Linear Precoding for MIMO Multiple Access Channels with Finite Discrete Inputs

Mingxi Wang, *Student Member, IEEE*, Weiliang Zeng, *Student Member, IEEE*, and Chengshan Xiao, *Fellow, IEEE*

***Abstract*—In this paper, we study linear precoding for multiple-input multiple-output (MIMO) multiple access channels (MAC) with finite discrete inputs. We derive the constellation-constrained capacity region for the MIMO MAC with an arbitrary number of users and find that the boundary can be achieved by solving the problem of weighted sum rate maximization with constellation and individual power constraints. Due to the non-concavity of the objective function, we obtain a set of necessary conditions for the optimization problem through Karush-Kuhn-Tucker analysis. To find the optimal precoding matrices for all users, we propose an iterative algorithm utilizing alternating optimization strategy. In particular, each iteration of the algorithm involves the gradient descent update with back-tracking line search. Numerical results show that when inputs are digital modulated signals and the signal-to-noise ratio is in the medium range, our proposed algorithm offers considerably higher sum rate than non-precoding and the traditional method which maximizes Gaussian-input sum capacity. Furthermore, a low-density parity-check coded system with iterative detection and decoding for MAC is presented to evaluate the bit error rate (BER) performance of precoders. BER results also indicate that the system with the proposed linear precoder achieves significant gains over the non-precoding system and the precoder designed for Gaussian inputs.**

***Index Terms*—Multiple-input multiple-output (MIMO), multiple access channels (MAC), linear precoding, finite discrete inputs.**

I. Introduction

THE problem of linear precoding for multiple-input multiple-output (MIMO) multiple access channels (MAC) has been investigated in the literature in the last few years. The existing methods were proposed based on different criteria. Most of them employed information theoretical analysis which finds the capacity region of MIMO MAC. It is well known that the input signals achieving the boundary of MIMO MAC capacity region are Gaussian distributed, and the capacity region only depends on input covariance matrices [1]–[3]. The optimal input covariance matrices can be found by maximizing the weighted sum rate, which is a convex optimization problem with Gaussian inputs [4]–[6]. In particular, when only sum rate maximization is considered, an effective algorithm called iterative water-filling (WF) algorithm is developed in [3]. Other criteria in linear precoding design of MAC are also utilized. For instance, [7] minimizes the mean square error (MSE) assuming a linear receiver structure, and [8] maximizes the signal-to-interference and noise ratio (SINR) for an iterative linear receiver.

However, there are some drawbacks of the aforementioned methods. Regarding the optimization techniques using capacity with Gaussian signals, the transmitted signals in practical digital communication systems are not Gaussian distributed, but rather drawn from certain constellation sets such as phase-shift keying (PSK), pulse-amplitude modulation (PAM), or quadrature amplitude modulation (QAM). Therefore, precoders obtained based on Gaussian input assumptions may lead to performance loss when the inputs are actually replaced by finite discrete inputs. On the other hand, MSE minimization and SINR maximization approaches may not necessarily provide minimum bit error rate (BER) or maximum date rate.

To overcome the shortcomings of optimization through Gaussian capacity, MSE, SINR, etc., the mutual information with finite discrete inputs has been employed for precoding design recently. This approach conforms to such an information theoretical principle that the mutual information with certain input constraints determines the potential achievable data rate of a communication system. In point-to-point communication scenarios, the mercury/waterfilling power allocation and general linear precoders are developed [9]–[15]. Similar problems are also investigated in relay networks [16], [17] and broadcast channels [18], [19]. For 2-user single-input single-output (SISO) MAC, [20] found the optimal angle of rotation and designed the code pairs based on trellis coded modulation. These papers have shown that adopting mutual information with finite discrete inputs for precoder design can achieve considerable performance gains compared with existing methods which are based on Gaussian input assumptions.

To the best of our knowledge, little research has been done on MIMO MAC precoding based on mutual information with finite discrete inputs. In this paper, the maximum mutual information with uniformly distributed finite discrete inputs is referred to as constellation-constrained capacity [20], [21], while the maximized mutual information with Gaussian inputs is called Gaussian capacity. We derive the constellation-constrained capacity region for the MIMO MAC with an arbitrary number of users and find that the boundary can be achieved by solving the problem of weighted sum rate maximization with constellation and individual power constraints. Since the weighted sum rate is no longer a concave function of precoding matrices as opposed to the case of Gaussian inputs

Manuscript received March 22, 2011; revised June 28, 2011; accepted August 15, 2011. The associated editor coordinating the review of this paper and approving it for publication was W. Zhang.

This work was supported in part by the National Science Foundation under Grant CCF-0915846.

M. Wang and C. Xiao are with the Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla, MO 65409, USA (e-mail: mw9zd@mail.mst.edu, xiaoc@mst.edu).

W. Zeng is with the Tsinghua National Laboratory for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: zengwl07@mails.tsinghua.edu.cn).

Digital Object Identifier 10.1109/TWC.2011.091911.110508

[3], we obtain a set of necessary conditions through Karush-Kuhn-Tucker (KKT) analysis [22]. To find optimal linear precoders for all users, we propose an iterative algorithm utilizing alternating optimization strategy with gradient descent update method. Furthermore, the backtracking line search is adopted to determine the step size for fast convergence. Our method is guaranteed to local optimum, and we resort to multiple run of the algorithm with random initializations to search for a best possible final solution. Numerical results show that the proposed algorithm converges fast under various signal-to-noise ratios (SNRs). In addition, when inputs are digital modulated signals, and the SNR is in the medium regime, our proposed algorithm offers much higher sum rate than non-precoding and the traditional power allocation method which maximizes Gaussian-input sum capacity.

Besides the sum rate and constellation-constrained capacity region, bit error rate of a system over MAC is of significant interest in practice. We thus present a multiuser system with low-density parity-check (LDPC) coding and linear precoding for all transmitters. At the receiver, the soft maximum a posteriori (MAP) multiuser detector and LDPC channel decoders are adopted to iteratively exchange the soft information. Simulations show that the system with the proposed precoder achieves significant SNR gains over the non-precoding system and the system with the precoder designed under Gaussian assumptions.

Throughout the paper, we denote matrices with boldface upper-case letters, and vectors with boldface lower-case letters. The superscripts $(\cdot)^t$ and $(\cdot)^h$ represent transpose and conjugate transpose operations, respectively. In addition, $\|\mathbf{a}\|$ and $\|\mathbf{A}\|_F$ means the Euclidean norm of vector $\mathbf{a}$ and the Frobenius norm of matrix $\mathbf{A}$, respectively. The determinant of matrix $\mathbf{A}$ is $|\mathbf{A}|$, and $\log$ stands for the logarithm with base 2. The symbol $\mathbb{C}$ denotes the complex number field, while $\mathbb{E}$ takes the expectation of a random variable or function.

The rest of the paper is organized as follows. Section II describes the model of MIMO MAC and a brief overview of the existing results on capacity region and optimal linear precoding with Gaussian input signals. The constellation-constrained capacity region of MIMO MAC with finite discrete inputs is derived in Section III. Section IV discusses necessary condition of the weighed sum rate maximization problem and the details of the iterative algorithm. Section V presents the MIMO system over MAC with iterative detection and decoding. Numerical results are provided in Section VI, and Section VII draws the conclusions.

II. System Model and Existing Results of MIMO MAC with Gaussian Inputs

Consider a K-user communication system with multiple antennas at transmitters and the receiver over multiple access channels. The signal model is given by

$$\begin{aligned}\mathbf{y} &= \mathbf{H}_1\mathbf{G}_1\mathbf{x}_1 + \mathbf{H}_2\mathbf{G}_2\mathbf{x}_2 + \cdots + \mathbf{H}_K\mathbf{G}_K\mathbf{x}_K + \mathbf{v}\\ &= \mathbf{H}\mathbf{G}\mathbf{x} + \mathbf{v}\end{aligned} \tag{1}$$

where $\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \cdots, \mathbf{H}_K]$, and $\mathbf{x} = [\mathbf{x}_1^t, \cdots, \mathbf{x}_K^t]^t$. $\mathbf{G} = \text{Bdiag}\{\mathbf{G}_1, \cdots, \mathbf{G}_K\}$, where Bdiag means a block diagonal matrix. Therefore, $\mathbf{H}$ and $\mathbf{G}$ can be viewed as the equivalent channel matrix and block diagonal precoding matrix for all users. Suppose there are N_r antennas at the receiver, and each user has N_t transmit antennas. The symbol $\mathbf{H}_i \in \mathbb{C}^{N_r \times N_t}$ represents the complex channel matrix between the i-th transmitter and the receiver. We assume that the receiver knows the channels of all users, and each transmitter knows its own channel state information. Throughout this paper, we constrain each user's precoding matrix to be a square matrix, which is denoted as $\mathbf{G}_i \in \mathbb{C}^{N_t \times N_t}$. The vector $\mathbf{x} \in \mathbb{C}^{N_t K \times 1}$ contains signals of all transmitters, and $\mathbf{y} \in \mathbb{C}^{N_r \times 1}$ is the received signal. The receiver noise $\mathbf{v} \in \mathbb{C}^{N_r \times 1}$ is a zero mean circularly symmetric complex Gaussian vector with covariance matrix $\sigma^2\mathbf{I}$, *i.e.*, $\mathbf{v} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$.

