Joint Power Control and Beamforming Codebook Design for MISO Channels under the Outage Criterion

Behrouz Khoshnevis and Wei Yu

Abstract

This paper investigates the joint design of power control and beamforming codebooks for limited-feedback multiple-input single-output (MISO) wireless systems. The problem is formulated as the minimization of the outage probability subject to the transmit power constraint and cardinality constraints on the beamforming and power codebooks. We show that the two codebooks need to be designed jointly in this setup, and provide a numerical method for the joint optimization. For independent and identically distributed (i.i.d.) Rayleigh channel, we also propose a low-complexity approach of fixing a uniform beamforming codebook and optimizing the power codebook for that particular beamformer, and show that it performs very close to the optimum. Further, this paper investigates the optimal tradeoffs between beamforming and power codebook sizes. We show that as the outage probability decreases, optimal joint design should use more feedback bits for power control and fewer feedback bits for beamforming. The jointly optimized beamforming and power control modules combine the power gain of beamforming and diversity gain of power control, which enable it to approach the performance of the system with perfect channel state information as the feedback link capacity increases — something that is not possible with either beamforming or power control alone.

Index Terms

Beamforming, limited feedback, multiple antennas, outage probability, power control.

The authors are with the Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, Ontario, Canada M5S 3G4 (email: bkhoshnevis@comm.utoronto.ca; weiyu@comm.utoronto.ca). The material in this paper is presented in part at IEEE Global Communications Conference (Globecom) 2009, Hawaii, USA.

I. INTRODUCTION

It is well established that the use of multiple antennas can considerably improve the performance of wireless systems in terms of reliability and capacity. A complete realization of these benefits, however, requires channel state information at the transmitter (CSIT). This motivates the study of *limited feedback* systems where the receiver quantizes and sends back the channel state information needed by the transmitter through a rate-limited feedback link. The study of limited feedback schemes is especially relevant for frequency division duplex (FDD) systems, where downlink and uplink transmissions use different frequency bands; hence the transmitter cannot learn the channel via reciprocity. During the past decade, a great amount of research has been done on limited feedback systems, both for single-user [1]–[16] and multiuser scenarios [17]–[26].

From a broader perspective, the limited-feedback communication systems can be categorized as control systems, where the transmitter uses the feedback information provided by the receiver in order to optimize a certain objective, e.g. maximize the transmission rate or minimize the outage probability. In this sense, the study of limited-feedback systems can be related to the analysis of control systems with limited communication capacity between the sensors and controllers [27]–[29]. The readers are referred to [30] for a review of the literature on limited-feedback communication systems.

This paper focuses on optimal design of single-user limited-feedback systems over multiple-input single-output (MISO) fading channels. In this regard, the authors of [31]–[33] show that, in order to maximize the mutual information in each fading block, the transmitter should use Gaussian inputs, which are completely characterized by their covariance matrices. Therefore, the optimal feedback strategy is to share a codebook of transmit covariance matrices between the transmitter and the receiver, where the receiver chooses the best covariance matrix based on the current channel realization and sends its index to the transmitter.

Unfortunately, the covariance codebook design is a large optimization problem and requires rather complicated numerical design algorithms [31], [32]. To simplify the design process, researchers usually resort to rank-one covariance matrices, which can be implemented by a power control module followed by a beamforming module [1]–[5], [8]–[11]. The rank-one simplification is justified by the fact that with perfect CSIT and a single-antenna receiver, the optimal transmit covariance matrix is a rank-one matrix, i.e., joint beamforming and power control is asymptotically optimal as the number of feedback bits increases and the limited-feedback system approaches perfect-CSIT system.

In spite of the rank-one simplification, the joint design of beamforming and power control modules is still a complicated optimization problem itself. As a result, most of the existing literature treats the beamforming and power control aspects of the problem separately and focuses on *independent* design of these modules. For instance, the papers that focus on power control assume isotropic beamforming at transmitter and investigate the optimal structure of the power control codebooks [8]–[11]. The papers that focus on beamforming, on the other hand, fix the transmission power and investigate the optimal beamforming codebook [1]–[5].

The independent design of beamforming and power control modules however introduces a significant performance loss that is overlooked by the earlier literature. Furthermore, as described in the next section, such an independent design, even with infinite CSI feedback capacity, will have a non-zero performance gap with respect to a perfect-CSIT system. In order to address this issues, this paper takes a fresh look at joint optimization of beamforming and power control modules for limited-feedback MISO systems. In particular, we study the design problem from an outage capacity perspective, which is the appropriate performance metric for delay-constrained real-time traffic [34], [35].

We formulate the optimization problem as minimization of the outage probability subject to an average power constraint at the transmitter. The optimization is over the beamforming and power control codebooks as well as the corresponding codebook sizes. We first fix the codebook sizes and express the constrained optimization in a Lagrangian formulation. The resulting unconstrained problem is then solved using a combination of Lloyd's algorithm, Broyden-Fletcher-Goldfarb-Shanno (BFGS) method, and a sequential approximation of the outage probability function based on Monte Carlo integration. The main contributions of this paper are as follows:

1) We prove the necessity of jointly optimized beamforming and power control modules by showing the performance gap incurred by an independent design of these modules.
2) The paper provides an efficient algorithm for joint optimization of the beamforming and power control codebooks, which is then used to derive the corresponding optimal codebook sizes in terms of the CSI feedback capacity and the target outage probability.
3) It is shown that as the outage probability decreases, the optimal power codebook size increases and the optimal beamforming codebook size decreases. Furthermore, the resulting optimal codebook sizes are shown to be independent of the target rate.
4) Numerical results are provided to show that the jointly optimized feedback scheme combines the power gain of beamforming and diversity gain of power control. This enables the overall

system performance to approach the performance of a perfect-CSIT system as the feedback link capacity increases; something that is not possible with independent beamforming and power control design.

The remainder of this paper is organized as follows. Section II describes the system model and explains the motivations behind the joint optimization problem. Section III, justifies the necessity of a joint beamforming and power control codebook design and provides a suboptimal design solution. The joint codebook optimization problem is then formulated and solved in Section IV. The corresponding codebook size optimization problem is addressed in Section V. Finally, the numerical results are presented in Section VI followed by conclusions in Section VII.

Notations: $\mathbb{C}$ and $\mathbb{R}_+$ denote the set of complex numbers and nonnegative real numbers. Bold upper case and lower case letters denote matrices and vertical vectors. $\mathbf{I}_M$ denotes the $M \times M$ identity matrix. $\mathrm{Tr}(\cdot)$ denotes the trace operation. $\|\cdot\|$ denotes the Euclidean norm of a vector. $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^\dagger$ denote respectively the transpose, the complex conjugate, and the transposed complex conjugate of a vector or a matrix. $\mathcal{CN}(0, \mathbf{I}_M)$ represents a circularly symmetric complex Gaussian distribution with zero mean and covariance matrix $\mathbf{I}_M$. $\mathrm{E}[\cdot]$ denotes the expectation operation and Prob$[\cdot]$ denotes the probability of an event.

II. System Model

This paper considers limited-feedback single-user multiple-input single-output (MISO) wireless systems. We assume a block-fading channel model, where the channel realizations are i.i.d. over different fading blocks. The system is assumed to have perfect channel state information at the receiver (CSIR). A delay-free noiseless feedback link with a finite capacity of B bits per fading block is available from the receiver to the transmitter as illustrated in Fig. 1.

Let $\mathbf{h} \in \mathbb{C}^M$ denote the channel vector from the transmitter to the receiver, where M is the number of transmit antennas. In each fading block, the receiver perfectly estimates its channel $\mathbf{h}$, chooses an appropriate transmission power level $P(\mathbf{h})$ and beamforming vector $\mathbf{u}(\mathbf{h})$ from the corresponding codebooks, and sends the corresponding codeword indices back to the transmitter. The problem is to optimize these codebooks and the corresponding codebook sizes with the objective of minimizing the outage probability subject to a power constraint at the transmitter.

The exact formulation for the joint codebook optimization problem and our solution approach are presented in Sections IV and V. Here, we emphasize on the motivations behind studying such a problem. The main motive behind this joint optimization is two-fold:

1) Mere beamforming or power control, even with perfect CSIT, is not sufficient for the optimal performance. In other words, one needs both modules to be present and function in order to approach the optimal performance of a perfect-CSIT system as the feedback rate increases.
2) An independent optimization of the two modules incurs a significant penalty on the system performance, hence a joint optimization of the modules is necessary.

To verify the first point, Fig. 2 plots the system performance when one applies mere beamforming (with fixed transmission power), as in [1]–[5], or mere power control (with isotropic beamforming), as in [8]–[11]. It is well established that by applying beamforming, one gains a power gain of $10\log_{10}(M) - \kappa\alpha^{-B}$ in dB, asymptotically as $B \to \infty$, where $\kappa > 0$ and $\alpha > 1$ depend on the system setup and M is the number of transmit antennas [1]–[5]. On the other hand, by applying power control, it can be shown that the diversity order, i.e. the slope of outage probability vs. SNR, improves as $\frac{M}{M-1}\left(M^{2^B}-1\right) \approx M^{2^B}$ as B increases [9]. However, neither the power gain of beamforming nor the diversity gain of power control is sufficient by itself to approach the optimal performance of a perfect-CSIT system as B increases, i.e. to traverse the gap between the rightmost and leftmost curves in Fig. 2. To do so, both beamforming and power control modules should be present and the feedback bits should be appropriately divided between the two in order to achieve a combination of power and diversity gains.

