

Research Article

Optimal Power Allocation with Channel Inversion Regularization-Based Precoding for MIMO Broadcast Channels

Yang Xu and Tho Le-Ngoc

Department of Electrical and Computer Engineering, Faculty of Engineering, McGill University, 3480 University Street, Montréal, QC, Canada H3A 2A7

Correspondence should be addressed to Tho Le-Ngoc, tho.le-ngoc@mcgill.ca

Received 9 July 2008; Revised 24 October 2008; Accepted 31 December 2008

Recommended by Huaiyu Dai

Zero-forcing (ZF) precoding scheme can achieve the asymptotic sum capacity as dirty-paper coding (DPC) in multiple-input multiple-output broadcast (MIMO-BC) channel when the number of users, K , approaches infinity. However, the gap between ZF and DPC is not negligible in a practical range of K , that is, $K \leq 100$. The capacity loss is partly due to the excessive transmission power penalty incurred by ZF when the channel matrix of the selected user subset is poorly conditioned. To avoid this power penalty, we propose to use a variation of ZF, channel inversion regularization (CIR), as a precoding scheme in MIMO-BC channels. But, unlike the interference-free ZF, the problem of maximizing sum-rate capacity using CIR precoding becomes nonconvex, which cannot be solved by water-filling strategy. Thus, we propose an efficient algorithm based on gradient projection (GP) as the optimal power allocation strategy for selected users, and show that the proposed CIR precoding scheme can achieve asymptotically the optimum sum-rate of the DPC strategy. Moreover, simulation results show that the CIR precoding scheme with the proposed optimal power allocation scheme achieves better sum-rate performance than ZF for a wide range of K .

Copyright © 2008 Y. Xu and T. Le-Ngoc. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

Multiple-input multiple-output (MIMO) systems demonstrate considerable capacity gain compared to single-input single-output systems (SISO) in a single-user communication system [1, 2]. But the application of MIMO is not limited to the single-user case. Recently, more and more research has been conducted to investigate the application of MIMO to multiuser systems in wireless networks.

This paper considers a MIMO-broadcast channel (MIMO-BC), where a transmitter equipped with multiple antennas at the base station communicates with several users equipped with multiple antennas. The capacity region of MIMO-BC channels has been characterized in [3–6] recently. In [5, 6], the duality between MIMO-BC and MIMO multiaccess (MIMO-MA) is established and exploited to calculate the capacity of MIMO-BC in an efficient manner. Moreover, the sum capacity of MIMO-BC channels can be achieved theoretically by applying dirty-paper coding (DPC) [7] to serve multiple users simultaneously at the transmitter

[4]. Based on interference precancellation at the transmitter, the multiuser encoding DPC is the optimal capacity achieving strategy for MIMO-BC. Though several practical implementations of DPC were proposed in [8, 9], DPC strategy still remains to be very difficult to be implemented for practical systems due to its high computation complexity incurred by successive encoding and decoding.

To circumvent the complexity of DPC, a simple suboptimal linear zero-forcing (ZF) precoding scheme is investigated in [10] and demonstrates achieving asymptotically the sum capacity as DPC in MIMO-BC channel if the number of users K approaches infinity. However, achieving optimum ZF sum-rate requires the highly complicated exhaustive search of all possible user subsets to select the best user subset that yields the highest sum-rate, which is clearly computationally infeasible for large K . Hence, [10] proposed a low-complexity suboptimum algorithm based on semiorthogonal user selection (SUS). The basic idea of SUS algorithm is to find a subset of users with best semiorthogonal channel vectors. As the search space of SUS

is greatly reduced as compared to the exhaustive search, the complexity of SUS is significantly lower but it can still achieve the sum-rate of optimal ZF asymptotically as $K \rightarrow \infty$. However, the SUS scheme requires perfect channel state information (CSI) at the base station, demanding huge amount of feedback information from users to the base station. To alleviate the CSI feedback burden imposed by SUS, two other suboptimum schemes [11, 12] can approach the performance of SUS scheme but with much less CSI feedback and computation complexity.

All these schemes based on ZF, however, incur non-negligible performance loss compared to DPC when K is relatively small. This sum-rate performance loss is partly due to the excess transmission power penalty incurred at the transmitter by ZF. More specifically, when the number of users K is relatively small, for example, less than 100, it is not always possible to find a group of M users with nearly orthogonal channels and the channel matrix of selected user subset is thus often poorly conditioned. It is known that inverting a poorly conditioned matrix unavoidably results in the reduction of effective channel gain [10]. In addition, the sum-rate loss caused by this reduced effective channel gain is more prominent in the case of low SNR.

In this paper, we propose to use channel inversion regularization (CIR) [13] as an alternative precoding scheme to alleviate the reduction of effective channel gain for MIMO-BC channels. Rather than applying pseudoinverse directly, CIR scheme adds a multiple of identity matrix before inverting to regularize an inverse. Hence, unlike the interference-free ZF, the zero-interference condition is no longer satisfied with CIR precoding scheme. As a result, the achieved sum-rate capacity of CIR scheme cannot be reduced to a simple convex optimization problem as ZF scheme, and hence the well-known transmission power “water-filling” strategy may not be applicable.

