

# Conjugate Gradient Projection Approach for Multi-Antenna Gaussian Broadcast Channels

Jia Liu<sup>†</sup>   Y. Thomas Hou<sup>†</sup>   Hanif D. Sherali<sup>‡</sup>

<sup>†</sup> The Bradley Department of Electrical and Computer Engineering

<sup>‡</sup> The Grado Department of Industrial and Systems Engineering

Virginia Tech, Blacksburg, VA, USA

## Abstract

It has been shown recently that the dirty-paper coding is the optimal strategy for maximizing the sum rate of multiple-input multiple-output Gaussian broadcast channels (MIMO BC). Moreover, by the channel duality, the nonconvex MIMO BC sum rate problem can be transformed to the convex dual MIMO multiple-access channel (MIMO MAC) problem with a sum power constraint. In this paper, we design an efficient algorithm based on conjugate gradient projection (CGP) to solve the MIMO BC maximum sum rate problem. Our proposed CGP algorithm solves the dual sum power MAC problem by utilizing the powerful concept of Hessian conjugacy. We also develop a rigorous algorithm to solve the projection problem. We show that CGP enjoys provable convergence, nice scalability, and great efficiency for large MIMO BC systems.

## 1 Introduction

Recently, researchers have shown great interests in characterizing the capacity region for multiple-input multiple-output broadcast channels (MIMO BC) and MIMO multiple-access channels (MIMO MAC). In particular, although the general capacity region for MIMO BC remains an open problem [1], the sum rate region has been shown achievable by the dirty-paper coding strategy [2, 3]. Moreover, by the remarkable channel *duality* between MIMO BC and MIMO MAC established in [4–6], the nonconvex MIMO BC sum rate problem can be transformed to the convex dual MIMO MAC problem with a sum power constraint.

However, although the standard interior point convex optimization method can be used to solve the sum power MIMO MAC problem, its complexity is considerably higher than those methods that exploit the special structure of the sum power MIMO MAC problem. Such specifically designed

algorithms include the minimax method (MM) by Lan and Yu [7], the steepest descent (SD) method by Viswanathan *et al.* [8], the dual decomposition (DD) method by Yu [9], and two iterative water-filling methods (IWFs) by Jindal *et al.* [10]. Among these algorithms, MM is more complex than the others having the linear complexity. SD and DD have longer running time per iteration than IWFs due to line searches and the inner optimization, respectively. Both IWFs in [10], however, do not scale well as the number of users, denoted by  $K$ , increases. The reason is that in each iteration of IWFs, the most recently updated solution only accounts for a fraction of  $1/K$  in the effective channels' computation. The authors of [10] proposed a hybrid algorithm as a remedy. But the hybrid algorithm introduces additional complexity in implementation and its performance depends upon the empirical switch timing, which, in turn, depends upon specific problems. In addition, one of the IWFs in [10], although converges relatively faster than the other one, requires a total storage size for  $K^2$  input covariance matrices. These limitations of the existing algorithms motivate us to design an efficient and scalable algorithm with a modest storage requirement for solving large MIMO BC systems.

Our major contribution in this paper is that we design a fast algorithm based on the Conjugate Gradient Projection (CGP) approach. Our algorithm is inspired by [11], where a gradient projection method was used to heuristically solve another nonconvex maximum sum rate problem for single-hop MIMO-based ad hoc networks with mutual interference. However, unlike [11], we use the *conjugate* gradient directions instead of gradient directions to eliminate the “zigzagging” phenomenon. Also, we develop a rigorous algorithm to exactly solve the projection problem (in [11], the way of handling gradient projection is based on heuristic: The authors simply set the first derivative to zero to get the solution when solving the constrained Lagrangian dual of the projection problem). The attractive features of our proposed CGP are as follows:

1. CGP is extremely fast, and enjoys provable convergence as well as nice scalability. As opposed to IWFs, the number of iterations required for convergence in CGP is very insensitive to the increase of the number of users.

2. CGP has the desirable linear complexity. By adopting the inexact line search method called “Armijo’s Rule”, we show that CGP has a comparable complexity to IWFs per iteration, and requires much fewer iterations for convergence in large MIMO BC systems.
3. CGP has a modest memory requirement: It only needs the solution information from the previous step, as opposed to one of the IWFs, which requires the solution information from previous  $K - 1$  steps. Moreover, CGP is very intuitive and easy to implement.

The remainder of this paper is organized as follows. In Section 2, we discuss the network model and formulation. Section 3 introduces the key components in our CGP framework, including conjugate gradient computation and how to perform projection. We analyze and compare the complexity of CGP with other existing algorithms in Section 4. Numerical results are presented in Section 5. Section 6 concludes this paper.

## 2 System Model and Problem Formulation

We begin with introducing notation. We use boldface to denote matrices and vectors. For a complex-valued matrix  $\mathbf{A}$ ,  $\mathbf{A}^*$  and  $\mathbf{A}^\dagger$  denotes the conjugate and conjugate transpose of  $\mathbf{A}$ , respectively.  $\text{Tr}\{\mathbf{A}\}$  denotes the trace of  $\mathbf{A}$ . We let  $\mathbf{I}$  denote the identity matrix with dimension determined from context.  $\mathbf{A} \succeq 0$  represents that  $\mathbf{A}$  is Hermitian and positive semidefinite (PSD).  $\text{Diag}\{\mathbf{A}_1 \dots \mathbf{A}_n\}$  represents the block diagonal matrix with matrices  $\mathbf{A}_1, \dots, \mathbf{A}_n$  on its main diagonal.