The second point in the list of motives above argues that, not only the beamforming and power control modules need to be used jointly, they also need to be optimized jointly in the design process. The necessity of such a joint optimization is addressed in the next section.

III. Optimal Power Control for Fixed Beamforming Module

In this section, the beamforming module is fixed and the power control module is optimized for the given fixed beamforming module. The purpose of this optimization is two fold:

1) It is shown through an example that if the power control module is designed independent of the beamforming module, a considerable power loss is incurred as compared to the performance of a power controller that is specifically designed and optimized for the given beamforming module. This proves the necessity of the joint design of the modules.
2) The beamforming module along with the optimized power control module serve as a suboptimal solution for the joint optimization problem discussed in Section IV.

We first formulate the problem of *power control with limited feedback* in its general form and describe the optimization process. Next, we demonstrate the power gains associated with optimizing the power controller for the given beamforming module. Finally, we provide a suboptimal solution for the joint codebook design problem.

A. Power Control with Limited Feedback

Consider the MISO channel in Fig. 3 with M transmit antennas and channel vector $\mathbf{h} \in \mathbb{C}^M$. The vector $\mathbf{x}$ is the output of some fixed beamforming module constrained by $\mathrm{Tr}\left(\mathbf{Q}_{\mathbf{x}|\mathbf{h}}\right)=1$, where $\mathbf{Q}_{\mathbf{x}|\mathbf{h}}=\mathrm{E}\left[\mathbf{x}\mathbf{x}^\dagger|\mathbf{h}\right]$ is the transmit covariance matrix. The receiver knows $\mathbf{h}$ perfectly and chooses the power level, to be used by the transmitter, from the power codebook $\mathbb{P}=\left\{P_1, P_2, \cdots, P_{N_{pc}}\right\}$, where N_{pc} is the number of available power levels.

We want to optimize the power control module such that the probability of outage is minimized for a given target rate R. This involves optimizing the power codebook $\mathbb{P}$ and the quantization function $P(\mathbf{h}) : \mathbb{C}^M \to \mathbb{P}$ that maps the channel realizations to the power codebook:

$$\begin{aligned} \min_{\mathbb{P}, P(\mathbf{h})} \quad & \mathrm{Prob}\left[\log_2\left(1 + P(\mathbf{h})\mathbf{h}^T\mathbf{Q}_{\mathbf{x}|\mathbf{h}}\mathbf{h}^*\right) < R\right] \\ \text{s.t.} \quad & \mathrm{E}[P(\mathbf{h})] \leq \mathsf{SNR}. \end{aligned} \tag{1}$$

Here, SNR denotes the normalized transmitter power constraint. The function $P(\cdot)$ is also referred to as the *power control function* in this paper.

Define

$$\gamma = \mathbf{h}^T\mathbf{Q}_{\mathbf{x}|\mathbf{h}}\mathbf{h}^* \tag{2}$$

as the effective channel gain. The dependence of γ on $\mathbf{h}$ is solely determined by the beamforming module, e.g.:

- *Isotropic beamforming:* $\mathbf{Q}_{\mathbf{x}|\mathbf{h}}=\frac{1}{M}\mathbf{I}_M$ and $\gamma=\frac{1}{M}\|\mathbf{h}\|^2$.
- *Matched-channel beamforming:* $\mathbf{Q}_{\mathbf{x}|\mathbf{h}}=\hat{\mathbf{h}}^*\hat{\mathbf{h}}^T$, where $\hat{\mathbf{h}}=\frac{\mathbf{h}}{\|\mathbf{h}\|}$. For this case $\gamma=\|\mathbf{h}\|^2$.
- *Limited-feedback beamforming:* $\mathbf{Q}_{\mathbf{x}|\mathbf{h}} = \mathbf{u}(\mathbf{h})\mathbf{u}(\mathbf{h})^\dagger$, where the unit vector $\mathbf{u}(\mathbf{h})$ belongs to some fixed beamforming codebook. For this case $\gamma = \left|\mathbf{h}^T\mathbf{u}(\mathbf{h})\right|^2$.

With this definition of the random variable γ, the quantization function $P(\mathbf{h})$ can be equivalently reformulated as $P(\gamma):\mathbb{R}_+\to\mathbb{P}$, and the problem (1) simplifies to the following:

$$\begin{aligned} \min_{\mathbb{P}, P(\gamma)} \quad & \mathrm{Prob}\left[P(\gamma) \cdot \gamma < c\right] \\ \text{s.t.} \quad & \mathrm{E}[P(\gamma)] \leq \mathsf{SNR}, \end{aligned} \tag{3}$$

where $c{=}2^R{-}1$, $\gamma{\geq}0$ is the quantization variable. It can be easily verified that changing the domain of the quantization function $P(\cdot)$ from $\mathbf{h}$ in (1) to γ in (3) does not change the problem solution, i.e. the two problems give the same optimal codebook and outage probability.

The general approach to codebook design problems, as in (3), is the Lloyd's algorithm, which calls for repeated updates of the codebook and the quantization function (or equivalently the quantization regions[1]) in an iterative process. The Lloyd's algorithm, however, is not needed for solving problem (3), because given the codebook $\mathbb{P}$, the structure of the optimal quantization (or power control) function $P(\gamma)$ can be derived using fairly simple arguments [10], [11]. These arguments are based on two facts: 1) when outage is inevitable, we should use the smallest power level in the codebook; 2) in order to prevent an outage, we should use the smallest power level needed to do so.

Fig. 4 shows the structure of the optimal power control function for power levels ordered in ascending order $0 \leq P_1 \leq P_2 \leq \cdots \leq P_{N_{pc}}$. The optimal power control function $P(\gamma)$ can, in fact, be considered as a step-like approximation of the optimal power control function with perfect CSIT, $P_{\text{CSIT}}(\gamma)$, which is shown in [35], [36] to be the *truncated channel inverting function* (Fig. 4). This justifies, for example, why $P(\gamma)$ uses the smallest power level both for very small and very large values of γ.

Having identified the power control function in Fig. 4, the outage probability and average power can be expressed as

$$\text{Prob}\left[P(\gamma)\cdot\gamma < c\right] = \text{Prob}[\gamma \in [0, c/P_{N_{pc}}]] = \text{Prob}[\gamma < c/P_{N_{pc}}] = F_\Gamma\left(c/P_{N_{pc}}\right), \tag{4}$$

$$\text{E}[P(\gamma)] = P_1 F_\Gamma\left(c/P_{N_{pc}}\right) + P_1\left(1 - F_\Gamma\left(c/P_1\right)\right) + \sum_{k=2}^{N_{pc}} P_k\left(F_\Gamma\left(c/P_{k-1}\right) - F_\Gamma\left(c/P_k\right)\right), \tag{5}$$

where $F_\Gamma(\cdot)$ is the cumulative distribution function (CDF) of γ. Combining (3), (4), and (5), the problem is now directly expressed in terms of the power levels. To solve this problem, however, we need the CDF of γ, which is not available in most cases. For example, for limited feedback beamforming case, it would be very difficult, if not impossible, to find a closed-form expression for the CDF of $\gamma{=}\left|\mathbf{h}^T\mathbf{u}(\mathbf{h})\right|^2$ given the CDF of $\mathbf{h}$ and definition of $\mathbf{u}(\mathbf{h})$.

To resolve this matter, this paper proposes an approach based on the interpolation of the CDF. In order to preserve the monotonicity and continuity of the first derivative, we use monotone piecewise cubic Hermite interpolation [37]. With the interpolation of $F_\Gamma(\cdot)$ in place, we then solve (3) using the primal-dual interior-point method [38].

[1]The quantization regions or cells are the inverse image of the codebook $\mathbb{P}$ under the mapping $P(\cdot)$.

In the following sections, the process of optimizing the power control module for a fixed beamforming module, as described in this section, will be referred to as *matching* the power control module to the beamforming module.

B. Gain of Matching the Power Control Module to the Beamforming Module

In this section, we present an example to show the gain associated with a matched power controller in comparison with a general (unmatched) power controller.

Consider the *matched-channel beamforming module* with $\gamma_1(\mathbf{h})=\|\mathbf{h}\|^2$ and a power controller matched to this module. Let us denote these modules as BF_1 and PC_1. Assume that we are bound to use PC_1 as the power controller no matter what the beamforming module is.

Consider, as the second beamformer, a *limited-feedback beamforming module* BF_2 with the beamforming codebook $\mathbb{U}=\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}$, the columns of $\mathbf{I}_4$, and beamforming function

$$\mathbf{u}(\mathbf{h}) = \arg\max_{\mathbf{u}\in\mathbb{U}} \left|\mathbf{h}^T\mathbf{u}\right|^2.$$

Denote the effective channel gain of BF_2 as $\gamma_2(\mathbf{h})$ and the corresponding matched power controller as PC_2.

