Summer 2012

Power allocation and linear precoding for wireless communications with finite-alphabet inputs

Mingxi Wang

Department: Electrical and Computer Engineering

POWER ALLOCATION AND LINEAR PRECODING FOR WIRELESS COMMUNICATIONS WITH FINITE-ALPHABET INPUTS

by

MINGXI WANG

A DISSERTATION

Presented to the Faculty of the Graduate School of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

ELECTRICAL ENGINEERING

2012

Approved by

Chengshan Xiao, Advisor
Yahong Rosa Zheng
Kurt Kosbar
Randy Moss
Xiaoqing (Frank) Liu

PUBLICATION DISSERTATION OPTION

This dissertation consists of the following four published or to be submitted papers, formatted in the style used by the Missouri University of Science and Technology, listed as follows:

Paper I, M. Wang, C. Xiao, and W. Zeng, "Linear Precoding for MIMO Multiple Access Channels with Finite Discrete Inputs," has been published in IEEE Trans. Wireless Commun., vol. 10, pp 3934-3942, Nov. 2011. (pages 6-36)

Paper II, W. Zeng, M. Wang, C. Xiao, and J. Lu, "On the Power Allocation for Relay Networks with Finite-alphabet Constraints," has been published in Proc. IEEE GLOBECOM, Miami, FL, Dec. 2010. (pages 37-55)

Paper III, W. Zeng, Y. R. Zheng, M. Wang, and J. Lu, "Linear Precoding for Relay Networks: A Perspective on Finite-Alphabet Inputs," has been accepted by IEEE Trans. on Wireless Commun., Dec. 2011. (pages 56-95)

Paper IV, M. Wang, Y. R. Zheng, and C. Xiao, "Practical Linear Precoder Design for Finite Alphabet MIMO-OFDM with Experiment Validation," to be submitted to IEEE Trans. Veh. Technol., 2012. (pages 96-127)

ABSTRACT

This dissertation proposes a new approach to maximizing data rate/throughput of practical communication system/networks through linear precoding and power allocation. First, the mutual information or capacity region is derived for finite-alphabet inputs such as phase-shift keying (PSK), pulse-amplitude modulation (PAM), and quadrature amplitude modulation (QAM) signals. This approach, without the commonly used Gaussian input assumptions, complicates the mutual information analysis and precoder design but improves performance when the designed precoders are applied to practical systems and networks.

Second, several numerical optimization methods are developed for multiple-input multiple-output (MIMO) multiple access channels, dual-hop relay networks, and point-to-point MIMO systems. In MIMO multiple access channels, an iterative weighted sum rate maximization algorithm is proposed which utilizes an alternating optimization strategy and gradient descent update. In dual-hop relay networks, the structure of the optimal precoder is exploited to develop a two-step iterative algorithm based on convex optimization and optimization on the Stiefel manifold. The proposed algorithm is insensitive to initial point selection and able to achieve a near global optimal precoder solution. The gradient descent method is also used to obtain the optimal power allocation scheme which maximizes the mutual information between the source node and destination node in dual-hop relay networks. For point-to-point

MIMO systems, a low complexity precoding design method is proposed, which maximizes the lower bound of the mutual information with discretized power allocation vector in a non-iterative fashion, thus reducing complexity.

Finally, performances of the proposed power allocation and linear precoding schemes are evaluated in terms of both mutual information and bit error rate (BER). Numerical results show that at the same target mutual information or sum rate, the proposed approaches achieve 3-10dB gains compared to the existing methods in the medium signal-to-noise ratio region. Such significant gains are also indicated in the coded BER systems.

ACKNOWLEDGMENTS

The research work reported in this dissertation was carried out with the Department of Electrical and Computer Engineering at Missouri University of Science and Technology, Rolla, Missouri, USA, from August 2009 to May 2011.

First of all, I would like to thank my Ph.D. advisor Professor Chengshan Xiao. I greatly appreciate his support and guidance for my course and research work. His attitudes and visions towards work and other things have been a role model for me during my study.

Besides my advisor Professor Xiao, I would like to thank Professor Yahong Rosa Zheng. During my Ph.D. study, she gave me valuable help with my course work and research. I also worked with her on several projects and papers from which I learned a lot.

I want to thank the other members of my advisory committee, Professors. Kurt Kosbar, Randy Moss, and Xiaoqing Liu. They have made various valuable comments and suggestions to help me improve my research work. They were also considerate in arranging their time for my comprehensive exams and final defense.

In addition, it is my pleasure to acknowledge my appreciation to some of my co-authors, especially Mr. Weiliang Zeng and Mr. Yongpeng Wu, with whom I worked on several theoretical problems. I enjoyed the collaborations and I am glad that we have made productive publications. I also cherish the time I spent with all the other members in the research lab.

Last but not least, I dedicate this dissertation to my parents and relatives in China. I would like to express my sincere gratitude for their love and support.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

PAPER IV

LIST OF TABLES

1 INTRODUCTION

1.1 BACKGROUND AND PROBLEM STATEMENT

Linear precoding for multiple-input multiple-output (MIMO) communications has been a popular research topic in the last decade. Traditional precoding methods can be mainly classified into three categories [1]: (a) diversity- driven designs [2]; (b) rate-driven designs [3, 4] and (c) mean squared error (MSE) minimization or signal-to-noise ratio (SNR) maximization [5]. The diversity-driven designs employs pairwise error probability analysis to maximize diversity order [2]. Yet it may not achieve the highest coding gain. The rate-driven designs utilize ergodic or outage capacity to optimize the precoder. Such approach rely on Gaussian input assumptions, whereas the transmitted signals in practical digital communication systems are non-Gaussian distributed, drawn from discrete 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. For the third categories, the MSE minimization and SNR maximization approaches [5] may not necessarily provide minimum bit error rate (BER) or maximum date rate for practical finite alphabet systems.

To overcome the aforementioned drawbacks, several works have recently reported that designing precoders for point to point MIMO system by maximizing mutual information with finite-alphabet inputs can achieve higher mutual information [1, 6, 7] and lower bit error rate (BER) [1] than employing other optimization criteria. The performance benefits of these approaches come from optimization of

mutual information formulated directly with finite-alphabet input constraints, compared to using other indirect methods such as maximizing channel capacity with Gaussian inputs, maximizing diversity order, minimizing SINR, or minimizing MSE.

To maximize the mutual information with finite-alphabet input, this dissertation develops several numerical optimization methods for multiple-input multiple-output (MIMO) multiple access channels, dual-hop relay networks, and point-to-point MIMO systems. In MIMO multiple access channels, an iterative weighted sum rate maximization algorithm is proposed which utilizes an alternating optimization strategy and gradient descent update. In dual-hop relay networks, the structure of the optimal precoder is exploited to develop a two-step iterative algorithm based on convex optimization and optimization on the Stiefel manifold. The proposed algorithm is insensitive to initial point selection and able to achieve a near global optimal precoder solution. The gradient descent method is also used to obtain the optimal power allocation scheme which maximizes the mutual information between the source node and destination node in dual-hop relay networks. For point-to-point MIMO systems, a low complexity precoding design method is proposed, which maximizes the lower bound of the mutual information with discretized power allocation vector in a non-iterative fashion, thus reducing complexity.

1.2 SUMMARY OF CONTRIBUTIONS

This dissertation consists of a couple of journal publications and conference papers as listed in the publication list. My contributions that are published or under review are:

1. *Linear precoding for MIMO multiple access channels with finite discrete inputs*: The constellation-constrained capacity region for the MIMO MAC is derived with an arbitrary number of users. Due to the non-concavity of the objective function, a set of necessary conditions for the optimization problem are obtained through

Karush-Kuhn-Tucker analysis. An iterative algorithm is proposed to utilize alternating optimization strategy. In particular, each iteration of the algorithm involves the gradient descent update with backtracking 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 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.

2. *On the power allocation for relay networks with finite-alphabet constraints*: An optimal power allocation scheme is proposed to maximize the mutual information for the relay networks under discrete-constellation input constraint. Numerical examples show that significant gain can be obtained compared to the conventional counterpart for nonfading channels and fading channels. At the same time, we show that the large performance gain on the mutual information will also represent the large gain on the bit-error rate, i.e., the benefit of the power allocation scheme predicted by the mutual information can indeed be harvested and can provide considerable performance gain in a practical system.

3. *Linear precoding for relay networks with a perspective on finite-alphabet inputs*: This paper exploits the structure of the optimal precoder that maximizes the mutual information and develops a two-step algorithm based on convex optimization and optimization on the Stiefel manifold. By doing so, the proposed algorithm is insensitive to initial point selection and able to achieve a near global optimal precoder solution. Besides, it converges fast and offers high mutual information gain. These

advantages are verified by numerical examples, which also show the large performance gain in mutual information also represents the large gain in the coded bit-error rate.

4. Practical linear precoder design for finite alphabet MIMO-OFDM with experiment validation. A low complexity precoding method is proposed for practical multiple-input multiple-output (MIMO) orthogonal frequency-division multiplexing (OFDM) systems. Based on the two-step optimal precoder design algorithm that maximizes the lower bound of the mutual information with finite-alphabet inputs, the proposed method simplifies the precoder design by fixing the right singular vectors of the precoder matrix, eliminating the iterative optimization between the two steps, and discretizing the search space of the power allocation vector. For a 4×4 channel, the computational complexity of the proposed precoder design is reduced to 3% and 6% of that required by the original two-step algorithm with Quadrature Phase Shift Keying (QPSK) and 8PSK, respectively. The proposed method achieves nearly the same mutual information as the two-step iterative algorithm for a large range of SNR region, especially for large MIMO size and/or high constellation systems. The proposed precoding design method is applied to a 2×2 MIMO-OFDM system with 2048 subcarriers by designing 1024 precoders for extended channel matrices of size 4×4. A transceiver test bed implements these precoding matrices in comparison with other existing precoding schemes. Indoor experiments are conducted for fixed-platform non-line-of-sight (NLOS) channels, and the data processing results show that the proposed precoding method achieves the lowest BER compared to maximum diversity, classic water-filling and channel diagonalization methods.

1.3 REFERENCES

[1] C. Xiao, Y. R. Zheng, and Z. Ding, "Globally optimal linear precoders for finite alphabet signals over complex vector Gaussian channels," IEEE Trans. on Signal Process., vol. 59, pp3301-3314, July 2011.

[2] Y. Xin, Z. Wang, and G. B. Giannakis, "Space-time diversity systems based on linear constellation precoding," *IEEE Trans. Wireless Commun.*, vol. 2, pp. 294-309, Mar. 2003.

[3] T. M. Cover and J. A. Thomas, Elements of Information Theory, 2nd ed. New York: Wiley, 2006.

[4] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," IEEE J. Seleted. Areas Commun., vol.21, pp.684-702, Jun. 2003.

[5] D. P. Palomar, J. Cioffi and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, pp.2381-2401, Sep 2003.

[6] A. Lozano, A. M. Tulino, and S. Verdu, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," IEEE Trans. Inform. Theory, vol.52, pp.3033-3051, July 2006.

[7] F. Perez-Cruz, M. R. D. Rodrigues, and S. Verdu, "MIMO Gaussian channels with arbitrary inputs: optimal precoding and power allocation," IEEE Trans. Inform. Theory, vol.56, pp.1070-1084, Mar. 2010.

PAPER

I. LINEAR PRECODING FOR MIMO MULTIPLE ACCESS CHANNELS WITH FINITE DISCRETE INPUTS

Mingxi Wang, Weiliang Zeng 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 backtracking 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.

1 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 [1,2,3]. The optimal input covariance matrices can be found by maximizing the weighted sum rate, which is a convex optimization problem with Gaussian inputs [4,5,6]. 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 procoding 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, 10, 12, 13, 14, 11]. Similar problems are also investigated in relay networks [15,16] and broadcast channels [17,18]. For 2-user single-input single-output (SISO) MAC, [19] 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 [19, 20], 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 [3], we obtain a set of necessary conditions through Karush-Kuhn-Tucker (KKT) analysis [21]. 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 2 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 3. Section 4 discusses necessary condition of the weighed sum rate maximization problem

and the details of the iterative algorithm. Section 5 presents the MIMO system over MAC with iterative detection and decoding. Numerical results are provided in Section 6, and Section 6 draws the conclusions.

2 SYSTEM MODEL AND EXISTING GAUSSIAN INPUTS RESULTS

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

$$\mathbf{y} = \mathbf{H}_1\mathbf{G}_1\mathbf{x}_1 + \mathbf{H}_2\mathbf{G}_2\mathbf{x}_2 + \cdots + \mathbf{H}_K\mathbf{G}_K\mathbf{x}_K + \mathbf{v} = \mathbf{HGx} + \mathbf{v} \tag{1}$$

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]:

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

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

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 [21, 22].

3 CONSTELLATION-CONSTRAINED CAPACITY REGION

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 $p\left(\mathbf{x}_1\right)$, $p\left(\mathbf{x}_2\right)$, $\cdots$, $p\left(\mathbf{x}_K\right)$.

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 [19].

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:*

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

where $\mathbf{v} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$*, 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 $I(\mathbf{x}_A; \mathbf{y}|\mathbf{x}_{A^c}) = H(\mathbf{y}|\mathbf{x}_{A^c}) - H(\mathbf{y}|\mathbf{x}_A, \mathbf{x}_{A^c})$, we can prove (5). Details can be found in Appendix 8.

According to [23], 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:

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

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

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

4 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.

4.1 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:*

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

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

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

$$\nu_i \geq 0 \tag{11}$$

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* $\mathbf{G}_{A_j} = Bdiag\{\mathbf{G}_1, \mathbf{G}_2, \cdots, \mathbf{G}_j\}$. *The symbol* $\mathbf{E}_{A_j}^i \in \mathbb{C}^{N_t j \times N_t}$ *stands for the i-th column block of the minimum mean square error (MMSE) matrix [24] of* $\mathbf{E}_{A_j}$, *which is defined as*

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

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

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

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 [25], then the KKT conditions are as follows:

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

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

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

$$\lambda_i\geq 0 \tag{17}$$

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

Due to the relation between the mutual information and MMSE [26, 27], the gradient of $f(\mathbf{G}_1,\cdots,\mathbf{G}_j)$ can be found as:

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

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

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

where $\mathbf{E}_{A_j}^i = \mathbf{E}_{A_j}(\mathbf{e}_i^h \otimes \mathbf{I}_{N_t}) \in \mathbb{C}^{N_t j \times N_t}$ 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).

4.2 ITERATIVE ALGORITHM FOR WSR 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 [28, 7, 29]. 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 [27]:

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

For fast convergence, we use backtracking line search [21] 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 1.

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 [13].

Table 1. 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.

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 [13, 14] 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

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

where

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

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:

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

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

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

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

5 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 4. 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 [30, 31, 32, 33].

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 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 4.

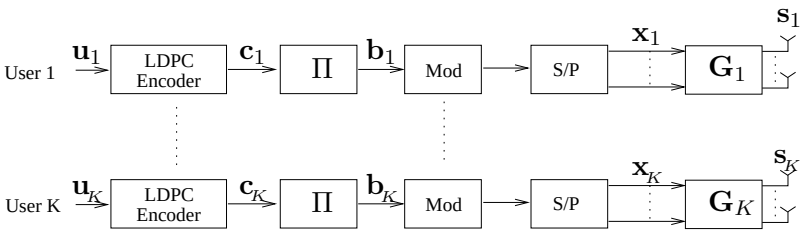


Figure 1. MIMO uplink transmitters with precoding.

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 [34], 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.

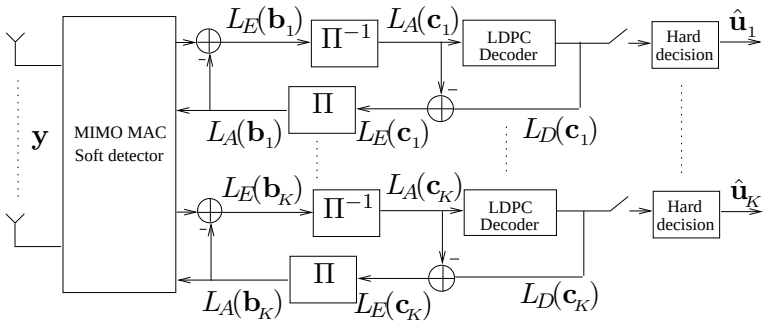


Figure 2. Iterative receiver of MIMO multiple access channel.

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 [35]. 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_tK 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 [30,36], soft interference cancelation [36], sphere decoding [19], and Markov chain Monte Carlo (MCMC) approach [37], 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 [30]

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

where $b_{i,j}$ means the j-th bit of the i-th user's bit vector $\mathbf{b}_i$, with $1 \leq i \leq K$, and $1 \leq j \leq M_c N_t$, 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 $\mathbf{L}_A\left(\mathbf{b}_{[k]}\right)$ 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

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

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

6 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 5. 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 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 $\mathrm{Tr}\left(\mathbf{H}_i\mathbf{H}_i^h\right) = N_r = 2$, as in [10].

Fig. 3 plots the convergence behavior of the sum rate maximization algorithm in Table 1 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 $\beta \in (0.1, 0.8)$ [21]. We choose $\alpha = 0.1$ and $\beta = 0.5$, for the algorithm in Section 4. 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 degree of estimate errors. This is the reason why there are small ripples of sum rate in Fig. 3.

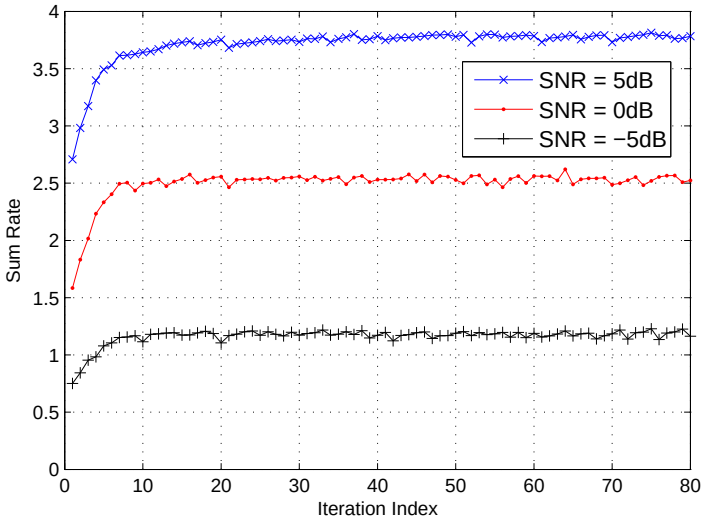


Figure 3. Convergence of sum rate maximization algorithm with BPSK inputs.

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 [22] 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 $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = diag\{100, 0\}$, 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.

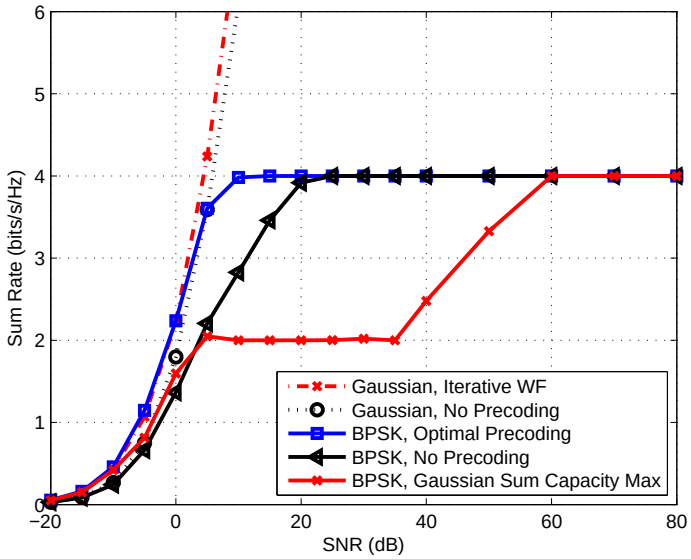


Figure 4. Sum rate of 2-user MAC with BPSK inputs.

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 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.

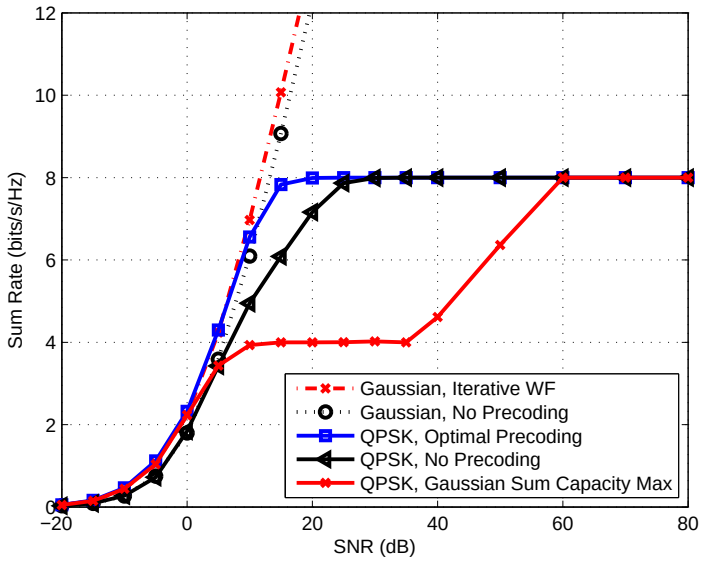


Figure 5. Sum rate of 2-user MAC with QPSK inputs.

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 4. 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.

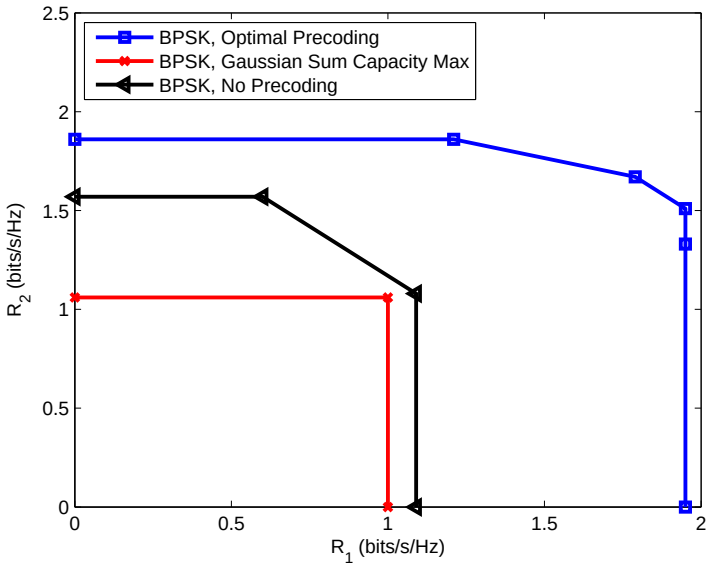


Figure 6. Capacity region of 2-user MAC with BPSK inputs, when SNR = 5dB.

