_1 + \beta_1}$$

$$\int d\omega_2 \frac{e^{-j\omega_2}}{\det((j\omega_1 + \beta_1)(\mathbf{A} - \mathbf{x}\mathbf{B}) - j\omega_2 I)}$$

We turn our attention now to the integral with respect to ω_2. To evaluate this integral, we use partial fraction expansion to represent the determinant as

$$\frac{1}{\det((j\omega_1 + \beta_1)(\mathbf{A} - \mathbf{x}\mathbf{B}) - j\omega_2 I)} \tag{3.20}$$

$$= \frac{1}{\prod_{i=1}^{M}((a_i - b_i x)(j\omega_1 + \beta_1) - j\omega_2)} \tag{3.21}$$

$$= \frac{1}{(j\omega_1 + \beta_1)^{M-1}} \sum_{i=1}^{M} \frac{\eta_i}{((a_i - b_i x)(j\omega_1 + \beta_1) - j\omega_2)} \tag{3.22}$$

where

$$\eta_i = \frac{1}{\prod_{k \neq i}((a_k - a_i) - (b_k - b_i)x)}$$

This expansion is valid assuming that $(a_k - a_i)^2 + (b_k - b_i)^2 \neq 0$. We can now residue theory to evaluate the integral with respect to ω_2 as

$$\frac{1}{2\pi} \int d\omega_2 \frac{e^{-j\omega_2}}{\det((j\omega_1 + \beta_1)(\mathbf{A} - \mathbf{x}\mathbf{B}) - j\omega_2 I)} =$$

$$\sum_{i=1}^{M} \eta_i e^{(a_i - b_i x)(j\omega_1 + \beta_1)} u(a_i - b_i x)$$

45

We can thus write

$$
\begin{aligned}
P\{Z \le x\} \quad = \quad & \frac{\Gamma(M)}{4\pi^{M+2}} \int_{-\infty}^{\infty} d\omega_1 \frac{1}{(j\omega_1 + \beta_1)^M} \\
& \sum_{i=1}^{M} \eta_i e^{(a_i - b_i x)(j\omega_1 + \beta_1)} u(a_i - b_i x)
\end{aligned}
$$

We can now use residue theory to show that

$$
\begin{aligned}
P\{Z \le x\} \quad = \quad & \frac{\Gamma(M)}{4\pi^{M+2}} \sum_{i=1}^{M} \eta_i u(a_i - b_i x) \\
& \int_{-\infty}^{\infty} d\omega_1 \frac{e^{(j\omega_1 + \beta_1)(a_i - b_i x)}}{(j\omega_1 + \beta_1)^M} \\
= \quad & \sum_{i=1}^{M} \eta_i (a_i - b_i x)^{M-1} u(a_i - b_i x) u(a_i - b_i x) \\
= \quad & \sum_{i=1}^{M} \eta_i (-a_i + b_i x)^{M-1} u(-a_i + b_i x)
\end{aligned}
$$

Chapter 4

Conclusions, Recommendations, Outcomes, and Publications

4.1 Conclusions

Broadcast channels are getting increased attention as downlink scheduling is a major bottleneck for current wideband wireless systems. It is thus important to devise new scheduling techniques and study the performance of existing ones under various non-idealities.

In this report, we considered 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 $\mathbf{R}$ which is assumed to be non-singular and $Tr(\mathbf{R}) = M$. In this case, the sum-rate of DPC and beamforming schemes will be different. It turns out that in this 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 $\mathbf{R}$.

Our theoretical results and simulations demonstrate that random beamforming is no more able to match DPC for large number of users in the correlated case and that this deviation increases with correlations. The second part of the report thus studied the effect of precoding on random beamforming. Various forms of precoding were introduced to reduce the hit that results from antenna correlation. It was shown that a technique like zero forcing that attempts to whiten the channel worsens the performance of the random beamforming, while MMSE and diagonal beamforming improve performance. An approximate precoding technique was also introduced that results in closed form solution of the precoding matrix and performs as good as diagonal precoding.

One important by-product of our study is that it introduced a new technique for calculating the CDF of ratios of weighted norms of Gaussian random variables. The technique was also extended to the case where the variables are isotropically distributed. (see Sections 2.7 and 3.6.)