Assume all signal vectors $\mathbf{x}_i$ of different users are independent from one another, and elements of each $\mathbf{x}_i$ are independent and identically distributed (i.i.d) with unit energy, *i.e.*, $\mathbb{E}[\mathbf{x}_i\mathbf{x}_i^h] = \mathbf{I}_{N_t}$. The symbols in $\mathbf{x}_i$ can be digital modulated signal points such as PSK or QAM signals. The covariance matrix of transmitted signal of user i is $\mathbf{Q}_i = \mathbb{E}[\mathbf{G}_i\mathbf{x}_i\mathbf{x}_i^h\mathbf{G}_i^h] = \mathbf{G}_i\mathbf{G}_i^h$.

We now briefly review existing results of MIMO MAC based on Gaussian inputs. For the K-user MAC, it is well known that the capacity region is the convex hull of the union of capacity pentagons, and the boundary of the capacity region can be fully characterized by maximizing the weighted sum rate $\sum_{i=1}^K \mu_i R_i$ for all nonnegative μ_i. Assuming that $\mu_1 \geq \cdots \geq \mu_K \geq 0$, and $\sum_{i=1}^K \mu_i = K$, then the optimal covariance matrices which maximize the capacity region can be found through solving the following optimization problem [2], [3]:

$$\begin{aligned}\max_{\mathbf{Q}_1,\cdots,\mathbf{Q}_K} \ & \mu_K \log\left|\mathbf{I} + \sum_{i=1}^K \mathbf{H}_i\mathbf{Q}_i\mathbf{H}_i^h\right| + \\ & \sum_{i=1}^{K-1}(\mu_i - \mu_{i+1})\log\left|\mathbf{I} + \sum_{l=i}^{l}\mathbf{H}_l\mathbf{Q}_l\mathbf{H}_l^h\right| \end{aligned} \tag{2}$$

$$\text{subject to} \qquad \text{Tr}(\mathbf{Q}_i) \leq P_i, \mathbf{Q}_i \succeq \mathbf{0}, i = 1, \cdots, K, \tag{3}$$

in which $\mathbf{Q}_i$ is hermitian and positive semidefinite. The above problem is a convex optimization problem, which can be solved by efficient numerical methods [22], [23].

III. Constellation-Constrained Capacity Region of MAC with Finite Discrete Inputs

We derive the constellation-constrained capacity region with finite discrete inputs for MIMO MAC in this section. Let the set A and its complement A^c partition all users into two groups, where $A = \{i_1, i_2, \cdots, i_{K_1}\} \subseteq \{1, 2, \cdots, K\}$, and $A^c = \{j_1, j_2, \cdots, j_{K_2}\}$, $K_1 + K_2 = K$. With the assumptions of $\mathbf{x}_A = [\mathbf{x}_{i_1}^t, \mathbf{x}_{i_2}^t, \cdots, \mathbf{x}_{i_{K_1}}^t]^t$, and $\mathbf{x}_{A^c} = [\mathbf{x}_{j_1}^t, \mathbf{x}_{j_2}^t, \cdots, \mathbf{x}_{j_{K_2}}^t]^t$, it is known that the achievable rate region of K-user MAC is the closure of the convex hull of the rate vectors $(R_1, R_2 \cdots, R_K)$, which satisfies the following constraints [1]:

$$\sum_{i \in A} R_i \leq I\left(\mathbf{x}_A; \mathbf{y}|\mathbf{x}_{A^c}\right), \quad \forall A \subseteq \{1, 2, \cdots, K\} \tag{4}$$

for some independent input distributions $p(\mathbf{x}_1)$, $p(\mathbf{x}_2)$, $\cdots$, $p(\mathbf{x}_K)$.

In practical digital communication systems over multiple access channels, transmitted signals are often equiprobably drawn from certain discrete constellations such as PSK, PAM, or QAM. Assuming that M_i is the number of constellation points in each component of $\mathbf{x}_i$, then the number of all possible vectors of $\mathbf{x}_i$ is $N_i = M_i^{N_t}$. Assuming that $\mathbf{H}_A = [\mathbf{H}_{i_1}, \mathbf{H}_{i_2}, \cdots, \mathbf{H}_{i_{K_1}}]$ and $\mathbf{G}_A = \text{Bdiag}\{\mathbf{G}_{i_1}, \mathbf{G}_{i_2}, \cdots, \mathbf{G}_{i_{K_1}}\}$, we have the following proposition, which generalizes the achievable rates of 2-user MAC in [20].

Proposition 1: When the discrete signals $\mathbf{x}_i$ of all users are independent and uniformly distributed, $I(\mathbf{x}_A; \mathbf{y}|\mathbf{x}_{A^c})$ is given as follows:

$$I\left(\mathbf{x}_A; \mathbf{y}|\mathbf{x}_{A^c}\right) = \log N_A - \frac{1}{N_A}\sum_{i=1}^{N_A}\mathbb{E}_{\mathbf{v}}$$
$$\left[\log\sum_{k=1}^{N_A}\exp\left(\frac{-\|\mathbf{H}_A\mathbf{G}_A\left(\mathbf{x}_A^i - \mathbf{x}_A^k\right)+\mathbf{v}\|^2 + \|\mathbf{v}\|^2}{\sigma^2}\right)\right] \quad (5)$$

where $\mathbf{v} \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2\mathbf{I}\right)$, and $N_A = \prod_{i\in A} N_i$. The symbol $\mathbf{x}_A^i$ represents one possible signal vector from $\mathbf{x}_A$, and the number of all possible constellation points in vector $\mathbf{x}_A$ is N_A.

Proof: From the definition $I\left(\mathbf{x}_A; \mathbf{y}|\mathbf{x}_{A^c}\right) = H\left(\mathbf{y}|\mathbf{x}_{A^c}\right) - H\left(\mathbf{y}|\mathbf{x}_A, \mathbf{x}_{A^c}\right)$, we can prove (5). Details can be found in Appendix A. ■

According to [24], the boundary of the constellation-constrained capacity region can be characterized by the solution of the weighted sum rate optimization problem. Without loss of generality, we assume the weights $\mu_1 \geq \cdots \geq \mu_K \geq \mu_{K+1} = 0$, *i.e.*, decoding user K first and user 1 last. Then, the weighted sum rate maximization with finite discrete inputs is equivalent to the following optimization problem:

$$\max_{\mathbf{G}_1,\cdots,\mathbf{G}_K} \quad g(\mathbf{G}_1, \cdots, \mathbf{G}_K) = \sum_{i=1}^{K}\Delta_i f(\mathbf{G}_1, \cdots, \mathbf{G}_i). \quad (6)$$

$$\text{subject to} \quad \text{Tr}(\mathbf{G}_i\mathbf{G}_i^h) \leq P_i, \quad i = 1, 2, \cdots, K \quad (7)$$

where $\Delta_i = \mu_i - \mu_{i+1}$; $i = 1, \cdots, K$; and $f(\mathbf{G}_1, \cdots, \mathbf{G}_i)$ is equal to $I\left(\mathbf{x}_1, \cdots, \mathbf{x}_i; \mathbf{y}|\mathbf{x}_{i+1}, \cdots, \mathbf{x}_K\right)$, which can be obtained by (5). When $\mu_1 = \cdots = \mu_K = 1$, the weighted sum rate maximization problem reduces to sum rate maximization.

IV. Weighted Sum Rate Maximization

In this section, we solve the problem of weighted sum rate maximization with finite discrete inputs described in (6) and (7). We obtain a set of necessary conditions for the optimization problem and then propose an iterative algorithm to find optimal precoding matrices.