In addition to the sum rate and rate region, we also simulate the BER performances of the 2-user MAC system described in Section 5. The LDPC encoder and decoder simulation package [38] 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 [34] 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

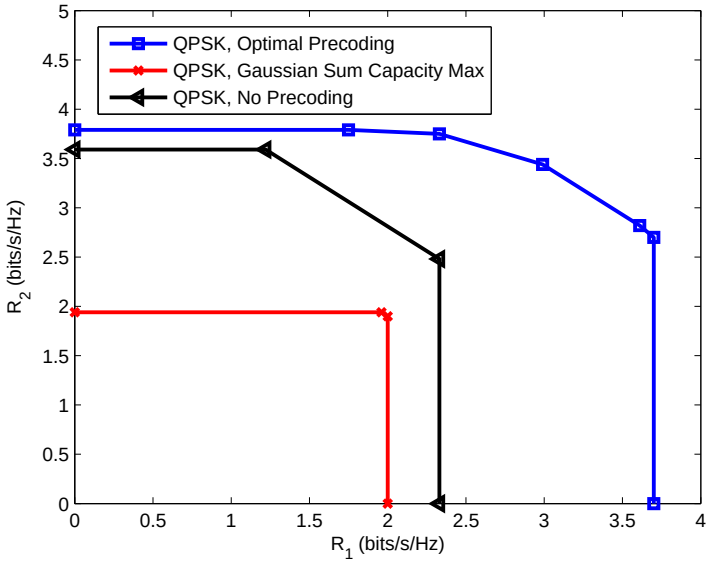


Figure 7. Capacity region of 2-user MAC with QPSK inputs, when SNR = 10dB.

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,

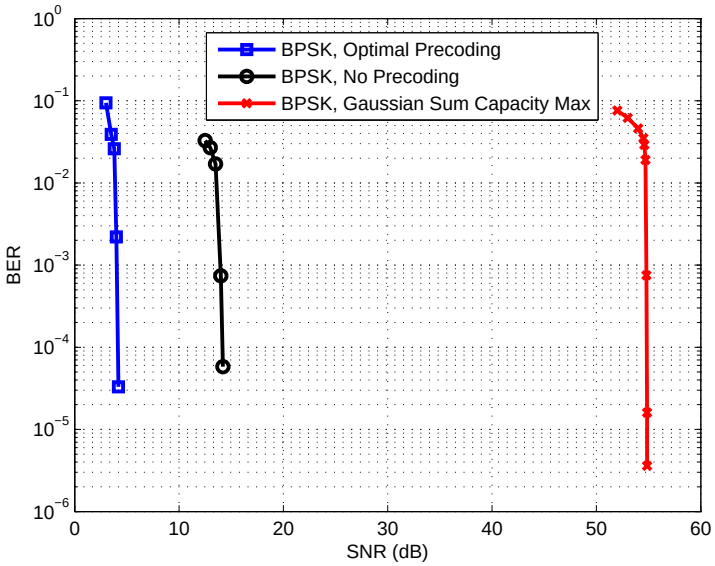


Figure 8. BER of 2-user MAC with BPSK inputs.

which is nearly the same amount of SNR gain when the sum rate is 6 bits/s/Hz in Fig. 5.

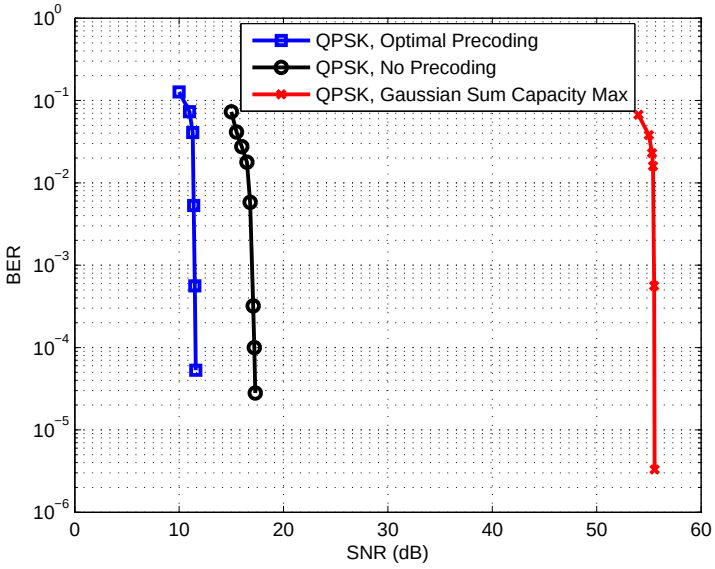


Figure 9. BER of 2-user MAC with QPSK inputs.

7 CONCLUSION

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

8 APPENDIX: 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

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

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

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

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

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

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

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

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

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

9 REFERENCES

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley, 2006.

[2] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Seleted. Areas Commun.*, vol.21, pp.684-702, Jun. 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 (SPAWC 2006)*, Cannes, 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. Inform. 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 (Globecom)*, New Orleans, USA, Nov. 30 - Dec. 4, 2008.

[11] M. Payaró and D. P. Palomar, "On optimal precoding in linear vector Gaussian channels with arbitrary input distribution, " in *Proc. of 2009 IEEE international conference on Symposium on Information Theory (ISIT)*, Seoul, Korea, June, 2009.

[12] M. Lamarca, "Linear precoding for mutual information maximization in MIMO systems," in *Proc. of 6th International Symposium on Wireless Communication Systems*, Siena, Italy, 2009.

[13] F. Pérez-Cruz, M. R. D. Rodrigues, and S. Verdú, "MIMO Gaussian channels with arbitrary inputs: optimal precoding and power allocation," *IEEE Trans. Inform. Theory*, vol.56, pp.1070-1084, Mar. 2010.

[14] C. Xiao, Y. R. Zheng, and Z. Ding, "Globally optimal linear precoders for finite alphabet signals over complex vector Gaussian channels," *IEEE Trans. on Signal Process.*, vol. 59, pp3301-3314, July 2011.

[15] 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 (Globecom)*, Miami, USA, Dec. 2010.

[16] 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. on Wireless Commun.*, Dec. 2010.

[17] N. Deshpande and B. S. Rajan, "Constellation constrained capacity of two-user broadcast channels," in *Proc. IEEE Glocal Telecommunications Conference (Globecom)*, Honolulu, HI, USA, 30 Nov. - 4 Dec. 2009.

[18] R. Ghaffar and R. Knopp, "Near optimal linear precoder for multiuser MIMO for discrete alphabets," in *Proc. IEEE International Conference on Communications (ICC)*, Cape Town, South Africa, May 23-27, 2010.

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

[20] E. Biglieri, *Coding for wireless channels*, Springer-Verlag New York, Inc, 2005.

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

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

[23] D. N. C. Tse and S. Hanly, "Multi-access fading channels: Part I: Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2796-2815, Nov. 1998.

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

[25] A. Hjørungnes, "Complex-valued matrix derivatives: with applications in signal processing and communications," Cambridge University Press, 2011.

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

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

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

[29] 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, Sept. 2007.

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

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

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

[33] 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.

[34] 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 (Globecom)*, San Antonio, USA, Nov. 25 - 29, 2001.

[35] L. Rasmussen, "Linear detection in iterative joint multiuser decoding,", in *Proc. of 3rd International Symposium on Communications, Control and Signal Processing (ISCCSP)*, St Julians, Malta, Mar. 12-14, 2008.

[36] 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. on Wireless Commun.*, vol.7, pp.2039-2044, Jun. 2008.

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

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

II. ON THE POWER ALLOCATION FOR RELAY NETWORKS WITH FINITE ALPHABET CONSTRAINT

Weiliang Zeng, Mingxi Wang, Chengshan Xiao, *Fellow, IEEE*, and Jianhua Lu

ABSTRACT—In this paper, we investigate the optimal power allocation scheme for relay networks with finite-alphabet constraints. It has been shown that the previous work utilizing various design criteria with the Gaussian inputs assumption may lead to significant loss for a practical system with finite constellation set constraint, especially when signal-to-noise ratio (SNR) is in medium-to-high regions, or when the channel coding rate is medium to high. An optimal power allocation scheme is proposed to maximize the mutual information for the relay networks under discrete-constellation input constraint. Numerical examples show that significant gain can be obtained compared to the conventional counterpart for nonfading channels and fading channels. At the same time, we show that the large performance gain on the mutual information will also represent the large gain on the bit-error rate (BER), i.e., the benefit of the power allocation scheme predicted by the mutual information can indeed be harvested and can provide considerable performance gain in a practical system.

1 INTRODUCTION AND RELATED WORK

Cooperative relaying has been shown to provide reliable high data rate services in wireless networks without the need of multiple antennas at each node. These benefits can be further exploited by utilizing judicious cooperative strategies see [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11] and the references therein.

The existing design methods may be categorized into two groups: i) diversity oriented designs; and ii) transmission rate oriented designs. The first group usually achieves the steepest asymptotic slope (the highest diversity order) on the outage probability versus SNR curve, however, it may not obtain the highest possible coding gain, such as the distributed space-time coding (DST) in [2, 3] and the relaying selection scheme in [4, 5]. The second group often optimizes the performance with the Gaussian inputs assumption, for example, maximizing output SNR [6, 7, 8, 9], minimizing mean square error (MSE) [10, 6] and maximizing channel capacity [6, 7, 11].

Although Gaussian inputs are capacity-achieving signaling, they can never be realized in practice. Rather, the inputs must be drawn from a finite constellation set (such as pulse amplitude modulation (PAM), quadrature amplitude modulation (QAM) and phase shift keying (PSK) modulation) in a practical communication systems, which may significantly depart from the Gaussian idealization [12, 13, 14]. Yet, no solution has been found in the above work for the power allocation that maximizes the potential transmission rate, i.e., mutual information, with non-Gaussian inputs over the relay networks, and this is exactly the concern of this study.

For multiple-input multiple-output (MIMO) systems, it has been shown that the design from the standpoint of finite alphabet can result in significant performance improvement [12]. In this paper, we will see similar performance gains achieved in relay networks over the existing schemes utilizing the design criteria such as SNR, MSE,

and channel capacity. At the same time, it has been validated that the large performance gain predicted by the mutual information with finite-alphabet constraints can indeed be harvested and will lead to considerable performance improvement in a practical system.

The rest of the paper is organized as follows. In Sec. II, we introduce the relay network model and the main problem. In Sec. III, we analyze the solution structure of the power allocation problem from the information theoretic and optimization theoretic point of view. We also utilize the receiver structure that has the near-capacity performance to validate the effectiveness of the proposed scheme. Sec. IV presents numerical results, followed by the conclusions in Sec. V.

Notation: Throughout this paper, we use boldface upper-case letters to denote matrices, boldface lower-case letters to denote column vectors, and italics to denote scalars. The superscripts $(\cdot)^T$ and $(\cdot)^H$ stand for transpose and conjugate transpose, respectively; $[\mathbf{A}]_{i,j}$ and $[\mathbf{A}]_{:,j}$ denote the $(i\text{th}, j\text{th})$ element and jth column of matrix $\mathbf{A}$, respectively; and $\|\mathbf{c}\|$ denotes Euclidean norm of vector $\mathbf{c}$. $\mathbf{I}$ denotes the identity matrix with the appropriate dimensions; $\text{diag}(\mathbf{c})$ denotes the diagonal matrix with diagonal elements given by the vector $\mathbf{c}$; $\text{Tr}\,(\mathbf{A})$ denotes the trace operation; $\Re$ denotes the real part of complex number; and $\mathbb{E}$ denotes statistical expectation. Likewise, all logarithms are to the base 2.

2 SYSTEM MODEL AND PRELIMINARIES

Consider a relay network with one transmit-and-receive pair, where the source node (s) attempts to communicate to the destination node (d) with the assistance of k relays $(r_1, r_2, \cdots, r_k)$. The link from the source to the ith relay is denoted as h_i, the link from the ith relay to the destination is denoted as g_i, and the direct link from the source to the destination is denoted as h_0.

We model the source-relay (S-R), source-destination (S-D) and relay-destination (R-D) links as the quasi-static flat-fading channels, which are applicable to the scenarios of narrow-band transmissions in a low-mobility environment. We assume that the ith relay knows its own channels h_i and g_i, and the destination obtains full knowledge of S-D, S-R and R-D channels.

We consider the average power constraint of each node for each time slot, e.g., the power used at the source and the ith relay should be less than P_s and P_r, respectively.

The data transmission is over two time slots using two hops. The symbols transmitted by the source node in the first and second time slot are denoted as x_1 and x_2, respectively. They may be chosen from some complex-valued finite constellation $\mathcal{C}$. We assume that $\mathbb{E}[x_i] = 0$ and $\mathbb{E}[|x_i|^2] = 1$ for $i = 1, 2$. During the first time slot, the source node sends $\alpha_0\sqrt{P_s}x_1$. Let y_{r_i} and $y_{d,1}$ be received signals at the ith relay node and the destination, respectively, which are given by

$$y_{r_i} = \alpha_0\sqrt{P_s}h_i x_1 + n_{r_i}, \tag{1}$$

$$y_{d,1} = \alpha_0\sqrt{P_s}h_0 x_1 + n_{d,1}, \tag{2}$$

where n_{r_i} and $n_{d,1}$ are complex additive white Gaussian noise at the ith relay and the destination with zero mean and unit variance $\sim \mathcal{CN}(0,1)$.

The ith relay node normalizes the received signal by a factor of $\sqrt{\mathbb{E}\left[|y_{r_i}|^2\right]}$ (so that the average energy is unity) and retransmits the signal

$$t_i = \sqrt{\frac{P_r}{\mathbb{E}\left[|y_{r_i}|^2\right]}}\alpha_i y_{r_i},\ i = 1, 2, \cdots, k \tag{3}$$

during the second time slot. At the same time, the source node sends $\alpha_{k+1}\sqrt{P_s}x_2$. Then the destination node receives a superposition of the relay transmissions and the source transmission during the second time slot according to

$$\begin{aligned} y_{d,2} &= \sum_{i=1}^{k} g_i t_i + \alpha_{k+1}\sqrt{P_s}h_0 x_2 + n_{d,2} \\ &= \sum_{i=1}^{k} \sqrt{\frac{P_s P_r}{1 + P_s\left|\alpha_0 h_i\right|^2}}\alpha_0\alpha_i h_i g_i x_1 + \sqrt{P_s}\alpha_{k+1}h_0 x_2 + v, \end{aligned} \tag{4}$$

where the effective noise $v \sim \mathcal{CN}(0, N_d)$ with $N_d = 1 + \sum_{i=1}^{k}\frac{P_r|\alpha_i g_i|^2}{1+P_s|\alpha_0 h_i|^2}$. We normalize $y_{d,2}$ by a factor $w = N_d^{1/2}$ in order to simplify the presentation. Finally, the effective input-output relation for the two-hop transmission can be summarized as

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{n}, \tag{5}$$

where $\mathbf{y} = [y_{d,1}\ y_{d,2}/w]^T$ is the received signal vector, and $\mathbf{G}$ is the effective channel matrix given by

$$\mathbf{G} = \begin{bmatrix} \sqrt{P_s}\alpha_0 h_0 & 0 \\ \sum_{i=1}^{k}\sqrt{\frac{P_s P_r}{(1+P_s|\alpha_0 h_i|^2)w^2}}\alpha_0\alpha_i h_i g_i & \sqrt{\frac{P_s}{w^2}}\alpha_{k+1}h_0 \end{bmatrix}. \tag{6}$$

$\mathbf{x} = [x_1\ x_2]^T$ is the transmitted signal vector, and $\mathbf{n} \in \mathbb{C}^{2\times 1}$ is the channel noise vector, assumed independent and identically distributed (i.i.d.) complex Gaussian with zero mean and unit variance, i.e., $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. Equivalently, we can rewrite the effective channel $\mathbf{G}$ as $\mathbf{HP}$ with the channel related matrix

$$\mathbf{H} = \begin{bmatrix} \sqrt{P_s}h_0 & 0 & \cdots & 0 & 0 \\ 0 & \sqrt{P_sP_r}h_1g_1 & \cdots & \sqrt{P_sP_r}h_kg_k & \sqrt{P_s}h_0 \end{bmatrix}, \tag{7}$$

and the power allocation related matrix

$$\mathbf{P} = \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_k & 0 \\ 0 & 0 & \cdots & 0 & \gamma_{k+1} \end{bmatrix}^T, \tag{8}$$

where $\gamma_0 = \alpha_0$; $\gamma_i = \alpha_0\alpha_i / \sqrt{\left(1 + P_s\left|\alpha_0 h_i\right|^2\right) w^2}$, $\forall i = 1, \cdots, k$; and $\gamma_{k+1} = \alpha_{k+1}/w$.

We should note that the coefficients $\alpha_0, \alpha_1, \cdots, \alpha_{k+1}$ are complex value, and the rationale of introducing them in the model is in fact quite intuitive. First, they can be used to control the average transmit power of each node at each time slot, which requires that

$$\alpha_i\alpha_i^* \leq 1 \quad \forall i = 0, \cdots, k+1. \tag{9}$$

Hence the average power used at the ith relay node is $\left|\alpha_i\right|^2 P_r$. Second, the choice of the angles can be used to cancel the phases introduced from the channels and ensure that the signal components are added constructively at the receiver, i.e., $\arg\alpha_i = -\left(\arg h_i + \arg g_i\right)$, $i = 1, \cdots, k$, [7, 8, 10]. We also set $\arg\alpha_0 = \arg\alpha_{k+1} = -\arg h_0$, since the optimal choice of both angles can be realized by rotating the input constellations equivalently. What is left is the choice of their magnitude, so we will treat the power allocation coefficients $\alpha_0, \alpha_1, \cdots, \alpha_{k+1}$ as a set of real numbers in the sequel.

Our power allocation scheme is thus the design of coefficients $\alpha_0, \alpha_1, \cdots, \alpha_{k+1}$ to maximize the mutual information with finite-alphabet constraints. Note that for

the proposed algorithm to be effectively implemented in practice, a low-rate feedback should be allowed from the destination to the source and relay nodes. The feedback is needed in the cooperative system since antennas are not located at a single terminal as in a MIMO system. This may result in small penalty on system performance, but the cost is often compensated by a significant performance gain at high SNR [1, 14].

3 POWER ALLOCATION FOR FINITE ALPHABET INPUTS

We consider the conventional equiprobable discrete signaling constellations such as M-ary PSK, PAM, or QAM, where M is the number of points in the signal constellation. The mutual information between $\mathbf{x}$ and $\mathbf{y}$, with $\mathbf{H}$ and $\mathbf{P}$ known at the receiver, is $\mathcal{I}(\mathbf{x};\mathbf{y})$ given by (10) [12],

$$\mathcal{I}(\mathbf{x};\mathbf{y}) = \log M - \frac{1}{2M^2}\sum_{m=1}^{M^2}\mathbb{E}_{\mathbf{n}}\left\{\log\sum_{j=1}^{M^2}\exp\left[-\|\mathbf{HP}(\mathbf{x}_m-\mathbf{x}_k)+\mathbf{n}\|^2+\|\mathbf{n}\|^2\right]\right\}. \tag{10}$$

where $\mathbf{x}$ contains two symbols, taken independently from the M-ary signal constellation.

The problem that we pose is the determination of the coefficients $\alpha_0, \alpha_1, \cdots, \alpha_{k+1}$ that maximizes the mutual information $\mathcal{I}(\mathbf{x};\mathbf{y})$ with given input distributions while satisfying individual power constraints, i.e.,

$$\max_{\alpha_0,\cdots,\alpha_{k+1}} \mathcal{I}(\mathbf{x};\mathbf{y}) \tag{11}$$

subject to:

$$\alpha_i \leq 1 \quad \forall i = 0, \cdots, k+1. \tag{12}$$

3.1 ANALYSIS FROM INFORMATION THEORETIC PERSPECTIVE

Applying the chain rule for mutual information [15], we have

$$\mathcal{I}(\mathbf{x};\mathbf{y}) = \mathcal{I}(x_2;\mathbf{y}) + \mathcal{I}(x_1;\mathbf{y}|x_2), \tag{13}$$

where $\mathcal{I}(x_2;\mathbf{y})$ is the mutual information between x_2 and $\mathbf{y}$, and $\mathcal{I}(x_1;\mathbf{y}|x_2)$ is the conditional mutual information between x_1 and $\mathbf{y}$ given x_2. The vector $\mathbf{y}$ is defined

in (5), which can also be written as:

$$\mathbf{y} = [\mathbf{G}]_{:,1}\, x_1 + [\mathbf{G}]_{:,2}\, x_2 + \mathbf{n}. \tag{14}$$

From (14), we can verify that

$$\mathcal{I}(\mathbf{x};\mathbf{y}) = \mathcal{I}(x_2;\mathbf{y}) + \mathcal{I}(x_1;\mathbf{y}|\alpha_{k+1}=0), \tag{15}$$

where $\mathcal{I}(x_1;\mathbf{y}|\alpha_{k+1}=0)$ is the mutual information between x_1 and the received signal $\mathbf{y}$ if the source node does not transmit at the second time-slot.

Since $\mathcal{I}(x_2;\mathbf{y}) > 0$ for $h_0, \alpha_{k+1} \neq 0$, it follows that to maximize the mutual information $\mathcal{I}(\mathbf{x};\mathbf{y})$, the source should always transmit at the second time slot [11]. For the same reason, the source should transmit at the first time slot. Based on the above discussions, we can state the following lemma:

Lemma 1. *The power allocation in the two time slots at the source node is nonzero, i.e.,* $\alpha_0, \alpha_{k+1} \neq 0$*, if the channel between the source and the destination node* h_0 *is nonzero.*

3.2 ANALYSIS FROM OPTIMIZATION THEORETIC PERSPECTIVE

We should note that the constraint (12) is convex in the coefficient matrix $\mathbf{C} = \mathrm{diag}([\alpha_0\, \alpha_1\, \cdots \alpha_{k+1}])$. The cost function (11), however, is nonconcave in the power allocation coefficient α_i for the general case [16]. In the sequel, we capitalize on the relationship between the mutual information and the minimum mean square error (MMSE) matrix to obtain the power allocation scheme for arbitrary input distributions.