Suppose that a MIMO Gaussian broadcast channel has  $K$  users, each of which is equipped with  $n_r$  antennas, and the transmitter has  $n_t$  antennas. The channel matrix for user  $i$  is denoted as  $\mathbf{H}_i \in \mathbb{C}^{n_r \times n_t}$ .

In [2, 4–6], it has been shown that the maximum sum rate capacity of MIMO BC is equal to the dirty-paper coding region, which can be computed by solving the optimization problem as follows:

$$\begin{aligned}
& \text{Maximize} && \sum_{i=1}^K \log \frac{\det(\mathbf{I} + \mathbf{H}_i (\sum_{j=1}^i \mathbf{\Gamma}_j) \mathbf{H}_i^\dagger)}{\det(\mathbf{I} + \mathbf{H}_i (\sum_{j=1}^{i-1} \mathbf{\Gamma}_j) \mathbf{H}_i^\dagger)} \\
& \text{subject to} && \mathbf{\Gamma}_i \succeq 0, \quad i = 1, 2, \dots, K \\
& && \sum_{i=1}^K \text{Tr}(\mathbf{\Gamma}_i) \leq P,
\end{aligned} \tag{1}$$

where  $\mathbf{\Gamma}_i \in \mathbb{C}^{n_t \times n_t}$ ,  $i = 1, \dots, K$ , are the downlink input covariance matrices. It is evident that (1) is a nonconvex optimization problem. However, the authors in [4, 6] showed that due to the duality between MIMO BC and MIMO MAC, (1) is equivalent to the following MIMO MAC problem with a sum power constraint:

$$\begin{aligned} & \text{Maximize} \quad \log \det \left( \mathbf{I} + \sum_{i=1}^K \mathbf{H}_i^\dagger \mathbf{Q}_i \mathbf{H}_i \right) \\ & \text{subject to} \quad \mathbf{Q}_i \succeq 0, \quad i = 1, 2, \dots, K \\ & \quad \quad \quad \sum_{i=1}^K \text{Tr}(\mathbf{Q}_i) \leq P, \end{aligned} \tag{2}$$

where  $\mathbf{Q}_i \in \mathbb{C}^{n_r \times n_r}$ ,  $i = 1, \dots, K$  are the uplink input covariance matrices. For convenience, we use the matrix  $\mathbf{Q} = [\mathbf{Q}_1 \quad \mathbf{Q}_2 \quad \dots \quad \mathbf{Q}_K]$  to denote the set of all uplink input covariance matrices, and let  $F(\mathbf{Q}) = \log \det \left( \mathbf{I} + \sum_{i=1}^K \mathbf{H}_i^\dagger \mathbf{Q}_i \mathbf{H}_i \right)$  represent the objective function of (2). After solving (2), we can recover the solutions of (1) by the mapping proposed in [4].

### 3 Conjugate Gradient Projection for MIMO BC

In this paper, we propose an efficient algorithm based on conjugate gradient projection (CGP) to solve (2). CGP utilizes the important and powerful concept of Hessian conjugacy to deflect the gradient direction appropriately so as to achieve the superlinear convergence rate [12], which is similar to that of the well-known quasi-Newton methods (e.g., BFGS method). CGP follows the same idea of gradient projection which was originally proposed by Rosen [13]. In each iteration, CGP projects the conjugate gradient direction to find an improving feasible direction. Its convergence proof can be found in [12]. The framework of CGP for solving (2) is shown in Algorithm 1.

Due to the complexity of the objective function in (2), we adopt the inexact line search method called ‘‘Armijo’s Rule’’ to avoid excessive objective function evaluations, while still enjoying provable convergence [12]. The basic idea of Armijo’s Rule is that at each step of the line search, we sacrifice accuracy for efficiency as long as we have sufficient improvement. According to Armijo’s Rule, in the  $k^{th}$  iteration, we choose  $\sigma_k = 1$  and  $\alpha_k = \beta^{m_k}$  (the same as in [11]), where  $m_k$  is the first

---

**Algorithm 1** Gradient Projection Method
 

---

**Initialization:**

Choose the initial conditions  $\mathbf{Q}^{(0)} = [\mathbf{Q}_1^{(0)}, \mathbf{Q}_2^{(0)}, \dots, \mathbf{Q}_K^{(0)}]^T$ . Let  $k = 0$ .