Instead, the problem of CIR scheme becomes a nonlinear nonconvex optimization problem. To solve globally this nonconvex optimization problem, we can resort to the global difference of convex (d.c.) optimization technique via recognizing that its objective function can be written as the difference of two convex functions. In [14], a similar global d.c. optimization approach has been proposed to find the global optimum of the power allocation problem in digital subscriber lines (DSLs) interference channels with much less complexity than the existing exhaustive search method. But this global optimization approach is not the ideal choice for designing the highly efficient algorithm required by wireless fading channel, where the algorithm for power allocation has to be done for every fading block. As an alternative, we then propose to use a low-complexity gradient-projection (GP) method, which is an extension of the unconstrained steepest descent method, as the power allocation scheme at the transmitter for the selected user subset. Because the feasible region of our particular power allocation optimization problem is a geometrically simple simplex, GP method is well suited as a candidate for solving it efficiently. The asymptotic convexity analysis in Section 5 shows that the nonconvex optimization problem reduces to a convex optimization problem when the number of users

in the system is sufficiently large. Thus, local GP method can find the global optimum solution when K is sufficiently large. Moreover, simulation results in various wireless scenarios also indicate that the sum-rate performance loss of local GP method is negligible compared to that of global d.c. approach, justifying the use of local GP method for MIMO-BC channels.

The remainder of this paper is organized as follows. Section 2 begins with the system model of MIMO-BC and describes the ZF precoding scheme. Section 3 presents the proposed scheme based on CIR with optimal power control, and considers both global d.c. optimization approach and local GP method for solving the nonconvex optimization problem. The analysis of asymptotic sum-rate performance and convexity is presented in Section 4, showing that the proposed CIR precoding scheme can achieve asymptotically the optimum sum-rate offered by the DPC strategy. Section 5 provides simulation results on the performance of the proposed scheme under various conditions. Finally, concluding remarks are made in Section 6.

2. SYSTEM MODEL AND ZF PRECODING SCHEME

2.1. System model

In this paper, we consider a MIMO broadcast system using one transmitter to serve K users in a single-cell wireless scenario. The transmitter at the base station is equipped with M transmit antennas, and user k with N_k receiver antennas. The channel between user k and base station is modeled as a zero-mean circularly symmetric Gaussian matrix \mathbf{H}_k . The received signal vector of user k is written as

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{n}_k, \quad k = 1, \dots, K, \quad (1)$$

where $\mathbf{x} \in \mathbb{C}^{M \times 1}$ is the transmitted signal from the base station, $\mathbf{H}_k \in \mathbb{C}^{N_k \times M}$ the channel gain matrix between base station and user k , $\mathbf{n}_k \in \mathbb{C}^{N_k \times 1}$ the additive white Gaussian noise (AWGN) at the k th user with covariance matrix $E\{\mathbf{n}_k \mathbf{n}_k^H\} = \mathbf{I}_{N_k}$, and \mathbf{y}_k the received signal vector of user k .

Assume that the transmitter at the base station has an average power constraint P over M antennas, that is, $E\{\text{tr}(\mathbf{x} \mathbf{x}^H)\} \leq P$. The entries of \mathbf{H}_k are assumed to be independent, and \mathbf{H}_k is assumed to be unchanged for the duration of a frame. Also the channel matrix and AWGN are normalized such that the entries of \mathbf{H}_k and \mathbf{n}_k all have unit variance. For ease of presentation, $N_k = 1$, $k = 1, \dots, K$ is assumed in the following discussions without loss of generality in this paper. As argued in [10, Section VIII], the most straightforward way of extension to multiple receive antennas is to treat each antenna of the k th user with $N_k > 1$ as a separate user, assuming N_k receiver antennas do not coordinate. This uncoordinated receiver strategy results in $\sum_{k=1}^K N_k$ single-antenna users. Thus, in the limit of $K \rightarrow \infty$, we can extend all the algorithm development and asymptotic analysis presented in this paper to multiple receive antennas without loss of generality.

Using transmitter linear precoding at the base station, the transmitted signal vector \mathbf{x} can be expressed as a

linear combination of transmitted symbols for each user. User streams are separated by different precoding vectors, namely, beamforming directions. Let \mathbf{s}_k , \mathbf{w}_k , and P_k be the data symbol, precoding column vector, and allocated power for user k , respectively. The transmitted signal of linear precoding schemes can thus be expressed as

$$\mathbf{x} = \sum_{k=1}^K \sqrt{P_k} \mathbf{w}_k s_k, \quad \sum_{k=1}^K P_k \leq P. \quad (2)$$

As a result, the received signal of user k can be written as

$$y_k = \sqrt{P_k} \mathbf{h}_k \mathbf{w}_k s_k + \sum_{j \neq k} \sqrt{P_j} \mathbf{h}_k \mathbf{w}_j s_j + n_k, \quad (3)$$

where \mathbf{h}_k is a row vector referring to the k th user's channel when $N_k = 1$, and the second term represents the interference caused by all other users to user k .

Treating the interference from all other users as an additive Gaussian white noise (AWGN) and assuming single-user detection is employed, the achievable sum-rate of linear precoding for MIMO-BC channels is given by [10]

$$R = \max_{\mathbf{w}_k, P_k} \sum_{k=1}^K \log_2 \left(1 + \frac{P_k |\mathbf{h}_k \mathbf{w}_k|^2}{1 + \sum_{j=1, j \neq k}^K P_j |\mathbf{h}_k \mathbf{w}_j|^2} \right), \quad (4)$$

$$P_k \geq 0, \quad \sum_{k=1}^K \|\mathbf{w}_k\|^2 P_k \leq P.$$

In contrast, the sum-rate capacity of MIMO-BC channel achieved by DPC strategy was shown in [3, 6]

$$C_{\text{BC}} = R_{\text{DPC}} = \max_{P_k \geq 0, \sum_{k=1}^K P_k < P} \log_2 \left(\left| \mathbf{I} + \sum_{k=1}^K P_k \mathbf{h}_k^H \mathbf{h}_k \right| \right), \quad (5)$$

where $|\cdot|$ denotes the determinant of the matrix inside.