A. Necessary Conditions

In general, the objective function $g(\mathbf{G}_1, \cdots, \mathbf{G}_K)$ is nonconcave on precoding matrices $\{\mathbf{G}_1, \cdots, \mathbf{G}_K\}$. Thus, the weighted sum rate maximization with finite discrete inputs is not concave, and we can only find a set of necessary conditions for this optimization problem, as given in the following proposition.

Proposition 2: The solution for the weighted sum rate maximization described in (6) and (7) satisfies:

$$\nu_i\mathbf{G}_i = \sum_{j=i}^{K}\Delta_j\mathbf{H}_i^h\mathbf{H}_{A_j}\mathbf{G}_{A_j}\mathbf{E}_{A_j}^i \quad (8)$$

$$\nu_i\left[\text{Tr}\left(\mathbf{G}_i\mathbf{G}_i^h\right) - P_i\right] = 0 \quad (9)$$

$$\text{Tr}(\mathbf{G}_i\mathbf{G}_i^h) - P_i \leq 0 \quad (10)$$

$$\nu_i \geq 0 \quad (11)$$

for all $i = 1, 2, \cdots, K$. Since the set $A_j = \{1, 2, \cdots, j\}$, we have $\mathbf{H}_{A_j} = [\mathbf{H}_1, \mathbf{H}_2, \cdots, \mathbf{H}_j]$ and $\mathbf{G}_{A_j} = \text{Bdiag}\{\mathbf{G}_1, \mathbf{G}_2, \cdots, \mathbf{G}_j\}$. The symbol $\mathbf{E}_{A_j}^i \in \mathbb{C}^{N_tj\times N_t}$ stands for the i-th column block of the minimum mean square error (MMSE) matrix [25] of $\mathbf{E}_{A_j}$, which is defined as

$$\mathbf{E}_{A_j} = \mathbb{E}\left[\left(\mathbf{x}_{A_j} - \mathbb{E}\left[\mathbf{x}_{A_j}|\mathbf{y}, \mathbf{x}_{A_j^c}\right]\right)\left(\mathbf{x}_{A_j} - \mathbb{E}\left[\mathbf{x}_{A_j}|\mathbf{y}, \mathbf{x}_{A_j^c}\right]\right)^h\right]. \quad (12)$$

Proof: The Lagrangian for (6) and (7) is given by

$$\mathcal{L}(\mathbf{G}, \lambda) = -g(\mathbf{G}_1, \cdots, \mathbf{G}_K) + \sum_{i=1}^{K}\lambda_i\left[\text{Tr}\left(\mathbf{G}_i\mathbf{G}_i^h\right) - P_i\right] \quad (13)$$

in which $\lambda_i \geq 0, i = 1, \cdots, K$. Define the gradient $\nabla_{\mathbf{G}_i}\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\mathbf{G}^*}$ as in [26], then the KKT conditions are as follows:

$$\nabla_{\mathbf{G}_i}\mathcal{L} = -\nabla_{\mathbf{G}_i}g(\mathbf{G}_1, \cdots, \mathbf{G}_K) + \lambda_i\mathbf{G}_i = 0 \quad (14)$$

$$\lambda_i\left[\text{Tr}\left(\mathbf{G}_i\mathbf{G}_i^h\right) - P_i\right] = 0 \quad (15)$$

$$\text{Tr}(\mathbf{G}_i\mathbf{G}_i^h) - P_i \leq 0 \quad (16)$$

$$\lambda_i \geq 0 \quad (17)$$

for all $i = 1, 2, \cdots, K$.

Due to the relation between the mutual information and MMSE [27], [28], the gradient of $f(\mathbf{G}_1, \cdots, \mathbf{G}_j)$ can be found as:

$$\nabla_{\mathbf{G}_{A_j}}f(\mathbf{G}_1, \cdots, \mathbf{G}_j) = \frac{\log e}{\sigma^2}\mathbf{H}_{A_j}^h\mathbf{H}_{A_j}\mathbf{G}_{A_j}\mathbf{E}_{A_j}. \quad (18)$$

For $j \geq i$, $\mathbf{G}_i$ is the i-th block of diagonal submatrices of $\mathbf{G}_{A_j}$. We can write $\mathbf{G}_i$ as $(\mathbf{e}_i \otimes \mathbf{I}_{N_t})\mathbf{G}_{A_j}(\mathbf{e}_i^h \otimes \mathbf{I}_{N_t})$, in which $\mathbf{e}_i$ is the i-th row of the $j-$dimensional identity matrix $\mathbf{I}$. Then, we have

$$\begin{aligned}\nabla_{\mathbf{G}_i}&f(\mathbf{G}_1, \cdots, \mathbf{G}_j)\\ &= (\mathbf{e}_i \otimes \mathbf{I}_{N_t})\nabla_{\mathbf{G}_{A_j}}f(\mathbf{G}_1, \cdots, \mathbf{G}_j)(\mathbf{e}_i^h \otimes \mathbf{I}_{N_t})\\ &= \frac{\log e}{\sigma^2}\mathbf{H}_i^h\mathbf{H}_{A_j}\mathbf{G}_{A_j}\mathbf{E}_{A_j}^i \end{aligned} \quad (19)$$

where $\mathbf{E}_{A_j}^i = \mathbf{E}_{A_j}(\mathbf{e}_i^h \otimes \mathbf{I}_{N_t}) \in \mathbb{C}^{N_tj\times N_t}$ is the i-th column block of the MMSE matrix $\mathbf{E}_{A_j}$. Substituting (19) to (14) and letting $\nu_i = \lambda_i\sigma^2/\log e$, we can prove (8). ■

B. Iterative Algorithm For Weighted Sum Rate Maximization

From (8), it can be seen that the optimal precoders of different users depend on one another. A common approach to multidimensional optimization problem is the alternating optimization method which iteratively optimizes one variable at a time with others fixed [7], [29], [30]. We adopt this method to maximize the weighed sum rate with finite discrete inputs. During each iteration of the algorithm, only one user's precoding matrix $\mathbf{G}_i$ is updated while others are fixed. For i-th user at n-th iteration, we first generate $\widetilde{\mathbf{G}}_i^{(n)}$ based on the gradient of $g(\mathbf{G}_1, \cdots, \mathbf{G}_K)$ with respect to $\mathbf{G}_i$ as follows

$$\widetilde{\mathbf{G}}_i^{(n)} = \mathbf{G}_i^{(n)} + t\nabla_{\mathbf{G}_i} g(\mathbf{G}_1^{(n)}, \cdots, \mathbf{G}_K^{(n)}) \tag{20}$$

where t is the step size. If $\|\widetilde{\mathbf{G}}_i^{(n)}\|_F^2 > P_i$, we project $\widetilde{\mathbf{G}}_i^{(n)}$ to the feasible set $\mathrm{Tr}(\mathbf{G}\mathbf{G}^h) \leq P_i$ to obtain the update [28]:

$$\mathbf{G}_i^{(n+1)} = \left[\widetilde{\mathbf{G}}_i^{(n)}\right]^+_{\mathrm{Tr}(\mathbf{G}\mathbf{G}^h)\leq P_i} = \sqrt{P_i}\widetilde{\mathbf{G}}_i^{(n)}/\|\widetilde{\mathbf{G}}_i^{(n)}\|_F. \tag{21}$$

For fast convergence, we use backtracking line search [22] to determine the step size t in gradient update. The two parameters in backtracking line search are α, β with $\alpha \in (0, 0.5)$ and $\beta \in (0, 1)$. Detailed steps of the proposed algorithm are shown in Table I.