**Main Loop:**

1. Calculate the conjugate gradients  $\mathbf{G}_i^{(k)}$ ,  $i = 1, 2, \dots, K$ .
  2. Choose an appropriate step size  $s_k$ . Let  $\mathbf{Q}_i'^{(k)} = \mathbf{Q}_i^{(k)} + s_k \mathbf{G}_i^{(k)}$ , for  $i = 1, 2, \dots, K$ .
  3. Let  $\bar{\mathbf{Q}}^{(k)}$  be the projection of  $\mathbf{Q}'^{(k)}$  onto  $\Omega_+(P)$ , where  $\Omega_+(P) \triangleq \{\mathbf{Q}_i, i = 1, \dots, K | \mathbf{Q}_i \succeq 0, \sum_{i=1}^K \text{Tr}\{\mathbf{Q}_i\} \leq P\}$ .
  4. Choose appropriate step size  $\alpha_k$ . Let  $\mathbf{Q}_i^{(k+1)} = \mathbf{Q}_i^{(k)} + \alpha_k (\bar{\mathbf{Q}}_i^{(k)} - \mathbf{Q}_i^{(k)})$ ,  $i = 1, 2, \dots, K$ .
  5.  $k = k + 1$ . If the maximum absolute value of the elements in  $\mathbf{Q}_i^{(k)} - \mathbf{Q}_i^{(k-1)} < \epsilon$ , for  $i = 1, 2, \dots, L$ , then stop; else go to step 1.
- 

non-negative integer  $m$  that satisfies

$$\begin{aligned}
 F(\mathbf{Q}^{(k+1)}) - F(\mathbf{Q}^{(k)}) &\geq \sigma \beta^m \langle \mathbf{G}^{(k)}, \bar{\mathbf{Q}}^{(k)} - \mathbf{Q}^{(k)} \rangle \\
 &= \sigma \beta^m \sum_{i=1}^K \text{Tr} \left[ \mathbf{G}_i^{\dagger(k)} \left( \bar{\mathbf{Q}}_i^{(k)} - \mathbf{Q}_i^{(k)} \right) \right], \tag{3}
 \end{aligned}$$

where  $0 < \beta < 1$  and  $0 < \sigma < 1$  are fixed scalars.

Next, we will consider two major components in the CGP framework: 1) how to compute the conjugate gradient direction  $\mathbf{G}_i$ , and 2) how to project  $\mathbf{Q}'^{(k)}$  onto the set  $\Omega_+(P) \triangleq \{\mathbf{Q}_i, i = 1, \dots, K | \mathbf{Q}_i \succeq 0, \sum_{i=1}^K \text{Tr}\{\mathbf{Q}_i\} \leq P\}$ .

### 3.1 Computing the Conjugate Gradients

The gradient  $\bar{\mathbf{G}}_i \triangleq \nabla_{\mathbf{Q}_i} F(\mathbf{Q})$  depends on the partial derivatives of  $F(\mathbf{Q})$  with respect to  $\mathbf{Q}_i$ . By using the formula  $\frac{\partial \ln \det(\mathbf{A} + \mathbf{B}\mathbf{X}\mathbf{C})}{\partial \mathbf{X}} = [\mathbf{C}(\mathbf{A} + \mathbf{B}\mathbf{X}\mathbf{C})^{-1}\mathbf{B}]^T$  [11, 14], we can compute the partial derivative of  $F(\mathbf{Q})$  with respect to  $\mathbf{Q}_i$  as follows (by letting  $\mathbf{A} = \mathbf{I} + \sum_{j=1, j \neq i}^K \mathbf{H}_j^\dagger \mathbf{Q}_j \mathbf{H}_j$ ,  $\mathbf{B} = \mathbf{H}_i^\dagger$ ,  $\mathbf{X} = \mathbf{Q}_i$ , and  $\mathbf{C} = \mathbf{H}_i$ ):

$$\begin{aligned}
 \frac{\partial F(\mathbf{Q})}{\partial \mathbf{Q}_i} &= \frac{\partial}{\partial \mathbf{Q}_i} \log \det \left( \mathbf{I} + \sum_{j=1}^K \mathbf{H}_j^\dagger \mathbf{Q}_j \mathbf{H}_j \right) \\
 &= \left[ \mathbf{H}_i \left( \mathbf{I} + \sum_{j=1}^K \mathbf{H}_j^\dagger \mathbf{Q}_j \mathbf{H}_j \right)^{-1} \mathbf{H}_i^\dagger \right]^T. \tag{4}
 \end{aligned}$$

Further, from the definition  $\nabla_z f(z) = 2(\partial f(z)/\partial z)^*$  [15], we have

$$\bar{\mathbf{G}}_i = 2 \left( \frac{\partial F(\mathbf{Q})}{\partial \mathbf{Q}_i} \right)^* = 2\mathbf{H}_i \left( \mathbf{I} + \sum_{j=1}^K \mathbf{H}_j^\dagger \mathbf{Q}_j \mathbf{H}_j \right)^{-1} \mathbf{H}_i^\dagger. \quad (5)$$

Then, the conjugate gradient direction can be computed as  $\mathbf{G}_i^{(k)} = \bar{\mathbf{G}}_i^{(k)} + \rho_k \mathbf{G}_i^{(k-1)}$ . In this paper, we adopt the Fletcher and Reeves' choice of deflection [12]. The Fletcher and Reeves' choice of deflection can be computed as

$$\rho_k = \frac{\|\bar{\mathbf{G}}_i^{(k)}\|^2}{\|\bar{\mathbf{G}}_i^{(k-1)}\|^2}. \quad (6)$$

The purpose of deflecting the gradient using (6) is to find  $\mathbf{G}_i^{(k)}$ , which is the Hessian-conjugate of  $\mathbf{G}_i^{(k-1)}$ . By doing this, we can eliminate the “zigzagging” phenomenon encountered in the conventional gradient projection method, and achieve the superlinear convergence rate [12] without actually storing the matrix of Hessian approximation as in quasi-Newton methods.