2.2. Zero-forcing precoding scheme (see [10])

ZF precoding scheme selects $S \leq M$ precoding vectors such that $\mathbf{h}_i \mathbf{w}_j = 0$, for all $i \neq j$, that is, no interference among S users. Since the dimension of \mathbf{h}_i is $1 \times M$, there exists at most M precoding vectors satisfying the zero-interference condition. Let $S \subset \{1, \dots, K\}$, $S = |S| \leq M$ be the subset of user indexes selected for transmission, where $|S|$ denotes the size of a set S . Define $\mathbf{H}(S)$ and $\mathbf{W}(S)$ as the corresponding submatrices of channel matrix $\mathbf{H} = [\mathbf{h}_1^T, \dots, \mathbf{h}_K^T]^T$ and precoding matrix $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_K]$ indexed by the subset of selected users S , respectively.

The Moore-Penrose pseudoinverse of $\mathbf{H}(S)$ generates $\mathbf{W}(S)$ that satisfies zero-interference among the following users:

$$\mathbf{W}(S) = \mathbf{H}(S)^\dagger = \mathbf{H}(S)^H (\mathbf{H}(S) \mathbf{H}(S)^H)^{-1}, \quad (6)$$

where \dagger denotes the Moore-Penrose pseudoinverse of the matrix.

Substituting the ZF precoding vector $\mathbf{w}_i \in \mathbf{W}(S)$ into (4), the achievable sum-rate of the ZF scheme for MIMO-BC channels is

$$R_{\text{ZF}}(S) = \max_{P_i \geq 0, \sum_{i \in S} P_i \leq P} \sum_{i \in S} \log_2 (1 + \gamma_i P_i), \quad (7)$$

where γ_i is the effective gain of subchannel i given by

$$\gamma_i = \|\mathbf{w}_i\|^{-2} = \frac{1}{[(\mathbf{H}(S) \mathbf{H}(S)^H)^{-1}]_{i,i}}. \quad (8)$$

Under the condition of zero-interference among parallel subchannels, water-filling strategy [15] gives the optimal power allocation P_i for the optimization problem (7). Since there are many possible choices of selected user subsets, the achievable sum-rate capacity of optimum ZF is defined in [10] as $R_{\text{ZF}} = \max_{S \subset \{1, \dots, K\}: |S| \leq M} R_{\text{ZF}}(S)$.

3. PROPOSED SCHEME

Despite its simplicity and asymptotic sum capacity optimality, as a suboptimal linear precoding scheme, the main disadvantage of ZF scheme is the excess transmission power penalty incurred at the transmitter. In particular, the effective channel gain is greatly reduced when the selected $\mathbf{H}(S)$ is poorly conditioned, which is often the case with small number of users in the system. A channel matrix is called poorly conditioned if the spread of its singular values is large.

In this paper, to alleviate the reduction of effective channel gain, we propose to use CIR [13] as an alternative linear precoding scheme in conjunction with user selection algorithms, for example, SUS [10]. Rather than performing pseudoinverse on the channel matrix of selected users $\mathbf{H}(S)$ directly as conducted in ZF scheme, the CIR scheme adds a multiple of identity matrix before inverting to regularize an inverse for MIMO-BC channels.

Instead of performing Moore-Penrose pseudoinverse to obtain $\mathbf{W}(S)$ using (6), we compute the CIR precoding matrix as

$$\mathbf{W}(S) = \mathbf{H}(S)^H (\mathbf{H}(S) \mathbf{H}(S)^H + \beta \mathbf{I})^{-1}, \quad (9)$$

where \mathbf{I} is an $S \times S$ identity matrix and $\beta \geq 0$ is a load factor introduced for controlling the amount of interference allowed. Clearly, by choosing β sufficiently large, the inverse of a poorly conditioned $\mathbf{H}(S)$ in (6) can be made to behave as well as desired.

Note that in the case of $\beta = 0$, (9) is equivalent to the simple channel inverse of ZF precoding scheme, given by pseudoinverse in (6). The amount of interference increases with β , and can be controlled by β as desired. One reasonable choice of β is to maximize the SINR at the receiver under the assumption of equal power for users is given in [13] as

$$\beta = \frac{M}{P}. \quad (10)$$

Substituting the CIR precoding vector $\mathbf{w}_i \in \mathbf{W}(S)$ in (9) into (4), we can obtain the achievable sum-rate of MIMO-BC channels with CIR as the linear precoding scheme as

$$\begin{aligned} R_{\text{CIR}}(S) &= \max_{P_i} R(\mathbf{P}) \\ &= \max_{P_i} \sum_{i \in S} \log_2 \left(1 + \frac{P_i |\mathbf{h}_i \mathbf{w}_i|^2}{1 + \sum_{j \in S, j \neq i} P_j |\mathbf{h}_i \mathbf{w}_j|^2} \right), \\ &\quad 0 \leq P_i, \quad i \in S, \quad \sum_{i \in S} \|\mathbf{w}_i\|^2 P_i \leq P. \end{aligned} \quad (11)$$