TABLE I
THE ALGORITHM FOR WEIGHTED SUM RATE MAXIMIZATION WITH FINITE DISCRETE INPUTS

initialize $\mathbf{G}_i^{(0)}$ with $\mathrm{Tr}\left(\mathbf{G}_i\mathbf{G}_i^h\right) = P_i$, $i = 1, 2, \cdots, K$.
repeat
  compute $g^{(n)} = g(\mathbf{G}_1^{(n)}, \cdots, \mathbf{G}_K^{(n)})$, and $\mathbf{E}_{A_j}^{(n)}$ for $j = 1, \cdots, K$.
  for $i = 1 : K$
    $\nabla_{\mathbf{G}_i} g(\mathbf{G}_1^{(n)}, \cdots, \mathbf{G}_K^{(n)})$
    $= \frac{\log e}{\sigma^2}\sum_{j=i}^{K} \Delta_j \mathbf{H}_i^h \mathbf{H}_{A_j}\mathbf{G}_{A_j}^{(n)}(\mathbf{E}_{A_j}^i)^{(n)}$.
    set step size $t := 1$.
    do
      $\widetilde{\mathbf{G}}_i^{(n)} = \mathbf{G}_i^{(n)} + t\nabla_{\mathbf{G}_i} g(\mathbf{G}_1^{(n)}, \cdots, \mathbf{G}_K^{(n)})$.
      $\mathbf{G}_i^{(n+1)} = \sqrt{P_i}\widetilde{\mathbf{G}}_i^{(n)}/\|\widetilde{\mathbf{G}}_i^{(n)}\|_F$, if $\|\widetilde{\mathbf{G}}_i^{(n)}\|_F^2 > P_i$.
      compute $g^{(n+1)}$ based on $\mathbf{G}^{(n+1)} =$ $\mathrm{Bdiag}\{\mathbf{G}_1^{(n)}, \cdots, \mathbf{G}_{i-1}^{(n)}, \mathbf{G}_i^{(n+1)}, \mathbf{G}_{i+1}^{(n)}, \cdots, \mathbf{G}_K^{(n)}\}$.
      $t := \beta t$.
    while $g^{(n+1)} < g^{(n)} + \alpha t\|\nabla_{\mathbf{G}_i} g(\mathbf{G}_1^{(n)}, \cdots, \mathbf{G}_K^{(n)})\|_F^2$.
  end
until the $g(\mathbf{G}_1^{(n)}, \cdots, \mathbf{G}_K^{(n)})$ converges or n reaches maximum iteration number.

Due to the non-concavity of the weighted sum rate $g(\mathbf{G}_1, \cdots, \mathbf{G}_K)$, the proposed algorithm can only find local optimum. To reduce the chance of being trapped in local maxima, we run the iterative algorithm with random initialization multiple times and choose the one with maximal weighted sum rate to be the final solution [14].

We note that the complexity of the proposed algorithm is mainly due to computations of $g(\mathbf{G}_1, \cdots, \mathbf{G}_K)$ and MMSE matrix $\mathbf{E}_{A_j}$, $j = 1, 2, \cdots, K$. When input signals are Gaussian, the weighted sum rate has a simple analytical expression as in (2). However, the computation of weighted sum rate with finite discrete inputs demands more consideration and higher complexity. From (5) and (6), we can see that the computation of $g(\mathbf{G}_1, \cdots, \mathbf{G}_K)$ involves summation of all possible transmitted vectors from all users, and thus its complexity grows exponentially with $N_t{\cdot}K$. Since it is generally very difficult to obtain a closed-form expression of $g(\mathbf{G}_1, \cdots, \mathbf{G}_K)$, we use Monte Carlo simulation method to estimate its value. Such an approach has been adopted in [14], [15] dealing with single user.

Similar to computing $g(\mathbf{G}_1, \cdots, \mathbf{G}_K)$, we argue that the MMSE matrix $\mathbf{E}_{A_j}$ can also be estimated via Monte Carlo simulation method. When $N_{A_j} = \prod_{i\in j} N_i$ and $p(\mathbf{x}_{A_j} = \mathbf{x}_{A_j}^i) = 1/N_{A_j}$, the MMSE estimate of $\mathbf{x}_{A_j}$ is given by

$$\begin{aligned}\hat{\mathbf{x}}_{A_j} &= \mathbb{E}(\mathbf{x}_{A_j}|\mathbf{y}, \mathbf{x}_{A_j^c}) = \sum_{l=1}^{N_{A_j}} \mathbf{x}_{A_j}^l p(\mathbf{x}_{A_j}^l|\mathbf{y}, \mathbf{x}_{A_j^c})\\ &= \frac{\sum_{l=1}^{N_{A_j}} \mathbf{x}_{A_j}^l p(\mathbf{y}|\mathbf{x}_{A_j} = \mathbf{x}_{A_j}^l, \mathbf{x}_{A_j^c})}{\sum_{i=1}^{N_{A_j}} p(\mathbf{y}|\mathbf{x}_{A_j} = \mathbf{x}_{A_j}^i, \mathbf{x}_{A_j^c})}\end{aligned} \tag{22}$$

where

$$p(\mathbf{y}|\mathbf{x}_{A_j} = \mathbf{x}_{A_j}^l, \mathbf{x}_{A_j^c}) = \frac{1}{(\pi\sigma^2)^{N_r}} \exp\left(-\frac{\|\mathbf{y} - \mathbf{H}_{A_j}\mathbf{G}_{A_j}\mathbf{x}_{A_j}^l - \mathbf{H}_{A_j^c}\mathbf{G}_{A_j^c}\mathbf{x}_{A_j^c}\|^2}{\sigma^2}\right). \tag{23}$$

Substitute (22) and (23) to (12), the MMSE matrix can be formulated to the expectation of a function of complex Gaussian vector $\mathbf{v}$ as:

$$\mathbf{E}_{A_j} = \mathbf{I}_{N_t j} - \frac{1}{N_A}\sum_{m=1}^{N_A} \mathbb{E}_{\mathbf{v}}\left\{\frac{\left[\sum_{l=1}^{N_A} \mathbf{x}_{A_j}^l q_{m,l}(\mathbf{v})\right]\left[\sum_{k=1}^{N_A} (\mathbf{x}_{A_j}^k)^h q_{m,k}(\mathbf{v})\right]}{\left[\sum_{i=1}^{N_A} q_{m,i}(\mathbf{v})\right]^2}\right\} \tag{24}$$

where the function $q_{m,l}(\mathbf{v})$ is defined as

$$q_{m,l}(\mathbf{v}) = \exp\left(-\frac{\|\mathbf{H}_{A_j}\mathbf{G}_{A_j}\left(\mathbf{x}_{A_j}^m - \mathbf{x}_{A_j}^l\right) + \mathbf{v}\|^2}{\sigma^2}\right). \tag{25}$$

Therefore, we can randomly generate Gaussian vectors $\mathbf{v}$ to obtain an estimate of $\mathbf{E}_{A_j}$ by (24) and (25).

V. ITERATIVE DETECTION AND DECODING FOR MAC

We have discussed the precoder design with finite discrete inputs from the information theoretical perspective in previous sections. Yet, another major concern in practical communication systems is the bit error rate or frame error rate. Therefore, we provide a transmission scheme for the multiple access channels with multiple antennas in this section. More specifically, all the transmitters adopt the LDPC channel coding and linear precoders discussed in Section IV. At the receiver side, the iterative processing technique involving the soft detection and channel decoder is employed to achieve good performance. This type of transceiver structure has been reported to have promising performance in various applications [31]–[34].

Fig. 1. shows a bank of parallel transmitters of K users. The i-th user transmits blocks of information bits, and each block $\mathbf{u}_i$ is encoded by the LDPC encoder. The coded bits $\mathbf{c}_i$ are then interleaved and fed into the modulator. Since a squared precoder is considered at each transmitter, we split the stream of the modulated symbols into N_t independent streams by a serial to parallel converter. Finally, the symbol $\mathbf{x}_i$ is multiplied by the individual precoding matrix $\mathbf{G}_i$, and transmitted to

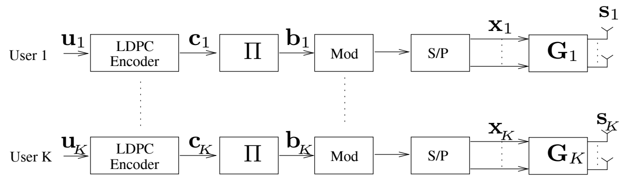

Fig. 1. MIMO uplink transmitters with precoding.

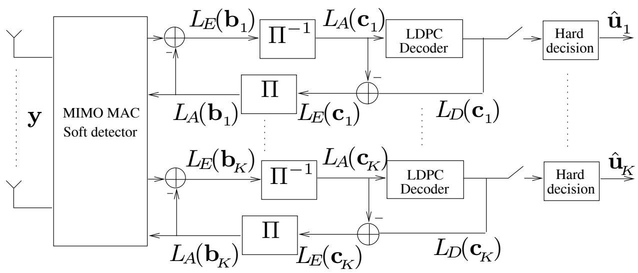

Fig. 2. Iterative receiver of MIMO multiple access channel.

the space through N_t antennas. We note that all users can separately use identical LDPC encoders and interleavers, while the linear precoders may differ from each other according to Section IV.