4.2   Recommendations

It is important to study the performance of various downlink scheduling techniques under various non-idealities. This report considered the effect of spatial correlation. One limitation of our study, however, is that we assumed that the users share a common correlation matrix. It would be interesting to extend this study to the case where the users have different correlation matrices (or correlation matrices that are a function of a random parameter that varies from one user to another).

The study also showed that in contrast to random beamforming, deterministic beamforming continues to match DPC inspite of correlation. However, the major drawback of deterministic beamforming is that it is not able to provide fairness (thus a strong user will always dominate weaker users). It would be thus interesting to devise hybrid deterministic/random beamforming techniques that exhibit less hit with correlation and yet is capable to schedule users in a fair manner.

Finally, it would be interesting to extend the contour integration technique that we introduced in this report to evaluate the CDF of ratios of quadratic forms to more general scenarios (e.g. to correlated (real) Gaussian variables with nonzero means, to joint distributions of ratios of Gaussian variables, ... etc.). This is a problem that we are currently considering [45].

4.3 Summary of the Outcome of the Project Results

The project resulted in the following outcomes

1. The project studied the effect of spatial correlation on the scaling of dirty paper coding for large number of users.

2. The project studied the effect of spatial correlation on the scaling of deterministic and random beamforming.

3. The project showed that correlation results in an SNR hit on the sum-rate that depends on the eigenvalues of the correlation matrix. The functional dependence of the hit on the eigenvalues varies according to the scheduling technique used.

4. The project concluded that DPC and deterministic beamforming experience the least hit. Moreover, random beamforming is no more able to match the sum-rate of DPC.

5. The project introduced precoding as a technique that could improve the performance of beamforming. While zero-forcing worsens the performance of beamforming, MMSE and the more general diagonal precoding improve random beamforming performance.

6. As a by-product of our study, the project introduced a technique that evaluates the CDF of (indefinite) quadratic forms in Gaussian random variables. The approach by passes the characteristic function approach that is usually used and evaluates the CDF directly in closed form.

4.4 Publications that Resulted from the Project

Here is a summary of the publications that resulted from this work

1. T. Y. Al-Naffouri, M. Sharif, and B. Hassibi " How much does transmit correlation affect the sum-rate of MIMO downlink channels?" *to appear in IEEE Transactions on Communications.*

2. T. Y. Al-Naffouri and B. Hassibi, "On the distribution of indefinite Hermitian quadratic forms in Gaussian random variables," *under preparation for submission to IEEE Transactions on Information Theory*

3. T. Y. Al-Naffouri and B. Hassibi, "On the distribution of indefinite Hermitian quadratic forms in Gaussian random variables," *under preparation for submission to the International Symposium on Information Theory,* 2009.

4. T. Y. Al-Naffouri "Opportunistic beamforming with precoding for spatially correlated channels," *submitted to IEEE Communication Letters.*

5. T. Y. Al-Naffouri, M. Sharif, and Bnn. Hassibi "How much does transmit correlation affect the sum-rate of MIMO downlink channels?" *International Symposium on Information Theory,* Seattle, OR, Jul. 2006.

6. T. Y. Al-Naffouri and M Eltayeb, "Opportunistic beamforming with precoding for spatially correlated channels," *to be submitted to Vehicular Technology Conference,* 2009.

4.5   Talks that Resulted from the Project

This project also resulted in the following talks

1. "Indefinite quadratic forms in Gaussian random variables: Distribution, scaling, and application to the broadcast channel," *Electrical Engineering Department, University of Texas at Dallas,* TX, Sep. 4, 2008.

2. "Indefinite quadratic forms in Gaussian random variables: Distribution, scaling, and application to the broadcast channel," *Electrical Engineering Department, Smart Antenna Research Group, Stanford University,* CA, Aug. 22, 2008.

3. "Scaling laws of multiple antenna (group) broadcast channels," *Electrical Engineering Department, University of California at Irvine,* CA, Jun. 18, 2008.

4. "Scaling laws of multiple antenna (group) broadcast channels," *Electrical Engineering Department, University of Southern California,* CA, Feb. 20, 2008.