### 3.2 Projection onto $\Omega_+(P)$

Noting from (5) that  $\mathbf{G}_i$  is Hermitian. We have that  $\mathbf{Q}_i'^{(k)} = \mathbf{Q}_i^{(k)} + s_k \mathbf{G}_i^{(k)}$  is Hermitian as well. Then, the projection problem becomes how to simultaneously project a set of  $K$  Hermitian matrices onto the set  $\Omega_+(P)$ , which contains a constraint on sum power for all users. This is different to [11], where the projection was performed on individual power constraint. In order to do this, we construct a block diagonal matrix  $\mathbf{D} = \text{Diag}\{\mathbf{Q}_1 \dots \mathbf{Q}_K\} \in \mathbb{C}^{(K \cdot n_r) \times (K \cdot n_r)}$ . It is easy to recognize that if  $\mathbf{Q}_i \in \Omega_+(P)$ ,  $i = 1, \dots, K$ , we have  $\text{Tr}(\mathbf{D}) = \sum_{i=1}^K \text{Tr}(\mathbf{Q}_i) \leq P$ , and  $\mathbf{D} \succeq 0$ . In this paper, we use Frobenius norm, denoted by  $\|\cdot\|_F$ , as the matrix distance criterion. The distance between two matrices  $\mathbf{A}$  and  $\mathbf{B}$  is defined as  $\|\mathbf{A} - \mathbf{B}\|_F = (\text{Tr}[(\mathbf{A} - \mathbf{B})^\dagger(\mathbf{A} - \mathbf{B})])^{\frac{1}{2}}$ . Thus, given a block diagonal matrix  $\mathbf{D}$ , we wish to find a matrix  $\tilde{\mathbf{D}} \in \Omega_+(P)$  such that  $\tilde{\mathbf{D}}$  minimizes  $\|\tilde{\mathbf{D}} - \mathbf{D}\|_F$ . For more convenient algebraic manipulations, we instead study the following equivalent optimization

problem:

$$\begin{aligned} & \text{Minimize} \quad \frac{1}{2} \|\tilde{\mathbf{D}} - \mathbf{D}\|_F^2 \\ & \text{subject to} \quad \text{Tr}(\tilde{\mathbf{D}}) \leq P, \tilde{\mathbf{D}} \succeq 0. \end{aligned} \quad (7)$$

In (7), the objective function is convex in  $\tilde{\mathbf{D}}$ , the constraint  $\tilde{\mathbf{D}} \succeq 0$  represents the convex cone of positive semidefinite matrices, and the constraint  $\text{Tr}(\tilde{\mathbf{D}}) \leq P$  is a linear constraint. Thus, the problem is a convex minimization problem and we can exactly solve this problem by solving its Lagrangian dual problem. Associating Hermitian matrix  $\mathbf{X}$  to the constraint  $\tilde{\mathbf{D}} \succeq 0$ ,  $\mu$  to the constraint  $\text{Tr}(\tilde{\mathbf{D}}) \leq P$ , we can write the Lagrangian as

$$\begin{aligned} g(\mathbf{X}, \mu) = \min_{\tilde{\mathbf{D}}} & \left\{ (1/2) \|\tilde{\mathbf{D}} - \mathbf{D}\|_F^2 - \text{Tr}(\mathbf{X}^\dagger \tilde{\mathbf{D}}) \right. \\ & \left. + \mu (\text{Tr}(\tilde{\mathbf{D}}) - P) \right\}. \end{aligned} \quad (8)$$

Since  $g(\mathbf{X}, \mu)$  is an unconstrained convex quadratic minimization problem, we can compute the minimizer of (8) by simply setting the derivative of (8) (with respect to  $\tilde{\mathbf{D}}$ ) to zero, i.e.,  $(\tilde{\mathbf{D}} - \mathbf{D}) - \mathbf{X}^\dagger + \mu \mathbf{I} = 0$ . Noting that  $\mathbf{X}^\dagger = \mathbf{X}$ , we have  $\tilde{\mathbf{D}} = \mathbf{D} - \mu \mathbf{I} + \mathbf{X}$ . Substituting  $\tilde{\mathbf{D}}$  back into (8), we have

$$\begin{aligned} g(\mathbf{X}, \mu) &= \frac{1}{2} \|\mathbf{X} - \mu \mathbf{I}\|_F^2 - \mu P + \text{Tr}[(\mu \mathbf{I} - \mathbf{X})(\mathbf{D} + \mathbf{X} - \mu \mathbf{I})] \\ &= -\frac{1}{2} \|\mathbf{D} - \mu \mathbf{I} + \mathbf{X}\|_F^2 - \mu P + \frac{1}{2} \|\mathbf{D}\|^2. \end{aligned} \quad (9)$$