The iterative receiver is given in Fig. 2. The information intended for all users is iteratively exchanged between a MIMO MAC soft detector and a bank of LDPC soft channel decoders. Within each iteration, the soft multiuser detector, dealing with both cross-antenna and multiuser interferences, generates the extrinsic information $L_E(\mathbf{b}_i), i = 1, \cdots, K$, based on the received signal $\mathbf{y}$ and the priori information $L_A(\mathbf{b}_i), i = 1, \cdots, K$. The log likelihood ratio (LLR) $L_E(\mathbf{b}_i)$ is then interleaved and fed into the i-th user's LDPC decoder as the intrinsic information $L_A(\mathbf{c}_i)$. Those soft decoding methods of LDPC codes, such as the log domain sum product algorithm [35], should be adopted to exploit the redundancy among coded bits $\mathbf{c}_i$ and compute the intrinsic information $L_D(\mathbf{c}_i)$. After channel decoding, $L_D(\mathbf{c}_i)$ is subtracted by $L_A(\mathbf{c}_i)$, and interleaved to become the priori knowledge $L_A(\mathbf{b}_i)$ of the MAC detector for use in the next iteration. At the final iteration, hard decisions are made upon $L_D(\mathbf{c}_i), i = 1, \cdots, K$ to obtain estimate of information of all users.

There are two main categories of soft multiuser detection methods: linear and non-linear detection. An recent overview of iterative linear detection can be found in [36]. Although the linear processing has the advantage of low complexity and easy implementation, its performance is usually non-optimal. In addition, when the number of total transmit antennas $N_t K$ exceeds the number of receiver antennas N_r, the linear detection may not have the sufficient capability to tackle the MAC interference.

On the other hand, the non-linear detections include the MAP method [31], [37], soft interference cancelation [37], sphere decoding [20], and Markov chain Monte Carlo (MCMC) approach [38], etc. Among these methods, the MAP detection can achieve the optimal performance while others are sub-optimal or near-optimal. Since our main goal in this section is to verify that the precoder designed via weighed sum rate maximization can also provide excellent BER performance, we choose to implement the MAP method for the MIMO MAC soft detector in Fig. 2 to obtain optimal BER results. For each received vector $\mathbf{y}$, the extrinsic LLR $L_E(\mathbf{b}_i)$ can be given as [31]

$$L_E(b_{i,j}) = \ln \frac{\sum_{\mathbf{b}\in\mathbb{B}_{k,+1}} p(\mathbf{y}|\mathbf{b}) \exp\left[\frac{1}{2}\mathbf{b}_{[k]}^t \mathbf{L}_A(\mathbf{b}_{[k]})\right]}{\sum_{\mathbf{b}\in\mathbb{B}_{k,-1}} p(\mathbf{y}|\mathbf{b}) \exp\left[\frac{1}{2}\mathbf{b}_{[k]}^t \mathbf{L}_A(\mathbf{b}_{[k]})\right]} \quad (26)$$

where $b_{i,j}$ means the j-th bit of the i-th user's bit vector $\mathbf{b}_i$, with $1 \le i \le K$, and $1 \le j \le M_c N_t$, assuming that all users employ the same modulation and the number of the constellation points of each modulated symbol is M_c. The vector $\mathbf{b} = [\mathbf{b}_1^t, \cdots, \mathbf{b}_K^t]^t$ with length $M_c N_t K$ contains the interleaved bits from all users. The vector $\mathbf{b}_{[k]}^t$ denotes the subvector of $\mathbf{b}$ with the k-th element omitted, in which $k = (i-1) M_c N_t + j$. The vector $\mathbf{L}_A(\mathbf{b}_{[k]})$ with $(M_c N_t K - 1)$ elements represents the priori information of $\mathbf{b}_{[k]}$. The sets $\mathbb{B}_{k,+1}$ and $\mathbb{B}_{k,-1}$ denote the sets of $2^{M_c N_t K - 1}$ bit vectors $\mathbf{b}$ with the k-th element equaling to $+1$ and -1, respectively. The channel likelihood function in (26) is given by

$$\begin{aligned} p(\mathbf{y}|\mathbf{b}) &= p[\mathbf{y}|\mathbf{x} = map(\mathbf{b})] \\ &= \frac{1}{(\pi\sigma^2)^{N_r}} \exp\left(-\frac{\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2}{\sigma^2}\right) \end{aligned} \quad (27)$$

where $\mathbf{x} = map(\mathbf{b})$ means the mapping from the bit vector $\mathbf{b}$ to symbol vectors $\mathbf{x}$, including modulation and S/P conversions of all users.

VI. Numerical Results

In this section, we provide numerical results of the constellation-constrained capacity region and sum rate with finite discrete inputs for 2-user MAC. We also simulate BER performance of the system described in Section V. Assume that there are two receiver antennas and two transmit antennas for each user. Suppose the maximum individual power $P_1 = P_2 = P$, and all users adopt the same modulation scheme. Then, the signal to noise ratio can be defined as $\text{SNR} = P/\sigma^2$, when the channels are normalized. In our simulations, we choose the noise power $\sigma^2 = 1$.

For illustrative purpose, we consider an example of two fixed channel matrices for two users, which are given by

$$\mathbf{H}_1 = \begin{bmatrix} 1.3898 & 0.1069j \\ -0.1069j & 0.2138 \end{bmatrix}, \mathbf{H}_2 = \begin{bmatrix} 1.2247 & 0 \\ 0 & 0.707 \end{bmatrix}.$$

Each channel matrix has normalized power with $\text{Tr}\left(\mathbf{H}_i \mathbf{H}_i^h\right) = N_r = 2$, as in [10].

Fig. 3 plots the convergence behavior of the sum rate maximization algorithm in Table I with BPSK inputs. At each SNR, we run the algorithm with random initializations 10 times and choose the one with the largest sum rate at the end of iterations. From the figure, we can see that the proposed algorithm usually converges after 15 iterations under different SNRs. For backtracking line search, the typical range of parameters α and β are $\alpha \in (0.01, 0.3)$, and $\beta \in (0.1, 0.8)$ [22]. We choose $\alpha = 0.1$ and $\beta = 0.5$, for the algorithm in Section IV. The Monte Carlo simulation number for both the sum rate and MMSE matrix is set to 500. In general, a limited simulation number in Monte Carlo method leads to a certain

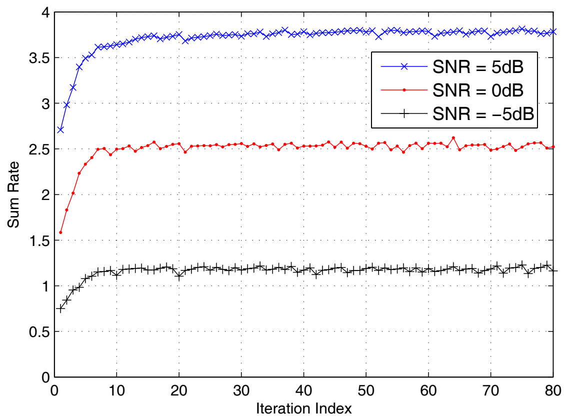

Fig. 3. Convergence of sum rate maximization algorithm with BPSK inputs.

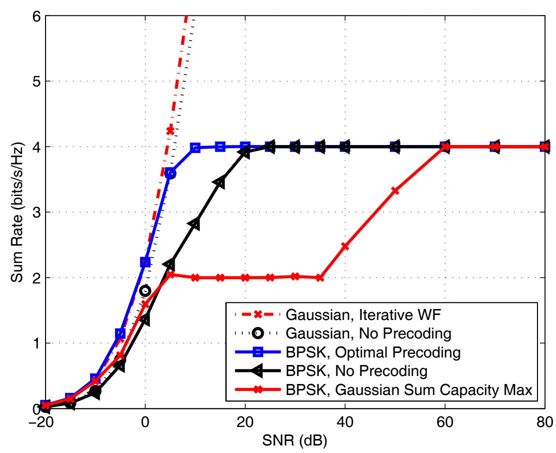

Fig. 4. Sum rate of 2-user MAC with BPSK inputs.

degree of estimate errors. This is the reason why there are small ripples of sum rate in Fig. 3.