5. "How much does correlation affect the sum-rate of MIMO downlink channels? *Institute Eurcom, Sophia-Antipolis,* France, June 21, 2007.

6. "Broadcasting data to multiple user groups: Information theoretic investigation of the wide band case," *Electrical Engineering Department, King Fahd University of Petroleum and Minerals,* Dhahran, Saudi Arabia, May 1st, 2007.

7. "Opportunistic scheduling in wireless networks: An overview of issues and design considerations," (jointly with Dr. Yahya Al-Harthi (KFUPM) and Dr. Mohamed-Slim Alouini (Texas A & M Qatar), Tutorial at the *International Symposium on Signal Processing and its Applications (ISSPA 2007),* Sharjah, UAE, Feb 11, 2007.

8. "The effect of spatial correlation on the capacity of MIMO broadcast channels with partial side information," *Electrical Engineering Department, King Fahd University of Petroleum and Minerals,* Dhahran, Saudi Arabia, Jan. 13, 2007.

9. "How much does correlation affect the sum-rate of MIMO downlink channels? *Electrical Engineering Department, Imperial College,* London, UK, Nov. 23, 2006.

10. "How much does correlation affect the sum-rate of MIMO downlink channels?" *Research Department, Intel Corporation,* Santa Clara, CA, Aug. 22, 2006.

4.6   Master Thesis Related to the Project

Part of this project also contributed to the following master thesis

"Opportunistic Scheduling with Limited Feedback in Wireless Communications Systems" - in progress.

Bibliography

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

[2] E. Telatar, "Capacity of multi-antenna Gaussian channel," *European Trans. Telecommunications*, vol. 10, pp. 585–595, Nov. 1999.

[3] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, Mar. 1998.

[4] L. Zheng and D. N. Tse, "Communications on the Grassman manifold: a geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Info.*, vol. 48, no. 2, pp. 359–384, Feb. 2002.

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

[6] B. Hochwald, T. Marzetta, and V. Tarokh, "Multi-antenna channel-hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inform.*, vol. 50, no. 9, pp. 1893–1909, Sep. 2004.

[7] B. Hochwald, B. Peel, and A. L. Swindlehurst, "A Vector perturbation technique for near capacity multiantenna multiuser communication–Part II: Perturbation," *IEEE Trans. Comm.*, vol. 53, no. 3, pp. 537–545, March 2005.

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

[9] T. Yoo and A. Goldsmith, "On the Optimality of Multi-Antenna Broadcast Scheduling Using Zero-Forcing Beamforming" *IEEE JSAC Special Issue on 4G Wireless Systems*, VOL. 24, NO. 3, MARCH 2006.

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

[11] N. Jindal, S. Jafar, S. Vishwanath, and A. Goldsmith, "Sum power waterfilling for Gaussian broadcast channels," in *Proc. 36th Asilomar Conf. on Sig. and Syst.*, 2002.

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

[13] N. Jindal and A. Goldsmith, "Dirty paper coding vs. TDMA for MIMO broadcast channel," *to appear in Proc. IEEE Inter. Conf. Commu., June 2004.*

[14] M. Fakhereddin, M. Sharif and B. Hassibi, "Throughput analysis in wideband MIMO broadcast channels with partial feedback, *Proc. IEEE Workshop on Signal Proc. for Wireless Comm.,* June 2005.

[15] A. Vakili, M. Sharif, B. Hassibi, "The Effect of Channel Estimation Error on the Throughput of Broadcast Channels," *Proc. of IEEE ICASSP,* Toulouse, France, May 2006.

[16] M. Kountouris and D. Gesbert, "Robust Multi-User Opportunistic Beamforming for Sparse Networks", *Proc. of IEEE Workshop on Signal Processing Advances in Wireless Communications, (SPAWC) 2005*, New York, NY.

[17] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian MIMO broadcast channel," "*Proc. IEEE ISIT*, June 2004.

[18] P. Viswanath and D. N. Tse, "Sum capacity of the vector Gaussian broadcast channel and downlink-uplink duality," *IEEE Trans. Inform.*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.