We want to compare the performance of the matched pair $(\mathrm{BF}_2, \mathrm{PC}_2)$ with the performance of the unmatched pair $(\mathrm{BF}_2, \mathrm{PC}_1)$. In order to make PC_1 applicable to BF_2, the transmitter should compensate for the power loss of limited-feedback beamforming BF_2 with respect to the matched-channel beamforming BF_1. Since the exact channel is not known at the transmitter, it should compensate for the maximum possible loss, which is $L=\max_{\mathbf{h}} \gamma_1(\mathbf{h})/\gamma_2(\mathbf{h})$. Therefore, when PC_1 is used with BF_2, the output of the beamforming codebook needs to be multiplied by $\sqrt{L}$. In this example $L=4\approx 6$dB.

Fig. 5 compares the performance of matched and unmatched modules for Rayleigh i.i.d. channel $\mathbf{h}\sim\mathcal{CN}(0,\mathbf{I}_M)$, $M=4$ antennas, $B=2$ bits, and $R=1$ bits/sec/Hz. As Fig. 5 shows, we gain almost 3dB by using the matched power controller for BF_2, i.e. using PC_2 instead of PC_1. This example shows that there is a considerable gain associated with optimizing the power controller for the beamforming module and this illustrates the necessity of the joint design of the beamforming and power control modules.

C. Suboptimal Joint Codebook Design

The discussion above motivates a suboptimal joint codebook design, where we fix an appropriate beamforming codebook and match the power control codebook to it.

As an example, for i.i.d. Rayleigh channel, $\mathbf{h} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$, the channel direction $\hat{\mathbf{h}} = \frac{\mathbf{h}}{\|\mathbf{h}\|}$ is uniformly distributed on the unit complex hypersphere. Therefore, roughly speaking, the beamforming vectors should be uniformly spread on the hypersphere[2].

Define the uniform codebook as $\mathbb{U}^{(\text{uni})} = \left\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{N_{bf}}\right\}$ that maximizes the expected value of received signal-to-noise ratio (with fixed transmission power):

$$\mathrm{E}\left[\left|\mathbf{h}^T \mathbf{u}(\mathbf{h})\right|^2\right],$$

where

$$\mathbf{u}(\mathbf{h}) = \arg\max_{\mathbf{u}\in\mathbb{U}^{(\text{uni})}} \left|\mathbf{h}^T \mathbf{u}\right|^2. \tag{6}$$

Lloyd's algorithm can then be easily used to obtain such a uniform codebook, since there is a closed-form expression for the optimum beamforming vectors for fixed quantization regions[3]. Alternatively, one could use any other meaningful criterion, as in [1]–[5], for the definition of the uniform codebook.

Let $\mathbf{H}$ be a training set with S realizations of the channel vector and let N_{bf} and N_{pc} denote the beamforming and power control codebook sizes. We now propose a suboptimal algorithm for joint power control and beamforming codebook design as follows:

Algorithm 1:

1) Generate a uniform beamforming codebook $\mathbb{U}^{(\text{uni})}$ of size N_{bf}.
2) Generate S values of the variable $\gamma = \left|\mathbf{h}^T \mathbf{u}(\mathbf{h})\right|^2$ for $\mathbf{h} \in \mathbf{H}$ and $\mathbf{u}(\mathbf{h})$ defined in (6).
3) Interpolate $F_\Gamma(\gamma)$, the CDF of γ.
4) Solve problem (3) with the objective and the constraint functions defined in (4) and (5) with multiple random start points.

The cubic Hermite interpolating function in step 3 is a monotonically increasing piecewise cubic function with continuous derivatives at the extremes of the interpolation intervals.

By using Algorithm 1 with different values of SNR, we can derive the suboptimal curve of outage probability vs. SNR. As it is shown in Section IV, this algorithm performs very close to the joint

[2]This, in part, justifies why the beamforming codebook design problems in the literature with different performance criteria lead to similar design criteria [2]–[4].

[3]If the channel vectors of a quantization region are placed into the columns of a matrix $\mathbf{A}$, the optimum beamforming vector for that region is the dominant eigenvector of $\mathbf{A}^*\mathbf{A}^T$.

optimization algorithm. This result is attractive numerically, since the joint optimization is much more complex than merely optimizing the power codebook for a fixed chosen beamforming codebook.

IV. Jointly Optimal Power Control and Beamforming

In this section, we formulate the general problem of joint optimization of the power control and beamforming modules and present our numerical solution. The codebook sizes are assumed fixed in this section. The optimization of the codebook sizes is deferred to Section V.

A. Problem Formulation

Consider the limited-feedback system with M transmit antennas in Fig. 6. The transmitter and receiver share a power codebook

$$\mathbb{P} = \left\{P_1, P_2, \cdots, P_{N_{pc}}\right\}$$

and a beamforming codebook

$$\mathbb{U} = \left\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{N_{bf}}\right\} \subset \mathcal{U}_M,$$

where $\mathcal{U}_M$ is the unit hypersphere in $\mathcal{C}^M$ and N_{bf} and N_{pc} are the beamforming and power codebook sizes, respectively.

The receiver has perfect CSIR and chooses the appropriate beamforming vector $\mathbf{u}(\mathbf{h}) \in \mathbb{U}$ and power level $P(\mathbf{h}) \in \mathbb{P}$ and sends the index of the corresponding $(P(\mathbf{h}), \mathbf{u}(\mathbf{h}))$ pair back to the transmitter. The transmitter multiplies the output of its scalar encoder by $\sqrt{P(\mathbf{h})}\mathbf{u}(\mathbf{h})$ and transmits it through its antennas. Note that we need $N_{bf}N_{pc} \leq 2^B$ so that the transmitter can distinguish between different $(P(\mathbf{h}), \mathbf{u}(\mathbf{h}))$ pairs.

The problem is to optimize the beamforming codebook $\mathbb{U}$, the power codebook $\mathbb{P}$, the beamforming function $\mathbf{u}(\mathbf{h}) : \mathbb{C}^M \to \mathbb{U}$, and the power control function $P(\mathbf{h}) : \mathbb{C}^M \to \mathbb{P}$, such that the outage probability is minimized. Following the same line as problem (1), this problem can be formulated as

$$\begin{aligned} \min_{\mathbb{U},\mathbb{P},\mathbf{u}(\mathbf{h}),P(\mathbf{h})} \quad & \text{Prob}\left[P(\mathbf{h}) \cdot \left|\mathbf{h}^T\mathbf{u}(\mathbf{h})\right|^2 < c\right] \\ \text{s.t.} \quad & \mathrm{E}[P(\mathbf{h})] \leq \mathsf{SNR}, \end{aligned} \tag{7}$$

where $c = 2^R - 1$ and the beamforming vectors are constrained by $\|\mathbf{u}_i\| = 1$, $1 \leq i \leq N_{bf}$.

Our approach for solving (7) is based on the Lloyd's algorithm applied to a Lagrangian formulation of the constrained problem. The norm constraint on the beamforming vectors can be eliminated by the change of variables $\mathbf{u}_i = \frac{\mathbf{v}_i}{\|\mathbf{v}_i\|}$ for $1 \leq i \leq N_{bf}$, where $\mathbf{v}_i$'s are unconstrained vectors. In order to incorporate the transmit power constraint into the objective we introduce a Lagrange multiplier λ and rewrite the problem as

$$\min_{\mathbb{U},\mathbb{P},\mathbf{u}(\mathbf{h}),P(\mathbf{h})} \text{Prob}\left[P(\mathbf{h})\left|\mathbf{h}^T\mathbf{u}(\mathbf{h})\right|^2 < c\right] + \lambda \mathrm{E}[P(\mathbf{h})]. \tag{8}$$

This is a natural approach for constrained quantization, e.g. see entropy/memory-constrained vector quantization in [39], [40]. The validity of this approach depends on the convexity structure of the optimization problem (7). For the Lagrangian approach to work, the minimized outage probability in the objective function in (7) needs to be a convex function of the constraint SNR.

Unfortunately, the convexity structure for (7) appears to be difficult to establish. Nevertheless, it is possible to prove that convexity does hold as $B \to \infty$. This is shown in the Appendix and it justifies the asymptotic optimality of the proposed approach.