It is worth noting that when $\beta > 0$, the zero-interference condition $\mathbf{h}_i \mathbf{w}_j = 0$, for all $i \neq j$ is no longer satisfied with the CIR precoding scheme. In other words, the received signal at user i contains the interference from other users, denoted by $\sum_{j \in S, j \neq i} P_j |\mathbf{h}_i \mathbf{w}_j|^2$. This residual interference from other users makes the optimization problem (11) dramatically different from the convex optimization problem (7) resulting from ZF precoding scheme. In fact, the optimization problem in (11) cannot be reduced to a simple convex optimization problem, and is essentially a nonlinear nonconvex optimization problem with many local optima. Unlike the closed-form water-filling solution of convex optimization problem (7), the optimum power allocation for the nonlinear nonconvex optimization problem (11) defies a closed-form solution, and requires a constrained numerical optimization algorithm to find the optimal power allocation for selected user subset S .

The global optimum of nonconvex optimization problem generally cannot be found by conventional optimization techniques which are only capable of finding one of the local optima. To solve the nonconvex optimization problem (11) globally, we can resort to global optimization techniques that are developed to solve globally multiextreme optimization problems arising from important practical applications [16]. Yet the successful application of global optimization methods to solve practical problems depends heavily on exploiting the specific structure of a problem.

3.1. Global d.c. optimization approach

The objective function in optimization problem (11) can be decomposed into the following form:

$$\begin{aligned} -R(\mathbf{P}) &= -\sum_{i \in S} \log_2 \left(1 + \frac{P_i |\mathbf{h}_i \mathbf{w}_i|^2}{1 + \sum_{j \in S, j \neq i} P_j |\mathbf{h}_i \mathbf{w}_j|^2} \right) \\ &= -\sum_{i \in S} \log_2 \left(\frac{1 + \sum_{j \in S} P_j |\mathbf{h}_i \mathbf{w}_j|^2}{1 + \sum_{j \in S, j \neq i} P_j |\mathbf{h}_i \mathbf{w}_j|^2} \right) \\ &= -\sum_{i \in S} \log_2 \left(1 + \sum_{j \in S} P_j |\mathbf{h}_i \mathbf{w}_j|^2 \right) \\ &\quad + \sum_{i \in S} \log_2 \left(1 + \sum_{j \in S, j \neq i} P_j |\mathbf{h}_i \mathbf{w}_j|^2 \right) \\ &= f(\mathbf{P}) - g(\mathbf{P}), \end{aligned} \quad (12)$$

where

$$\begin{aligned} f(\mathbf{P}) &= -\sum_{i \in S} \log_2 \left(1 + \sum_{j \in S} P_j |\mathbf{h}_i \mathbf{w}_j|^2 \right), \\ g(\mathbf{P}) &= -\sum_{i \in S} \log_2 \left(1 + \sum_{j \in S, j \neq i} P_j |\mathbf{h}_i \mathbf{w}_j|^2 \right), \end{aligned} \quad (13)$$

where $\mathbf{P} = [P_1, \dots, P_S]^T$ is power allocation vector for the selected users.

Because the sum of convex functions is convex, $g(\mathbf{P}) = -\sum_{i \in S} \log_2(u_i(\mathbf{P}))$ is convex if $-\log_2(u_i(\mathbf{P}))$ is convex, that is, $\log_2(u_i(\mathbf{P}))$ is concave, where $u_i(\mathbf{P}) = 1 + \sum_{j \in S, j \neq i} P_j |\mathbf{h}_i \mathbf{w}_j|^2$. Since $\log_2 x$ is a nondecreasing concave function, and $u_i(\mathbf{P})$ is a concave function as it is a linear function of \mathbf{P} , $\log_2(u_i(\mathbf{P}))$ is concave because the composition function $w = v \circ u$ is concave if v is concave and nondecreasing, and u is concave [17]. Similarly, we can show that $f(\mathbf{P})$ is also convex. Hence, $-R(\mathbf{P}) = f(\mathbf{P}) - g(\mathbf{P})$ is a direct connection function. It follows that

$$\max R(\mathbf{P}) \equiv \min(-R(\mathbf{P})) = \min[f(\mathbf{p}) - g(\mathbf{p})], \quad \mathbf{P} \in D, \quad (14)$$

where $D \triangleq \{\mathbf{P} \in \mathfrak{R}^N : 0 \leq P_i, i \in S \text{ and } \sum_{i \in S} \|\mathbf{w}_i\|^2 P_i \leq P\}$ are a convex set in \mathfrak{R}^N . Thus, $\max R(\mathbf{P})$ in (11) is a global d.c. optimization problem [16] similar to that formulated for optimal spectrum balancing (OSB) in digital subscriber lines (DSLs) interference channels [14]. This particular class of d.c. optimization problems with only linear constraints can be transformed into their equivalent global concave minimization problems [18]. The prismatic branch and bound (PBnB) algorithm was introduced for solving the concave minimization formulation for OSB in DSL in [14], which can find the global optimum efficiently via only solving a sequence of linear programming (LP) subproblems. The detailed description of PBnB algorithm was presented in [14] and is omitted in this paper for the sake of brevity.