Theorem 1. *The optimal power allocation coefficients* $\left[\alpha_0^\star\, \alpha_1^\star\, \cdots \alpha_{k+1}^\star\right]$ *that solves (11) subject to (12) satisfy:*

$$\left.\frac{\partial \mathcal{I}\left(\mathbf{x};\mathbf{y}\right)}{\partial \alpha_i}\right|_{\alpha_i=\alpha_i^\star} = \lambda_i \tag{16}$$

$$\lambda_i\left(\alpha_i^\star - 1\right) = 0 \tag{17}$$

$$\lambda_i \geq 0 \tag{18}$$

with

$$\frac{\partial \mathcal{I}\left(\mathbf{x};\mathbf{y}\right)}{\partial \alpha_i} = \sum_{j=0}^{k+1} \frac{\partial \mathcal{I}\left(\mathbf{x};\mathbf{y}\right)}{\partial \gamma_j}\frac{\partial \gamma_j}{\partial \alpha_i}, \tag{19}$$

and

$$\frac{\partial \mathcal{I}\left(\mathbf{x};\mathbf{y}\right)}{\partial \gamma_i} = \Re\left[\mathbf{H}^H\mathbf{H}\mathbf{P}\mathbf{E}\right]_{i+1,1}, \quad \forall i = 0, \cdots, k \tag{20}$$

$$\frac{\partial \mathcal{I}\left(\mathbf{x};\mathbf{y}\right)}{\partial \gamma_{k+1}} = \Re\left[\mathbf{H}^H\mathbf{H}\mathbf{P}\mathbf{E}\right]_{k+2,2} \tag{21}$$

where $\mathbf{E}$ *is the MMSE matrix defined by*

$$\mathbf{E} = \mathbb{E}\left\{\left[\mathbf{x} - \mathbb{E}\left(\mathbf{x}|\mathbf{y},\mathbf{H},\mathbf{P}\right)\right]\left[\mathbf{x} - \mathbb{E}\left(\mathbf{x}|\mathbf{y},\mathbf{H},\mathbf{P}\right)\right]^H\right\} \tag{22}$$

Proof. The possible solution to (11) subject to (12) is characterized by the *Karush-Kuhn-Tucker* theorem [17], which gives necessary conditions, known as the KKT or first order conditions. To investigate stationary points of the problem (11) we formulate the Lagrangian

$$\mathcal{L}\left(\mathbf{P},\lambda\right) = -\mathcal{I}\left(\mathbf{x};\mathbf{y}\right) + \sum_{i=0}^{k+1}\lambda_i\left(\alpha_i - 1\right), \tag{24}$$

in which the Lagrangian multipliers λ_i, $i = 1, \cdots, k+1$, are chosen to satisfy the power constraints. Then the first order conditions are given by (16) to (18). The partial

derivative $\partial\mathcal{I}/\partial\alpha_i$, $i = 0, \cdots, k+1$, can be proved by employing the techniques developed in [18, 19] for derivatives of mutual information, the techniques developed in [20] for matrix differentiation, and the chain rule for multiple variables.

Typically, it is involved to calculate the MMSE matrix (22), especially for large input dimensions M. But we have been able to estimate the matrix $\mathbf{E}$ using Monte Carlo methods. Hence, we can solve this problem using gradient-based methods in Table 1 according to the gradient of mutual information (19). Since the cost function (11) is nonconcave for the general case, it is possible that (11) has local maxima. Therefore, we should perform the algorithm with multiple initial points and keep the power allocation coefficients offering the largest mutual information.

Table 1. The algorithm for optimal power allocation coefficients

initialize $0 \leq \alpha_i \leq 1$, $i = 0, 1, \cdots, k+1$.
repeat
  for $i = 1 : k$
    set step size $t = 1$.
    compute $g^{(n)}(\alpha_i^{(n)}) = \frac{\partial\mathcal{I}\left(\alpha_i^{(n)}\right)}{\partial\alpha_i}$, in (19).
    $\alpha_i^{(n+1)} = \alpha_i^{(n)} + t \cdot g^{(n)}(\alpha_i^{(n)})$.
    set $\alpha_i^{(n+1)} = 1$, if $\alpha_i^{(n+1)} > 1$.
    choose t by backtracking line search.
  end
until $\mathcal{I}\left(\alpha_0^{(n)}, \alpha_1^{(n)}, \cdots, \alpha_k^{(n)}\right)$ converges or n reaches maximum iteration number.

Finally, the destination node will notify the source and each relay node of its assigned transmission power. In this way, the instantaneous mutual information is maximized for each set of channel realizations. Notice that the resulting optimal power allocation scheme is significantly different from the existing ones in the conventional setting both due to the presence of finite-alphabet constraints and the

multiplexing structure capitalized in the relay networks. The results in section 4 show the significant gains obtained compared to the existing methods.

3.3 ITERATIVE DETECTION AND DECODING

To evaluate the advantage of the proposed method in a more practical may, we utilize the "turbo principle" at the destination node [21,22], which is illustrated in Fig. 1. The signal sequence $\mathbf{b}$ at the source node is encoded by the capacity achievable codes, e.g., low-density parity-check (LDPC) codes, interleaved, and mapped according to the conventional equiprobable discrete signaling constellations. Then it is divided into x_1 and x_2, and transmitted at the two time slots, respectively.

At the receiver, the maximum a posteriori (MAP) detector takes channel observations $\mathbf{y}$ and *a priori* knowledge $L_e(b)$ from the decoder and computes new information $L_e(c)$ for each of the coded bits. In this way, the extrinsic information between the MAP detector and decoder is exchanged in an iterative fashion until desired performance is achieved. It has been shown that the iterative processing is very effective that can achieve near-capacity performance.

4 NUMERICAL RESULTS

Computer simulation was carried out to validate the performance of the proposed scheme. For the sake of completeness, in all the figures, we show the performance corresponding to MMSE strategy with local power constraint in [10] and network beamforming in [7]. To ensure a fair comparison, we also show the performance of modified MMSE strategy and modified network beamforming (sending different symbols at the second time slot, rather than sending the same symbols or sleeping). We consider a three-relay network with the same transmit power at the source and each relay node, i.e., $P_s = P_r = P$, which is indicated by the horizontal axis in the following figures.

We look first at a fixed (non-fading) system with the channel coefficient $h_0 = 0.5$, $h = [0.7, -0.7, 1j]$ and $g = [0.9j, 2.1, 0.3]$. The instantaneous mutual information that can be achieved by different schemes is shown in Fig. 2, in which the

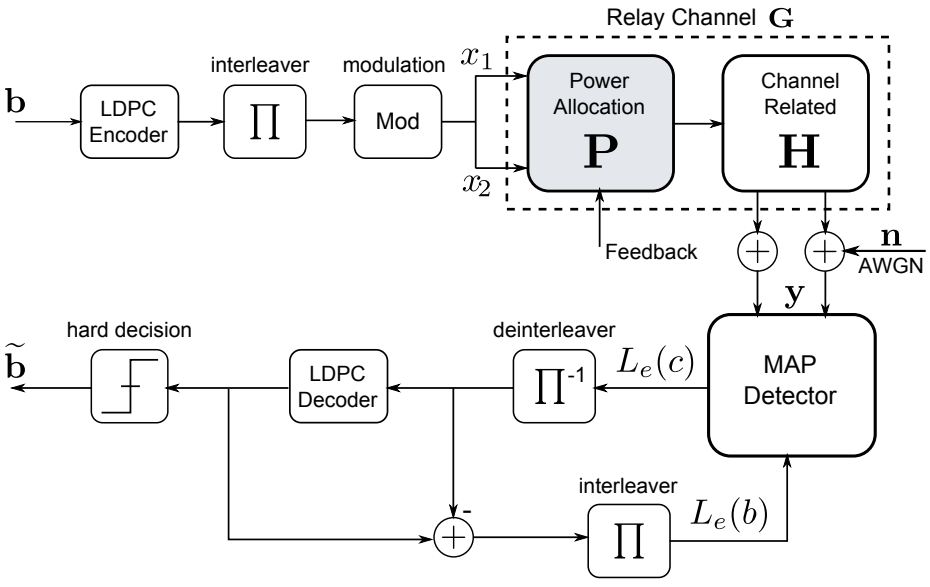


Figure 1. Block diagram of the relay network with iterative receiver at the destination.

information symbol $\mathbf{x}$ is modulated as quadrature phase shift keying (QPSK). From Fig. 2, we have several observations. First, the performance loss of the MMSE and network beamforming is small in the low SNR region, and large in the high SNR region. This is because both schemes maximize the power gain, which is much impressive compared to the degree-of-freedom gain at low SNR [23]. At high SNR, however, a degree-of-freedom gain is much more important, which is not provided by the original MMSE and network beamforming schemes [7, 10], since the source node transmits the same symbols for the two time slots. Hence, the mutual information is bounded by 1 bit/s/Hz. Second, the modified MMSE and modified network beamforming are not bounded by 1 bit/s/Hz, since we remove the constraint that sends the same symbols at the two time slots. Moreover, the proposed power allocation method results in significant gain on mutual information when SNR is in medium-to-high regions, or when the channel coding rate is medium to high. For example, it is about 4dB and 8dB better than the modified network beamforming and modified MMSE scheme when the channel coding rate is 3/4.

In Fig. 3, we show that the large performance gain on the mutual information will also represent the large performance gain on the BER. To validate the advantage of the power allocation scheme, we realize the simulation model illustrated in Sec. 3.3. The coding length is 1800; the coding rate is 3/4; and the iteration between the MAP detector and the LDPC decoder is 5. Then we compare the optimal power allocation with modified MMSE and modified network beamforming. It is worth noting that the benefit of the power allocation scheme predicted by the mutual information can indeed be realized, and the power allocation coefficients that are "blessed" by the mutual information formula (10) can provide considerable performance gain in a practical system.

Then we work on the Rayleigh fading channel, and consider the average mutual information achieved by different methods. We assume the channels of S-R

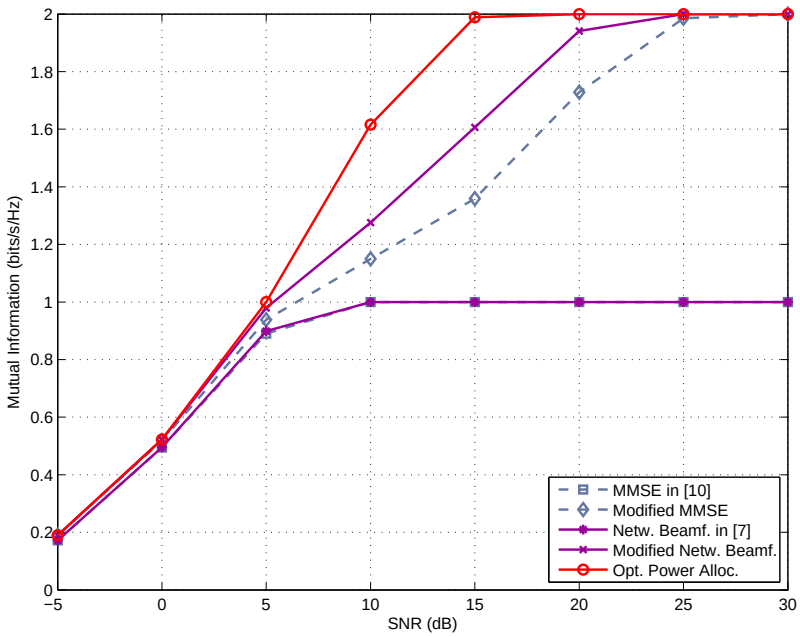


Figure 2. Mutual Information for a fixed channel with QPSK inputs.

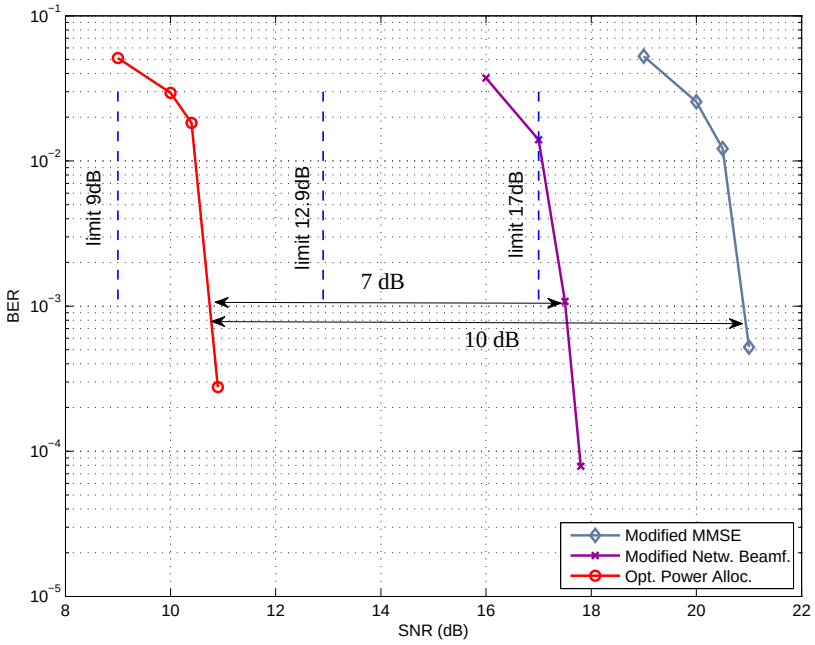


Figure 3. Bit error performance comparison for optimal power allocation, modified MMSE and modified network beamforming.

and R-D have the same average SNR, and are 10dB better than the S-D channel, considering the practical deployment of the relay nodes. Fig. 4 depicts the average mutual information of the relay network with QPSK inputs. The MMSE and network beamforming saturates very quickly, while the modified ones perform much better. However, they still have about 3dB to 15dB loss compared to the optimal power allocation when the channel coding rate is 3/4.

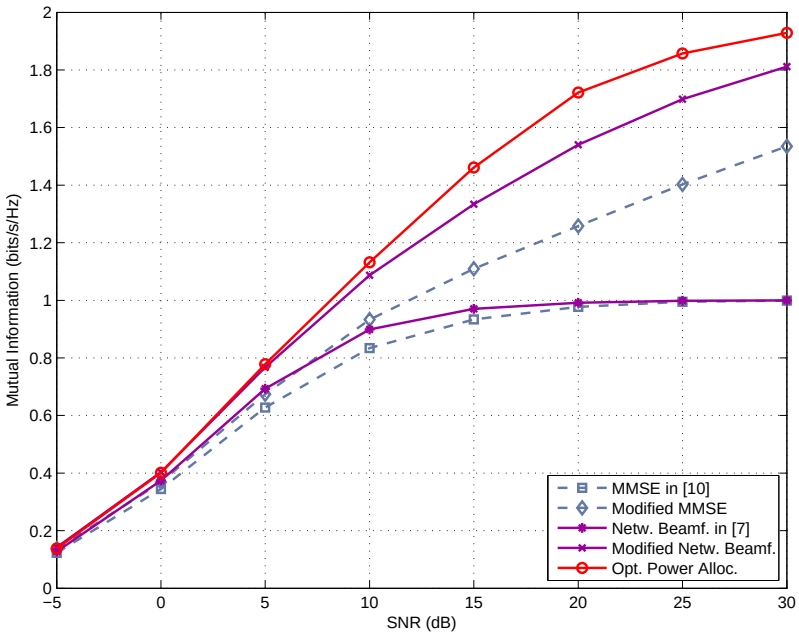


Figure 4. Average mutual Information over Rayleigh fading channels with QPSK inputs.

5 CONCLUSION

In this paper, we have studied the optimal power allocation for dual-hop wireless relay networks. In contrast with the previous work utilizing various design criteria with the unrealistic Gaussian inputs assumption, the proposed scheme attempts to maximize the mutual information for the relay networks from the standpoint of discrete-constellation inputs. To determine the optimal power allocation policy, we capitalized on the relationship between mutual information and MMSE. Numerical examples have shown that significant gains can be obtained compared to the conventional counterpart for nonfading channels and fading channels, especially when SNR is in medium-to-high regions, or when the channel coding rate is medium to high. Likewise, it has been shown that the large performance gain on the mutual information can represent the large gain on the bit-error rate, i.e., the benefit of the power allocation scheme predicted by the mutual information can indeed be realized and can provide considerable performance gain in a practical system.

6 REFERENCES

[1] Y. Hong, W. Huang, F. Chiu, and C. Kuo, "Cooperative communications in resource-constrained wireless networks," *IEEE Signal Process Mag.*, vol. 24, no. 3, p. 47, 2007.

[2] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[3] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, p. 3524, 2006.

[4] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 659–672, March 2006.

[5] W. Zeng, C. Xiao, Y. Wang, and J. Lu, "Opportunistic relaying for multi-antenna cooperative decode-and-forward relay networks," in *Proc. IEEE ICC*, 2010.

[6] A. S. Behbahani, R. Merched, and A. M. Eltawil, "Optimizations of a MIMO relay network," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5062–5073, 2008.

[7] Y. Jing and H. Jafarkhani, "Network beamforming using relays with perfect channel information," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2499–2517, Jun. 2009.

[8] G. Zheng, K. Wong, A. Paulraj, and B. Ottersten, "Collaborative-relay beamforming with perfect CSI: Optimum and distributed implementation," *IEEE Signal Process Lett.*, vol. 16, no. 4, pp. 257–260, 2009.

[9] A. Gershman, N. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming: From receive to transmit and network designs," *IEEE Signal Process Mag., Special Issue on Convex Optim. for Signal Process*, 2010.

[10] N. Khajehnouri and A. Sayed, "Distributed MMSE relay strategies for wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 55, no. 7, p. 3336, 2007.

[11] R. Nabar, H. Bolcskei, and F. Kneubuhler, "Fading relay channels: Performance limits and space-time signal design," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, 2004.

[12] C. Xiao and Y. R. Zheng, "On the mutual information and power allocation for vector gaussian channels with finite discrete inputs," in *Proc. IEEE Globecom*, Nov. 2008, pp. 1–5.

[13] A. Lozano, A. Tulino, and S. Verdu, "Optimum power allocation for parallel gaussian channels with arbitrary input distributions," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3033–3051, 2006.

[14] G. Caire and K. Kumar, "Information theoretic foundations of adaptive coded modulation," *Proc. IEEE*, vol. 95, no. 12, pp. 2274–2298, 2007.

[15] T. Cover and J. Thomas, *Elements of information theory*, 2nd ed. Wiley, 2006.

[16] F. Pérez-Cruz, M. Rodrigues, and S. Verdú, "Optimal precoding for digital subscriber lines," in *Proc. IEEE ICC*, 2008.

[17] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[18] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, 2005.

[19] D. Palomar and S. Verdú, "Gradient of mutual information in linear vector gaussian channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 141–154, 2006.

[20] J. Magnus and H. Neudecker, *Matrix differential calculus with applications in statistics and econometrics*, 2nd ed. Wiley, 1999.

[21] B. Hochwald and S. Ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, 2003.

[22] R. Koetter, A. Singer, and M. Tuchler, "Turbo equalization," *IEEE Signal Process Mag.*, vol. 21, no. 1, pp. 67–80, 2004.

[23] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

III. LINEAR PRECODING FOR RELAY NETWORKS: A PERSPECTIVE ON FINITE-ALPHABET INPUTS

Weiliang Zeng, Yahong Rosa Zheng, Mingxi Wang, and Jianhua Lu

ABSTRACT—This paper considers the precoder design for dual-hop amplify-and-forward relay networks and formulates the design from the standpoint of finite-alphabet inputs. In particular, the mutual information is employed as the utility function, which, however, results in a nonlinear and nonconcave problem. This paper exploits the structure of the optimal precoder that maximizes the mutual information and develops a two-step algorithm based on convex optimization and optimization on the Stiefel manifold. By doing so, the proposed algorithm is insensitive to initial point selection and able to achieve a near global optimal precoder solution. Besides, it converges fast and offers high mutual information gain. These advantages are verified by numerical examples, which also show the large performance gain in mutual information also represents the large gain in the coded bit-error rate.

1 INTRODUCTION

Relaying technology is promising to provide reliable communication, high throughput, and broad coverage for wireless networks. These benefits can be achieved with judicious designs that exploit network configurations and/or relaying strategies (see [1, 2, 3, 4, 5, 6, 7, 8, 9] and references therein). The existing design methods for improving the end-to-end transmission rate and reliability of the relay networks may be categorized into two groups: diversity oriented designs and data-rate oriented designs. The diversity oriented designs try to obtain the highest diversity order, by reaching the steepest asymptotic slope of either the outage capacity or uncoded bit error rate (BER) versus signal-to-noise ratio (SNR) curve, but are not necessary to achieve the highest coding gain. Examples of such designs include the distributed space-time coding in [1, 2] and the precoder design for maximal diversity in [3, 4]. The data-rate oriented designs try to obtain the highest data rate of each end-to-end source-destination pair by maximizing the output SNR [5, 6] or maximizing the network capacity [5, 7, 8, 9] with Gaussian input assumption.

This paper considers two-hop relay networks employing simple relays that are equipped with single antenna and adopt the amplify-and-forward (AF) strategy [1]. It focuses on the design of precoder and the selection of relay node that maximize the mutual information of one source-destination pair. The channel model of such a relay network becomes very similar to that of a standard multiple-input multiple-output (MIMO) channel and the existing precoding methods developed for MIMO settings are applicable to the relay networks. However, most of the existing works in precoder design use extensively the Gaussian input assumption so that the mutual information between the transmit and receive signals is a simple and elegant function of the precoder and channel matrix and the optimization problem becomes easier to solve

than that of finite-alphabet inputs. A practical system usually utilizes finite-alphabet constellations, such as pulse amplitude modulation (PAM), phase shift keying (PSK) modulation, or quadrature amplitude modulation (QAM). The mutual information of the discrete-input systems depart significantly from that of the Gaussian inputs [10, 11, 12, 13]. Therefore, applying the precoder/relaying schemes designed for Gaussian inputs to discrete-input systems results in a significant performance loss from those designed directly for finite-alphabet inputs [14].