B. *Numerical Solution*

Let us define the outage probability and average power functions as

$$p_{out} = \text{Prob}\left[P(\mathbf{h})\left|\mathbf{h}^T\mathbf{u}(\mathbf{h})\right|^2 < c\right], \tag{9}$$

$$P_{ave} = \mathrm{E}[P(\mathbf{h})]. \tag{10}$$

In order to apply the Lloyd's algorithm to problem (8), we need to express the objective in the form of an *average distortion function*:

$$p_{out} + \lambda P_{ave} = \mathrm{E}\left[D(\mathbf{h})\right], \tag{11}$$

where

$$D(\mathbf{h}) = \mathbf{1}_c\Big(P(\mathbf{h})\left|\mathbf{h}^T\mathbf{u}(\mathbf{h})\right|^2\Big) + \lambda P(\mathbf{h}) \tag{12}$$

and the indicator function $\mathbf{1}_c(\cdot)$ is defined as

$$\mathbf{1}_c(x) = \begin{cases} 1 & \text{if } x < c, \\ 0 & \text{if } x \geq c. \end{cases} \tag{13}$$

The Lloyd's algorithm starts with a random codebook and iteratively updates the quantization regions and the quantization codebook. Assume that we have a total of S realizations of the channel vector and denote the set of realizations by $\mathbf{H}$. The two steps of the Lloyd's algorithm are described in the following:

1. Updating the regions: The beamforming and power codebooks $\mathbb{U}$, $\mathbb{P}$ are fixed. For each $\mathbf{h} \in \mathbf{H}$, the beamforming vector and the power level are chosen such that the distortion function is minimized:

$$\mathbf{u}(\mathbf{h}) = \arg\max_{\mathbf{u}\in\mathbb{U}} \left|\mathbf{h}^T\mathbf{u}\right|^2, \tag{14}$$

$$P(\mathbf{h}) = \arg\min_{P\in\mathbb{P}} \mathbf{1}_c\Big(P\left|\mathbf{h}^T\mathbf{u}(\mathbf{h})\right|^2\Big) + \lambda P, \tag{15}$$

and based on this, the quantization regions are formed as

$$\mathcal{H}_{ij} = \left\{\mathbf{h} \in \mathbf{H} \middle|\, \mathbf{u}(\mathbf{h}) = \mathbf{u}_i \text{ and } P(\mathbf{h}) = P_j\right\}, \tag{16}$$

where $1 \le i \le N_{bf}$, $1 \le j \le N_{pc}$, and $\mathbf{H}$ is the set of channel vector realizations.

2. Updating the codebooks: In this step, the regions are fixed and the beamforming and power codebooks are optimized such that the average distortion in (11) is minimized. We use the Monte Carlo integration to approximate the average distortion:

$$\begin{aligned} p_{out} + \lambda P_{ave} &\approx \frac{1}{S}\sum_{\mathbf{h}\in\mathbf{H}} \mathbf{1}_c\Big(P(\mathbf{h})\left|\mathbf{h}^T\mathbf{u}(\mathbf{h})\right|^2\Big) + \lambda P(\mathbf{h}) \\ &= \frac{1}{S}\sum_{i=1}^{N_{bf}}\sum_{j=1}^{N_{pc}}\sum_{\mathbf{h}\in\mathcal{H}_{ij}} \mathbf{1}_c\left(P_j\left|\mathbf{h}^T\mathbf{u}_i\right|^2\right) + \lambda P_j. \end{aligned} \tag{17}$$

In order to minimize (17) in terms of $\mathbf{u}_i$'s and P_j's, we replace the indicator function with a differentiable approximation:

$$\mathbf{1}_c(x) \approx \sigma_{k,c}(x) \overset{\text{def}}{=} \sigma(k(x-c)), \tag{18}$$

where $\sigma(x) = \frac{1}{1+\exp(x)}$ is the *sigmoid* function. The parameter k determines the dropping slope of $\sigma_{k,c}(x)$ and controls the sharpness of the approximation, i.e. the higher the k the better the approximation. Fig. 7 shows the effect of k on the approximation.

Using (18), the objective function in (17) can be approximated by the following sequence of functions as $k \to \infty$:

$$f_k(\mathbb{U},\mathbb{P}) = \frac{1}{S}\sum_{i=1}^{N_{bf}}\sum_{j=1}^{N_{pc}}\sum_{\mathbf{h}\in\mathcal{H}_{ij}} \sigma_{k,c}\left(P_j\left|\mathbf{h}^T\mathbf{u}_i\right|^2\right) + \lambda P_j. \tag{19}$$

Let us define

$$(\mathbb{U}_k,\mathbb{P}_k) = \arg\min_{\mathbb{U},\mathbb{P}} f_k(\mathbb{U},\mathbb{P}). \tag{20}$$

In order to minimize the approximate average distortion function in (17), we start with a small k and minimize $f_k(\mathbb{U},\mathbb{P})$ and use the resulting optimum $(\mathbb{U}_k,\mathbb{P}_k)$ as a start point for a larger k. This is repeated until increasing k does not make a considerable change in the objective. The following subroutine provides the details:

Subroutine 1:

Inputs: Initial codebooks $(\mathbb{U}_{in}, \mathbb{P}_{in})$, Quantization regions $\mathcal{H}_{ij}$'s.

1) Choose k_0, $r_k > 1$, $\epsilon_1 \ll 1$, $r_\epsilon < 1$, $\epsilon_2 \ll 1$, $f_{old} \gg 1$, and $f_{new} = 0$.
2) Set $k = k_0$ and $(\mathbb{U}_{strt}, \mathbb{P}_{strt}) = (\mathbb{U}_{in}, \mathbb{P}_{in})$.
3) Apply a numerical unconstrained optimization method such as Newton's method with BFGS updates of Hessian matrix [41] with the start point $(\mathbb{U}_{strt}, \mathbb{P}_{strt})$ and the stopping criterion $\|\nabla f_k\| < \epsilon_1$ to solve (20).

   Set $f_{new} \leftarrow f_k(\mathbb{U}_k, \mathbb{P}_k)$.
4) If $|f_{new} - f_{old}|/f_{old} > \epsilon_2$, then

   Set $(\mathbb{U}_{strt}, \mathbb{P}_{strt}) \leftarrow (\mathbb{U}_k, \mathbb{P}_k)$.

   $k \leftarrow r_k k$, $\epsilon_1 \leftarrow r_\epsilon \epsilon_1$, $f_{old} \leftarrow f_{new}$.

   Go to step 3.

   Otherwise, stop.

Output: $(\mathbb{U}_{out}, \mathbb{P}_{out}) = (\mathbb{U}_k, \mathbb{P}_k)$.

For our numerical results, we use k_0=20, r_k=1.5, ϵ_1=0.1, r_ϵ=0.6, and ϵ_2=0.005.

Subroutine 1 has the same flavor of the interior-point method for constrained optimization using barrier functions. Although a large value of k would give a more exact approximation of the distortion function, it also increases the magnitude of the derivatives and can make the numerical convergence more difficult. We therefore start with a small k and increase it gradually until convergence. It should also be noted that, for small k, we do not need an exact minimization of f_k. Therefore, we start with a loose stopping criterion and tighten it by reducing ϵ_1 as k increases. Loosely speaking, as k increases to infinity, Subroutine 1 converges to a local minimum of (17). This concludes the second step of the Lloyd's algorithm, i.e. updating the codebooks.

The overall algorithm for minimizing the average distortion function (11) for fixed λ works as follows. We start with T random starting points, run the Lloyd's algorithm on each starting point, and choose the best among them:

Subroutine 2:

- For $t = 1, 2, \cdots, T$:
  1) Generate a random start point $(\mathbb{U}_t, \mathbb{P}_t)$.
  2) Repeat until convergence:
     a) Update the regions:
        Find $\mathcal{H}_{ij}$'s for $(\mathbb{U}_t, \mathbb{P}_t)$ using (14), (15), and (16).
     b) Update the codebooks:
        Run Subroutine 1 with $\mathcal{H}_{ij}$'s and $(\mathbb{U}_{in}, \mathbb{P}_{in}) = (\mathbb{U}_t, \mathbb{P}_t)$ as input.
        Set $(\mathbb{U}_t, \mathbb{P}_t) \leftarrow (\mathbb{U}_{out}, \mathbb{P}_{out})$.
- Choose the codebook pair $(\mathbb{U}_s, \mathbb{P}_s)$, $1 \leq s \leq T$, with minimum average distortion $p_{out} + \lambda P_{ave}$ given by (17).

An example of a solution sequence generated by Subroutine 2 is presented in the numerical results in Section VI.

The final step is to vary λ and run Subroutine 2 for each λ to derive the optimal curve of outage probability vs. SNR:

Algorithm 2:

1) Choose λ_0, $r_\lambda < 1$, and $q \ll 1$.
2) Set $\lambda = \lambda_0$ and $p^\star_{out} = 1$.
3) Repeat until $p^\star_{out} < q$:
   a) Run Subroutine 2 (with multiple starting points) for current λ and record the optimum point $(P^\star_{ave}, p^\star_{out})$.
   b) Set $\lambda \leftarrow r_\lambda \lambda$.
4) Take the convex hull of the $(P^\star_{ave}, p^\star_{out})$ points.

The performance of Algorithm 2 and its comparison with Algorithm 1 in Section III-C is studied numerically in Section VI.

V. Optimal Power Control and Beamforming Codebook Sizes

Section IV studies the joint beamforming and power codebook design problem with fixed codebook sizes N_{bf} and N_{pc}. In order to derive the optimal design given a fixed feedback link capacity B, we also need to find the optimal values of N_{bf} and N_{pc}. In this section, we optimize the codebook sizes numerically by searching over all integer pairs (N_{bf}, N_{pc}) satisfying $N_{bf}N_{pc}{\le}2^B$, then choosing the pair with the best performance.

First we describe some constraints that can be imposed on the search set $\left\{(N_{bf}, N_{pc})\,\middle|\, N_{bf}N_{pc}{\le}2^B\right\}$. One constraint to consider is $N_{bf}{\ge}M$. This is justified by noting that rank-one beamforming is not appropriate when the beamforming codebook size is less than the number of transmit antennas, at least for i.i.d. Rayleigh fading channels. To see this, consider a codebook with $N_{bf}{<}M$ beamforming vectors, which spans an N_{bf}-dimensional subspace in the M-dimensional complex space. Since the channel direction is uniformly distributed in space, this codebook should have the same performance as any rotated version of it. By appropriately rotating the codebook, we can get a codebook, the vectors of which are all perpendicular to say $\mathbf{e}_M = [0, 0, \cdots, 0, 1]^T$. This means that none of the vectors use the M^{th} antenna or equivalently, this antenna is permanently turned off. The joint codebook design problem, therefore, reduces to a problem with smaller number of antennas. This loss of degrees of freedom considerably reduces the diversity gain, which should be avoided.