3.2. Local gradient projection approach

To solve the resulting nonlinear optimization problem of CIR scheme efficiently, we propose to use local gradient-projection (GP) method [19], which is an extension of the unconstrained steepest descent method particularly suitable for constrained optimization problem with convex feasible region. Since the feasible region of our particular power allocation optimization problem (11) is a geometrically simple simplex, GP method can find the optimum solution of optimization problem (11) very efficiently.

Recall that the feasible region of optimization problem (11) is defined by

$$\mathfrak{S}(P) = \left\{ (P_1, \dots, P_M) \in \mathfrak{R}^M \mid P_i \geq 0, \sum_{i=1}^M \|\mathbf{w}_i\|^2 P_i \leq P \right\}. \quad (15)$$

In order to apply the GP method to our particular optimization problem in (11), we have to first show how

Initialization:

$$l = 0; \mathbf{P}^0 = \left[\|\mathbf{w}_1\|^2 P / \sum_{i=1}^M \|\mathbf{w}_i\|^2, \dots, \|\mathbf{w}_M\|^2 P / \sum_{i=1}^M \|\mathbf{w}_i\|^2 \right]$$

Iteration:

do

(1) Compute the gradient $g_i^l = \frac{\partial R_{\text{CIR}}(\mathbf{P}^l)}{\partial P_i}$, $i \in S$ by (18)

(2) $\bar{P}_i^l = P_i^l + s_i g_i^l$, $i \in S$

(3) \hat{P}_i^l obtained by projecting \bar{P}_i^l onto $\mathbb{S}(P)$, $i \in S$

(4) Applying Armijo rule along feasible direction to compute δ^l

(5) $P_i^{l+1} = P_i^l + \delta^l (\hat{P}_i^l - P_i^l)$, $i \in S$

(6) $l = l + 1$

until $\max_i (|P_i^l - P_i^{l-1}|) \leq \varepsilon$

ALGORITHM 1: Optimal power allocation based on GP method.

to perform projection onto the constraint set given by the simplex $\mathbb{S}(P)$ and the gradient of the objective function in (11), respectively.

The projection of point outside a simplex, that is, $\sum_{i=1}^M \|\mathbf{w}_i\|^2 P_i \geq P$, is straightforward. The distance from a point $\mathbf{X} = (x_1, \dots, x_M)$ to the hyperplane defined by $\sum_{i=1}^M \|\mathbf{w}_i\|^2 P_i = P$ is

$$D(\mathbf{X}) = \frac{\left| \sum_{i=1}^M \|\mathbf{w}_i\|^2 x_i - P \right|}{\sum_{i=1}^M \|\mathbf{w}_i\|^2}. \quad (16)$$

Then, the projection of a point X onto this hyperplane is

$$\hat{x}_i = x_i - \frac{\|\mathbf{w}_i\|^2 D(\mathbf{X})}{\sum_{i=1}^M \|\mathbf{w}_i\|^2}, \quad i = 1, \dots, M. \quad (17)$$

The gradient g_i is derived as

$$\begin{aligned} g_i &= \frac{\partial R_{\text{CIR}}}{\partial P_i} \\ &= \frac{1}{\log_e 2} \left(\frac{1 + \sum_{j \in S, j \neq i} P_j |\mathbf{h}_i \mathbf{w}_j|^2}{1 + \sum_{j \in S} P_j |\mathbf{h}_i \mathbf{w}_j|^2} \right) |\mathbf{h}_i \mathbf{w}_i|^2 \\ &\quad - \frac{1}{\log_e 2} \sum_{j \in S, j \neq i} \left(\frac{|\mathbf{h}_j \mathbf{w}_i|^2}{1 + \sum_{l \in S, l \neq j} P_l |\mathbf{h}_j \mathbf{w}_l|^2} \right). \end{aligned} \quad (18)$$

The proposed algorithm using GP method is presented in Algorithm 1.

The GP algorithm is guaranteed to converge to a stationary point [19]. In most of the numerical simulations we conducted for the CIR precoding with optimal power control, the GP algorithm can converge successfully to a local maximum solution in less than 40 iterations.

4. ASYMPTOTIC ANALYSIS

In this section, we first study the asymptotic sum-rate performance of our proposed CIR precoding in conjunction

with SUS user selection algorithm in MIMO-BC channels in the limit of $K \rightarrow \infty$. Then, we show that the proposed local GP algorithm can achieve the global optimum of the nonconvex optimization problem (11) as $K \rightarrow \infty$. Throughout the proof in this section, we assume that $|S| = M$ is almost surely true with sufficiently large K as argued in [10].

4.1. Asymptotic sum-rate analysis

It has been shown in [10] that the ZF precoding scheme with SUS algorithm has the same asymptotic sum-rate as that of DPC. Thus, we only need to prove that the CIR precoding vectors $\mathbf{W}(S)$ are equivalent to the ZF precoding vectors as $K \rightarrow \infty$.

We define the measure of orthogonality between two arbitrary vectors \mathbf{a} and \mathbf{b} as

$$\rho(\mathbf{a}, \mathbf{b}) \triangleq \frac{|\mathbf{a}^H \mathbf{b}|}{\|\mathbf{a}\| \|\mathbf{b}\|}. \quad (19)$$

Clearly, as \mathbf{a} and \mathbf{b} become nearly orthogonal to each other, $\rho(\mathbf{a}, \mathbf{b})$ tends to closely approach zero.

To maximize the sum-rate using ZF precoding scheme for MIMO-BC channels, the user selection algorithms SUS in [10] chooses a group of users with nearly orthogonal channel responses and also with sufficiently large magnitude of the channel responses.