Fig. 4 shows the sum rate of various precoding schemes with BPSK modulation. We implement the Gaussian-input sum capacity maximization method [2] and replace the inputs with BPSK inputs, which is denoted as "BPSK, Gaussian sum capacity max". In this method, the optimal input covariance matrices $\{\mathbf{Q}_1, \cdots, \mathbf{Q}_K\}$ are obtained by maximizing Gaussian-input sum capacity described in (2) and (3). After using standard convex optimization tool [23] to solve this problem, we decompose each covariance matrix as $\mathbf{Q}_i = \mathbf{V}_i \mathbf{\Sigma}_i \mathbf{V}_i^h$, and choose the precoder to be $\mathbf{G}_i = \mathbf{V}_i \mathbf{\Sigma}_i^{\frac{1}{2}}$. Then, we replace Gaussian inputs to BPSK signals and calculate the sum rate of this precoding scheme using (5). A noticeable result of Gaussian-input sum capacity maximization method is that within certain SNR range, each user's covariance matrix $\mathbf{Q}_i$ has only one positive eigenvalue due to the water-filling policy. For example, when SNR is 20 dB, the covariance matrices obtained by the traditional method are

$$\mathbf{Q}_1 = \begin{bmatrix} 98.63 & 11.61j \\ -11.61j & 1.37 \end{bmatrix}, \mathbf{Q}_2 = \begin{bmatrix} 2.64 & -16.04j \\ 16.04j & 97.36 \end{bmatrix}.$$

With eigenvalue decomposition $\mathbf{Q}_i = \mathbf{V}_i \mathbf{\Sigma}_i \mathbf{V}_i^h$, we find that

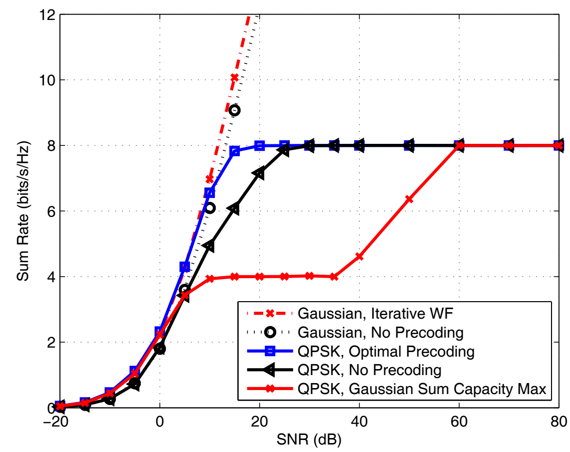

Fig. 5. Sum rate of 2-user MAC with QPSK inputs.

$\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = diag\,\{100, 0\}$, and the precoding matrices are as follows:

$$\mathbf{G}_1 = \begin{bmatrix} -9.93 & 0 \\ 1.17j & 0 \end{bmatrix}, \mathbf{G}_2 = \begin{bmatrix} -1.63 & 0 \\ -9.87j & 0 \end{bmatrix}.$$

In this case the precoder acts as beamforming by allowing each user to transmit only one modulated symbol in vector $\mathbf{x}_i$. Therefore, although the traditional method can achieve sum capacity with the ideal Gaussian inputs assumption, usually it fails to serve as the optimal strategy for practical finite discrete inputs.

For comparison purpose, we plot the Gaussian-input sum capacity achieved by iterative water-filling [3] and sum rate of Gaussian inputs without precoding. From the numerical results, we have several observations: our proposed algorithm, denoted as the "optimal precoding", outperforms the non-precoding (*i.e.*, $\mathbf{G}_i = \sqrt{\frac{P_i}{N_t}\mathbf{I}_{N_t}}$) with finite discrete inputs for a wide SNR range from -15 dB to 20 dB. For instance, the SNR gain of the optimal precoding over non-precoding to achieve the sum rate 3 bits/s/Hz is about 8 dB. When SNR approaches infinity, the sum rate of both methods saturate at 4 bits/s/Hz, which is determined by the constellation size, the number of users and transmit antennas. In addition, we find that when the SNR is less than 0 dB, the optimal precoding with BPSK inputs obtains the same sum rate as the iterative WF with Gaussian inputs. For the SNR range below 5 dB, our method also has performance gains compared to sum rate of Gaussian inputs without precoding or power allocation.

The precoding matrices obtained via optimal precoding method when SNR $= 5$ dB are

$$\mathbf{G}_1^{opt} = \begin{bmatrix} 1.22 - 0.048i & -0.21 + 1.20i \\ 0.12 - 0.17i & 0.37 - 0.10i \end{bmatrix}$$

and

$$\mathbf{G}_2^{opt} = \begin{bmatrix} -0.56 - 0.81i & 0.66 - 0.04i \\ 0.44 + 0.36i & 0.47 + 1.10i \end{bmatrix}.$$

The sum rate results regarding QPSK inputs are shown in Fig. 5. We have similar observations as the BPSK inputs, in the sense that the proposed precoding with QPSK inputs achieves higher sum rate than non-precoding for the SNR

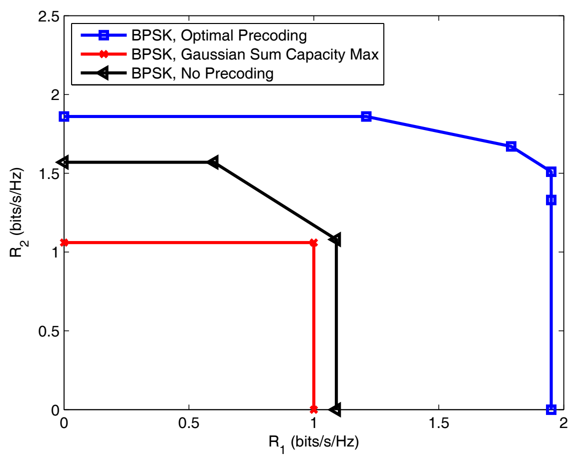

Fig. 6. Capacity region of 2-user MAC with BPSK inputs, when SNR = 5dB.

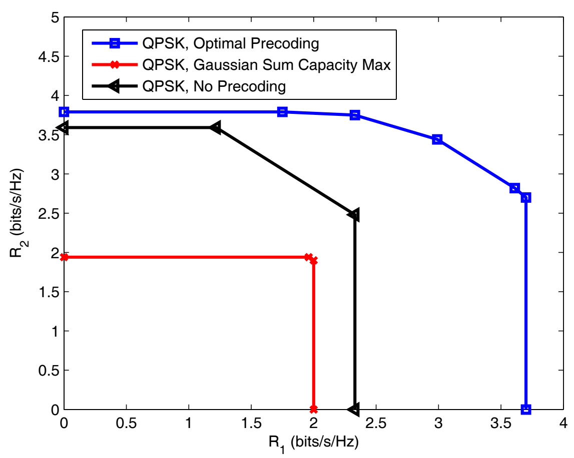

Fig. 7. Capacity region of 2-user MAC with QPSK inputs, when SNR = 10dB.

range from 0 dB to 25 dB, and obtains the same sum rate of iterative WF with Gaussian inputs when SNR is less than 5 dB. At the target sum rate of 3 bits/s/Hz, the optimal precoding achieves SNR gains of about 5 dB and 40 dB, compared to non-precoding and Gaussian-input sum capacity maximization method, respectively.

The constellation-constrained capacity region of optimal precoding with BPSK inputs when SNR = 5 dB is illustrated in Fig. 6. When inputs are BPSK signals, the rate regions achieved by non-precoding and Gaussian-input sum capacity maximization schemes are also plotted. The pentagons are determined by equations in Proposition 1. The curve of optimal precoding is obtained by varying μ_1 and μ_2 using weighted sum rate maximization algorithm with discrete inputs in Section IV. We can see that the constellation-constrained capacity region of optimal precoding is much larger than the rate regions of non-precoding and Gaussian-input sum capacity maximization method.

Fig. 7 plots regions of QPSK signals when SNR = 10 dB. Similar to the results of BPSK signals, the rate regions achieved by non-precoding and Gaussian-input sum capacity maximization are inside the constellation-constrained capacity region with QPSK inputs, which is obtained by the proposed precoding algorithm.

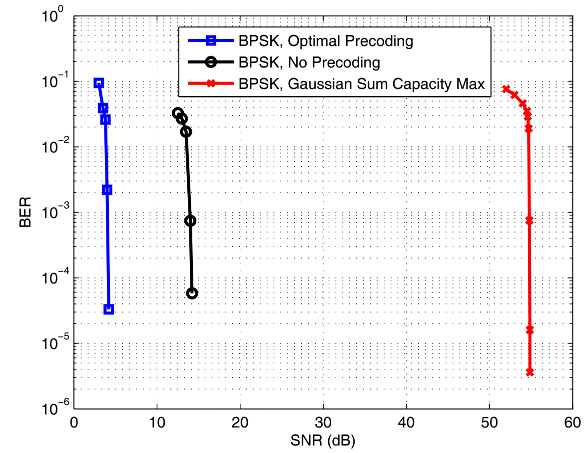

Fig. 8. BER of 2-user MAC with BPSK inputs.