The search set can be further restricted by noting that if $(m_1, n_1){\le}(m_2, n_2)$ element-wise, then the optimal beamforming and power codebooks with sizes m_2 and n_2 would clearly outperform the optimal codebooks with sizes m_1 and n_1.

Definiton 1: We say the integer pair (m_2, n_2) ***dominates*** the pair (m_1, n_1) if $(m_1, n_1){\le}(m_2, n_2)$ element-wise.

Definiton 2: For any integer number N, let

$$\mathcal{B}_N = \left\{ (m, n) \in \mathbb{N}^2 \middle|\, mn{\le}N \right\}.$$

We define the *maximal subset* $\mathcal{A}_N$ as a subset of $\mathcal{B}_N$ such that any pair in $\mathcal{B}_N$ is dominated by a pair in $\mathcal{A}_N$ and no pair in $\mathcal{A}_N$ can be dominated by another pair in $\mathcal{A}_N$.

The following is a characterization of the maximal subset. The proof is omitted for brevity.

Proposition 1: For any integer N, the maximal subset of $\mathcal{B}_N$ is given by

$$\mathcal{A}_N = \left\{ (i, \lfloor N/i \rfloor), (\lfloor N/i \rfloor, i) \middle| 1 \le i \le \lfloor \sqrt{N} \rfloor \right\}, \tag{21}$$

where $\lfloor \cdot \rfloor$ is the floor function.

Considering these constraints, we can now restrict the search of the optimal codebook sizes to the following set:

$$\mathcal{C}(B, M) = \left\{ (N_{bf}, N_{pc}) \in \mathcal{A}_{2^B} \middle|\, N_{bf} \ge M \right\} \cup \left\{ \left(0, 2^B\right) \right\}. \tag{22}$$

Here we have added the pair $\left(0, 2^B\right)$ to represent the case where all the feedback bits are used for power control — there is no beamforming codebook for this case, i.e. the transmission is isotropic or $\mathbf{Q}_{\mathbf{x}|\mathbf{h}} = \frac{1}{M}\mathbf{I}_M$ (see Section III). As an example, for $B = 5$ and $M = 4$, we have

$$\mathcal{C}(5,4) = \\ \left\{(32,1),(16,2),(10,3),(8,4),(6,5),(5,6),(4,8),(0,32)\right\}.$$

The optimal codebook sizes, found by searching over $\mathcal{C}(B, M)$ for different values of B and M, as well as the corresponding system performance results are presented in the next section.

VI. Numerical Results

This section presents the numerical results for the joint beamforming and power control optimization problem and the corresponding optimal codebook sizes.

A. Performance of Subroutine 2

We start by showing an example of the solution sequence generated by Subroutine 2 in Section IV-B. Fig. 8 shows the performance of Subroutine 2 for a fixed λ and a single start point shown by the filled circle on the (P_{ave}, p_{out}) plane. The solution sequence converges to a point on the optimal curve, where the slope of the tangent line is equal to $-\lambda$.

B. Performance of the Joint Optimization Algorithms

We start with Algorithm 2 in Section IV-B. For our numerical results, we set $\lambda_0 = 2$, $r_\lambda = 0.8$, $q = 10^{-4}$, and use 10 random starts for each λ. It should be noted that, one needs to increase the number of channel realizations, S, as the outage probability decreases. In order to ensure the reliability of computed outage probabilities, for each λ, we set $S = \frac{100}{p_{out}^\star}$, where $p_{out}^\star$ is the outage probability for the previous λ. Note that the convexity structure of the problem implies that the optimum curve (in linear scale) is a convex curve asymptotically. Therefore, we take the convex hull of the points in the last step.

Algorithm 2 is a general algorithm in the sense that it can be applied to any arbitrary channel statistics if we are provided with sufficient number of channel realizations. This algorithm, in spite of its complexity, works well with modest values of outage probability, number of antennas, and codebook sizes. For small values of outage probability, however, the number of channel realizations (training size) must be large and this increases the cost of updating the quantization regions. Moreover, for updating the codebooks in Subroutine 1, we have to go through the summations in (19) to compute the gradient of the objective and repeat this for different values of k. Although the codebook design process is done offline with no real-time implication in the actual system implementation, the design process can be time-consuming when the training size S, and the codebook sizes, N_{bf}, N_{pc}, are large.

For i.i.d. Rayleigh channels, however, we can use the less complex Algorithm 1 in Section III. The main advantage of Algorithm 1 is that its speed is controlled by the number of interpolation points, i.e. the complexity of the interpolating function, rather the training size. In our numerical results for Algorithm 1, we use 100 random starting points, 100 interpolation points, and training size S in the range of 10^6-10^7.

Fig. 9 compares the performance of Algorithms 1 and 2 for $M{=}3$, $N_{bf}{=}5$, $N_{pc}{=}3$, and $R{=}1$ (or $c{=}1$). Algorithm 1 slightly outperforms Algorithm 2, possibly because of more start points used. The figure also shows the performance of Algorithm 2 when it relies on the output of Algorithm 1 as its starting point. This only slightly improves the performance of Algorithm 1, suggesting that the output of Algorithm 1 is already close to a local optimum of the joint optimization problem. We therefore rely on Algorithm 1 for deriving the optimal beamforming and power control codebook sizes as explained in the next section.

C. Optimal Beamforming and Power Control Codebook Sizes

In order to find the optimal codebook sizes, we use Algorithm 1 in Section III to compare the performance of different codebook size pairs in (22). Fig. 10 shows the comparison results for $M{=}4$ antennas, $B{=}5$ bits, and $R{=}1$ bits/sec/Hz. The minimum outage probability recorded is 10^{-6}. For this case, the optimum codebook size pair is $(8, 4)$ for outage probabilities $p_{out}{<}6{\times}10^{-3}$ and $(4, 8)$ otherwise. The minimum of the outage curves associated with these two pairs outperforms all other pairs in $\mathcal{C}(5, 4)$ and therefore the performance of the other pairs is not included in the figure.

The figure also includes the performance of the $(0, 32)$ pair (no-beamforming case), which shows that we can gain a considerable gain by joint beamforming and power control with optimal codebook sizes, e.g. almost 2.5dB for $p_{out}{=}10^{-3}$.

Fig. 11 shows the performance of joint beamforming and power control with optimal codebook sizes for $M = 4$ and different values of B. The different line widths on each curve imply that different codebook size pairs are optimal for different ranges of the outage probability. The figure shows that the joint design provides a combined power and diversity gain which enables it to approach the performance of the perfect CSIT case, as the number of feedback bits increases to infinity.

Table I summarizes the optimal codebook sizes for the outage probability range 10^{-1} to 10^{-6} and different values of M and B. In each cell of the table, the first row is the optimal size pair (N_{bf}, N_{pc}) and the second row is the range of outage probabilities over which this size pair is optimal. It should be noted that the optimal codebook sizes and the corresponding outage ranges in Table I, although originally derived for the target rate $R{=}1$, hold for any rate. This is justified by considering the original joint design problem (7), where $c{=}2^R{-}1$ and SNR is the power constraint. If we scale $P(\mathbf{h})$ and SNR by c, we get the joint design problem with $c{=}1$ (or $R{=}1$). This means that we can apply all our results, originally derived for $R{=}1$, for any rate R, as long as we horizontally shift the outage vs. SNR curves by $10\log_{10}\left(2^R{-}1\right)$dB, and this would not change the outage probability ranges and the corresponding codebook sizes.

The results in Table I also show that as the outage probability decreases, the optimal size of the power codebook increases and the optimal size of the beamforming codebook decreases. This is to be expected based on the discussion in Section II, according to which power control provides diversity gain, which is the dominant factor for higher SNR values (small outage probabilities), while beamforming provides power gain which is an important factor for lower SNR values (higher outage probabilities). Joint beamforming and power control realizes both the power gain of the beamforming and the diversity gain of the power control, enabling the limited-feedback system to approach the optimal perfect-CSIT system performance as the feedback capacity increases.

VII. Conclusions

A limited-feedback system requires both beamforming and power control modules in order to approach the performance of the perfect-CSIT system as the feedback link capacity increases. This paper shows that the two modules should also be designed and optimized jointly. Based on a convexity structure of the problem, we propose a joint design of the beamforming and power control codebooks using Monte Carlo integration, Lloyd's algorithm, and BFGS optimization method. For i.i.d. Rayleigh channels, we propose a less complex algorithm, where a uniform beamforming codebook is chosen and fixed and only the power codebook is optimized. The two algorithms are shown to have a close performance in terms of the outage probability vs. SNR.

We further investigate the optimal beamforming and power control codebook sizes, given a fixed feedback link capacity constraint. The results provide the optimal codebook sizes as a function of target outage probability and independent of the target rate. The optimal performance curves show that the joint beamforming and power control provides a combined power and diversity gain, which enables the system to approach the performance of a perfect-CSIT system as the feedback link capacity increases.