Therefore, the Lagrangian dual problem can be written as

$$\begin{aligned} & \text{Maximize} \quad -\frac{1}{2} \|\mathbf{D} - \mu \mathbf{I} + \mathbf{X}\|_F^2 - \mu P + \frac{1}{2} \|\mathbf{D}\|^2 \\ & \text{subject to} \quad \mathbf{X} \succeq 0, \mu \geq 0. \end{aligned} \quad (10)$$

After solving (10), we can have the optimal solution to (7) as:

$$\tilde{\mathbf{D}}^* = \mathbf{D} - \mu^* \mathbf{I} + \mathbf{X}^*, \quad (11)$$

where  $\mu^*$  and  $\mathbf{X}^*$  are the optimal dual solutions to Lagrangian dual problem in (10). Although the Lagrangian dual problem in (10) has a similar structure as that in the primal problem in (7) (having a positive semidefinite matrix constraint), we find that the positive semidefinite matrix constraint can indeed be easily handled. To see this, we first introduce Moreau Decomposition Theorem from convex analysis.

**Theorem 1** (*Moreau Decomposition [16]*) *Let  $\mathcal{K}$  be a closed convex cone. For  $\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{C}^p$ , the two properties below are equivalent:*

1.  $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$  with  $\mathbf{x}_1 \in \mathcal{K}$ ,  $\mathbf{x}_2 \in \mathcal{K}^\circ$  and  $\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = 0$ ,

2.  $\mathbf{x}_1 = p_{\mathcal{K}}(\mathbf{x})$  and  $\mathbf{x}_2 = p_{\mathcal{K}^\circ}(\mathbf{x})$ ,

where  $\mathcal{K}^\circ \triangleq \{\mathbf{s} \in \mathbb{C}^p : \langle \mathbf{s}, \mathbf{y} \rangle \leq 0, \forall \mathbf{y} \in \mathcal{K}\}$  is called the polar cone of cone  $\mathcal{K}$ ,  $p_{\mathcal{K}}(\cdot)$  represents the projection onto cone  $\mathcal{K}$ .

In fact, the projection onto a cone  $\mathcal{K}$  is analogous to the projection onto a subspace. The only difference is that the orthogonal subspace is replaced by the polar cone.

Now we consider how to project a Hermitian matrix  $\mathbf{A} \in \mathbb{C}^{n \times n}$  onto the positive and negative semidefinite cones. First, we can perform eigenvalue decomposition on  $\mathbf{A}$  yielding  $\mathbf{A} = \mathbf{U} \text{Diag}\{\lambda_i, i = 1, \dots, n\} \mathbf{U}^\dagger$ , where  $\mathbf{U}$  is the unitary matrix formed by the eigenvectors corresponding to the eigenvalues  $\lambda_i, i = 1, \dots, n$ . Then, we have the positive semidefinite and negative semidefinite projections of  $\mathbf{A}$  as follows:

$$\mathbf{A}_+ = \mathbf{U} \text{Diag}\{\max\{\lambda_i, 0\}, i = 1, 2, \dots, n\} \mathbf{U}^\dagger, \quad (12)$$

$$\mathbf{A}_- = \mathbf{U} \text{Diag}\{\min\{\lambda_i, 0\}, i = 1, 2, \dots, n\} \mathbf{U}^\dagger. \quad (13)$$

The proof of (12) and (13) is a straightforward application of Theorem 1 by noting that  $\mathbf{A}_+ \succeq 0$ ,  $\mathbf{A}_- \preceq 0$ ,  $\langle \mathbf{A}_+, \mathbf{A}_- \rangle = 0$ ,  $\mathbf{A}_+ + \mathbf{A}_- = \mathbf{A}$ , and the positive semidefinite cone and negative semidefinite cone are polar cones to each other.



We now consider the term  $\mathbf{D} - \mu\mathbf{I} + \mathbf{X}$ , which is the only term involving  $\mathbf{X}$  in the dual objective function. We can rewrite it as  $\mathbf{D} - \mu\mathbf{I} - (-\mathbf{X})$ , where we note that  $-\mathbf{X} \preceq 0$ . Finding a negative semidefinite matrix  $-\mathbf{X}$  such that  $\|\mathbf{D} - \mu\mathbf{I} - (-\mathbf{X})\|_F$  is minimized is equivalent to finding the projection of  $\mathbf{D} - \mu\mathbf{I}$  onto the negative semidefinite cone. From the previous discussions, we immediately have

$$-\mathbf{X} = (\mathbf{D} - \mu\mathbf{I})_- . \quad (14)$$

Since  $\mathbf{D} - \mu\mathbf{I} = (\mathbf{D} - \mu\mathbf{I})_+ + (\mathbf{D} - \mu\mathbf{I})_-$ , substituting (14) back to the Lagrangian dual objective function, we have

$$\min_{\mathbf{X}} \|\mathbf{D} - \mu\mathbf{I} + \mathbf{X}\|_F = (\mathbf{D} - \mu\mathbf{I})_+ . \quad (15)$$

Thus, the matrix variable  $\mathbf{X}$  in the Lagrangian dual problem can be removed and the Lagrangian dual problem can be rewritten to

$$\begin{aligned} & \text{Maximize } \psi(\mu) \triangleq -\frac{1}{2} \|(\mathbf{D} - \mu\mathbf{I})_+\|_F^2 - \mu P + \frac{1}{2} \|\mathbf{D}\|^2 \\ & \text{subject to } \mu \geq 0. \end{aligned} \quad (16)$$