SUS algorithm proposed in [10] introduced a small positive constant α to control the degree of semiorthogonality among selected users. More specifically, at iteration number i , user k is considered in the next iteration if and only if $\rho(\mathbf{g}_i, \mathbf{h}_k) < \alpha$ is satisfied, where \mathbf{g}_i is the orthogonal vector derived from \mathbf{h}_k by using the Gram-Schmidt procedure [10]. Using some simple but tedious algebraic operations, we can show that $\rho(\mathbf{g}_i, \mathbf{h}_k) < \alpha$ implies an α -orthogonal $\mathbf{H}(S)$. A set of selected channel vectors $\mathbf{H}(S)$ is called ε -orthogonal if $\rho(\mathbf{h}_i, \mathbf{h}_j) \leq \varepsilon$, for all $i, j \in S$, $i \neq j$.

It can be shown that for an arbitrary small positive ε , we can construct an ε -orthogonal $\mathbf{H}(S)$ by SUS algorithm with probability one with sufficiently larger number of user K [10]. Hence, the channel vectors of selected users obtained by SUS algorithm become increasingly more orthogonal as K increases, and eventually these M vectors form an orthogonal basis for the space of \mathbb{C}^M in the limit of $K \rightarrow \infty$.

Recall that the CIR precoding matrix is computed by $\mathbf{W}(S) = \mathbf{H}(S)^H (\mathbf{H}(S) \mathbf{H}(S)^H + \beta \mathbf{I})^{-1}$. In the limit of $K \rightarrow \infty$, we have

$$\begin{aligned} \mathbf{W}_{\text{CIR}}(S) &= \mathbf{H}(S)^H (\text{diag}(\|\mathbf{h}_1\|^2, \dots, \|\mathbf{h}_M\|^2) + \beta \mathbf{I})^{-1} \\ &= \mathbf{H}(S)^H (\text{diag}((\|\mathbf{h}_1\|^2 + \beta)^{-1}, \dots, (\|\mathbf{h}_M\|^2 + \beta)^{-1})). \end{aligned} \quad (20)$$

Compared with the ZF precoding matrix,

$$\begin{aligned} \mathbf{W}_{\text{ZF}}(S) &= \mathbf{H}(S)^H (\mathbf{H}(S) \mathbf{H}(S)^H)^{-1} \\ &= \mathbf{H}(S)^H (\text{diag}(\|\mathbf{h}_1\|^{-2}, \dots, \|\mathbf{h}_M\|^{-2})), \end{aligned} \quad (21)$$

it is evident that the normalized precoding vectors of both the CIR and ZF schemes are exactly the same. Therefore, the expected sum-rate performance of the CIR is equivalent to that of ZF precoding scheme as $K \rightarrow \infty$. Since ZF precoding scheme was shown in [10] to achieve asymptotically the optimum sum-rate, the proposed CIR precoding with SUS can thus achieve asymptotically the optimum sum-rate capacity equal to that of the optimum DPC strategy as $K \rightarrow \infty$, that is,

$$E\{R_{\text{CIR}}\} \sim M \log_2 \left(1 + \frac{P}{M} \log K \right) \sim E\{R_{\text{DPC}}\}. \quad (22)$$

4.2. Asymptotic convexity analysis

The asymptotic convexity analysis for the nonconvex optimization problem (11) reveals that the objective function can be regarded as a convex function of users' power allocation vector \mathbf{P} when $K \rightarrow \infty$.

The d.c. representation in (14) establishes the equivalence between the original sum-rate maximization problem (11) and its d.c. formulation, that is, $\max R(\mathbf{P}) \equiv \min(-R(\mathbf{P})) = \min[f(\mathbf{p}) - g(\mathbf{p})]$. As previously discussed, $\mathbf{H}(S)$ obtained by SUS can be made as arbitrarily orthogonal as desired with sufficient large K . In other words, the CIR precoding vectors $\mathbf{W}_{\text{CIR}}(S)$ given in (20) are just a scaled rotation of $\mathbf{H}(S)$ in the limit of $K \rightarrow \infty$. Hence, we have

$$\begin{aligned} & \lim_{K \rightarrow \infty} -g(\mathbf{P}) \\ &= \lim_{K \rightarrow \infty} \sum_{i \in S} \log_2 \left(1 + \sum_{j \in S, j \neq i} P_j |\mathbf{h}_i \mathbf{w}_j|^2 \right) \\ &\stackrel{a}{\approx} \lim_{K \rightarrow \infty} \sum_{i \in S} \log_2 \left(1 + \sum_{j \in S, j \neq i} P_j |\mathbf{h}_i \mathbf{h}_j|^2 (\|\mathbf{h}_j\|^2 + \beta)^{-2} \right) \\ &\stackrel{b}{\approx} \lim_{K \rightarrow \infty} \log_2 e \sum_{i \in S} \sum_{j \in S, j \neq i} P_j |\mathbf{h}_i \mathbf{h}_j|^2 (\|\mathbf{h}_j\|^2 + \beta)^{-2} \\ &\stackrel{c}{\leq} P \log_2 e \lim_{K \rightarrow \infty} \sum_{i \in S} \sum_{j \in S, j \neq i} \epsilon \|\mathbf{h}_i\|^2 \|\mathbf{h}_j\|^{-2} \\ &\stackrel{d}{\leq} PM^2 C \log_2 e \lim_{K \rightarrow \infty} \epsilon \end{aligned} \quad (23)$$