APPENDIX

In this Appendix, we first present a sufficient condition under which a general constrained optimization problem possesses certain convexity structure. Such a structure allows us to derive the solution of the optimization problem by optimizing the corresponding Lagrange formulation parameterized by a dual variable. Next, we show that the joint beamforming and power control codebook design problem possesses this convexity structure in the asymptotic case of $B \to \infty$.

Consider the following optimization problem parameterized by $\theta \in \mathbb{R}$:

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & g(\mathbf{x}) \leq \theta, \end{aligned} \tag{23}$$

where $\mathbf{x} \in \mathbb{R}^n$, and $f(\mathbf{x})$ and $g(\mathbf{x})$ are real scalar functions. Let $\mathbf{x}_\theta^\star$ denote the optimum of (23).

Definiton 3: The problem (23) is said to be a *convex-like* problem, if the minimized objective function $f(\mathbf{x}_\theta^\star)$ is a strictly convex function of the constraint parameter θ.

Theorem 1: For a *convex-like* problem (23), the inequality constraint is active at the optimum, i.e. $g(\mathbf{x}_\theta^\star) = \theta$. Moreover, for any value of θ, there exists a real number λ, such that $\mathbf{x}_\theta^\star$ is the optimum point for the following problem:

$$\min_{\mathbf{x}} f(\mathbf{x}) + \lambda g(\mathbf{x}). \tag{24}$$

Proof: The proof results from a direct application of Definition 3 and some convexity arguments as in [42]. ■

Theorem 1 implies that the solution set $\{(\theta, f(\mathbf{x}_\theta^\star)) | \, \theta\}$ is the same as $\{(g(\mathbf{x}_\theta^\star), f(\mathbf{x}_\theta^\star)) | \, \theta\}$ and that the latter set can be found by solving (24) for the corresponding values of λ. Assuming that the problem (7) is *convex-like*, this implies that we can fully derive the optimal curve of outage probability (objective of (7)) vs. SNR by solving (8) for appropriate values of λ.

In the following we prove that the joint beamforming and power control codebook design problem is *convex-like* asymptotically as $B \to \infty$. This guarantees that minimizing the Lagrangian formulation (8), and therefore Algorithm 2 in Section IV, is asymptotically optimal. In order to avoid the degenerate case of zero outage probability, an upper bound is imposed on SNR.

Theorem 2: The joint codebook design problem (7) is *convex-like* asymptotically as $B \to \infty$ provided that

$$\mathsf{SNR} < \int_0^\infty \frac{c}{\gamma} \;\; \mathrm{d}F_\Gamma(\gamma) \tag{25}$$

and $F_\Gamma(\gamma)$ is an strictly increasing function. Here $F_\Gamma(\gamma)$ is the CDF of the effective channel gain γ.

Proof: The asymptotic case of $B \to \infty$ is equivalent to perfect CSIT and, according to the discussion in Section III, for this case $\gamma = \|\mathbf{h}\|^2$, the optimal beamforming function is matched-channel beamforming $\mathbf{u}(\mathbf{h}){=}\mathbf{h}^*/\|\mathbf{h}\|$, and the optimal power control function is given by

$$P(\gamma) = \begin{cases} 0 & \text{if} \quad \gamma < \gamma_0 \\ \frac{c}{\gamma} & \text{if} \quad \gamma \geq \gamma_0 \end{cases}, \tag{26}$$

where γ_0 is the transmission threshold.

The corresponding outage probability and average transmission power are given by

$$p_{out} = \mathrm{Prob}[P(\gamma)\cdot\gamma < c] = F_\Gamma(\gamma_0), \tag{27}$$

$$P_{ave} = \mathrm{E}[P(\gamma)] = \int_{\gamma_0}^\infty \frac{c}{\gamma} \;\; \mathrm{d}F_\Gamma(\gamma) \leq \mathsf{SNR}. \tag{28}$$

In order to minimize p_{out}, the transmission threshold γ_0 should satisfy (28) with equality and therefore $\mathsf{SNR}{=}P_{ave}$. Note that the condition (25) guarantees that $\gamma_0{>}0$ and $p_{out}{>}0$ and therefore the degenerate case of zero outage is excluded from discussion.

In order to prove the convexity structure, we have to show that

$$\frac{\partial^2 p_{out}}{\partial \mathsf{SNR}^2} = \frac{\partial^2 p_{out}}{\partial {P_{ave}}^2} > 0.$$

Define $f_\Gamma(\gamma) = \partial F_\Gamma(\gamma)/\partial\gamma > 0$. From (28) we have

$$\frac{\partial \gamma_0}{\partial P_{ave}} = \frac{1}{\partial P_{ave}/\partial \gamma_0} = \frac{-\gamma_0}{c f_\Gamma(\gamma_0)}, \tag{29}$$

By using (27), (28), (29), and the chain rule, we have

$$\frac{\partial p_{out}}{\partial P_{ave}} = \frac{\partial F_\Gamma(\gamma_0)}{\partial \gamma_0} \cdot \frac{\partial \gamma_0}{\partial P_{ave}} = -\frac{\gamma_0}{c}.$$

Finally,

$$\frac{\partial^2 p_{out}}{\partial P_{ave}^2} = -\frac{1}{c}\frac{\partial \gamma_0}{\partial P_{ave}} = \frac{\gamma_0}{c^2 f_\Gamma(\gamma_0)} > 0.$$

■

REFERENCES

[1] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1423-1436, Oct. 1998.

[2] D. J. Love and R. W. Heath, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2735-2747, Oct. 2003.

[3] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple-antenna systems," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2562-2579, Oct. 2003.

[4] P. Xia and G. B. Giannakis, "Design and analysis of transmit-beamforming based on limited-rate feedback," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1853-1863, May 2006.

[5] J. C. Roh and B. D. Rao, "Transmit beamforming in multiple-antenna systems with finite rate feedback: a VQ-based approach," *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 1101-1112, Mar. 2006.

[6] W. Santipach and M. L. Honig, "Asymptotic performance of MIMO wireless channels with limited feedback," in *Proc. IEEE Mil. Comm. Conf.*, vol. 1, pp. 141-146, Oct. 2003.

[7] R. S. Blum, "MIMO with limited feedback of channel state information," in *Proc. Int. Conf. on Acoust., Speech, and Signal Processing (ICASSP)*, vol. 4, pp. 89-92, Hong Kong, Apr. 2003.

[8] S. Bhashyam, A. Sabharwal, and B. Aazhang, "Feedback gain in multiple antenna systems," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 785-798, May 2002.

[9] A. Khoshnevis and A. Sabharwal, "On the asymptotic performance of multiple antenna channels with quantized feedback." *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3869-3877, Oct. 2008.

[10] T. T. Kim and M. Skoglund, "Outage behavior of MIMO channels with partial feedback and minimum multiplexing gains," in *Proc. IEEE Int. Symp. Information Theory*, Seattle, WA, Jul. 2006.

[11] T. T. Kim and M. Skoglund, "Diversity-multiplexing tradeoff in MIMO channels with partial CSIT," *IEEE Trans. Inform. Theory*, vol. 53, no. 8, pp. 2743-2759, Aug. 2007.

[12] B. Khoshnevis and W. Yu, "Joint power control and beamforming codebook design for MISO channels with limited feedback," in *Proc. IEEE Global Communications Conference (Globecom 2009)*, Hawaii, U.S.A., Dec. 2009.

[13] B. Khoshnevis and W. Yu, "High resolution quantization codebook design for multiple-antenna fading channels," in *Proc. 25th Biennial Symposium on Communications*, Kingston, Ontario, Canada, May 2010.

[14] B. Khoshnevis and W. Yu, "Bit allocation law for multi-antenna channel feedback quantization: single-user case," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2270-2283, May 2011.

[15] R. Mulholland, A. Joshi, and J. Chu," A measure of efficiency for a class of decision feedback communication systems," *Journal of the Franklin Institute*, vol. 285, no. 3, pp. 175-186, Mar. 1968.

[16] T. Başar, T. Başar, "A Bandwidth expanding scheme for communication channels with noiseless feedback in the presence of unknown jamming noise," *Journal of the Franklin Institute*, vol. 317, no. 2, pp. 73-88, Feb. 1984.

[17] N. Jindal, "MIMO broadcast channels with finite rate feedback," *IEEE Trans. Inform. Theory*, vol. 52, no. 11, pp. 5045-5059, Nov. 2006.

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

[19] J. Chen, R. Berry, and M. Honig, "Limited feedback schemes for downlink OFDMA based on sub-channel groups," *IEEE J. Select. Areas Commun.*, vol. 26, no. 8, pp. 1451-1461, Oct. 2008.

[20] B. Khoshnevis, W. Yu, and R. S. Adve, "Grassmannian beamforming for MIMO amplify-and-forward relaying," *IEEE J. Select. Areas Commun.*, vol. 26, no. 8, pp. 1397-1407, Oct. 2008.

[21] N. Ahmed, M. Khojastepour, A. Sabharwal, and B. Aazhang, “Outage minimization with limited feedback for the fading relay channel,” *IEEE Trans. Commun.*, vol. 54, no. 4, pp. 659-669, Apr. 2006.