Suppose that after performing eigenvalue decomposition on  $\mathbf{D}$ , we have  $\mathbf{D} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\dagger$ , where  $\mathbf{\Lambda}$  is the diagonal matrix formed by the eigenvalues of  $\mathbf{D}$ ,  $\mathbf{U}$  is the unitary matrix formed by the corresponding eigenvectors. Since  $\mathbf{U}$  is unitary, we have  $(\mathbf{D} - \mu\mathbf{I})_+ = \mathbf{U}(\mathbf{\Lambda} - \mu\mathbf{I})_+\mathbf{U}^\dagger$ . It then follows that

$$\|(\mathbf{D} - \mu\mathbf{I})_+\|_F^2 = \|(\mathbf{\Lambda} - \mu\mathbf{I})_+\|_F^2. \quad (17)$$

We denote the eigenvalues in  $\mathbf{\Lambda}$  by  $\lambda_i$ ,  $i = 1, 2, \dots, K \cdot n_r$ . Suppose that we sort them in non-increasing order such that  $\mathbf{\Lambda} = \text{Diag}\{\lambda_1 \lambda_2 \dots \lambda_{K \cdot n_r}\}$ , where  $\lambda_1 \geq \dots \geq \lambda_{K \cdot n_r}$ . It then follows that

$$\|(\mathbf{\Lambda} - \mu\mathbf{I})_+\|_F^2 = \sum_{j=1}^{K \cdot n_r} (\max\{0, \lambda_j - \mu\})^2. \quad (18)$$

From (18), we can rewrite  $\psi(\mu)$  as

$$\psi(\mu) = -\frac{1}{2} \sum_{j=1}^{K \cdot n_r} (\max\{0, \lambda_j - \mu\})^2 - \mu P + \frac{1}{2} \|\mathbf{D}_n\|^2. \quad (19)$$

It is evident from (19) that  $\psi(\mu)$  is continuous and (piece-wise) concave in  $\mu$ . Generally, piece-wise concave maximization problems can be solved by using the subgradient method. However, due to the heuristic nature of its step size selection strategy, subgradient algorithm usually does not perform well. In fact, by exploiting the special structure, (16) can be efficiently solved. We can search the optimal value of  $\mu$  as follows. Let  $\hat{I}$  index the pieces of  $\psi(\mu)$ ,  $\hat{I} = 0, 1, \dots, K \cdot n_r$ . Initially we set  $\hat{I} = 0$  and increase  $\hat{I}$  subsequently. Also, we introduce  $\lambda_0 = \infty$  and  $\lambda_{K \cdot n_r + 1} = -\infty$ . We let the endpoint objective value  $\psi_{\hat{I}}(\lambda_0) = 0$ ,  $\phi^* = \psi_{\hat{I}}(\lambda_0)$ , and  $\mu^* = \lambda_0$ . If  $\hat{I} > K \cdot n_r$ , the search stops. For a particular index  $\hat{I}$ , by setting

$$\frac{\partial}{\partial \mu} \psi_{\hat{I}}(\nu) \triangleq \frac{\partial}{\partial \mu} \left( -\frac{1}{2} \sum_{i=1}^{\hat{I}} (\lambda_i - \mu)^2 - \mu P \right) = 0, \quad (20)$$

we have

$$\mu_{\hat{I}}^* = \frac{\sum_{i=1}^{\hat{I}} \lambda_i - P}{\hat{I}}. \quad (21)$$

Now we consider the following two cases:

1. If  $\mu_{\hat{I}}^* \in [\lambda_{\hat{I}+1}, \lambda_{\hat{I}}] \cap \mathbb{R}_+$ , where  $\mathbb{R}_+$  denotes the set of non-negative real numbers, then we have found the optimal solution for  $\mu$  because  $\psi(\mu)$  is concave in  $\mu$ . Thus, the point having zero-value first derivative, if exists, must be the unique global maximum solution. Hence, we can let  $\mu^* = \mu_{\hat{I}}^*$  and the search is done.
2. If  $\mu_{\hat{I}}^* \notin [\lambda_{\hat{I}+1}, \lambda_{\hat{I}}] \cap \mathbb{R}_+$ , we must have that the local maximum in the interval  $[\lambda_{\hat{I}+1}, \lambda_{\hat{I}}] \cap \mathbb{R}_+$  is achieved at one of the two endpoints. Note that the objective value  $\psi_{\hat{I}}(\lambda_{\hat{I}})$  has been

computed in the previous iteration because from the continuity of the objective function, we have  $\psi_{\hat{I}}(\lambda_{\hat{I}}) = \psi_{\hat{I}-1}(\lambda_{\hat{I}})$ . Thus, we only need to compute the other endpoint objective value  $\psi_{\hat{I}}(\lambda_{\hat{I}+1})$ . If  $\psi_{\hat{I}}(\lambda_{\hat{I}+1}) < \psi_{\hat{I}}(\lambda_{\hat{I}}) = \phi^*$ , then we know  $\mu^*$  is the optimal solution; else let  $\mu^* = \lambda_{\hat{I}+1}$ ,  $\phi^* = \psi_{\hat{I}}(\lambda_{\hat{I}+1})$ ,  $\hat{I} = \hat{I} + 1$  and continue.