Several recent works target on the maximization of mutual information with finite-alphabet inputs and complex precoding matrices. Lozano *et. al.* [10] constrains the channel and precoding matrices to be diagonal such that the mutual information becomes a concave function of the squared precoder and the global maximal can be solved. The work in [15] discovers the linear relationship between the gradients of mutual information and the minimum mean squared error (MMSE) versus SNR curves and proposes a gradient descent (GD) method to solve for the precoder iteratively. The GD method is then further explored in [16]. However, as will be shown later in this paper, the GD method is very sensitive to initial point selection and can easily get stuck in a local maximal because the mutual information is a non-concave function of the precoder. More recently, the work in [13] derives the asymptotic mutual information expressions for relay networks in large-system regime with several parameters approaching infinity and provides necessary conditions that the optimal precoder satisfies. Hessian and concavity results are developed in [17] for real-valued channels, but with no precoder design. A possible design is presented in [18] that uses the analysis with real-valued channel assumption, without rigorous proof, in the complex-valued channel case. The work in [19] proposes a two-layer iterative algorithm that finds the precoding matrix by iterations between the bottom layer (*i.e.*, the precoding matrix) and top layer (*i.e.*, the compound channel-precoding matrix) using the concavity of

the mutual information on the compound channel-precoding matrix. The global optimal solutions can be achieved for some combinations of discrete constellations and MIMO configurations.

This paper proposes a novel two-step iterative algorithm and a new framework for linear precoder design with finite-alphabet inputs by exploring the structure of the precoder that maximizes the mutual information. The proposed method first separates precoder and channel matrices, by the singular value decomposition (SVD), into product of the left singular vectors, diagonal power allocation matrix, and right singular vectors. Then making use of the results that the left singular vectors of the optimal precoder coincide with the right singular vectors of the effective channel matrix [20, Appendix 3.B] and that the mutual information is a concave function on the squared singular values of the linear precoder, the proposed algorithm maximizes the mutual information by first designing the power allocation matrix with a given set of the right singular vectors then optimizing the right singular vectors with the obtained power allocation matrix. The algorithm iterates through the two steps until the mutual information is maximized.

The success of the two-step iteration is based on the result that the mutual information is a concave function on the squared singular values of the linear precoder for a complex-valued channel. Although optimizing the right singular vectors is extremely difficult even for real-valued channel and precoder matrices [21], we reformulate the complex-valued problem on the complex Stiefel manifold and solve it by the gradient method with projection. The proposed two-step algorithm is applied to a two-hop relay network and the maximization of the end-to-end mutual information with multiple relay nodes is also considered by relay selection. Our simulation examples achieve 68% and 38% of mutual information improvement over no precoding for

Binary PSK (BPSK) and Quadrature PSK (QPSK) systems, respectively, and similar performance gain is expected to be achievable when applied to standard MIMO channels.

The proposed precoder design algorithm has several advantages over the exiting works. First, the proposed algorithm is applicable to general AF relay networks and complex-valued MIMO systems with arbitrary combinations of constellations and antenna configurations. It contains the *unitary* matrix based precoder design method [3, 4]) as its special cases. It can also handle the special combinations of the discrete constellations and MIMO configurations addressed in [19] and achieve almost the same performance. Second, the proposed algorithm is insensitive to initial point selection and it converges much faster than the GD method. Third, the proposed algorithm, although suboptimal, can reach most of the maximal capacity predicted by Gaussian inputs at high probability when using finite-alphabet inputs and random initial points.

The remainder of this paper is organized as follows: Section II introduces the system model for the two-hop relaying scheme. The properties of mutual information are addressed in Section III. Section IV proposes the precoding scheme to maximize the mutual information, which includes the design of the left singular vectors, the power allocation matrix by convex optimization, and the right singular vectors using optimization on the complex Stiefel manifold. Section V presents several numerical examples to demonstrate the performance gain of this scheme over the existing ones. Finally, Section VI offers conclusions.

Notation: Real and complex spaces are denoted by $\mathbb{R}$ and $\mathbb{C}$, respectively. Boldface uppercase letters denote matrices, boldface lowercase letters denote column vectors, and italics denote scalars. The superscripts $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$, and $(\cdot)^+$ stand for transpose, complex conjugate, Hermitian, and Moore-Penrose pseudoinverse operations, respectively. The scalar with subscript c_i denotes the i-th element of vector

$\mathbf{c}$, whereas $[\mathbf{A}]_{i,j}$ and $[\mathbf{A}]_{:,j}$ denote the (i,j)-th element and j-th column of matrix $\mathbf{A}$, respectively; $\mathbf{Diag}(\mathbf{a})$ represent a diagonal matrix whose nonzero elements are given by the elements of vector $\mathbf{a}$; $\mathbf{vec}(\mathbf{A})$ represents the vector obtained by stacking the columns of $\mathbf{A}$; $\mathbf{I}$ and $\mathbf{0}$ represent identity matrix and zero matrix of appropriate dimensions, respectively. The Kronecker matrix product is represented by $\mathbf{A} \otimes \mathbf{B}$; $\mathrm{Tr}(\mathbf{A})$ denotes the trace operation; $\mathbb{E}$ denotes statistical expectation; log and ln are used, respectively, for the base two logarithm and natural logarithm.

2 SYSTEM MODEL

Consider a relay network with one transmit-and-receive pair, where the source node communicates to the destination node with the assistance of k relays, $r_1, r_2, \cdots, r_k$. All nodes are equipped with a single antenna and operate in the half-duplex mode. The relaying transmission system is assumed to be flat fading. The channel gain from the source to the destination is denoted by h_0, and those from the source to the i-th relay and from the i-th relay to the destination are denoted as h_i and g_i, respectively. These channel gains are assumed to remain unchanged during a period of observation. The relay system adopts the two-hop AF protocols [7] combined with single-relay selection [22]. The signals are transmitted in blocks with block length $2L$, where $L \geq 1$. The selected relay node receives in the first period of length L and transmits in the second period of length L.

The original signal at the source node is denoted by $\mathbf{x} = \begin{bmatrix} \mathbf{x}_a^T & \mathbf{x}_b^T \end{bmatrix}^T$, where $\mathbf{x}_a = [x_1, \cdots, x_L]^T$ and $\mathbf{x}_b = [x_{L+1}, \cdots, x_{2L}]^T$ with x_l being the symbol at the l-th time slot for $l = 1, \cdots, 2L$. It is assumed to be equally probable from a discrete constellation set, such as PSK, PAM, or QAM, with a unit covariance matrix, *i.e.*, $\mathbb{E}\left[\mathbf{x}\mathbf{x}^H\right] = \mathbf{I}$.

The original signal is processed by a precoding matrix before being transmitted from the source node. The precoded data $\mathbf{s} = \begin{bmatrix} \mathbf{s}_a^T & \mathbf{s}_b^T \end{bmatrix}^T$ can be written as

$$\mathbf{s} = \begin{bmatrix} \mathbf{s}_a \\ \\ \mathbf{s}_b \end{bmatrix} = \mathbf{P} \begin{bmatrix} \mathbf{x}_a \\ \\ \mathbf{x}_b \end{bmatrix} \tag{1}$$

where $\mathbf{P} \in \mathbb{C}^{2L \times 2L}$ is a matrix to be designed to improve the end-to-end performance.

The source node transmits the signal $\sqrt{P_s}\mathbf{s}_a$ with power P_s during the first L time slots. Let $\mathbf{y}_i$ and $\mathbf{y}_a$ be the received signals at the i-th relay node and the

destination, respectively. We have

$$\mathbf{y}_i = \sqrt{P_s} h_i \mathbf{s}_a + \mathbf{n}_i \tag{2}$$

$$\mathbf{y}_a = \sqrt{P_s} h_0 \mathbf{s}_a + \mathbf{n}_a \tag{3}$$

where $\mathbf{n}_i$ and $\mathbf{n}_a$ denote the independent and identically distributed (i.i.d.) zero-mean circularly Gaussian noise vector with unit variance at the i-th relay and the destination, respectively.

Assume that the i-th relay node is selected for the information forwarding in the second time slot and that the relay knows only the second-order statistics of h_i. The selected relay node scales the received signal by a factor b, which guarantees that the average transmit power of the i-th relay $b^2 \mathrm{Tr}\left(\mathbb{E}\left[\mathbf{y}_i \mathbf{y}_i^H\right]/L\right)$ is less than or equal to the power constraint P_r. If the channel gains are assumed to have unit variance, then b can be chosen as $\sqrt{P_r/\left(1+2P_s\right)}$.

At the same time, the source node sends the signal $\sqrt{P_s}\mathbf{s}_b$. Hence, the destination node receives the superposition of the relay transmission and the source transmission during the second time slot:

$$\mathbf{y}_b = b g_i \mathbf{y}_i + \sqrt{P_s} h_0 \mathbf{s}_b + \mathbf{n}_b = \sqrt{P_s} b h_i g_i \mathbf{s}_a + \sqrt{P_s} h_0 \mathbf{s}_b + \mathbf{n}_e \tag{4}$$

where $\mathbf{n}_b$ denotes the noise vector of the destination at the second time slot, and $\mathbf{n}_e$ denotes the effective end-to-end noise with complex Gaussian distribution $\mathcal{CN}\left(\mathbf{0}, N_d \mathbf{I}\right)$ and $N_d = 1 + b^2 |g_i|^2$.

For convenience of presentation, $\mathbf{y}_b$ is normalized by a factor $w = 1/\sqrt{N_d}$, and the received signal vector for the two time slots is denoted as $\mathbf{y} = \left[\mathbf{y}_a^T \;\; w\mathbf{y}_b^T\right]^T$. Thus, the effective input-output relationship for the two-hop transmission with precoding

is summarized as

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n} = \mathbf{H}\mathbf{P}\mathbf{x} + \mathbf{n} \tag{5}$$

where $\mathbf{x}$ is the original transmitted signal vector; $\mathbf{n} = \left[\mathbf{n}_a^T \;\; w\mathbf{n}_e^T\right]^T \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$; $\mathbf{H}$ is the effective channel matrix of the two-hop relay channel,

$$\mathbf{H} = \sqrt{P_s} \begin{bmatrix} h_0\mathbf{I} & \mathbf{0} \\ wbh_i g_i \mathbf{I} & wh_0\mathbf{I} \end{bmatrix} \tag{6}$$

which is full rank for any nonzero channel gain h_0.

The precoding matrix $\mathbf{P}$ is thus designed to maximize the mutual information with finite-alphabet inputs. Note that to effectively implement the proposed algorithm in practice, a low-rate feedback shall be allowed from the destination to the relays and to the source, respectively, for delivering information of node selection and precoding. The feedback will be discussed in Section 4.4.

3 MUTUAL INFORMATION FOR RELAY NETWORKS

Consider conventional equiprobable discrete constellations such as M-ary PSK, PAM, or QAM, where M is the number of points in the constellation set. The mutual information between the input $\mathbf{x}$ and the output $\mathbf{y}$, with the equivalent channel matrix $\mathbf{H}$ and the precoding matrix $\mathbf{P}$ known at the receiver, is $\mathcal{I}(\mathbf{x};\mathbf{y})$ given by [12]

$$\mathcal{I}(\mathbf{x};\mathbf{y}) = \log M - \frac{1}{2LM^{2L}} \sum_{m=1}^{M^{2L}} \mathbb{E}_{\mathbf{n}} \log \sum_{k=1}^{M^{2L}} \exp(-d_{mk}) \tag{7}$$

where d_{mk} is $\|\mathbf{HP}(\mathbf{x}_m - \mathbf{x}_k) + \mathbf{n}\|^2 - \|\mathbf{n}\|^2$ and $\|\cdot\|$ denotes the Euclidean norm of a vector. Both $\mathbf{x}_m$ and $\mathbf{x}_k$ contain $2L$ symbols, taken independently from the M-ary signal constellation.

The mutual information $\mathcal{I}(\mathbf{x};\mathbf{y})$ is fully determined by the distribution of $\|\mathbf{HP}(\mathbf{x}_m - \mathbf{x}_k) + \mathbf{n}\|^2$ for $m, k \in \{1, \cdots, M^{2L}\}$, which remain unchanged when a unitary transform $\mathbf{U}$ is applied on the output signal $\mathbf{y}$ because a unitary matrix is an *isometry* for the Euclidean norm. That is,

$$\mathcal{I}(\mathbf{x};\mathbf{y}) = \mathcal{I}(\mathbf{x};\mathbf{U}\mathbf{y}). \tag{8}$$

However, if the linear transform is applied on the input signal, then $\mathcal{I}(\mathbf{U}\mathbf{x};\mathbf{y})$ may change to a different value, even though the transmit power is not altered. That is,

$$\mathcal{I}(\mathbf{U}\mathbf{x};\mathbf{y}) \neq \mathcal{I}(\mathbf{x};\mathbf{y}). \tag{9}$$

Note that the property of mutual information for the discrete input vector is different from the case of Gaussian inputs. For Gaussian inputs, the mutual information $\mathcal{I}^{\mathcal{G}}(\mathbf{x};\mathbf{y})$ is unchanged when either the transmitted signal $\mathbf{x}$ or the received

signal $\mathbf{y}$ is rotated by a unitary matrix:

$$\mathcal{I}^{\mathcal{G}}(\mathbf{x};\mathbf{y}) = \mathcal{I}^{\mathcal{G}}(\mathbf{U}\mathbf{x};\mathbf{y}) = \mathcal{I}^{\mathcal{G}}(\mathbf{x};\mathbf{U}\mathbf{y}). \tag{10}$$

The case of finite inputs does not follow the same rule; thus, a new opportunity is available here to improve system performance by transforming the input signal.

The $2L \times 2L$ complex precoding matrix $\mathbf{P}$ implements linear transform on $\mathbf{x}$, thus can exploit property (9) to maximize the mutual information of relay systems with finite-alphabet inputs. The optimization problem is formulated as:

$$\begin{aligned} &\text{maximize} \quad \mathcal{I}(\mathbf{x};\mathbf{y}) \\ &\text{subject to} \quad \mathrm{Tr}\left\{\mathbb{E}\left[\mathbf{s}\mathbf{s}^H\right]\right\} = \mathrm{Tr}\left(\mathbf{P}\mathbf{P}^H\right) \leq 2L \end{aligned} \tag{11}$$

which is difficult to solve because it is nonconcave over $\mathbf{P}$ [21]. The next section provides a new algorithm for this problem.

4 PRECODER DESIGN TO MAXIMIZE MUTUAL INFORMATION

This section examines the properties of the precoding matrix $\mathbf{P}$ under finite-alphabet inputs and presents several new results for complex-valued relay channels. These results are the foundation for the development of the proposed algorithm, which maximizes the mutual information.

4.1 OPTIMAL LEFT SINGULAR VECTORS

The first step is to characterize the dependence of mutual information $\mathcal{I}(\mathbf{x};\mathbf{y})$ on the precoding matrix $\mathbf{P}$. Given the signal constellation and the SNR, $\mathcal{I}(\mathbf{x};\mathbf{y})$ is a function of the following variable:

$$\|\mathbf{HP}(\mathbf{x}_m - \mathbf{x}_k) + \mathbf{n}\|^2 = \mathrm{Tr}\Big[\hat{\mathbf{e}}_{mk}\hat{\mathbf{e}}_{mk}^H\mathbf{P}^H\mathbf{H}^H\mathbf{HP} + 2\Re\left(\hat{\mathbf{e}}_{mk}^H\mathbf{P}^H\mathbf{H}^H\mathbf{n}\right) + \mathbf{nn}^H\Big] \quad (12)$$

where $\hat{\mathbf{e}}_{mk} = \mathbf{x}_m - \mathbf{x}_k$, and $\Re$ denotes the real part of a complex number; $\mathcal{I}(\mathbf{x};\mathbf{y})$ changes based on the distribution of $\|\mathbf{HP}(\mathbf{x}_m - \mathbf{x}_k) + \mathbf{n}\|^2$, which depends on $\mathbf{P}$ through $\mathbf{P}^H\mathbf{H}^H\mathbf{HP}$: the first term of the right-hand side depends on $\mathbf{P}$ through $\mathbf{P}^H\mathbf{H}^H\mathbf{HP}$; the second term $\hat{\mathbf{e}}_{mk}^H\mathbf{P}^H\mathbf{H}^H\mathbf{n}$ is a Gaussian random variable determined by its zero mean and its variance $\hat{\mathbf{e}}_{mk}^H\mathbf{P}^H\mathbf{H}^H\mathbf{HP}\hat{\mathbf{e}}_{mk}$, which also depends on $\mathbf{P}$ through $\mathbf{P}^H\mathbf{H}^H\mathbf{HP}$; the last term is independent of the precoding matrix. Therefore, $\mathcal{I}(\mathbf{x};\mathbf{y})$ depends on $\mathbf{P}$ through $\mathbf{P}^H\mathbf{H}^H\mathbf{HP}$, which is called the compound channel-precoding matrix.

Consider the SVD of the $2L \times 2L$ channel matrix $\mathbf{H} = \mathbf{U_H}\mathbf{Diag}(\boldsymbol{\sigma})\mathbf{V_H}^H$, where $\mathbf{U_H}$ and $\mathbf{V_H}$ are unitary matrices, and the vector $\boldsymbol{\sigma}$ contains nonnegative entries in decreasing order. We also decompose the precoding matrix by SVD as

$\mathbf{P} = \mathbf{U_P}\mathbf{Diag}(\sqrt{\boldsymbol{\lambda}})\mathbf{V}_\mathbf{P}^H$, where the vector $\boldsymbol{\lambda}$ is nonnegative constrained by the transmit power. To simplify notation, we define $\mathbf{U} = \mathbf{U_P}$ and $\mathbf{V} = \mathbf{V}_\mathbf{P}^H$, where $\mathbf{U}$ and $\mathbf{V}$ are called left and right singular vectors of $\mathbf{P}$, respectively.

For a given matrix $\mathbf{P}^H\mathbf{H}^H\mathbf{HP}$ and a specific mutual information value, the structure of the precoding matrix can be used to minimize the transmit power $\mathrm{Tr}(\mathbf{PP}^H)$, as shown in Appendix 3.B of [20], by letting the left singular vectors of $\mathbf{P}$ coincide with the right singular vectors of $\mathbf{H}$, that is, $\mathbf{P} = \mathbf{V_H}\mathbf{Diag}(\sqrt{\boldsymbol{\lambda}})\mathbf{V}$. Similar results based on real-valued channels are reported in [21, 18]. In other words, for a given matrix $\mathbf{P}^H\mathbf{H}^H\mathbf{HP}$ and a specific power constraint, letting the left singular vectors of $\mathbf{P}$ equal to the right singular vectors of $\mathbf{H}$ maximizes the mutual information for general channel conditions and arbitrary inputs.

Adopting the optimal left singular vectors $\mathbf{U} = \mathbf{V_H}$, the channel matrix (5) can be further simplified by (8) to

$$\mathbf{y} = \mathbf{Diag}(\boldsymbol{\sigma})\mathbf{Diag}(\sqrt{\boldsymbol{\lambda}})\mathbf{Vx} + \mathbf{n}. \tag{13}$$

It is clear from (13) that the mutual information is now dependent only on the squared singular values of the precoder $\boldsymbol{\lambda}$ and $\mathbf{V}$. We use the notations $\mathcal{I}(\boldsymbol{\lambda})$ and $\mathcal{I}(\mathbf{V})$ to develop a two-step optimization algorithm that first maximizes the mutual information via optimal power allocation $\boldsymbol{\lambda}$, and then via optimal right singular vectors $\mathbf{V}$.

4.2 OPTIMAL POWER ALLOCATION

Given the right singular vectors of the precoder, the optimization problem (11) is addressed over $\boldsymbol{\lambda}$:

$$\begin{aligned} \text{maximize} \quad & \mathcal{I}(\boldsymbol{\lambda}) \\ \text{subject to} \quad & \mathrm{Tr}\left(\mathbf{PP}^H\right) = \mathbf{1}^T\boldsymbol{\lambda} \leq 2L \\ & \boldsymbol{\lambda} \succeq \mathbf{0} \end{aligned} \tag{14}$$

where $\mathbf{1}$ and $\mathbf{0}$ denote the column vector with all entries being one and zero, respectively.