[22] B. Khoshnevis and W. Yu, “Structure of channel quantization codebook for multiuser spatial multiplexing systems,” in *Proc. IEEE International Conference on Communications (ICC 2010)*, Cape Town, South Africa, May 2010.

[23] B. Khoshnevis and W. Yu, “High resolution channel quantization rules for multiuser spatial multiplexing systems,” in *Proc. Conference on Information Sciences and Systems (CISS 2010)*, Princeton, NJ, U.S.A., Mar. 2010.

[24] B. Khoshnevis and W. Yu, “Bit allocation laws for multi-antenna channel quantization: multi-user case,” *IEEE Transactions on Signal Processing*, 2011, to be published.

[25] B. Khoshnevis and W. Yu, “A limited-feedback scheduling and beamforming scheme for multi-user multi-antenna systems,” in *Proc. IEEE Global Communications Conference (Globecom 2011)*, Houston, TX, 5-9 Dec. 2011.

[26] B. Khoshnevis and W. Yu, “Two-stage channel feedback for beamforming and scheduling in network MIMO systems,” submitted to *IEEE International Conference on Communications (ICC 2012)*, Ottawa, Canada, Jun. 2012.

[27] J. Dai, “A delay system approach to networked control systems with limited communication capacity,” *Journal of the Franklin Institute*, vol. 347, no. 7, pp. 1334-1352, Sep. 2010.

[28] M. Liu and G. Sun, “Observer-based sliding mode control for Itô stochastic time-delay systems with limited capacity channel,” *Journal of the Franklin Institute*, 2011, to be published.

[29] B. Hu and A. Michel, “Robustness analysis of digital feedback control systems with time-varying sampling periods,” *Journal of the Franklin Institute*, vol. 337, no. 2-3, pp. 117-130, Mar. 2000.

[30] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, “An overview of limited feedback in wireless communication systems,” *IEEE J. Select. Areas Commun.*, vol. 26, no. 8, pp. 1341-1365, Oct. 2008.

[31] V. Lau, Y. Liu, and T. Chen, “On the design of MIMO block-fading channels with feedback-link capacity constraint,” *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 62-70, Jan. 2004.

[32] M. Skoglund and G. Jöngren, “On the capacity of a multiple-antenna communication link with channel side information,” *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 395-405, Apr. 2003.

[33] G. Caire and S. Shamai, “On the capacity of some channels with channel state information,” *IEEE Trans. Inform. Theory*, vol. 45, no. 6, pp. 2007-2019, Sep. 1999.

[34] I. E. Telatar, “Capacity of multi-antenna Gaussian channels,” *European Trans. Telecomm.*, vol. 10, no. 6, pp. 585-595, Dec. 1999.

[35] G. Caire, G. Taricco, and E. Biglieri, “Optimum power control over fading channels,” *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1468-1489, Jul. 1999.

[36] E. Biglieri, G. Caire, and G. Taricco, “Limiting performance of block fading channels with multiple antennas,” *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1273-1289, May 2001.

[37] F. N. Fritsch and R. E. Carlson, “Monotone piecewise cubic interpolation,” *SIAM J. Numer. Anal.*, vol. 17, no. 2, pp. 238-246, Apr. 1980.

[38] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[39] R. M. Gray and J. T. Gill, “A Lagrangian formulation of fixed-rate and entropy/memory constrained quantization,” in *Proc. IEEE Data Compression Conference*, pp. 261-269, Mar. 2005.

[40] P. A. Chou, T. Lookabaugh, and R. M. Gray, “Entropy-constrained vector quantization,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 1, pp. 31-42, Jan. 1989.

[41] J. E. Dennis and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Englewood Cliffs, NJ: Prentice-Hall, 1983.

[42] W. Yu and R. Lui, “Dual methods for nonconvex spectrum optimization of multicarrier systems,” *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310-1322, Jul. 2006.

TABLE I

THE OPTIMAL BEAMFORMING AND POWER CONTROL CODEBOOK SIZES (N_{bf}, N_{pc}) FOR DIFFERENT VALUES OF FEEDBACK LINK CAPACITY B AND NUMBER OF TRANSMIT ANTENNAS M.

<table>
<tr><th></th><th>$M=2$</th><th>$M=3$</th><th>$M=4$</th></tr>
<tr><td>B=1</td><td>$(0,2)$
10^{-1}–10^{-6}</td><td>$(0,2)$
10^{-1}–10^{-6}</td><td>$(0,2)$
10^{-1}–10^{-6}</td></tr>
<tr><td rowspan="2">B=2</td><td>$(2,2)$
10^{-1}–3.1×10^{-2}</td><td rowspan="2">$(0,4)$
10^{-1}–10^{-6}</td><td rowspan="2">$(0,4)$
10^{-1}–10^{-6}</td></tr>
<tr><td>$(0,4)$
3.1×10^{-2}–10^{-6}</td></tr>
<tr><td rowspan="2">B=3</td><td>$(2,4)$
10^{-1}–3.7×10^{-4}</td><td>$(4,2)$
10^{-1}–5.8×10^{-3}</td><td>$(4,2)$
10^{-1}–1.3×10^{-3}</td></tr>
<tr><td>$(0,8)$
3.7×10^{-4}–10^{-6}</td><td>$(0,8)$
5.8×10^{-3}–10^{-6}</td><td>$(0,8)$
1.3×10^{-3}–10^{-6}</td></tr>
<tr><td rowspan="2">B=4</td><td>$(3,5)$
10^{-1}–1.9×10^{-2}</td><td>$(4,4)$
10^{-1}–1.2×10^{-3}</td><td rowspan="2">$(4,4)$
10^{-1}–10^{-6}</td></tr>
<tr><td>$(2,8)$
1.9×10^{-2}–10^{-6}</td><td>$(3,5)$
1.2×10^{-3}–10^{-6}</td></tr>
<tr><td rowspan="3">B=5</td><td>$(4,8)$
10^{-1}–3.0×10^{-3}</td><td rowspan="2">$(6,5)$
10^{-1}–1.6×10^{-2}</td><td rowspan="2">$(8,4)$
10^{-1}–6.1×10^{-3}</td></tr>
<tr><td>$(3,10)$
3.0×10^{-3}–7.3×10^{-5}</td></tr>
<tr><td>$(2,16)$
7.3×10^{-5}–10^{-6}</td><td>$(4,8)$
1.6×10^{-2}–10^{-6}</td><td>$(4,8)$
6.1×10^{-3}–10^{-6}</td></tr>
</table>

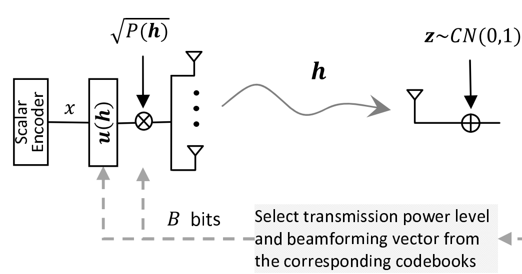


Fig. 1. Limited-feedback MISO system; the capacity of the feedback link is B bits per fading block.

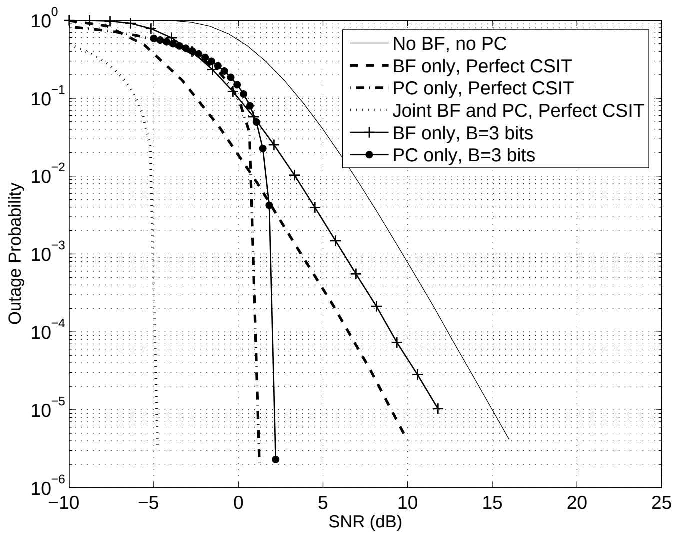


Fig. 2. Outage probability vs. SNR for $M = 4$ transmit antennas and target rate of $R = 1$ bit/sec/Hz. For high enough SNR values, power control (PC) outperforms beamforming (BF).

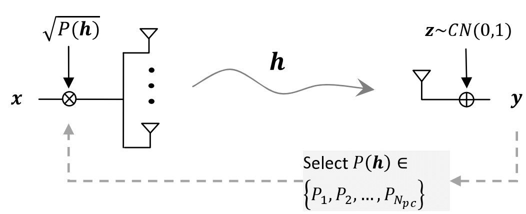
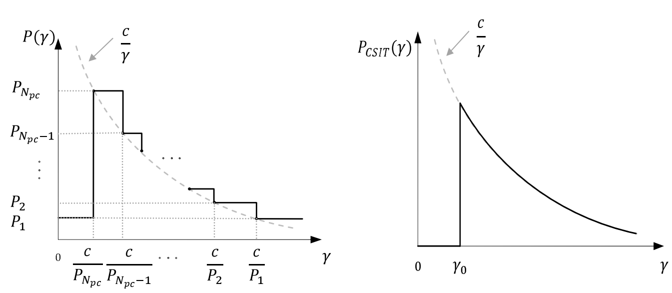


Fig. 3. Power control with limited feedback. The vector $\mathbf{x}$ is the output of an arbitrary (fixed) beamforming module.