Since there are  $K \cdot n_r + 1$  intervals in total, the search process takes at most  $K \cdot n_r + 1$  steps to find the optimal solution  $\mu^*$ . Hence, this search is of polynomial-time complexity  $O(n_r K)$ .

After finding  $\mu^*$ , we can compute  $\tilde{\mathbf{D}}^*$  as

$$\tilde{\mathbf{D}}^* = (\mathbf{D} - \mu^* \mathbf{I})_+ = \mathbf{U} (\mathbf{\Lambda} - \mu^* \mathbf{I})_+ \mathbf{U}^\dagger. \quad (22)$$

That is, the projection  $\tilde{\mathbf{D}}$  can be computed by adjusting the eigenvalues of  $\mathbf{D}$  using  $\mu^*$  and keeping the eigenvectors unchanged. The projection of  $\mathbf{D}_n$  onto  $\Omega_+(P)$  is summarized in Algorithm 2.

---

**Algorithm 2** Projection onto  $\Omega_+(P)$

---

**Initiation:**

1. Construct a block diagonal matrix  $\mathbf{D}$ . Perform eigenvalue decomposition  $\mathbf{D} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^\dagger$ , sort the eigenvalues in non-increasing order.
2. Introduce  $\lambda_0 = \infty$  and  $\lambda_{K \cdot n_r + 1} = -\infty$ . Let  $\hat{I} = 0$ . Let the endpoint objective value  $\psi_{\hat{I}}(\lambda_0) = 0$ ,  $\phi^* = \psi_{\hat{I}}(\lambda_0)$ , and  $\mu^* = \lambda_0$ .

**Main Loop:**

1. If  $\hat{I} > K \cdot n_r$ , go to the final step; else let  $\mu_{\hat{I}}^* = (\sum_{j=1}^{\hat{I}} \lambda_j - P) / \hat{I}$ .
2. If  $\mu_{\hat{I}}^* \in [\lambda_{\hat{I}+1}, \lambda_{\hat{I}}] \cap \mathbb{R}_+$ , then let  $\mu^* = \mu_{\hat{I}}^*$  and go to the final step.
3. Compute  $\psi_{\hat{I}}(\lambda_{\hat{I}+1})$ . If  $\psi_{\hat{I}}(\lambda_{\hat{I}+1}) < \phi^*$ , then go to the final step; else let  $\mu^* = \lambda_{\hat{I}+1}$ ,  $\phi^* = \psi_{\hat{I}}(\lambda_{\hat{I}+1})$ ,  $\hat{I} = \hat{I} + 1$  and continue.

**Final Step:** Compute  $\tilde{\mathbf{D}}$  as  $\tilde{\mathbf{D}} = \mathbf{U} (\mathbf{\Lambda} - \mu^* \mathbf{I})_+ \mathbf{U}^\dagger$ .

---

## 4 Complexity Analysis

In this section, we compare our proposed CGP with other existing methods for solving MIMO BC. Similar to IWFs, SD [8], and DD [9], CGP has the desirable linear complexity property. Although CGP also needs to compute gradients in each iteration, the computation is much easier than that in SD due to the different perspectives in handling MIMO BC. Thus, in this paper, we only compare CGP with IWF (Algorithms 1 and 2 in [10]), which appear to be the simplest in the literature so

far. For convenience, we will refer to Algorithm 1 and Algorithm 2 in [10] as IWF1 and IWF2, respectively.

To better illustrate the comparison, we list the complexity per iteration for each component of CGP and IWFs in Table 1. In both CGP and IWFs, it can be seen that the most time-consuming

Table 1: Per Iteration Complexity Comparison between CGP and IWFs

	CGP	IWFs
Gradient/Effec. Channel	$K$	$2K$
Line Search	$O(mK)$	N/A
Projection/Water-Filling	$O(n_r K)$	$O(n_r K)$
Overall	$O((m + 1 + n_r)K)$	$O((2 + n_r)K)$

part (increasing with respect to  $K$ ) is the additions of the terms in the form of  $\mathbf{H}_i^\dagger \mathbf{Q}_i \mathbf{H}_i$  when computing gradients and effective channels. Since the term  $(\mathbf{I} + \sum_{i=1}^K \mathbf{H}_i^\dagger \mathbf{Q}_i \mathbf{H}_i)$  is common to all gradients, we only need to compute this sum once in each iteration. Thus, the number of such additions per iteration for CGP is  $K$ . In IWF1 and IWF2, the number of such additions can be reduced to  $2K$  by a clever way of maintaining a running sum of  $(\mathbf{I} + \sum_{j \neq i}^K \mathbf{H}_j^\dagger \mathbf{Q}_j \mathbf{H}_j)$ . But the running sum, which requires  $K^2$  additions for IWF1, still needs to be computed in the initialization step.