We extend the Hessian and concavity results of real-valued case in [17, Theorem 5] to a complex-valued channel and precoder

Proposition 1. *Based on the simplified channel model in* (13)*, the mutual information is a concave function of* $\boldsymbol{\lambda}$*; that is, the Hessian of mutual information satisfies* $\mathcal{H}_{\boldsymbol{\lambda}}\mathcal{I}(\boldsymbol{\lambda}) \preceq \mathbf{0}$*. Moreover, the gradient and Hessian of the mutual information are given, respectively, as*

$$\nabla_{\boldsymbol{\lambda}}\mathcal{I}(\boldsymbol{\lambda}) = \mathbf{R}\cdot\mathbf{vec}\left(\mathbf{Diag}^2(\boldsymbol{\sigma})\mathbf{V}\mathbf{E}\mathbf{V}^H\right) \tag{15}$$

and

$$\begin{aligned}\mathcal{H}_{\boldsymbol{\lambda}}\mathcal{I}(\boldsymbol{\lambda}) = &-\frac{1}{2}\mathbf{R}\left[\mathbf{I}\otimes\mathbf{Diag}^2(\boldsymbol{\sigma})\right]\mathbb{E}\left\{\tilde{\boldsymbol{\Phi}}(\mathbf{y})^*\otimes\tilde{\boldsymbol{\Phi}}(\mathbf{y})\right\}\\ &\left[\mathbf{Diag}(\sqrt{\boldsymbol{\lambda}})\mathbf{Diag}^2(\boldsymbol{\sigma})\otimes\mathbf{Diag}(\sqrt{\boldsymbol{\lambda}})\right]\mathbf{R}^T\mathbf{Diag}^{-1}(\boldsymbol{\lambda})\\ &-\frac{1}{2}\mathbf{R}\left[\mathbf{I}\otimes\mathbf{Diag}^2(\boldsymbol{\sigma})\right]\mathbb{E}\left\{\tilde{\boldsymbol{\Psi}}(\mathbf{y})^*\otimes\tilde{\boldsymbol{\Psi}}(\mathbf{y})\right\}\\ &\left[\mathbf{Diag}(\sqrt{\boldsymbol{\lambda}})\otimes\mathbf{Diag}(\sqrt{\boldsymbol{\lambda}})\mathbf{Diag}^2(\boldsymbol{\sigma})\right]\mathbf{K}\mathbf{R}^T\mathbf{Diag}^{-1}(\boldsymbol{\lambda})\end{aligned} \tag{16}$$

where $\mathbf{E}$ *is the minimum mean square error (MMSE) matrix, defined as*

$$\mathbf{E}\triangleq\mathbb{E}\left\{\left[\mathbf{x}-\mathbb{E}(\mathbf{x}|\mathbf{y})\right]\left[\mathbf{x}-\mathbb{E}(\mathbf{x}|\mathbf{y})\right]^H\right\};$$

and $\tilde{\boldsymbol{\Phi}}(\mathbf{y}) = \mathbf{V}\boldsymbol{\Phi}(\mathbf{y})\mathbf{V}^H$ *and* $\tilde{\boldsymbol{\Psi}}(\mathbf{y}) = \mathbf{V}\boldsymbol{\Psi}(\mathbf{y})\mathbf{V}^T$ *with* $\boldsymbol{\Phi}(\mathbf{y})$ *and* $\boldsymbol{\Psi}(\mathbf{y})$ *being the MMSE matrix and companion MMSE matrix conditioned on a realization of the received signal* $\mathbf{y}$*, defined by*

$$\boldsymbol{\Phi}(\mathbf{y})\triangleq\mathbb{E}\left\{\left[\mathbf{x}-\mathbb{E}\{\mathbf{x}|\mathbf{y}\}\right]\left[\mathbf{x}-\mathbb{E}\{\mathbf{x}|\mathbf{y}\}\right]^H|\mathbf{y}\right\}$$

and

$$\mathbf{\Psi}(\mathbf{y}) \triangleq \mathbb{E}\left\{[\mathbf{x} - \mathbb{E}\{\mathbf{x}|\mathbf{y}\}]\,[\mathbf{x} - \mathbb{E}\{\mathbf{x}|\mathbf{y}\}]^T\,|\mathbf{y}\right\}$$

respectively. And $\mathbf{R} \in \mathbb{R}^{2L\times 4L^2}$ *is a reduction matrix with entries given by* $[\mathbf{R}]_{i,2L(j-1)+k} = \delta_{ijk}$.

Proof. See Appendix A. □

The concavity property in Proposition 1 ensures that a global optimal power allocation vector can be found given right singular vectors $\mathbf{V}$. In addition, the gradient and Hessian results in (15) and (16) permit either the steepest descent or Newton-type algorithms to solve for the global optimal power allocation vector. However, the existing general-purpose solvers for convex problem (*e.g.*, CVX [23]) fail to address this problem because of the complexity of the objective function $\mathcal{I}(\mathbf{x};\mathbf{y})$. Therefore, a specialized interior-point method is developed here.

The first step is to re-write problem (14), making the inequality constraints implicit in the objective function:

$$\text{minimize} \quad f(\boldsymbol{\lambda}) = -\mathcal{I}(\boldsymbol{\lambda}) + \sum_{i=1}^{2L} \phi(-\lambda_i) + \phi(\mathbf{1}^T\boldsymbol{\lambda} - 2L) \tag{17}$$

where $\phi(u)$ is the logarithmic barrier function approximating an indicator as whether constraints are violated:

$$\phi(u) = \begin{cases} -(1/t)\ln(-u), & u < 0 \\ +\infty, & u \geq 0 \end{cases}$$

in which the parameter $t > 0$ sets the accuracy of the approximation [24].

Based on Proposition 1, the gradient of the objective function (17) is written as

$$\nabla_{\boldsymbol{\lambda}} f(\boldsymbol{\lambda}) = -\mathbf{R} \cdot \mathbf{vec}\left(\mathbf{Diag}^2(\boldsymbol{\sigma})\mathbf{V}\mathbf{E}\mathbf{V}^H\right) - \frac{1}{t}\left(\mathbf{q} - \frac{1}{2L - \mathbf{1}^T\boldsymbol{\lambda}}\right) \tag{18}$$

where $q_i = 1/\lambda_i$ is the i-th element of vector $\mathbf{q}$. Thus the steepest descent direction is chosen as

$$\Delta\boldsymbol{\lambda} = -\nabla_{\boldsymbol{\lambda}} f(\boldsymbol{\lambda}).$$

Combining this search direction with the backtracking line search conditions [24], we establish Algorithm 1 for the optimal power allocation vector, which ensures the convergence because of concavity. The details of the algorithm are shown in Table 1.

Table 1. Algorithm 1: Optimization of power allocation vector

Step 1. Given a feasible vector $\boldsymbol{\lambda}$, $t := t^{(0)} > 0$, $\alpha > 1$, tolerance $\epsilon > 0$.
Step 2. Compute the gradient $\nabla_{\boldsymbol{\lambda}} f(\boldsymbol{\lambda})$ as (18) and the descent direction $\nabla_{\boldsymbol{\lambda}} f(\boldsymbol{\lambda})$.
Step 3. If $\|\Delta\boldsymbol{\lambda}\|^2$ is sufficiently small, then go to Step 6; else go to Step 4.
Step 4. Choose γ so that $f(\boldsymbol{\lambda} + \gamma\Delta\boldsymbol{\lambda}) < f(\boldsymbol{\lambda})$ by backtracking line search.
Step 5. Set $\boldsymbol{\lambda} := \boldsymbol{\lambda} + \gamma\Delta\boldsymbol{\lambda}$. Go to Step 2.
Step 6. Stop if $1/t < \epsilon$, else $t := \alpha t$, and go to step 2.

4.3 OPTIMIZATION OVER RIGHT SINGULAR VECTORS

Now we consider maximizing the mutual information over the right singular vectors $\mathbf{V}$ for a given $\boldsymbol{\lambda}$:

$$\begin{aligned} &\text{maximize} && \mathcal{I}(\mathbf{V}) \\ &\text{subject to} && \mathbf{V}^H\mathbf{V} = \mathbf{V}\,\mathbf{V}^H = \mathbf{I}. \end{aligned} \tag{19}$$

This unitary-matrix constrained problem can be formulated as an unconstrained optimization in a constrained search space:

$$\text{minimize} \quad g(\mathbf{V})$$

with domain restricted to the Stiefel manifold $\mathrm{St}(n)$ [25]:

$$\text{dom } g = \{\mathbf{V} \in \mathrm{St}(n)\}$$

and

$$\mathrm{St}(n) = \left\{\mathbf{V} \in \ \mathbb{C}^{n\times n} | \mathbf{V}^H\mathbf{V} = \mathbf{I}\right\}$$

where the function $g(\mathbf{V})$ is defined as $-\mathcal{I}(\mathbf{V})$. For each point $\mathbf{V} \in \mathrm{St}(n)$, the search direction $\Delta\mathbf{V}$ on the tangent space has been suggested in [26] to minimize the objective function,

$$\Delta\mathbf{V} = -\nabla_{\mathbf{V}} g(\mathbf{V}) = \nabla_{\mathbf{V}}\mathcal{I}(\mathbf{V}) - \mathbf{V}(\nabla_{\mathbf{V}}\mathcal{I}(\mathbf{V}))^H\mathbf{V} \tag{20}$$

where $\nabla_{\mathbf{V}}\mathcal{I}(\mathbf{V})$ is the gradient of mutual information with respect to $\mathbf{V}$, given by $\mathbf{Diag}^2(\boldsymbol{\sigma})\mathbf{Diag}(\boldsymbol{\lambda})\mathbf{VE}$.

Note that moving along the descent direction on the tangent space may cause the unitary property being lost. Therefore, it needs to be restored in each step via projection. For an arbitrary matrix $\mathbf{W} \in \ \mathbb{C}^{n\times n}$, its projection $\pi(\mathbf{W})$ on the Stiefel manifold is defined as the point closest to $\mathbf{W}$ in the Euclidean norm:

$$\pi(\mathbf{W}) = \arg\min_{\mathbf{Q}\in\mathrm{St}(n)} \|\mathbf{W} - \mathbf{Q}\|^2.$$

If the SVD of $\mathbf{W}$ is $\mathbf{W} = \mathbf{U}_\mathbf{W}\mathbf{\Sigma}\mathbf{V}_\mathbf{W}$, the projection can be expressed by $\mathbf{U}_\mathbf{W}\mathbf{V}_\mathbf{W}$ [27, Sec. 7.4.8].

Combining the search direction and the projection with the backtracking line search condition, Algorithm 2 in Table 2 is developed to maximize the mutual information over the right singular vectors $\mathbf{V}$.

Table 2. Algorithm 2: Optimization of right singular vectors

Step 1. Given a feasible $\mathbf{V} \in \mathbb{C}^{n\times n}$ such that $\mathbf{V}^H\mathbf{V} = \mathbf{I}$.
Step 2. Compute the descent direction $\Delta\mathbf{V}$ as (20). Set the step size $\gamma := 1$.
Step 3. If $\|\Delta\mathbf{V}\|^2 = \mathrm{Tr}\{(\Delta\mathbf{V})^H\Delta\mathbf{V}\}$ is small, then stop; else go to Step 4.
Step 4. Choose γ so that $g(\pi(\mathbf{V} + \gamma\Delta\mathbf{V})) < g(\mathbf{V})$ by backtracking line search.
Step 5. Set $\mathbf{V} := \pi(\mathbf{V} + \gamma\Delta\mathbf{V})$. Go to Step 2.

4.4 TWO-STEP APPROACH TO OPTIMIZE PRECODER

Table 3 shows a complete two-step approach named Algorithm 3, which maximizes the mutual information over a generalized precoding matrix $\mathbf{P}$ by combining Algorithm 1 and Algorithm 2.

Table 3. Algorithm 3: Two-step optimization algorithm

Step 1. Initialization: Set the left singular vectors of the precoder $\mathbf{U} := \mathbf{V_H}$. Specify a feasible $\boldsymbol{\lambda}$ and $\mathbf{V}$.
Step 2. Update power allocation vector. Run Algorithm 1 given $\mathbf{V}$.
Step 3. Update right singular vectors: Run Algorithm 2 given $\boldsymbol{\lambda}$.
Step 4. Repeat Step 2 and Step 3 until convergence.

The proposed algorithm converges to the globally optimal solution in the low SNR region because mutual information is maximized by optimizing $\boldsymbol{\lambda}$ [16]. For

medium to high SNR, the proposed method, theoretically, converges to a local maximum. However, extensive numerical examples show that different initial points have limited effect on the algorithm (see Sec. 5); that is, the two-step method achieves near global optimal performance.

The destination node applies Algorithm 3 to each relay node and calculates the corresponding achievable mutual information and precoder. Denote the mutual information of the i-th relay node as $\mathcal{I}_i$. The best relay node is then selected by

$$R_s = \arg \max_{i=1,\cdots,k} \mathcal{I}_i.$$

Next, the index of the selected node and the corresponding precoder are transmitted via a feedback channel from the destination to the relay and to the source node, respectively.

5 SIMULATION RESULTS

This section examines the efficacy of the linear precoder on relay networks by several examples. These examples consider a single-relay network with a block length $L = 1$ and channel coefficients $h_0 = 0.4$, $h_1 = 1.2$, and $g_1 = -0.9\jmath$. The same transmit power is assumed at the source and relay node (*i.e.*, $P_s = P_r = P$).

The proposed two-step iterative algorithm is first tested from different feasible starting points. A general 2×2 unitary matrix group can be expressed as [28]

$$\mathbf{V} = \underbrace{\begin{bmatrix} e^{\jmath\alpha_1} & 0 \\ 0 & e^{\jmath\alpha_2} \end{bmatrix}}_{\mathbf{V}_1} \cdot \underbrace{\begin{bmatrix} \cos\psi & e^{-\jmath\theta}\sin\psi \\ -e^{\jmath\theta}\sin\psi & \cos\psi \end{bmatrix}}_{\mathbf{V}_2}.$$

For the simplified channel model in (13), the mutual information remains unchanged under the rotation of the diagonal unitary matrix $\mathbf{V}_1$. Thus, only the structure of $\mathbf{V}_2$ is used to generate feasible right singular vectors $\mathbf{V}$.

5.1 MUTUAL INFORMATION PERFORMANCE

We first consider two different starting points: Case A: $\boldsymbol{\lambda} = [0.5; 0.5]$, $\psi = \pi/6$, and $\theta = \pi/4$; Case B: $\boldsymbol{\lambda} = [0.2; 0.8]$, $\psi = \pi/10$, and $\theta = \pi/10$. Figure 1 illustrates the convergence for BPSK inputs when the SNR is 0 dB. For comparison, the figure also shows the mutual information corresponding to the cases of no precoding, optimal precoder with Gaussian inputs [5, 8], and optimal precoding with the GD method [15, 16]. The GD method is influenced by its initial point selection and converges to different mutual information levels with different initial points. The proposed two-step iteration of Algorithm 3, however, is insensitive to the initial points and converges to almost the same value, which increases the mutual information to 0.52 bps/Hz or

68% over that of no precoding. It also approaches the upper bound that is achieved by Gaussian inputs. Moreover, the progress of the proposed method exhibits a staircase shape, where each stair is associated with either the shift between the optimizations of the power allocation vector and the right singular vector or the iteration for the parameter t within Algorithm 1.

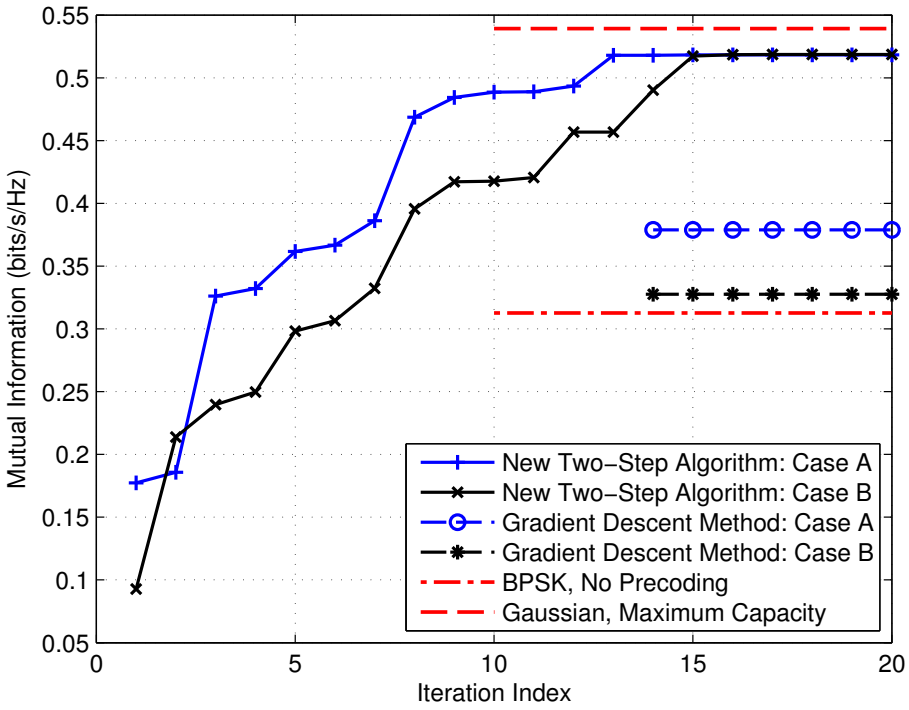


Figure 1. Evolution of the mutual information with BPSK inputs when SNR is 0 dB.

The cumulative distribution of the optimized mutual information from different initial points for the proposed two-step algorithm and the GD method are depicted in Fig. 2, which is obtained by maximizing the mutual information via 5,000 uniformly random initial points that are feasible to the considered problem. From Fig. 2, the two-step algorithm has a narrow and sharp curve, and the difference between the highest and lowest optimized mutual information is less than 0.003 bps/Hz; that

is, the iterative algorithm is able to obtain a near global optimal solution even if the problem is nonconcave. In contrast, the GD method depends highly on the initial point selection, and the difference between the highest and lowest optimized mutual information is as wide as 0.3 bps/Hz. It means the GD method may provide a solution even much worse than that of no precoding if the initial point is not chosen carefully.

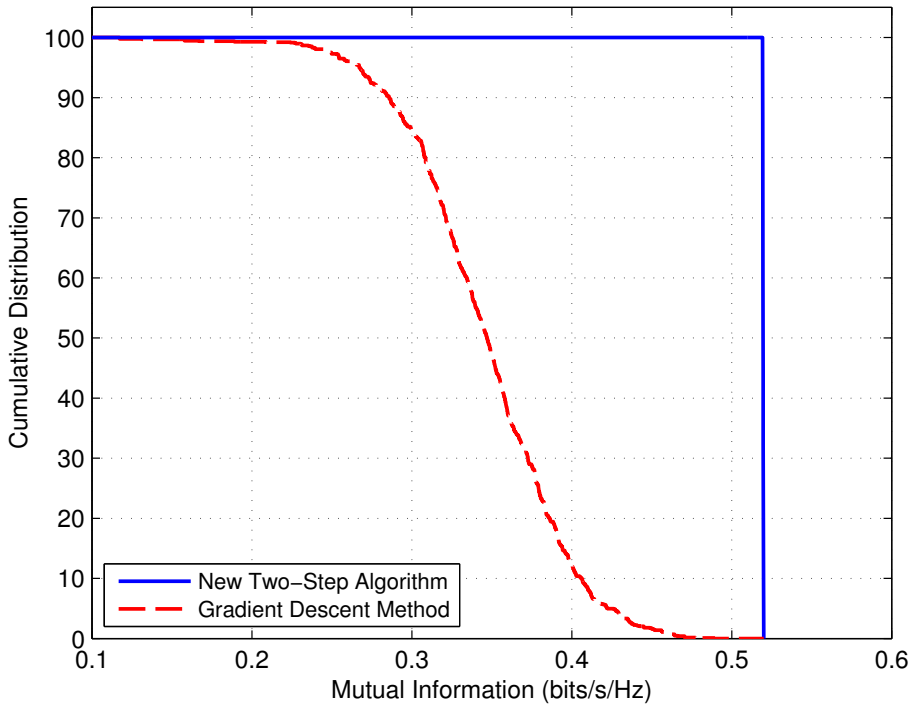


Figure 2. Cumulative distribution of the optimized mutual information with BPSK inputs when SNR is 0 dB.

Similar performance is also observed for QPSK with SNR given by 5 dB, as shown in Fig. 3 and 4. The convergence of the proposed two-step algorithm is achieved in 15 iterations, similar to the case of BPSK. It also achieves 38% improvement of mutual information over no precoding. From Fig. 4, the proposed algorithm

obtains mutual information within 98% of the maximum capacity of Gaussian inputs with 69% of the initial point selections, while the GD method reaches within 70% of the maximum capacity of Gaussian inputs with only 5% of the initial point selections.

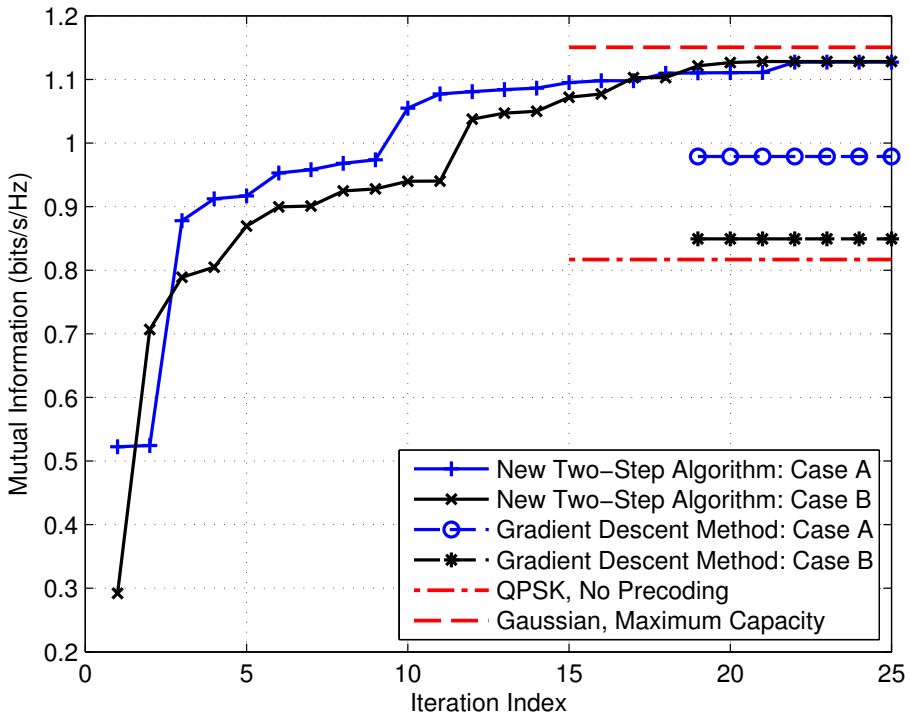


Figure 3. Evolution of the mutual information with QPSK inputs when SNR is 5 dB.

Figures 5 and 6 show the mutual information of the proposed algorithm versus SNR for BPSK and QPSK inputs, respectively, in comparison with other schemes such as no precoding, maximum diversity design in [3], maximum coding gain design in [3, 4], and maximum capacity design assuming Gaussian inputs in [5, 8]. Figures 5 and 6 suggest the following observations:

First, when the elements of the transmitted signal $\mathbf{x}$ are drawn from BPSK or QPSK, the mutual information for relay networks is bounded by 1 bps/Hz and 2 bps/Hz, respectively, which are achieved by all precoding schemes when SNR is high.