(a) (b)

Fig. 4. (a) Optimal power control (or quantization) function with N_{pc} power levels; outage only occurs when $\gamma < \frac{c}{P_{N_{pc}}}$. (b) Optimal power control with perfect CSIT: truncated channel inversion. The threshold value γ_0 is determined by the transmit power constraint.

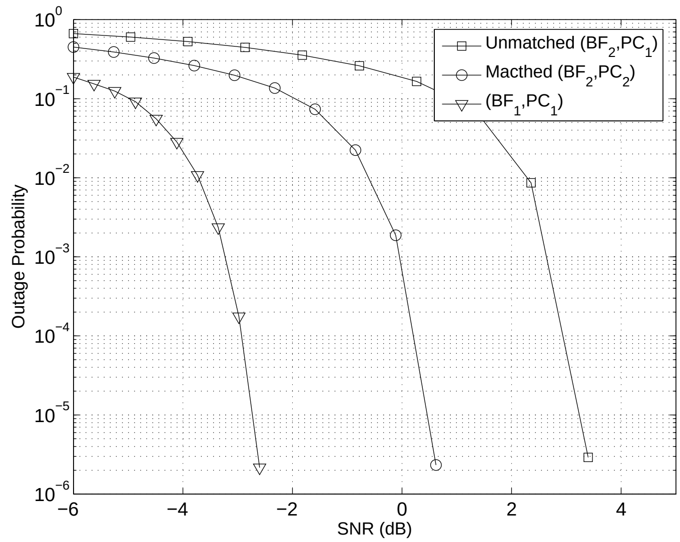


Fig. 5. Matched module pair $(\mathrm{BF}_2, \mathrm{PC}_2)$ vs. unmatched pair $(\mathrm{BF}_2, \mathrm{PC}_1)$ for $M{=}4$ antennas, $B{=}2$ bits, and $R{=}1$ bits/sec/Hz.

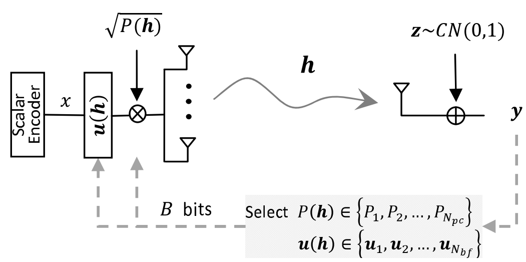


Fig. 6. Beamforming and power control with limited feedback; $N_{bf}N_{pc} \leq 2^B$. The scalar encoder is constrained by $\mathrm{E}\left[|x|^2\right] = 1$.

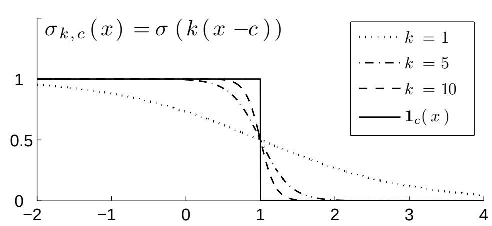


Fig. 7. $\mathbf{1}_c(x)$ and $\sigma_{k,c}(x)$ for $c = 1$ and $k = 1,\ 5,\ 10$.

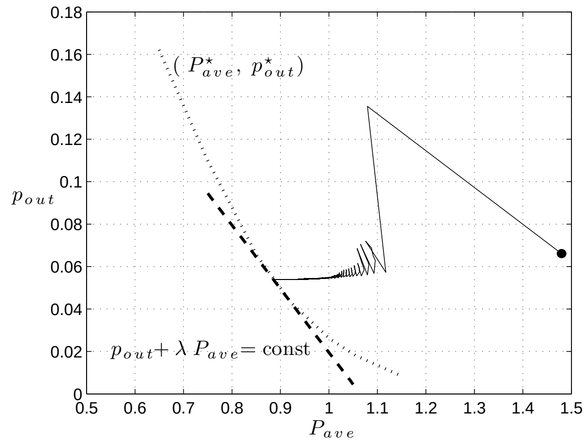


Fig. 8. The solution sequence of Subroutine 2 for $\lambda = 0.3$, $M = 3$, $N_{bf} = 5$, $N_{pc} = 3$, $R = 1$, and $S = 10^4$ realizations of the channel $\mathbf{h} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$. The filled circle shows the random start point on (P_{ave}, p_{out}) plane. Both axes are in linear scale. The optimum curve $(P^\star_{ave}, p^\star_{out})$ is generated by Algorithm 2.

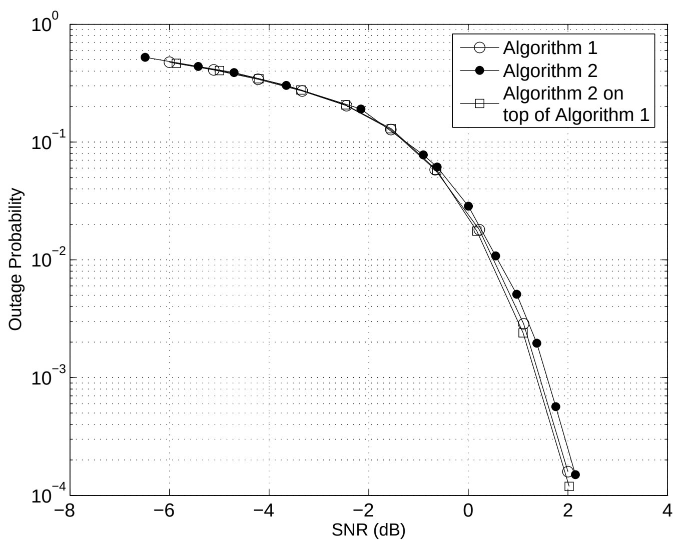


Fig. 9. The performance of Algorithms 1 and 2 for $M = 3$ antennas, $N_{bf} = 5$ beamforming vectors, $N_{pc} = 3$ power levels, and $R = 1$ bits/sec/Hz.

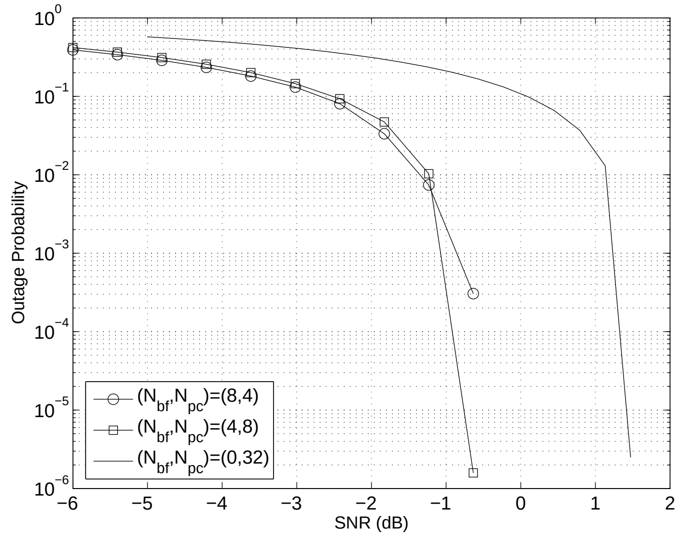


Fig. 10. The performance of different codebook size pairs for $M=4$ antennas, $B=5$ bits, and $R=1$ bits/sec/Hz.

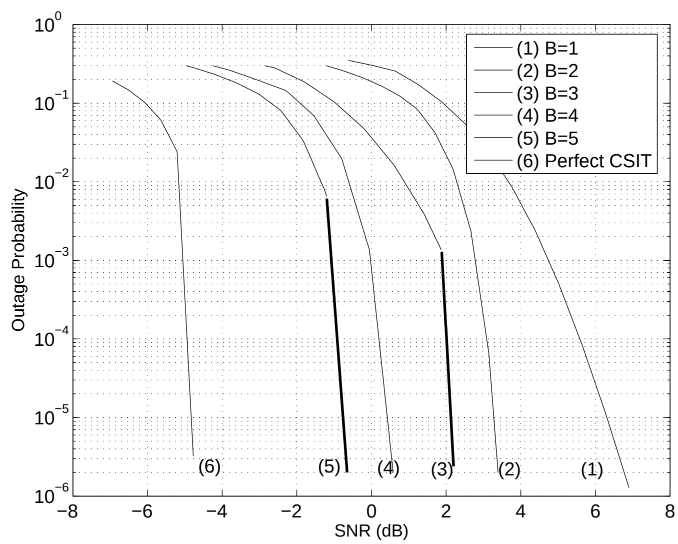


Fig. 11. Outage probability vs. SNR for joint beamforming and power control with optimal codebook sizes for $M = 4$ antennas. The optimal codebook sizes differ for different outage probability ranges; this is shown by changing the line width on each curve.