[19] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates and sum rate capacity of Gaussian MIMO broadcast channel," *IEEE Trans. Inform.*, vol. 49, no. 10, pp. 2658–2668, 2003.

[20] H. Lütkepohl, *Handbook of Matrices,* John Wiley & Sons, 1996.

[21] D. Park and S Y. Park, "Effect of Transmit Antenna Correlation on Multiuser Diversity," in *Proc. of IEEE ISIT*, 2005.

[22] W. Yu and J. M. Cioffi, "Trellis precoding for the broadcast channel," in *Proc. IEEE Glob. Comm. Conf.*, 2001.

[23] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Info.*, vol. 48, no. 6, pp. 1250–1277, June 2002.

[24] C. B. Peel, B. Hochwald, and A. L. Swindlehurst, "A vector perturbation technique for near capacity multi-antenna multi-user communication- Part I: Channel inversion and regularization," *IEEE TRANSACTIONS ON COMMUNICATIONS* ,VOL. 53, NO. 1, JANUARY 2005.

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

[26] M. Sharif and B. Hassibi "On the capacity of MIMO broadcast channel with partial side information" *IEEE Trans. Info. Theory*, vol. 51, no. 2, Feb. 2005.

[27] M. Sharif and B. Hassibi, "Scaling laws of sum rate using time-sharing, DPC, and beamforming for MIMO broadcast channels," in *Proc. of IEEE ISIT*, 2004.

[28] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE Jour. Selec. Areas. Commu.*, vol. 21, no. 5, pp. 684–702, June 2003.

[29] J. Chul Roh and B. Rao, "Multiple antenna channels with partial feedback," in *Proc. of IEEE ICC*, 2003.

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

[31] 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, Academic Press, New York*, pp. 346–353, 1954.

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

[33] M. Sharif, A. F. Dana, A. Vakili, and B. Hassibi," *Differentiated rate scheduling for the downlink of cellular system*," *IEEE TRANSACTIONS ON COMMUNICATIONS*, ,VOL. 56, NO. 10, OCTOBER 2008.

[34] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press Inc., London, 1965.

[35] A. Edelman, "Eigenvalues and condition numbers of random matrices," *MIT PhD Dissertation*, 1989.

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

[37] Q. H. Spencer, et. al., "An introduction to the multiuser MIMO downlink," *IEEE Communications Magazine,* pp. 60-67, Oct. 2004.

[38] G. Tziritas, "On the distribution of positive-definite Gaussian quadratic forms," *IEEE Transactions on Information Theory,* no. 6, vol., pp. 895-906, Nov. 1987.

[39] D. Raphaeli, "Distribution of noncentral indefinite quadratic forms in complex normal variables," IEEE Transactions on Information Theory, no. 3, vol. 42, May 1996, pp. 1002–1007.

[40] N. Shah and **H**. Li, "Distribution of quadratic form in Gaussian mixture variables and an application in relay networks," *IEEE 6th Workshop on Signal Processing Advances in Wireless Communications,* pp. 490-494, June 2005.

[41] K. **H**. Biyari and W. C. Lindsey, "Statistical distributions of Hermitian quadratic forms in complex Gaussian variables," *IEEE Transactions on Information Theory,* no. 3, vol. 39, pp. 1076-1082, May 1993.

[42] M.K. Simon, M.-S. Alouini "On the difference of two chi-square variates with application to outage probability computation," *IEEE Transactions on Communications,* no. 11, vol. 49, pp. 1946–1954, Nov. 2001.

[43] T. Al-Naffouri and M. El-Tayeb "Opportunistic Beamforming with Precoding for Spatially Correlated Channels," *IEEE Transactions on Communications,* the 11th Canadian Workshop on Information Theory (CWIT), May, 2009.

[44] **H**. Holm and M.-S. Alouini, "Sum and difference of two squared correlated Nakagami variates in connection with the McKay distribution," *IEEE Transactions on Communications,* no. 8, vol. 52, pp. 1367-1376, Aug. 2004.

[45] T. Y. Al-Naffouri and B. Hassibi, "On the distribution of indefinite Hermitian quadratic forms in Gaussian random variables," *under preparation for submission to IEEE Transactions on Information Theory*