In addition to the sum rate and rate region, we also simulate the BER performances of the 2-user MAC system described in Section V. The LDPC encoder and decoder simulation package [39] is used. The length of each codeword is set to 9600, and coding rate is $3/4$. All users separately employ identical LDPC codes and pseudorandom interleavers. At the receiver side, the multiuser detector uses MAP method expressed in (26), and LDPC decoders adopt the sum-product algorithm [35] with 30 iterations. The soft information exchange between the MAP multiuser detector and LDPC channel decoders. The iteration number between MAP detector and LDPC decoder is set to 5.

Fig. 8 plots the BER curves of BPSK inputs under the fixed channels. The results include optimal precoding, non-precoding, and Gaussian-input sum capacity maximization method, which show that that the optimal precoding achieves significant gains over the other methods. From the sum rate results in Fig. 4, we can see that when the channel coding rate is $3/4$ (sum rate of 3 bits/s/Hz), the SNRs of the optimal precoding and non-precoding schemes are about 3 dB and 11 dB, respectively. From information theoretical perspective, these SNR limits are the minimum acceptable SNRs for error-free communication. At the target BER of 10^{-4}, the SNR of optimal precoding is about 4 dB, which is close to the limit predicted by the sum rate vs. SNR curve. In addition, the performance gains of optimal precoding over non-precoding and Gaussian-input sum capacity maximization well match the results observed in Fig. 4. The comparison of sum rate and BER results justifies that the precoding method of using sum rate with finite discrete inputs can not only maximize the achievable information rate, but also achieve excellent system performance in terms of bit error rates.

Fig. 9 shows the BER performance with QPSK inputs. We have similar observations as the results of BPSK signals. The simulations indicate that to achieve the BER of 10^{-4}, the optimal precoding method outperforms non-precoding by 6 dB, which is nearly the same amount of SNR gain when the sum rate is 6 bits/s/Hz in Fig. 5.

VII. Conclusion

In this paper, we studied the linear precoder design for MIMO MAC with finite discrete inputs. From the infor-

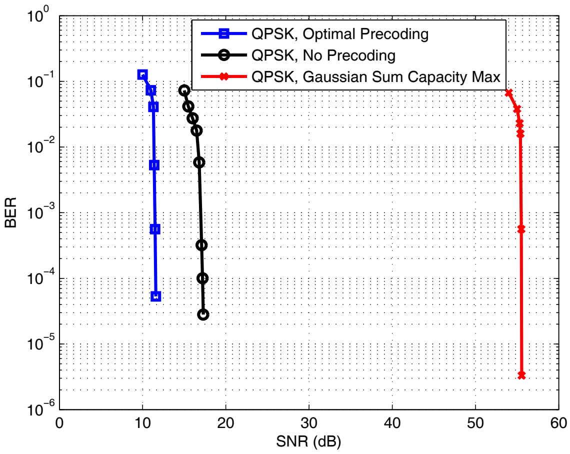

Fig. 9. BER of 2-user MAC with QPSK inputs.

mation theoretical perspective, we derived the constellation-constrained capacity region. We then found a set of necessary conditions of weighted sum rate maximization with individual power constraints and proposed an iterative algorithm to obtain the optimal precoding matrices for all users. The convergence behavior of our algorithm has been verified by simulations. We have shown that when inputs are digital modulated signals, and SNR is in the medium range, our precoding method offers significantly higher sum rate than non-precoding and the existing Gaussian-input sum capacity maximization approach. An LDPC coded system with iterative detection and decoding for MAC was further presented to evaluate the BER performance of such precoders. BER simulation results indicated that the optimal precoding achieves significant SNR gain against the non-precoding system and the system with Gaussian-input sum capacity maximization precoders.

Appendix A
Proof of Proposition 1

The *priori* probabilities of $\mathbf{x}_A$ and $\mathbf{x}_{A^c}$ are $p\left(\mathbf{x}_A = \mathbf{x}_A^i\right) = \frac{1}{N_A}$ and $p\left(\mathbf{x}_{A^c} = \mathbf{x}_{A^c}^k\right) = \frac{1}{N_{A^c}}$, where $N_A = \prod_{i \in A} N_i$ and $N_{A^c} = \prod_{i \in A^c} N_i$. Based on the Gaussian vector channel model $\mathbf{y} = \mathbf{H}_A \mathbf{G}_A \mathbf{x}_A + \mathbf{H}_{A^c} \mathbf{G}_{A^c} \mathbf{x}_{A^c} + \mathbf{v}$, the conditional probability density function of $\mathbf{y}$ can be written as

$$p\left(\mathbf{y}|\mathbf{x}_A = \mathbf{x}_A^{i_1}, \mathbf{x}_{A^c} = \mathbf{x}_{A^c}^{i_2}\right) = \frac{1}{(\pi\sigma^2)^{N_r}} \exp\left(-\frac{\|\mathbf{y} - \mathbf{H}_A \mathbf{G}_A \mathbf{x}_A^{i_1} - \mathbf{H}_{A^c} \mathbf{G}_{A^c} \mathbf{x}_{A^c}^{i_2}\|^2}{\sigma^2}\right). \quad (28)$$

The conditional entropy $H(\mathbf{y}|\mathbf{x}_{A^c})$ can be calculated as

$$\begin{aligned} H\left(\mathbf{y}|\mathbf{x}_{A^c}\right) &= \sum_{i_2=1}^{N_{A^c}} p\left(\mathbf{x}_{A^c} = \mathbf{x}_{A^c}^i\right) H\left(\mathbf{y}|\mathbf{x}_{A^c} = \mathbf{x}_{A^c}^{i_2}\right) \\ &= -\frac{1}{N_{A^c}} \sum_{i_2=1}^{N_{A^c}} \int p\left(\mathbf{y}|\mathbf{x}_{A^c} = \mathbf{x}_{A^c}^{i_2}\right) \log p\left(\mathbf{y}|\mathbf{x}_{A^c} = \mathbf{x}_{A^c}^{i_2}\right) d\mathbf{y} \\ &= -\frac{1}{N_{A^c}} \sum_{i_2=1}^{N_{A^c}} \int \left[\frac{1}{N_A} \sum_{i_1=1}^{N_A} p\left(\mathbf{y}|\mathbf{x}_A = \mathbf{x}_A^{i_1}, \mathbf{x}_{A^c} = \mathbf{x}_{A^c}^{i_2}\right)\right] \\ &\quad \log\left[\frac{1}{N_A} \sum_{k_1=1}^{N_A} p\left(\mathbf{y}|\mathbf{x}_A = \mathbf{x}_A^{k_1}, \mathbf{x}_{A^c} = \mathbf{x}_{A^c}^{i_2}\right)\right] d\mathbf{y}. \quad (29) \end{aligned}$$

Substituting (28) to the above equation and assuming $\mathbf{y} - \mathbf{H}_A \mathbf{G}_A \mathbf{x}_A^{i_1} - \mathbf{H}_{A^c} \mathbf{G}_{A^c} \mathbf{x}_{A^c}^{i_2} = \mathbf{v}$, we have

$$H(\mathbf{y}|\mathbf{x}_{A^c}) = \log N_A - \frac{1}{N_A} \sum_{i_1=1}^{N_A} \mathbb{E}_{\mathbf{v}} \left[\log \sum_{k_1=1}^{N_A} \frac{1}{(\pi\sigma^2)^{N_r}} \exp\left(-\frac{\|\mathbf{H}_A \mathbf{G}_A \left(\mathbf{x}_A^{i_1} - \mathbf{x}_A^{k_1}\right) + \mathbf{v}\|^2}{\sigma^2}\right)\right], \quad (30)$$

where $\mathbf{v}$ is a complex Gaussian vector with probability density function $p\left(\mathbf{v}\right) = \frac{1}{(\pi\sigma^2)^{N_r}} \exp\left(-\frac{\|\mathbf{v}\|^2}{\sigma^2}\right)$.

Similarly, we can get $H(\mathbf{y}|\mathbf{x}_A, \mathbf{x}_{A^c})$ as follows:

$$H\left(\mathbf{y}|\mathbf{x}_A, \mathbf{x}_{A^c}\right) = \mathbb{E}_{\mathbf{v}} \left[\log \frac{1}{(\pi\sigma^2)^{N_r}} \exp\left(-\frac{\|\mathbf{v}\|^2}{\sigma^2}\right)\right]. \quad (31)$$

From (30) and (31), we can prove Proposition 1.