Although the basic ideas of the projection in CGP and water-filling are different, the algorithm structure of them are very similar and they have exactly the same complexity of  $O(n_r K)$ . The only unique component in CGP is the line search step, which has the complexity of  $O(mK)$  (in terms of the additions of  $\mathbf{H}_i^\dagger \mathbf{Q}_i \mathbf{H}_i$  terms), where  $m$  is the number of trials in Armijo's Rule. Therefore, the overall complexity per iteration for CGP and IWFs are  $O((m + 1 + n_r)K)$  and  $O((2 + n_r)K)$ , respectively. According to our computational experience, the value of  $m$  usually lies in between two and four. Thus, when  $n_r$  is large (e.g.,  $n_r \geq 4$ ), the overall complexity per iteration for CGP and IWFs are comparable. However, as evidenced in the next section, the numbers of iterations required for convergence in CGP is much less than that in IWFs for large MIMO BC systems, and it is very insensitive to the increase of the number of users. Moreover, CGP has a modest memory requirement: it only requires the solution information from the previous step, as opposed to IWF1,

which requires previous  $K - 1$  steps.

## 5 Numerical Results

Due to the space limitation, we only give an example of a large MIMO BC system consisting of 100 users with  $n_t = n_r = 4$  in here. The convergence processes are plotted in Fig. 1. It is observed from Fig. 1 that CGP takes only 29 iterations to converge and it outperforms both IWFs. IWF1's convergence speed significantly drops after the quick improvement in the early stage. It is also seen in this example that IWF2's performance is inferior to IWF1, and this observation is in accordance with the results in [10]. Both IWF1 and IWF2 fail to converge within 100 iterations. The scalability problem of both IWFs is not surprising because in both IWFs, the most recently updated covariance matrices only account for a fraction of  $1/K$  in the effective channels' computation, which means it does not effectively make use of the most recent solution. In all of our numerical examples with different number of users, CGP always converges within 30 iterations.

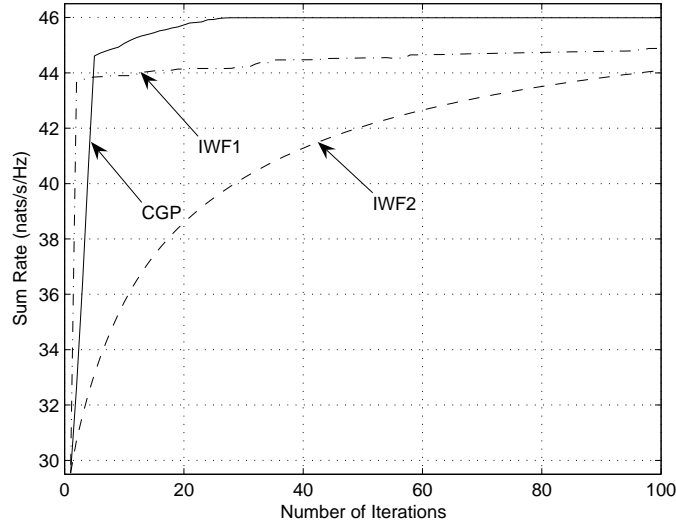


Figure 1: Comparison in a 100-user MIMO BC channel with  $n_t = n_r = 4$ .

## 6 Conclusion

In this paper, we developed an efficient algorithm based on conjugate gradient projection (CGP) for solving the maximum sum rate problem of MIMO BC. We theoretically and numerically analyzed its complexity and convergence behavior. The attractive features of CGP and encouraging results showed that CGP is an excellent method for solving the maximum sum rate problem for large MIMO BC systems.

## References

- [1] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 1, pp. 684–702, Jun. 2003.
- [2] G. Caire and S. S. (Shitz), "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.
- [3] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.
- [4] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.
- [5] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.
- [6] W. Yu, "Uplink-downlink duality via minimax duality," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 361–374, Feb. 2006.
- [7] T. Lan and W. Yu, "Input optimization for multi-antenna broadcast channels and per-antenna power constraints," in *Proc. IEEE GLOBECOM*, Dallas, TX, U.S.A., Nov. 2004, pp. 420–424.

- [8] H. Viswanathan, S. Venkatesan, and H. Huang, "Downlink capacity evaluation of cellular networks with known-interference cancellation," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 802–811, Jun. 2003.
- [9] W. Yu, "A dual decomposition approach to the sum power Gaussian vector multiple-access channel sum capacity problem," in *Proc. Conf. Information Sciences and Systems (CISS)*, Baltimore, MD, U.S.A., 2003.
- [10] N. Jindal, W. Rhee, S. Vishwanath, S. A. Jafar, and A. Goldsmith, "Sum power iterative water-filling for multi-antenna Gaussian broadcast channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1570–1580, Apr. 2005.
- [11] S. Ye and R. S. Blum, "Optimized signaling for MIMO interference systems with feedback," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2839–2848, Nov. 2003.
- [12] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, 3rd ed. New York, NY: John Wiley & Sons Inc., 2006.
- [13] J. B. Rosen, "The gradient projection method for nonlinear programming, Part I, Linear constraints," *SIAM J. Applied Mathematics*, vol. 8, pp. 181–217, 1960.
- [14] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Economics*. New York: Wiley, 1999.
- [15] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [16] J.-B. Hiriart-Urruty and C. Lemaréchal, *Fundamentals of Convex Analysis*. Berlin: Springer-Verlag, 2001.