where $x \approx y$ indicates that $\lim_{K \rightarrow \infty} x/y = 1$. In (a), we use $\mathbf{W}_{\text{CIR}}(S)$ in (20); (b) follows from $\lim_{x \rightarrow 0} \log_2(1+x) \approx x \log_2 e$; inequality (c) is obtained by $\rho(\mathbf{h}_i, \mathbf{h}_j) \leq \epsilon$, for all $i, j \in S, i \neq j$, $P_i \leq P$ and decreasing the denominator from $(\|\mathbf{h}_j\|^2 + \beta)^2$ to $(\|\mathbf{h}_j\|^2)^2$; inequality (d) follows the approximation of the number of combinations $\sum_{i \in S} \sum_{j \in S, j \neq i}$ by M^2 and the constant C is defined as $C = \max_{i \neq j \in S} \|\mathbf{h}_i\|^2 \|\mathbf{h}_j\|^{-2}$.

Therefore, the nonconvex component $-g(\mathbf{p})$ can be essentially ignored in the d.c. formulation $-R(\mathbf{P}) = f(\mathbf{p}) - g(\mathbf{p})$. As $f(\mathbf{p})$ was shown to be convex, the optimization problem $\max R(\mathbf{P}) \equiv \min(-R(\mathbf{P}))$ in (14) can be regarded as a convex optimization problem as $K \rightarrow \infty$. It implies that the local optimum solution obtained by GP method is actually the global optimum of the optimization problem

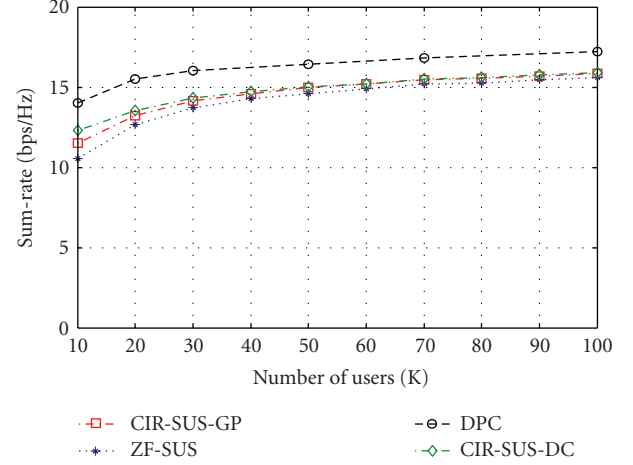


FIGURE 1: Sum-rate performance comparison of DPC, ZF-SUS, and CIR-SUS.

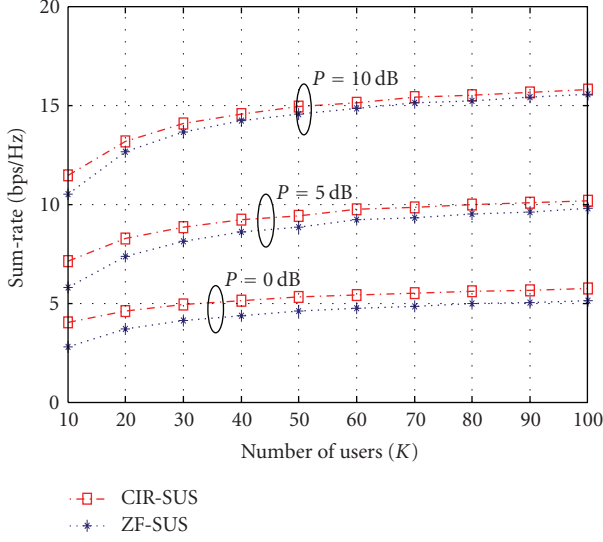
(11), because standard optimization techniques for finding local solutions yield the global optimum [17] for a convex optimization problem. Moreover, it provides the theoretical justification for using the local GP method to find the optimum solution of our sum-rate maximization problem, as the performance loss compared to that of the global d.c. approach is negligible whilst the computational complexity reduction is significant.

5. NUMERICAL RESULTS

In this section, illustrative numerical sum-rate results of the CIR-SUS scheme are presented to evaluate and compare its performance with that of the ZF-SUS scheme in MIMO-BC channels. In all simulations, the number of transmit antennas is $M = 4$, and $\alpha = 0.4$ for controlling semiorthogonality [10] in all SUS algorithms (ZF-SUS, CIR-SUS-GP, and CIR-SUS-DC). By doing this, all SUS algorithms have the same and fixed selected user subsets. All the plots are obtained by averaging over 500 independent channel realizations.

Figure 1 first compares the sum-rate performance versus the number of users in the system for DPC, ZF-SUS scheme, the proposed CIR-SUS scheme with local GP method, and the CIR-SUS with global d.c. method. The SNR is set to 10 dB which is equal to the transmit power $P = 10$ dB. It shows that the proposed scheme can achieve a higher sum-rate capacity than ZF with the same selected user subset obtained by the SUS algorithm. Note that the gain of CIR over ZF precoding decreases with the number of users. This reduction of gain is because channel matrix of selected user subset tends to be more near-orthogonal with a higher number of users. As a result, the effective channel gain reduction encountered by ZF would become less prominent when the number of users is larger.

Figure 1 also shows that the CIR-SUS with global d.c. optimization approach only leads to a slight ergodic sum-rate gain over CIR-SUS with local GP method in the case of small number of users is small. Furthermore, the gain offered by global d.c. approach quickly diminishes with increasing

FIGURE 2: Sum-rate comparison of ZF-SUS and CIR-SUS versus K .