References

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd edition. Wiley, 2006.
[2] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 684–702, June 2003.
[3] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, pp. 145–152, Jan. 2004.
[4] M. Kobayashi and G. Caire, "Iterative waterfilling for the weighted rate sum maximization in MIMO MAC," in *Proc. IEEE Signal Processing Advances in Wireless Communications*, July 2006.
[5] M. Kobayashi and G. Caire, "An iterative water-filling algorithm for maximum weighted sum-rate of Gaussian MIMO-BC," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1640–1646, Aug. 2006.
[6] 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, June 2006.
[7] S. Serbetli and A. Yener, "Transceiver optimization for multiuser MIMO systems," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 214–226, Jan. 2004.
[8] X. Yuan, C. Xu, and X. Lin, "Precoder design for multiuser MIMO ISI channels based on iterative LMMSE detection," *IEEE J. Sel. Topics Signal Process.*, vol. 3, pp. 1118–1128, Dec. 2009.
[9] A. Lozano, A. M. Tulino, and S. Verdú, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Trans. Inf. Theory*, vol. 52, pp. 3033–3051, July 2006.
[10] C. Xiao and Y. R. Zheng, "On the mutual information and power allocation for vector Gaussian channels with finite discrete inputs," in *Proc. IEEE Glocal Telecommunications Conference*, Nov. 2008.
[11] C. Xiao and Y. R. Zheng, "Transmit precoding for MIMO systems with partial CSI and discrete-constellation inputs," in *Proc. IEEE International Conference on Communications*, June 2009.
[12] M. Payaró and D. P. Palomar, "On optimal precoding in linear vector Gaussian channels with arbitrary input distribution," in *Proc. 2009 IEEE International Conference on Symposium on Information Theory*, June, 2009.
[13] M. Lamarca, "Linear precoding for mutual information maximization in MIMO systems," in *Proc. 6th International Symposium on Wireless Communication Systems*, 2009.
[14] F. Pérez-Cruz, M. R. D. Rodrigues, and S. Verdú, "MIMO Gaussian channels with arbitrary inputs: optimal precoding and power allocation," *IEEE Trans. Inf. Theory*, vol. 56, pp. 1070–1084, Mar. 2010.
[15] C. Xiao, Y. R. Zheng, and Z. Ding, "Globally optimal linear precoders for finite alphabet signals over complex vector Gaussian channels," *IEEE Trans. Signal Process.*, vol. 59, pp. 3301–3314, July 2011.
[16] W. Zeng, M. Wang, C. Xiao, and J. Lu, "On the power allocation for relay networks with finite-alphabet constraints," in *Proc. IEEE Glocal Telecommunications Conference*, Dec. 2010.
[17] W. Zeng, Y. R. Zheng, M. Wang, and J. Lu, "Linear precoding for relay networks: a perspective on finite-alphabet inputs," submitted to *IEEE Trans. Wireless Commun.*, Dec. 2010.

[18] N. Deshpande and B. S. Rajan, "Constellation constrained capacity of two-user broadcast channels," in *Proc. IEEE Glocal Telecommunications Conference*, Nov. 2009.

[19] R. Ghaffar and R. Knopp, "Near optimal linear precoder for multiuser MIMO for discrete alphabets," in *Proc. IEEE International Conference on Communications*, May 2010.

[20] J. Harshan and B. S. Rajan, "On two-user Gaussian multiple access channels with finite input constellations," *IEEE Trans. Inf. Theory*, vol. 57, pp. 1299–1327, Mar. 2011.

[21] E. Biglieri, *Coding for Wireless Channels*. Springer-Verlag New York, Inc., 2005.

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

[23] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming, version 1.21., http://cvxr.com/cvx, Jan. 2011.

[24] D. N. C. Tse and S. Hanly, "Multi-access fading channels—part I: polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inf. Theory*, vol. 44, pp. 2796–2815, Nov. 1998.

[25] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, 1993.

[26] A. Hjørungnes, *Complex-Valued Matrix Derivatives: With Applications in Signal Processing and Communications*. Cambridge University Press, 2011.

[27] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, pp. 1261–1282, Apr. 2005.

[28] D. P. Palomar and S. Verdú, "Gradient of mutual information in linear vector Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 52, pp. 141–154, Jan. 2006.

[29] D. Bertsekas and J. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods.* Prentice-Hall, 1989.

[30] A. Soysal and S. Ulukus, "Optimum power allocation for single-user MIMO and multi-user MIMO-MAC with partial CSI," *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 1402–1412, Sep. 2007.

[31] B. Hochwald and S. T. Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, pp. 389–399, Mar. 2003.

[32] M. Tüchler, R. Koetter, and A. Singer, "Turbo equalization: principles and new results," *IEEE Trans. Commun.*, vol. 50, pp. 754–767, May 2002.

[33] X. Wang and V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, pp. 1046–1061, July 1999.

[34] J. Ylioinas and M. Juntti, "Iterative joint detection, decoding, and channel estimation in Turbo-coded MIMO-OFDM," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 1784–1796, May. 2009.

[35] X-Y. Hu, E. Eleftherious, D-M. Arnold, and A. Dholakia, "Efficient implementaions of the sum-product algorithm for decodng LDPC codes," in *Proc. IEEE Glocal Telecommunications Conference*, Nov. 2001.

[36] L. Rasmussen, "Linear detection in iterative joint multiuser decoding," in *Proc. 3rd International Symposium on Communications, Control and Signal Processing*, Mar. 2008.

[37] L. Xu, S. Chen, and L. Hanzo, "EXIT chart analysis aided Turbo MUD designs for the rank-deficient multiple antenna assisted OFDM uplink," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 2039–2044, June 2008.

[38] S. Henriksen, B. Ninness, and S. R. Weller, "Convergence of Markov-chain Monte-Calro approaches to multiuser and MIMO detection," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 497–505, Apr. 2008.

[39] M. Valenti, *The Coded Modulation Library*, http://www.iterativesolutions.com.

Mingxi Wang received his B.S degree in telecommunication engineering from Beihang University (formerly: Beijing University of Aeronautics and Astronautics), China, in 2006, and M.S degree in signal and information processing from the same university in 2009. He is currently a Ph.D. candidate of department of electrical and computer engineering at Missouri University of Science and Technology (formerly: University of Missouri-Rolla), USA. His research interests include wireless communications, signal processing, MIMO, OFDM, linear precoding and iterative receiver.

Weiliang Zeng received the B.S. degree in electronic engineering from the University of Electronic Sciences and Technology of China (UESTC), Chengdu, China, in July 2007 with the highest honors. Since September 2007, he has been a Ph.D. candidate at the Tsinghua National Laboratory for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing, China. During his doctoral studies, he has conducted cooperative research at the Department of Electronic Engineering, Missouri University of Science and Technology, USA. His research interests include communications and information theory with special emphasis on wireless communications and signal processing.

Chengshan Xiao (M'99-SM'02-F'10) received the B.S. degree in electronic engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 1987, the M.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 1989, and the Ph.D. degree in electrical engineering from the University of Sydney, Sydney, Australia, in 1997.

From 1989 to 1993, he was with the Department of Electronic Engineering, Tsinghua University, where he was a member of Research Staff and then a Lecturer. From 1997 to 1999, he was a Senior Member of Scientific Staff with Nortel, Ottawa, ON, Canada. From 1999 to 2000, he was a Faculty Member with the University of Alberta, Edmonton, AB, Canada. From 2000 to 2007, he was with the University of Missouri, Columbia, where he was an Assistant Professor and then an Associate Professor. He is currently a Professor with the Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla (formerly University of Missouri, Rolla). His research interests include wireless communications, signal processing, and underwater acoustic communications. He is the holder of three U.S. patents. His algorithms have been implemented into Nortel's base station radios after successful technical field trials and network integration.

Dr. Xiao is the Editor-in-Chief of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, a Distinguished Lecturer of the IEEE Communications Society, a Distinguished Lecturer of the IEEE Vehicular Technology Society, and a Fellow of the IEEE. Previously, he served as the founding Area Editor for Transmission Technology of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS; an Associate Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS-I, and the *International Journal of Multidimensional Systems and Signal Processing*. He was the Technical Program Chair of the 2010 IEEE International Conference on Communications (ICC), the Lead Co-chair of the 2008 IEEE ICC Wireless Communications Symposium, and a Phy/MAC Program Co-chair of the 2007 IEEE Wireless Communications and Networking Conference. He served as the founding Chair of the IEEE Technical Committee on Wireless Communications, and the Vice-Chair of the IEEE Technical Committee on Personal Communications.