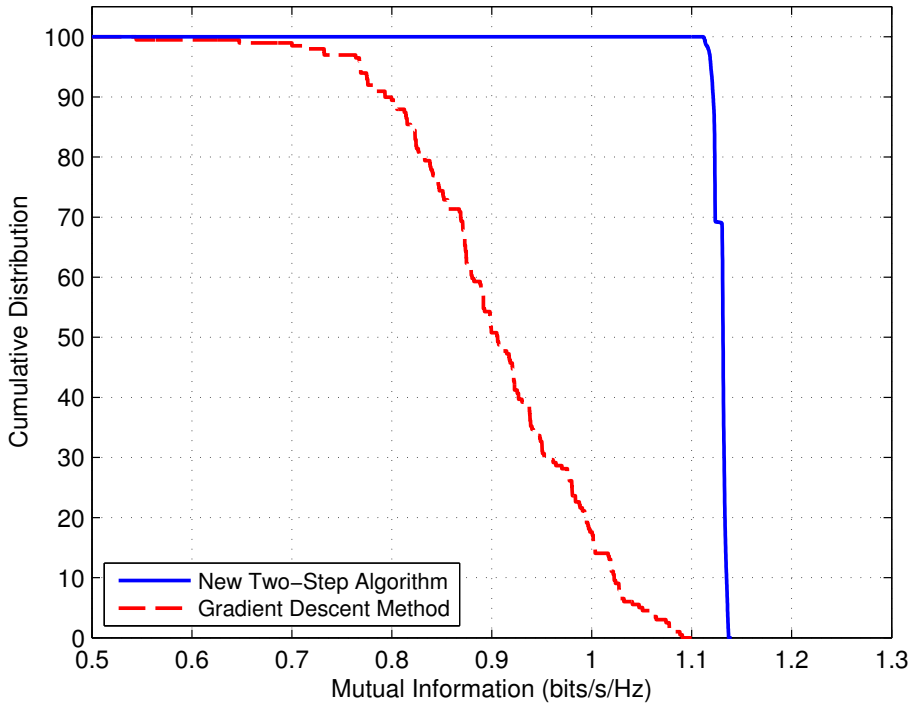


Figure 4. Cumulative distribution of mutual information with QPSK inputs when SNR is 5 dB.

Second, the precoder design based on maximizing capacity with the assumption of Gaussian inputs may result in a significant loss for systems employing discrete inputs. This loss comes from differences in designing the power allocation vector and the right singular vectors. For Gaussian inputs, allocating more power to the stronger subchannels and less to the weaker subchannels is helpful to maximize capacity. This strategy, however, does not work for finite-alphabet inputs, because the mutual information with finite inputs is bounded, therefore little incentive can be gained by allocating more power to the subchannels already close to saturation. Moreover, the right singular vectors for Gaussian-input systems is an arbitrary unitary matrix [see eq. (10)]. Systems with finite inputs, however, have to carefully select the right singular vectors to maximize the mutual information.

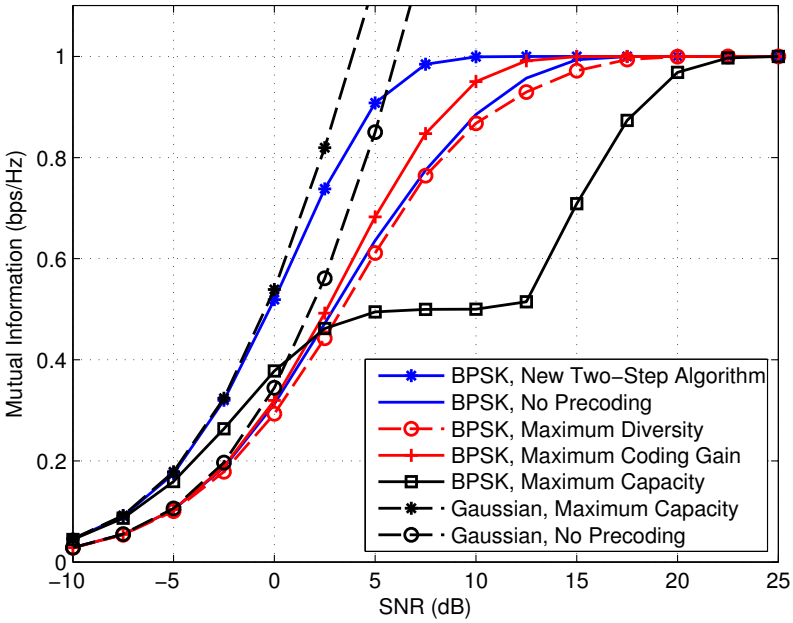


Figure 5. Mutual information versus the SNR with BPSK inputs.

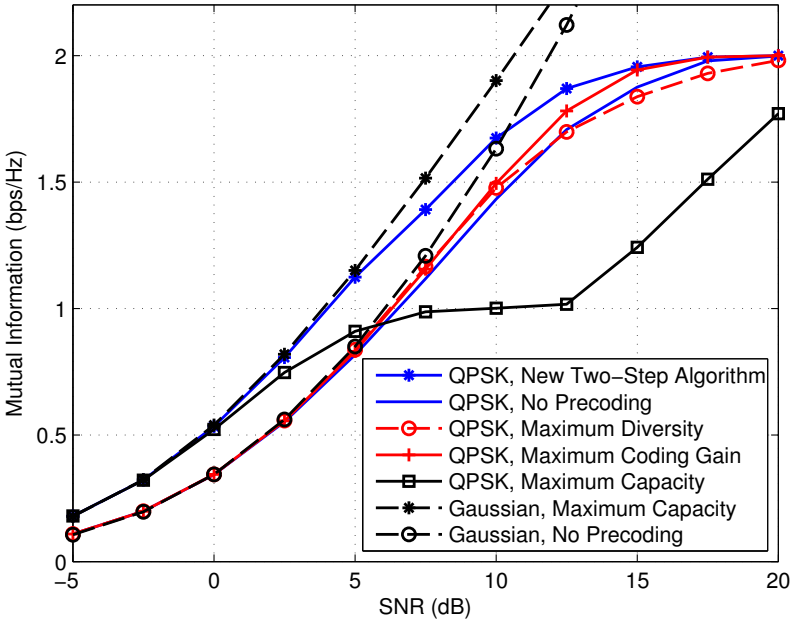


Figure 6. Mutual information versus the SNR with QPSK inputs.

Third, although the maximum coding gain method [3] performs better than the maximum diversity and no precoding methods, it is valid only for the case of block length $L = 1$ and QPSK inputs or the case of $L = 1$ and M-QAM inputs [4]. In comparison, the proposed two-step method can be used for an arbitrary block length L and input type. Our algorithm exploits the degrees of freedom in the optimal left singular vectors, the optimal power allocation vector, and the local optimal right singular vectors simultaneously. Therefore, providing significant gains of mutual information in a wide range of SNR. For example, with input BPSK and 3/4 channel coding rate, the performance of the proposed method is about 4 dB, 5.5 dB, and 6 dB better than those of the maximum coding gain, no precoding, and maximum capacity methods, respectively.

Last, the proposed method achieves mutual information very close to maximum capacity with Gaussian inputs when the channel coding rate is below 0.6 for both BPSK and QPSK; it also outperforms the case of Gaussian inputs with no precoding when the channel coding rate is below 0.9.

5.2 CODED BER PERFORMANCE

To evaluate the coded BER of the proposed method, channel coding is used at the source node and turbo principle [29, 30] is used at the destination node, as illustrated in Fig. 7. Note that the interleaver is not shown in the block diagram because of the usage of the low-density parity-check, or LDPC, codes [31]. The signal sequence $\mathbf{b}$ is encoded by the LDPC encoder and mapped according to the conventional equiprobable discrete signaling constellations. It is then divided into $\mathbf{x}_a$ and $\mathbf{x}_b$ and transmitted at the two time slots, respectively; the selected relay node simply amplifies and forwards the signal $\mathbf{x}_a$ in the second time slot.

At the receiver of the destination node, the maximum *a posteriori* (MAP) detector takes channel observations $\mathbf{y}$ and the *a priori* information $L_e(b)$ from the

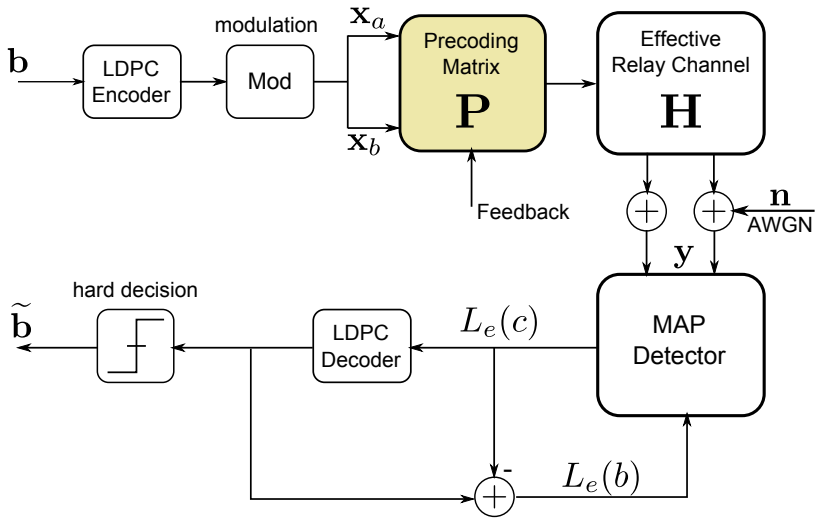


Figure 7. Block diagram of the relay network with precoding at the source and iterative receiver at the destination.

decoder and computes the extrinsic information $L_e(c)$ for each coded bit. Thus, the extrinsic information between the MAP detector and the decoder is exchanged iteratively until the desired performance is achieved [29].

In our simulation of the end-to-end system of Fig. 7, we use the block length for relaying transmission $L = 1$. The LDPC encoder and decoder modules are derived from [32] with coding rate 1/2. The channel coding length 2400, the coding rate 3/4 and 2/3 for BPSK and QPSK, respectively, and five iterations between the MAP detector and the LDPC decoder. Figures 8 and 9 show that the designed precoder that maximizes the mutual information provides large performance gains in the coded BER over other schemes including on precoding, the maximum coding gain, and maximum capacity methods. Note that the precoder designed by maximum capacity with Gaussian input offers worse BER performance than no precoding in practical BPSK and QPSK systems.

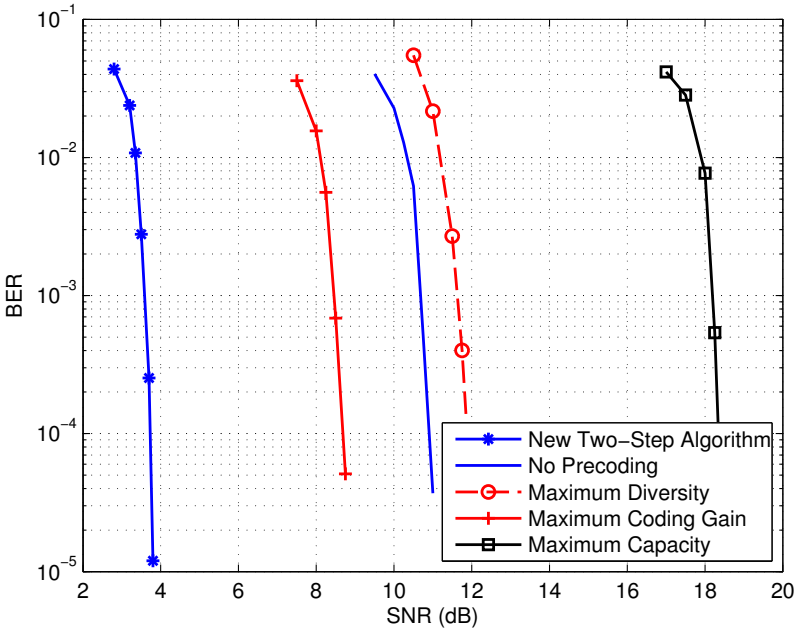


Figure 8. Bit error performance with BPSK inputs and 3/4 channel coding rate.

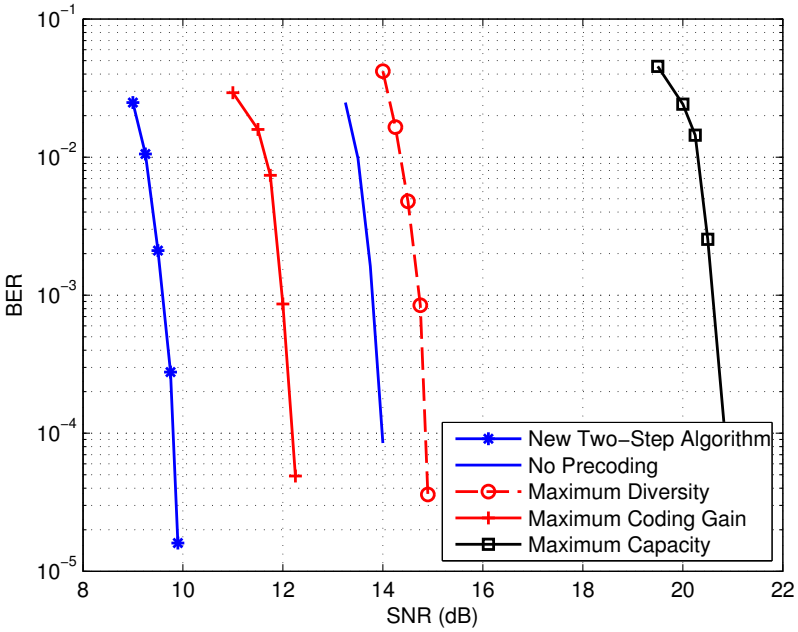


Figure 9. Bit error performance with QPSK inputs and 2/3 channel coding rate.

6 CONCLUSION

This paper has proposed a new two-step iterative algorithm for linear precoder design that maximizes mutual information with finite-alphabet inputs. By setting the left singular vectors of the precoding matrix equal to the right singular vectors of the channel matrix, the mutual information has proved to be a concave function on the squared singular values of the linear precoder for a complex-valued channel. Consequently, the proposed algorithm first optimizes the mutual information via the power allocation matrix or singular values of the linear precoder with a given set of right singular vectors, then solves the second maximization problem for right singular vectors with the obtained power allocation matrix, and iterates through the two steps until convergence. The second optimization problem is difficult and we have reformulated it on the complex Stiefel manifold and have solved it by the gradient method with projection. The proposed two-step linear precoder design algorithm has several advantages: being able to handle general complex-valued channels and system configurations, insensitive to initial point selection, fast convergence, and high mutual information gain. A numerical example of a two-hop relay network with amplify-and-forward strategy and Binary/Quadrature Phase Shift Keying (BPSK/QPSK) constellations is presented to demonstrate the performance gains of the proposed linear precoder algorithm.

7 APPENDIX A: PROOF OF PROPOSITION 1

To prove Proposition 1, we first introduce the definitions of gradient, Jacobian, and Hessian of a complex matrix [33].

7.1 DEFINITION

The gradient matrix with respect to a complex-valued matrix $\mathbf{Z}$ is defined as

$$\nabla_{\mathbf{Z}} f \triangleq \frac{\partial f}{\partial \mathbf{Z}^*}$$

where the (i,j)-th element of the gradient matrix is $[\nabla_{\mathbf{Z}} f]_{ij} = \partial f / \partial [\mathbf{Z}^*]_{ij}$.

Let $\mathbf{F}(\mathbf{Z}, \mathbf{Z}^*)$ be a complex matrix function of $\mathbf{Z}$ and $\mathbf{Z}^*$; the Jacobian matrices of $\mathbf{F}$ with respect to $\mathbf{Z}$ and $\mathbf{Z}^*$ are then given by:

$$\mathcal{D}_{\mathbf{Z}} \mathbf{F} \triangleq \frac{\partial \mathbf{vec}(\mathbf{F})}{\partial \mathbf{vec}^T(\mathbf{Z})} \qquad \text{and} \qquad \mathcal{D}_{\mathbf{Z}^*} \mathbf{F} \triangleq \frac{\partial \mathbf{vec}(\mathbf{F})}{\partial \mathbf{vec}^T(\mathbf{Z}^*)}. \tag{21}$$

Let $\mathbf{Z}_1$ and $\mathbf{Z}_2$ be two complex-valued matrices, and let f be a real-valued scalar function of $\mathbf{Z}_1$ and $\mathbf{Z}_2$. The complex Hessian matrix of f with respect to $\mathbf{Z}_1$ and $\mathbf{Z}_2$ is defined by:

$$\mathcal{H}_{\mathbf{Z}_1, \mathbf{Z}_2} f \triangleq \mathcal{D}_{\mathbf{Z}_1} \left(\mathcal{D}_{\mathbf{Z}_2} f \right) = \frac{\partial}{\partial \mathbf{vec}^T(\mathbf{Z}_1)} \left[\frac{\partial f}{\partial \mathbf{vec}^T(\mathbf{Z}_2)} \right]^T. \tag{22}$$

7.2 JACOBIAN OF THE MMSE MATRIX

The gradient of mutual information as the first-order characteristic has been derived in [15]:

$$\nabla_{\mathbf{P}} \mathcal{I}(\mathbf{x}; \mathbf{y}) = \mathbf{H}^H \mathbf{H} \mathbf{P} \mathbf{E} \tag{23}$$

where $\mathbf{E}$ is the MMSE matrix defined by

$$\mathbf{E} \triangleq \mathbb{E}\left\{ \left[\mathbf{x} - \mathbb{E}\left(\mathbf{x}|\mathbf{y}\right)\right] \left[\mathbf{x} - \mathbb{E}\left(\mathbf{x}|\mathbf{y}\right)\right]^H \right\} \tag{24}$$

The MMSE estimate of $\mathbf{x}$ (the conditional mean) is

$$\mathbb{E}\left(\mathbf{x}|\mathbf{y}\right) = \sum_{l=1}^{M^{2L}} \mathbf{x}_l p(\mathbf{x}_l|\mathbf{y}) = \mathbb{E}_{\mathbf{x}}\left[\mathbf{x}\frac{p(\mathbf{y}|\mathbf{x})}{p(\mathbf{y})}\right] \tag{25}$$

and the conditional probability density function of the received signal $\mathbf{y}$ is calculated as

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{\pi^{2L}} \exp\left(-\|\mathbf{y} - \mathbf{HPx}\|^2\right). \tag{26}$$

The Jacobian of the MMSE matrix $\mathcal{D}_{\mathbf{P}^*}\mathbf{E}$ is now introduced as a Lemma for deriving the second-order characteristic function of the mutual information. Note that Lemma 2 hold for the general complex-valued channel case, which agrees with [17, Theorem 3] for real-valued signal model.

Lemma 2. *The Jacobian of the MMSE matrix* $\mathbf{E}$ *with respect to* $\mathbf{P}^*$ *is given by*

$$\begin{aligned} \mathcal{D}_{\mathbf{P}^*}\mathbf{E} = &-\mathbb{E}\left\{\mathbf{\Phi}^*(\mathbf{y}) \otimes \mathbf{\Phi}(\mathbf{y})\right\} \mathbf{K} \left[\mathbf{I} \otimes \left(\mathbf{P}^T\mathbf{H}^T\mathbf{H}^*\right)\right] \\ &- \mathbb{E}\left\{\mathbf{\Psi}^*(\mathbf{y}) \otimes \mathbf{\Psi}(\mathbf{y})\right\} \left[\mathbf{I} \otimes \left(\mathbf{P}^T\mathbf{H}^T\mathbf{H}^*\right)\right] \end{aligned} \tag{27}$$

where $\mathbf{K}$ *is a unique* $4L^2 \times 4L^2$ *permutation matrix with* $\mathbf{vec}\left(\mathbf{\Phi}^T(\mathbf{y})\right) = \mathbf{K} \cdot \mathbf{vec}\left(\mathbf{\Phi}(\mathbf{y})\right)$.