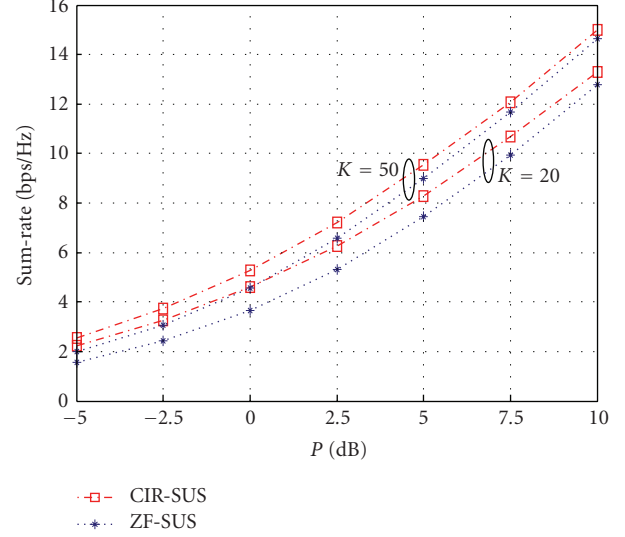
a large number of users. The simulation results support our claim in Section 4.2 that the nonconvex optimization problem can be essentially regarded as a convex optimization problem with a sufficiently large number of users. Hence, in the simulation results presented in Figures 2 and 3, the sum-rate performance of CIR-SUS is obtained by using local GP optimization method.

Figure 2 compares the sum-rate performance of CIR-SUS and ZF-SUS versus the number of users K at three different levels of the base station transmission power, that is, $P = 0$ dB, 5 dB, and 10 dB. Figure 2 shows that the gain of CIR over ZF increases with lower transmission power. This is not surprising because the rate function $\log(1+x) \approx x$ when x is small. In other words, when transmission power is low, the rate function behaves like a linear function rather than as a logarithmic function. Hence, any reduction of effective channel gain experienced by ZF is directly translated into a linear loss of sum-rate when SNR is low.

The sum-rate performance of CIR-SUS and ZF-SUS versus the transmission power is compared in Figure 3 for both the cases of 20 users and 50 users. Clearly, the gain of CIR over ZF increases when the number of users K is less or the transmit power is lower.

6. CONCLUSIONS

In this paper, for MIMO-BC channels, we proposed to use CIR as the precoding scheme in conjunction with user selection algorithm to alleviate the transmission power penalty incurred by ZF. Unlike the interference-free ZF, the optimization problem of maximizing sum-rate capacity using CIR precoding is nonlinear and nonconvex. A local GP method is then proposed to solve this optimization problem efficiently with negligible sum-rate performance loss compared to the global d.c. optimization technique, but with much less computational complexity. Asymptotic analysis shows that the proposed CIR precoding scheme in

FIGURE 3: Sum-rate comparison of ZF-SUS and CIR-SUS versus P .

conjunction with user selection algorithm can achieve the optimum ergodic sum-rate equal to that of DPC in the limit of a large number of users. Simulation results under various conditions show that the proposed CIR precoding scheme with optimal power allocation achieves better sum-rate performance than ZF in MIMO-BC channels, especially in the practical range of number of users, for example, $K \leq 100$, and at relatively low SNR.

REFERENCES

- [1] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.
- [2] 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, no. 3, pp. 311–335, 1998.
- [3] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, 2003.
- [4] W. Yu and J. M. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 1875–1892, 2004.
- [5] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2658–2668, 2003.
- [6] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Transactions on Information Theory*, vol. 49, no. 8, pp. 1912–1921, 2003.
- [7] M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, 1983.
- [8] W. Yu and J. M. Cioffi, "Treillis precoding for the broadcast channel," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '01)*, vol. 2, pp. 1344–1348, San Antonio, Tex, USA, November 2001.

- [9] J. Kusuma and K. Ramchandran, "Communicating by cosets and applications to broadcast," in *Proceedings of the Conference on Information Sciences and Systems (CISS '02)*, Princeton, NJ, USA, March 2002.
- [10] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 528–541, 2006.
- [11] H. R. Bahrami and T. Le-Ngoc, "A capacity achieving precoding scheme based on partial channel information for broadcast MIMO systems," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '07)*, pp. 1045–1050, Hong Kong, March 2007.
- [12] A. Bayesteh and A. K. Khandani, "On the user selection for MIMO broadcast channels," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '05)*, pp. 2325–2329, Adelaide, Australia, September 2005.
- [13] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication—part I: channel inversion and regularization," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 195–202, 2005.
- [14] Y. Xu, T. Le-Ngoc, and S. Panigrahi, "Global concave minimization for optimal spectrum balancing in multi-user DSL networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, part 1, pp. 2875–2885, 2008.
- [15] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, USA, 1991.
- [16] R. Horst and H. Tuy, *Global Optimization: Deterministic Approaches*, Springer, Berlin, Germany, 3rd edition, 1995.
- [17] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2003.
- [18] R. Horst, T. Q. Phong, Ng. V. Thoai, and J. de Vries, "On solving a D.C. programming problem by a sequence of linear programs," *Journal of Global Optimization*, vol. 1, no. 2, pp. 183–203, 1991.
- [19] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, Mass, USA, 1995.