Proof. The (i,j)-th element of the MMSE matrix $\mathbf{E}$ is given by $[\mathbf{E}]_{ij} = \mathbb{E}\left\{x_i x_j^*\right\} - \mathbb{E}\left\{\mathbb{E}\{x_i|\mathbf{y}\}\mathbb{E}\{x_j|\mathbf{y}\}\right\}$, from which it follows that

$$\begin{aligned}\frac{\partial[\mathbf{E}]_{ij}}{\partial[\mathbf{P}]_{kl}^*} = &-\int \frac{\partial p(\mathbf{y})}{\partial[\mathbf{P}]_{kl}^*}\mathbb{E}\{x_i|\mathbf{y}\}\mathbb{E}\{x_j^*|\mathbf{y}\}d\mathbf{y} \\ &- \int p(\mathbf{y})\frac{\partial\mathbb{E}\{x_i|\mathbf{y}\}}{\partial[\mathbf{P}]_{kl}^*}\mathbb{E}\{x_j^*|\mathbf{y}\}d\mathbf{y} \\ &- \int p(\mathbf{y})\mathbb{E}\{x_i|\mathbf{y}\}\frac{\partial\mathbb{E}\{x_j^*|\mathbf{y}\}}{\partial[\mathbf{P}]_{kl}^*}d\mathbf{y}. \end{aligned} \tag{28}$$

Employing $\mathcal{D}_{\mathbf{y}}\|\mathbf{y}-\mathbf{HPx}\|^2 = (\mathbf{y}-\mathbf{HPx})^H$ yields

$$\mathcal{D}_{\mathbf{y}}p(\mathbf{y}|\mathbf{x}) = -p(\mathbf{y}|\mathbf{x})(\mathbf{y}-\mathbf{HPx})^H \tag{29}$$

$$\mathcal{D}_{\mathbf{y}}p(\mathbf{y}) = \mathbb{E}\{\mathcal{D}_{\mathbf{y}}p(\mathbf{y}|\mathbf{x})\} = -\mathbb{E}\left\{p(\mathbf{y}|\mathbf{x})(\mathbf{y}-\mathbf{HPx})^H\right\}. \tag{30}$$

The gradient of probability density function $\nabla_{\mathbf{P}}p(\mathbf{y})$ can then be written as

$$\begin{aligned}\nabla_{\mathbf{P}}p(\mathbf{y}) &= \mathbb{E}\left[\nabla_{\mathbf{P}}p(\mathbf{y}|\mathbf{x})\right] \\ &= \mathbb{E}\left[p(\mathbf{y}|\mathbf{x})\mathbf{H}^H(\mathbf{y}-\mathbf{HPx})\mathbf{x}^H\right] \\ &= -\mathbf{H}^H\mathbb{E}\left\{[\mathcal{D}_{\mathbf{y}}p(\mathbf{y}|\mathbf{x})]^H\mathbf{x}^H\right\}\end{aligned} \tag{31}$$

and the gradient of conditional expectation $\nabla_{\mathbf{P}}\mathbb{E}\{x_i|\mathbf{y}\}$ becomes

$$\begin{aligned}\nabla_{\mathbf{P}}\mathbb{E}\{x_i|\mathbf{y}\} &= \nabla_{\mathbf{P}}\mathbb{E}\left\{x_i\frac{p(\mathbf{y}|\mathbf{x})}{p(\mathbf{y})}\right\} \\ &= -\mathbb{E}\left\{x_i\frac{\mathbf{H}^H[\mathcal{D}_{\mathbf{y}}p(\mathbf{y}|\mathbf{x})]^H\mathbf{x}^H}{p(\mathbf{y})}\right\} \\ &\quad + \mathbb{E}\left\{x_i\frac{p(\mathbf{y}|\mathbf{x})\mathbf{H}^H\mathbb{E}\left\{[\mathcal{D}_{\mathbf{y}}p(\mathbf{y}|\mathbf{x})]^H\mathbf{x}^H\right\}}{[p(\mathbf{y})]^2}\right\} \\ &= \frac{1}{p(\mathbf{y})}\mathbf{H}^H\Bigg[-\mathbb{E}\left\{x_i[\mathcal{D}_{\mathbf{y}}p(\mathbf{y}|\mathbf{x})]^H\mathbf{x}^H\right\} \\ &\quad + \mathbb{E}\{x_i|\mathbf{y}\}\mathbb{E}\left\{[\mathcal{D}_{\mathbf{y}}p(\mathbf{y}|\mathbf{x})]^H\mathbf{x}^H\right\}\Bigg].\end{aligned} \tag{32}$$

The Jacobian $\mathcal{D}_{\mathbf{y}}\mathbb{E}\{\mathbf{x}|\mathbf{y}\}$ is derived as follows:

$$\begin{aligned}
\mathcal{D}_{\mathbf{y}}\mathbb{E}\{\mathbf{x}|\mathbf{y}\} &= \mathcal{D}_{\mathbf{y}}\mathbb{E}\left\{\mathbf{x}\frac{p(\mathbf{y}|\mathbf{x})}{p(\mathbf{y})}\right\} \\
&= \mathbb{E}\left\{\mathbf{x}\frac{\mathcal{D}_{\mathbf{y}}p(\mathbf{y}|\mathbf{x})\cdot p(\mathbf{y}) - p(\mathbf{y}|\mathbf{x})\cdot\mathcal{D}_{\mathbf{y}}p(\mathbf{y})}{[p(\mathbf{y})]^2}\right\} \\
&= \mathbb{E}\left\{\mathbf{x}\frac{-p(\mathbf{y}|\mathbf{x})\cdot(\mathbf{y}-\mathbf{HPx})^H}{p(\mathbf{y})}\right\} + \mathbb{E}\left\{\mathbf{x}\frac{p(\mathbf{y}|\mathbf{x})\cdot(\mathbf{y}-\mathbf{HP}\mathbb{E}\{\mathbf{x}|\mathbf{y}\})^H}{p(\mathbf{y})}\right\} \\
&= \mathbb{E}\left\{\mathbf{x}\mathbf{x}^H\frac{p(\mathbf{y}|\mathbf{x})}{p(\mathbf{y})}\mathbf{P}^H\mathbf{H}^H\right\} - \mathbb{E}\left\{\mathbf{x}\frac{p(\mathbf{y}|\mathbf{x})}{p(\mathbf{y})}\mathbb{E}\{\mathbf{x}|\mathbf{y}\}^H\mathbf{P}^H\mathbf{H}^H\right\} \\
&= \left[\mathbb{E}\left\{\mathbf{x}\mathbf{x}^H|\mathbf{y}\right\} - \mathbb{E}\{\mathbf{x}|\mathbf{y}\}\mathbb{E}\{\mathbf{x}|\mathbf{y}\}^H\right]\mathbf{P}^H\mathbf{H}^H \\
&= \mathbf{\Phi}(\mathbf{y})\mathbf{P}^H\mathbf{H}^H. \qquad (33)
\end{aligned}$$

Following the similar steps in (33), the Jacobian can also be obtained:

$$\mathcal{D}_{\mathbf{y}^*}\mathbb{E}\{\mathbf{x}|\mathbf{y}\} = \mathbf{\Psi}(\mathbf{y})\mathbf{P}^H\mathbf{H}^H. \qquad (34)$$

Substituting (31) and (32) into (28) yields

$$\begin{aligned}
\frac{\partial[\mathbf{E}]_{ij}}{\partial[\mathbf{P}]^*_{kl}} = &-\int \mathbb{E}\{x_i|\mathbf{y}\}\mathbb{E}\{x_j^*|\mathbf{y}\}\mathbf{e}_k^T\mathbf{H}^H\mathbb{E}\left\{[\mathcal{D}_{\mathbf{y}}p(\mathbf{y}|\mathbf{x})]^H x_l^*\right\}d\mathbf{y} \\
&+\int \mathbf{e}_k^T\mathbf{H}^H\mathbb{E}\left\{x_i[\mathcal{D}_{\mathbf{y}}p(\mathbf{y}|\mathbf{x})]^H x_l^*\right\}\mathbb{E}\{x_j^*|\mathbf{y}\}d\mathbf{y} \\
&+\int \mathbb{E}\{x_i|\mathbf{y}\}\mathbf{e}_k^T\mathbf{H}^H\mathbb{E}\left\{x_j^*[\mathcal{D}_{\mathbf{y}}p(\mathbf{y}|\mathbf{x})]^H x_l^*\right\}d\mathbf{y} \qquad (35)
\end{aligned}$$

where $\mathbf{e}_k$ and $\mathbf{e}_l$ are, respectively, the k-th and l-th columns of an identity matrix with appropriate dimensions.

The first term in (35) can be rewritten as

$$
\begin{aligned}
&-\int \mathbb{E}\{x_i|\mathbf{y}\}\mathbb{E}\{x_j^*|\mathbf{y}\}\mathbf{e}_k^T\mathbf{H}^H\mathbb{E}\left\{[\mathcal{D}_{\mathbf{y}}p(\mathbf{y}|\mathbf{x})]^H x_l^*\right\}d\mathbf{y} \\
&\overset{(a)}{=} -\int \mathbb{E}\{x_i|\mathbf{y}\}\mathbb{E}\{x_j^*|\mathbf{y}\}\mathbf{e}_k^T\mathbf{H}^H\frac{\partial p(\mathbf{y})\mathbb{E}\{x_l^*|\mathbf{y}\}}{\partial \mathbf{vec}(\mathbf{y}^*)}d\mathbf{y} \\
&\overset{(b)}{=} \int p(\mathbf{y})\mathbb{E}\{x_l^*|\mathbf{y}\}\mathbf{e}_k^T\mathbf{H}^H\frac{\partial \mathbb{E}\{x_i|\mathbf{y}\}\mathbb{E}\{x_j^*|\mathbf{y}\}}{\partial \mathbf{vec}(\mathbf{y}^*)}d\mathbf{y} \\
&\overset{(c)}{=} \int p(\mathbf{y})\mathbb{E}\{x_l^*|\mathbf{y}\}\mathbb{E}\{x_j^*|\mathbf{y}\}\mathbf{e}_k^T\mathbf{H}^H\mathbf{H}\mathbf{P}\mathbf{\Psi}(\mathbf{y})\mathbf{e}_i d\mathbf{y} \\
&\quad + \int p(\mathbf{y})\mathbb{E}\{x_l^*|\mathbf{y}\}\mathbb{E}\{x_i|\mathbf{y}\}\mathbf{e}_k^T\mathbf{H}^H\mathbf{H}\mathbf{P}\mathbf{\Phi}(\mathbf{y})\mathbf{e}_j d\mathbf{y}
\end{aligned} \tag{36}
$$

where (a) follows from the definition of Jacobian matrix (21) and the fact $\mathbb{E}\{x_l^*|\mathbf{y}\} = \mathbb{E}\{x_l^* p(\mathbf{y}|\mathbf{x})/p(\mathbf{y})\}$; (b) is the result of integration by parts; (c) is from the result of (33) and (34).

Similarly, the second and third terms in (35) can be re-expressed, respectively, as

$$
\begin{aligned}
&\int \mathbf{e}_k^T\mathbf{H}^H\mathbb{E}\left\{x_i[\mathcal{D}_{\mathbf{y}}p(\mathbf{y}|\mathbf{x})]^H x_l^*\right\}\mathbb{E}\{x_j^*|\mathbf{y}\}d\mathbf{y} \\
&= -\int \mathbf{e}_k^T\mathbf{H}^H\mathbf{H}\mathbf{P}\mathbf{\Phi}(\mathbf{y})\mathbf{e}_j\mathbb{E}\{x_i x_l^*|\mathbf{y}\}p(\mathbf{y})d\mathbf{y}
\end{aligned} \tag{37}
$$

and

$$
\begin{aligned}
&\int \mathbb{E}\{x_i|\mathbf{y}\}\mathbf{e}_k^T\mathbf{H}^H\mathbb{E}\left\{x_j^*[\mathcal{D}_{\mathbf{y}}p(\mathbf{y}|\mathbf{x})]^H x_l^*\right\}d\mathbf{y} \\
&= -\int \mathbf{e}_k^T\mathbf{H}^H\mathbf{H}\mathbf{P}\mathbf{\Psi}(\mathbf{y})\mathbf{e}_j\mathbb{E}\{x_j^* x_l^*|\mathbf{y}\}p(\mathbf{y})d\mathbf{y}.
\end{aligned} \tag{38}
$$

Substituting (36), (38), and (37) into (35) yields

$$\begin{aligned}\frac{\partial[\mathbf{E}]_{ij}}{\partial[\mathbf{P}]^*_{kl}} &= -\int \mathbf{e}_k^T\mathbf{H}^H\mathbf{H}\mathbf{P}\mathbf{\Phi}(\mathbf{y})\mathbf{e}_j p(\mathbf{y})\left[\mathbb{E}\{x_i x_l^*|\mathbf{y}\} - \mathbb{E}\{x_i|\mathbf{y}\}\mathbb{E}\{x_l^*|\mathbf{y}\}\right]d\mathbf{y} \\ &\quad - \int \mathbf{e}_k^T\mathbf{H}^H\mathbf{H}\mathbf{P}\mathbf{\Psi}(\mathbf{y})\mathbf{e}_j p(\mathbf{y})\left[\mathbb{E}\{x_j^* x_l^*|\mathbf{y}\} - \mathbb{E}\{x_j^*|\mathbf{y}\}\mathbb{E}\{x_l^*|\mathbf{y}\}\right]d\mathbf{y} \\ &= -\mathbb{E}_{\mathbf{y}}\left\{\mathbf{e}_i^T\mathbf{\Phi}(\mathbf{y})\mathbf{e}_l\mathbf{e}_j^T\mathbf{\Phi}^T(\mathbf{y})\mathbf{G}^T\mathbf{H}^T\mathbf{H}^*\mathbf{e}_k\right\} \\ &\quad - \mathbb{E}_{\mathbf{y}}\left\{\mathbf{e}_j^T\mathbf{\Psi}^*(\mathbf{y})\mathbf{e}_l\mathbf{e}_i^T\mathbf{\Psi}(\mathbf{y})\mathbf{G}^T\mathbf{H}^T\mathbf{H}^*\mathbf{e}_k\right\} \\ &= -\mathbb{E}_{\mathbf{y}}\left\{\left[\mathbf{K}\big(\mathbf{\Phi}(\mathbf{y})\otimes\ (\mathbf{\Phi}^T(\mathbf{y})\mathbf{P}^T\mathbf{H}^T\mathbf{H}^*)\big)\right]_{i+(j-1)2L,k+(l-1)2L}\right\} \\ &\quad - \mathbb{E}_{\mathbf{y}}\left\{\left[\mathbf{\Psi}^*(\mathbf{y})\otimes\ \Big(\mathbf{\Psi}(\mathbf{y})\mathbf{P}^T\mathbf{H}^T\mathbf{H}^*\Big)\right]_{i+(j-1)2L,k+(l-1)2L}\right\} \end{aligned} \tag{39}$$

where $\mathbf{K}\in\mathbb{R}^{4L^2\times 4L^2}$ is a unique permutation matrix that satisfies [34]:

$$\mathbf{vec}\left(\mathbf{\Phi}^T(\mathbf{y})\right) = \mathbf{K}\cdot\mathbf{vec}\left(\mathbf{\Phi}(\mathbf{y})\right). \tag{40}$$

From $\partial[\mathbf{E}]_{ij}/\partial[\mathbf{P}]^*_{kl} = [\mathcal{D}_{\mathbf{P}^*}\mathbf{E}]_{i+(j-1)2L,k+(l-1)2L}$, we have

$$\begin{aligned}\mathcal{D}_{\mathbf{P}^*}\mathbf{E} &= -\mathbf{K}\mathbb{E}_{\mathbf{y}}\left\{\mathbf{\Phi}(\mathbf{y})\otimes\left[\mathbf{\Phi}^T(\mathbf{y})\mathbf{P}^T\mathbf{H}^T\mathbf{H}^*\right]\right\} \\ &\quad - \mathbb{E}_{\mathbf{y}}\left\{\mathbf{\Psi}^*(\mathbf{y})\otimes\left[\mathbf{\Psi}(\mathbf{y})\mathbf{P}^T\mathbf{H}^T\mathbf{H}^*\right]\right\} \\ &= -\mathbf{K}\mathbb{E}_{\mathbf{y}}\left\{\mathbf{\Phi}(\mathbf{y})\otimes\mathbf{\Phi}^*(\mathbf{y})\right\}\left[\mathbf{I}\otimes\mathbf{P}^T\mathbf{H}^T\mathbf{H}^*\right] \\ &\quad - \mathbb{E}_{\mathbf{y}}\left\{\mathbf{\Psi}^*(\mathbf{y})\otimes\mathbf{\Psi}(\mathbf{y})\right\}\left[\mathbf{I}\otimes\mathbf{P}^T\mathbf{H}^T\mathbf{H}^*\right]. \end{aligned} \tag{41}$$

Applying of the property of permutation matrix $\mathbf{K}$ [34], $\mathbf{K}\left(\mathbf{A}\otimes\mathbf{B}\right) = \left(\mathbf{B}\otimes\mathbf{A}\right)\mathbf{K}$, the proof is complete.

Proof of Proposition 1. Let us denote the remaining part of the precoder, $\mathbf{Diag}(\sqrt{\boldsymbol{\lambda}})\mathbf{V}$, as $\mathbf{G}$, which allows the power allocation vector $\boldsymbol{\lambda}$ to be rewritten as

$$\boldsymbol{\lambda} = \mathbf{R}\cdot\mathbf{vec}(\mathbf{G}\mathbf{G}^H). \tag{42}$$

According to (22), the Hessian of the mutual information can be obtained from

$$\mathcal{H}_{\boldsymbol{\lambda}}\mathcal{I}\left(\boldsymbol{\lambda}\right)=\mathcal{D}_{\boldsymbol{\lambda}}\left[\mathcal{D}_{\boldsymbol{\lambda}}\mathcal{I}\left(\boldsymbol{\lambda}\right)\right]. \tag{43}$$

Now $\mathcal{D}_{\boldsymbol{\lambda}}\mathcal{I}\left(\boldsymbol{\lambda}\right)$ can be calculated as

$$\begin{aligned}\mathcal{D}_{\boldsymbol{\lambda}}\mathcal{I}(\boldsymbol{\lambda})=\mathcal{D}_{\mathbf{G}\mathbf{G}^{H}}\mathcal{I}(\boldsymbol{\lambda})\cdot\mathbf{R}^{T}&\overset{(a)}{=}\mathbf{vec}^{T}\left(\mathbf{Diag}^{2}(\boldsymbol{\sigma})\mathbf{V}\mathbf{E}\mathbf{V}^{H}\right)\mathbf{R}^{T}\\&\overset{(b)}{=}\mathbf{vec}^{T}\left(\mathbf{E}\right)\left(\mathbf{V}^{H}\otimes\mathbf{V}^{T}\mathbf{Diag}^{2}(\boldsymbol{\sigma})\right)\mathbf{R}^{T}\end{aligned} \tag{44}$$

where (a) follows from the results in [15, Theorem 2], and (b) follows from the property [34]:

$$\mathbf{vec}\left(\mathbf{ATB}\right)=\left(\mathbf{B}^{T}\otimes\mathbf{A}\right)\mathbf{vec}(\mathbf{T}) \tag{45}$$

in which $\mathbf{A}$, $\mathbf{T}$, and $\mathbf{B}$ are matrices with appropriate dimensions. The result in (43) can then be written as

$$\mathcal{H}_{\boldsymbol{\lambda}}\mathcal{I}(\boldsymbol{\lambda})=\mathcal{D}_{\boldsymbol{\lambda}}\left[\mathbf{R}\left(\mathbf{V}^{*}\otimes\mathbf{Diag}^{2}(\boldsymbol{\sigma})\mathbf{V}\right)\mathbf{vec}\left(\mathbf{E}\right)\right]=\mathbf{R}\left(\mathbf{V}^{*}\otimes\mathbf{Diag}^{2}(\boldsymbol{\sigma})\mathbf{V}\right)\mathcal{D}_{\boldsymbol{\lambda}}\mathbf{E}. \tag{46}$$

From the chain rule of the Jacobian [33], it yields

$$\mathcal{D}_{\mathbf{G}^{*}}\mathbf{E}=\mathcal{D}_{\boldsymbol{\lambda}}\mathbf{E}\cdot\mathcal{D}_{\mathbf{G}^{*}}\boldsymbol{\lambda}+\mathcal{D}_{\boldsymbol{\lambda}^{*}}\mathbf{E}\cdot\mathcal{D}_{\mathbf{G}^{*}}\boldsymbol{\lambda}^{*}=2\mathcal{D}_{\boldsymbol{\lambda}}\mathbf{E}\cdot\mathcal{D}_{\mathbf{G}^{*}}\boldsymbol{\lambda} \tag{47}$$

where $\mathcal{D}_{\mathbf{G}^{*}}\boldsymbol{\lambda}$ can be derived from the definition of Jacobian (21):

$$\mathcal{D}_{\mathbf{G}^{*}}\boldsymbol{\lambda}=\mathcal{D}_{\mathbf{G}^{*}}\left[\mathbf{R}\cdot\mathbf{vec}\left(\mathbf{G}\mathbf{G}^{H}\right)\right]=\mathbf{R}\left(\mathbf{I}\otimes\mathbf{G}\right)\mathbf{K}. \tag{48}$$

Because $\mathcal{D}_{\mathbf{G}^*}\boldsymbol{\lambda}$ has full column rank, it is possible to invert the transformation in (47) to obtain

$$\mathcal{D}_{\boldsymbol{\lambda}}\mathbf{E} = \frac{1}{2}\left(\mathcal{D}_{\mathbf{G}^*}\mathbf{E}\right)\left(\mathcal{D}_{\mathbf{G}^*}\boldsymbol{\lambda}\right)^{+} = \frac{1}{2}\left(\mathcal{D}_{\mathbf{G}^*}\mathbf{E}\right)\cdot\left(\mathbf{G}^{H}\otimes\mathbf{I}\right)\mathbf{K}^{T}\mathbf{R}^{T}\mathbf{Diag}^{-1}\left(\boldsymbol{\lambda}\right). \tag{49}$$

Plugging (27) and (49) into (46) yields the Hessian result provided in (16), which can be readily identified as negative semi-definite.

8 REFERENCES

[1] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[2] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3524–3536, Dec. 2006.

[3] Y. Ding, J. Zhang, and K. Wong, "The amplify-and-forward half-duplex cooperative system: Pairwise error probability and precoder design," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 605–617, Feb. 2007.

[4] ——, "Optimal precoder for amplify-and-forward half-duplex relay system," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 2890–2895, Aug. 2008.

[5] A. S. Behbahani, R. Merched, and A. M. Eltawil, "Optimizations of a MIMO relay network," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5062–5073, Oct. 2008.

[6] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming: From receive to transmit and network designs," *IEEE Signal Process Mag.*, vol. 27, no. 3, pp. 62–75, Mar. 2010.

[7] R. Nabar, H. Bolcskei, and F. Kneubuhler, "Fading relay channels: Performance limits and space-time signal design," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.

[8] R. Mo and Y. Chew, "Precoder design for non-regenerative MIMO relay systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 5041–5049, Oct. 2009.

[9] W. Zeng, C. Xiao, Y. Wang, and J. Lu, "Opportunistic cooperation for multi-antenna multi-relay networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 10, pp. 3189–3199, Oct. 2010.

[10] A. Lozano, A. Tulino, and S. Verdú, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Trans. Inform. Theory*, vol. 52, no. 7, pp. 3033–3051, Jul. 2006.

[11] G. Caire and K. Kumar, "Information theoretic foundations of adaptive coded modulation," *Proc. IEEE*, vol. 95, no. 12, pp. 2274–2298, Dec. 2007.

[12] C. Xiao and Y. R. Zheng, "On the mutual information and power allocation for vector Gaussian channels with finite discrete inputs," in *Proc. IEEE Globecom*, New Orleans, LA, 2008.

[13] C. Wen, K. Wong, and C. Ng, "On the asymptotic properties of amplify-and-forward MIMO relay channels," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 590–602, Feb. 2009.

[14] W. Zeng, C. Xiao, M. Wang, and J. Lu, "Linear precoding for relay networks with finite-alphabet constraints," in *IEEE Int. Conf. Commun. (ICC)*, Kyoto, Japan, 2011.

[15] D. Palomar and S. Verdú, "Gradient of mutual information in linear vector Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 1, pp. 141–154, Jan. 2006.

[16] F. Pérez-Cruz, M. Rodrigues, and S. Verdú, "MIMO Gaussian channels with arbitrary inputs: Optimal precoding and power allocation," *IEEE Trans. Inform. Theory*, vol. 56, no. 3, pp. 1070 –1084, Mar. 2010.

[17] M. Payaró and D. P. Palomar, "Hessian and concavity of mutual information, differential entropy, and entropy power in linear vector Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 55, no. 8, pp. 3613–3628, Aug. 2009.

[18] M. Lamarca, "Linear precoding for mutual information maximization in MIMO systems," in *Proc. Int. Symp. on Wireless Commun. Syst. (ISWCS)*, 2009, pp. 26–30.

[19] 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, no. 7, pp. 3301–3314, Jul. 2011.

[20] D. Palomar and Y. Jiang, *MIMO Transceiver Design via Majorization Theory*. Now Publishers Inc., 2006.

[21] M. Payaró and D. P. Palomar, "On optimal precoding in linear vector Gaussian channels with arbitrary input distribution," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Seoul, Korea, 2009, pp. 1085–1089.

[22] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 659–672, Mar. 2006.

[23] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 1.21," retrieved Apr. 2011. [Online]. Available: `http://cvxr.com/cvx`

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

[25] P. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2008.

[26] J. H. Manton, "Optimization algorithms exploiting unitary constraints," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 635–650, Mar. 2002.

[27] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[28] F. Murnaghan, *Lectures on Applied Mathematics: The Unitary and Rotation Groups*. Washington, DC: Spartan Books, 1962, vol. 3.

[29] B. Hochwald and S. Ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.

[30] R. Koetter, A. Singer, and M. Tuchler, "Turbo equalization," *IEEE Signal Process Mag.*, vol. 21, no. 1, pp. 67–80, Jan. 2004.

[31] B. Lu, X. Wang, and K. Narayanan, "LDPC-based space-time coded OFDM systems over correlated fading channels: Performance analysis and receiver design," *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 74–88, Jan. 2002.

[32] M. Valenti, "Iterative solutions coded modulation library (ISCML)." [Online]. Available: `http://www.iterativesolutions.com/Matlab.htm`

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

[34] J. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, 3rd ed. John Wiley & Sons, 2007.

IV. PRACTICAL LINEAR PRECODER DESIGN FOR FINITE ALPHABET MIMO-OFDM WITH EXPERIMENT VALIDATION

Mingxi Wang, Yahong Rosa Zheng and Chengshan Xiao

ABSTRACT—A low complexity precoding method is proposed for practical Multiple-Input Multiple-Output (MIMO) Orthogonal Frequency-Division Multiplexing (OFDM) systems. Based on the two-step optimal precoder design algorithm that maximizes the lower bound of the mutual information with finite-alphabet inputs, the proposed method simplifies the precoder design by fixing the right singular vectors of the precoder matrix, eliminating the iterative optimization between the two steps, and discretizing the search space of the power allocation vector. For a 4×4 channel, the computational complexity of the proposed precoder design is reduced to 3% and 6% of that required by the original two-step algorithm with Quadrature Phase Shift Keying (QPSK) and 8PSK, respectively. The proposed method achieves nearly the same mutual information as the two-step iterative algorithm for a large range of SNR region, especially for large MIMO size and/or high constellation systems. The proposed precoding design method is applied to a 2×2 MIMO-OFDM system with 2048 subcarriers by designing 1024 precoders for extended channel matrices of size 4×4. A transceiver test bed implements these precoding matrices in comparison with other existing precoding schemes. Indoor experiments are conducted for fixed-platform non-line-of-sight (NLOS) channels, and the data processing results show that the proposed precoding method achieves the lowest BER compared to maximum diversity, classic water-filling and channel diagonalization methods.

1 INTRODUCTION

Linear precoding for multiple-input multiple-output (MIMO) communications has been a popular research topic in the last decade. Traditional precoding methods include maximizing the channel capacity with Gaussian inputs [1], maximizing the diversity order [2], maximizing the signal to interference-plus-noise ratio (SINR) [3], minimizing the mean square error (MSE) [3], etc. More recently, several works have found that designing precoders by maximizing mutual information with finite-alphabet inputs can achieve higher mutual information [4,5,6,7,8,9,10] and lower bit error rate (BER) [11] than employing other optimization criteria. The performance benefits of these approaches [4,5,6,7,8,9,10,10,11] come from optimization of mutual information with finite-alphabet input constraints, compared to using other indirect methods such as maximizing channel capacity with Gaussian inputs, maximizing diversity order, minimizing SINR, or minimizing MSE.

However, applying precoding to practical MIMO systems encounter several obstacles. First of all, the algorithms developed to find the optimal precoders [10, 11, 12, 13, 14] are computationally complicated because of their iterative algorithm and multiple computations of gradients through Monte Carlo simulations. For instance, [10] utilized the gradient descent algorithm to directly find the precoding matrix, and [14] proposed a two-step iterative algorithm to exploit the precoder structure. Even offline calculation of precoders is prohibitive for large MIMO dimensions and high constellations. Second, the optimal precoders are often designed for a specific SNR value. Their sensitivity to SNR estimation errors is largely ignored in the literature. Third, the soft maximum a posteriori (MAP) detector [15] is often employed for performance evaluation [11,12] and good BER performance have been demonstrated. However, MAP-based iterative receiver has the highest complexity and is thus difficult

to implement in practice. Whether the performance gain of precoders can be leveraged by practical suboptimal receivers is still questionable. In addition, the demonstrated performance gains in [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14] are all based on simulations over either fixed channels or statistical models of fading channels. The feasibility of employing precoders in practical communication systems and their performances under real-world wireless channels are still unknown.

In this paper, we first propose a low complexity precoding design method to simplify the two-step precoder design algorithm that maximizes the lower bound of the mutual information with finite-alphabet inputs [14]. The proposed method employs fixed right singular vectors that are modulation diversity matrices designed for different modulation schemes, thus eliminating the need for the iterative optimization between the two steps — one for the right singular vectors, one for the power allocation vector. Furthermore, the proposed method discretizes the search space of the power allocation vector and further reduces the complexity of the iterative design of the power allocation. For a 4×4 channel, the computational complexity of the proposed precoder design is reduced to 3% and 6% of that required by the two-step algorithm for Quadrature Phase Shift Keying (QPSK) and 8PSK, respectively. Moreover, the proposed method achieves nearly the same mutual information as the original two-step iterative algorithm for a large range of SNR region for high constellation or under large-sized MIMO channels. It also outperforms the maximum diversity, classic waterfilling, and channel diagonalization methods in most part of the SNR region.

The proposed precoding design method is applied to a 2×2 MIMO-OFDM system with 2048 subcarriers for QPSK, 8PSK, and 16QAM (Quadrature Amplitude Modulation). For each modulation scheme, we design 1024 precoders for the extended channel matrices of size 4×4 that combines the 2×2 MIMO with 2 subcarriers, thus

leveraging frequency diversity gain along spatial diversity gains. Moreover, the precoders are designed for one specified SNR value but are applied to all SNR scenarios, further reducing the computational cost on precoder design.

The designed precoding matrices are implemented in a Field Programable Gate Array (FPGA)-based transceiver test bed employing a rate 3/4 low-density parity-check (LDPC) code, where the baseband transmitter and digital Intermediate Frequency (IF) circuits are implemented in Altera's Stratix III FPGA. Two baseband receiver algorithms are realized: the soft MMSE linear equalizer with interference cancellation [16, 17] in Altera's Stratix IV FPGA, and the fixed-complexity list sphere decoding (FSD) algorithm [18, 19] in MATLAB. Although various MIMO-OFDM testbeds [20, 21, 22, 23, 24, 25, 26] are reported in the literature for wideband communications, they focus mainly on supporting spatial multiplexing [21, 22, 23, 24, 25] and transmit diversity schemes [20], or space time block codes [21, 23], or polarization diversity [26]. To the best of our knowledge, this is the first test bed that implements precoder designs in both transmitter and receiver.

With the MIMO-OFDM test bed, we conduct indoor non-line-of-sight (NLOS) experiment to acquire channels, design precoders offline, and verify performance with precoding. The data processing results of the experiments show that the proposed precoding method achieves the lowest BER compared to maximum diversity, classic water-filling and channel diagonalization methods. Besides, the maximum diversity outperforms the classic water-filling and channel diagonalization methods.

Throughout the paper, we denote vectors with boldface lower-case letters and matrices with boldface upper-case letters. The superscripts $(\cdot)^h$ and $(\cdot)^+$ represent conjugate transpose and pseudoinverse operations, respectively. In addition, the symbol $\mathbb{C}$ stands for the complex number field, and $diag\{\mathbf{a}\}$ denotes a diagonal matrix with elements of vector $\mathbf{a}$.

2 LINEAR PRECODER DESIGN FOR FINITE ALPHABET

Consider a MIMO system with baseband equivalent model

$$\mathbf{y} = \mathbf{HGx} + \mathbf{n}, \tag{1}$$

where the vector $\mathbf{x} \in \mathbb{C}^{N_i \times 1}$ contains transmitted signals of N_i symbols. The matrix $\mathbf{H} \in \mathbb{C}^{N_o \times N_i}$ is the complex channel matrix, and $\mathbf{G} \in \mathbb{C}^{N_i \times N_i}$ is the designed precoding matrix. The receiver noise $\mathbf{n} \in \mathbb{C}^{N_o \times 1}$ is a zero mean circularly symmetric complex Gaussian vector with covariance matrix $\sigma^2\mathbf{I}$, *i.e.*, $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$.

To design the precoding matrix $\mathbf{G}$, we first decomposite the channel matrix $\mathbf{H}$ via singular value decomposition (SVD), thus:

$$\mathbf{H} = \mathbf{U_H}\mathbf{\Sigma_H}\mathbf{V_H}^h \tag{2}$$

where $\mathbf{U_H}$ and $\mathbf{V_H}$ are unitary matrices, and $\mathbf{\Sigma_H}$ is a diagonal matrix.

Similarly, the precoding matrix $\mathbf{G}$ can also be decomposed into three components as:

$$\mathbf{G} = \mathbf{U_G}\mathbf{\Sigma_G}\mathbf{V_G}^h \tag{3}$$

in which $\mathbf{U_G}$ and $\mathbf{V_G}$ are left and right singular vectors of $\mathbf{G}$, respectively. The diagonal matrix $\mathbf{\Sigma_G}$ has all singular values. The decomposition in (3) is a general form that incorporates the four precoding schemes we discuss in this paper.

Scheme 1 is the channel diagonalization precoder, also known as parallel decomposition [27, 28], which sets $\mathbf{V_G}$ and $\mathbf{\Sigma_G}$ to identity matrices, and $\mathbf{U_G} = \mathbf{V_H}$.

The precoding matrix is then given by

$$\mathbf{G}_D = \mathbf{V}_{\mathbf{H}}, \tag{4}$$

which simply diagonalizes the channel.

Scheme 2 is the classic water-filling precoder which maximizes the channel capacity assuming that the input signal $\mathbf{x}$ is Gaussian distributed. The MIMO channel capacity with Gaussian input assumption is [27]

$$C = \max_{\mathrm{Tr}(\mathbf{Q})=P} \log \left| \mathbf{I} + \frac{1}{\sigma^2} \mathbf{H}\mathbf{Q}\mathbf{H}^h \right| \tag{5}$$

where P is the transmit power and $\mathbf{Q}$ is the covariance matrix after precoding such that

$$\mathbf{Q} = \mathbf{V}_{\mathbf{H}} \Sigma_{\mathbf{WF}} \mathbf{V}_{\mathbf{H}}^h \tag{6}$$

where the diagonal matrix $\Sigma_{\mathbf{WF}}$ can be solved by water-filling [1,27]. Since $\mathbf{Q} = \mathbf{G}\mathbf{G}^h$, the classic water-filling precoding matrix is

$$\mathbf{G}_{WF} = \mathbf{V}_{\mathbf{H}} \mathbf{\Sigma}_{WF}^{\frac{1}{2}} \tag{7}$$

Scheme 3 is the maximum diversity precoder which maximizes diversity and coding gains for two-dimensional constellations without Gaussian input assumptions [2]. It consists of two matrices in (3):

$$\mathbf{G}_{MD} = \mathbf{V}_{\mathbf{H}} \mathbf{V}_{MD} \tag{8}$$

in which $\mathbf{V}_{MD}$ is a matrix obtained through the algebraic construction and it is designed to be the same for all types of modulations. Specifically, when N_i is a power

of two, $\mathbf{V}_{MD}$ is a unitary matrix given by

$$\mathbf{V}_{MD} = \frac{1}{\sqrt{N_i}} \begin{bmatrix} 1 & \alpha_1 & \dots & \alpha_1^{N_i-1} \\ 1 & \alpha_2 & \dots & \alpha_2^{N_i-1} \\ \vdots & \vdots & & \vdots \\ 1 & \alpha_{N_i} & \dots & \alpha_{N_i}^{N_i-1} \end{bmatrix} \tag{9}$$

in which

$$\alpha_k = \exp\left[j\frac{4(k-1)+1}{2N_i}\right], \quad k = 1, 2, \cdots, N_i \tag{10}$$

Scheme 4 is the optimal precoding which maximizes the channel mutual information directly with finite-alphabet inputs. Based on the signal model in (1), the mutual information between the input $\mathbf{x}$ and output $\mathbf{y}$ is given by [11]

$$\mathcal{I}\left(\mathbf{x};\mathbf{y}\right)\left[\mathbf{H},\mathbf{G}\right] = N_i \log M - \frac{1}{M^{N_i}} \sum_{m=1}^{M^{N_i}} E_{\mathbf{n}}\left[\log \sum_{k=1}^{M^{N_i}} \exp\left(-d_{mk}\right)\right] \tag{11}$$

where $d_{mk} = (\|\mathbf{HGe}_{mk} + \mathbf{n}\|^2 - \|\mathbf{n}\|^2)/\sigma^2$, and $\mathbf{e}_{mk} = \mathbf{x}^m - \mathbf{x}^k$. The symbol $E_{\mathbf{n}}[\cdot]$ takes expectation over the noise $\mathbf{n}$. The signal $\mathbf{x}^m$ is a vector with each of its elements drawn from the M-ary signal constellation.

The calculation of the mutual information in (11) involves mathematical expectation, which is often estimated through Monte Carlo simulations [10] that is computationally expensive. To reduce the high computation complexity caused by Monte Carlo simulations of the expectation, a lower bound of mutual information is derived in [29] as

$$\mathcal{I}_L\left(\mathbf{H},\mathbf{G}\right) = N_i \log M - \left(1/\ln 2 - 1\right) N_o - \frac{1}{M^{N_i}} \sum_{m=1}^{M^{N_i}} \log \sum_{k=1}^{M^{N_i}} \exp\left(-\frac{\mathbf{c}_{mk}^H \mathbf{c}_{mk}}{2\sigma^2}\right) \tag{12}$$

where $\mathbf{c}_{mk} = \mathbf{HG}(\mathbf{x}^m - \mathbf{x}^k)$. It is shown that the lower bound (12) is a close approximation with a constant shift of the accurate mutual information (11) under various channel conditions [29].

Employing this lower bound and setting $\mathbf{U_G} = \mathbf{V_H}$ in the precoder design yields a two-step iterative algorithm [14]. After initializations, the algorithm alternatively updates $\mathbf{\Sigma_G}$ and $\mathbf{V_G}$ in the two steps until the convergence. During each iteration, the first step is to optimize the power allocation vector $\boldsymbol{\lambda} = diag(\mathbf{\Sigma_G})$ given $\mathbf{V_G}$:

$$\text{maximize} \quad \mathcal{I}_L(\boldsymbol{\lambda}) \tag{13}$$

$$\text{subject to} \quad \mathbf{1}^T\boldsymbol{\lambda} = N_i \tag{14}$$

$$\boldsymbol{\lambda} \succeq \mathbf{0} \tag{15}$$

where $\mathcal{I}_L(\boldsymbol{\lambda})$ is the lower bound in (12). Given the $\boldsymbol{\lambda}$ obtained in the first step, the second step is to optimize $\mathbf{V_G}$:

$$\text{maximize} \quad \mathcal{I}_L(\mathbf{V_G}) \tag{16}$$

$$\text{subject to} \quad \mathbf{V}_\mathbf{G}^h\mathbf{V_G} = \mathbf{I} \tag{17}$$

As shown later in Section 3, the computational complexity of the optimal precoder design using he lower bound is still prohibitively high. Moreover, the complexity grows exponentially with the modulation level M and the size of precoding matrix N_i. Thus, lowering the overall computational complexity further is necessary to make practical use of the precoder design.

3 PROPOSED SIMPLIFIED LINEAR PRECODER DESIGN

To reduce the complexity of the precoder design, we propose a non-iterative approach to design precoder for finite-alphabet inputs. The proposed precoder $\mathbf{G}$ contains three components as follows

$$\mathbf{G} = \mathbf{V_H} \mathbf{\Sigma_G} \mathbf{V}_{mod} \tag{18}$$

where $\mathbf{V}_{mod}$ is a unitary matrix designed for different modulation constellations and it is referred to as modulation diversity matrix in this paper. The matrix $\mathbf{\Sigma_G}$ is a diagonal matrix for power allocation policy depending on both the channel condition and modulation constellation. The non-iterative precoder design approach first fixes $\mathbf{V}_{mod}$ and then solves $\mathbf{\Sigma_G}$.

The modulation diversity matrix $\mathbf{V}_{mod}$ in (18) is different from the maximum diversity matrix [2], because $\mathbf{V}_{mod}$ is tailored for the constellation used for the transmitted signal vector $\mathbf{x}$. Specifically, we propose the following structure for the design of the modulation diversity matrix

$$\mathbf{V}_{mod} = \frac{1}{\sqrt{N_i}} \begin{bmatrix} 1 & \beta_1 & \dots & \beta_1^{N_i-1} \\ 1 & \beta_2 & \dots & \beta_2^{N_i-1} \\ \vdots & \vdots & & \vdots \\ 1 & \beta_{N_i} & \dots & \beta_{N_i}^{N_i-1} \end{bmatrix} \tag{19}$$

where

$$\beta_k = \exp\left[j\pi \frac{2(k-1) + q_{mod}}{N_i} \right], \quad k = 1, 2, \cdots, N_i \tag{20}$$

The parameter q_{mod} is a rotation angle depending on the modulation type. For MPSK modulations, q_{mod} is $(\frac{1}{2})^{M-1}$, where M is the constellation size. The rotation angle q_{mod} is $\frac{1}{2}$ for all QAM such as 4QAM, 16QAM and 64QAM. In this case, the constellation expansion is maximized. Table 1 lists values of q_{mod} for different modulations.

Table 1. The values of q_{mod} corresponding to different modulations

Modulation	BPSK	QPSK	8PSK	16PSK	16QAM	64 QAM
q_{mod}	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{2}$

Actually, the modulation diversity matrix for 2×2 MIMO system can be easily derived from the values of ν of Table I in [11]. The corresponding modulation diversity matrix for 4×4 MIMO system can be obtained with the extension of 2×2 MIMO systems.

Comparing (20) and (9), we see that the modulation diversity matrix is identical to the maximum diversity for QAM and QPSK modulations when N_i is a power of 2. However, these two methods are different from each other for other modulations. For instance, the 2×2 maximum diversity matrix for all types of modulations is given by

$$\mathbf{V}_{MD} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & e^{j\frac{\pi}{4}} \\ 1 & e^{j\frac{5\pi}{4}} \end{bmatrix} \tag{21}$$

On the other hand, the 2×2 modulation diversity matrices for BPSK and 8PSK modulations are

$$\mathbf{V}_{mod-BPSK} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & e^{j\frac{\pi}{2}} \\ 1 & e^{j\frac{3\pi}{2}} \end{bmatrix}, \quad \mathbf{V}_{mod-8PSK} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & e^{j\frac{\pi}{8}} \\ 1 & e^{j\frac{9\pi}{8}} \end{bmatrix}. \tag{22}$$

For QPSK and 16QAM modulations, the modulation diversity matrices are the same as the maximum diversity matrix.

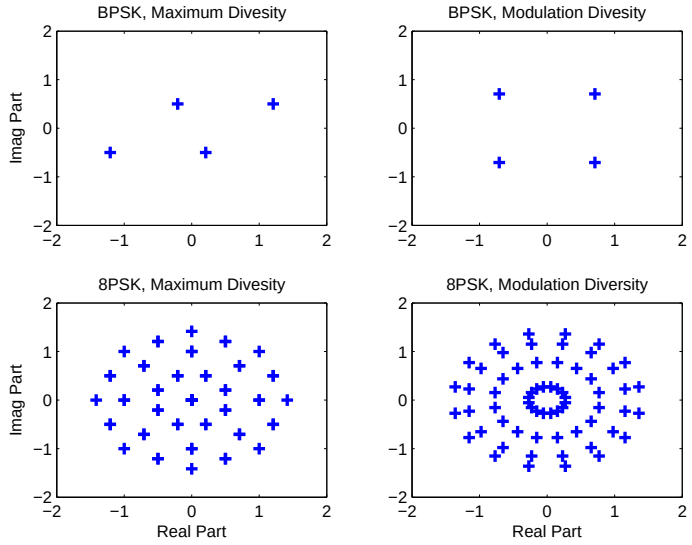


Figure 1. Scatter plot of the precoded BPSK and 8PSK signals using 2×2 maximum diversity and modulation diversity

In Fig. 1, we plot the constellations of precoded BPSK and 8PSK inputs using 2×2 modulation diversity and maximum diversity matrices. It is seen that the scatter plot of the modulation diversity for BPSK inputs is the same as the QPSK constellation, which is different from the result obtained by maximum diversity. Comparing the scatter plots of 8PSK inputs, we notice that there are 33 and 64 points for maximum diversity and modulations diversity, respectively. This indicates that

one of the shortcomings of the maximum diversity is that the original 64 possible signal points in $\mathbf{x}$ collapse to less points in the scatter plot after precoding. The modulation diversity, on the other hand, preserves all the signal points in the constellations and converts them to a complex Gaussian-like scatter plot.

After $\mathbf{V}_{mod}$ is fixed, $\mathbf{\Sigma_G}$ can be found by exhaustive search based on the lower bound of the mutual information. Specifically, we partition $[0, 1]$ into K uniform segments. Assume that the diagonal matrix $\mathbf{\Sigma_G} = diag\{\sigma_1, \ldots, \sigma_{N_i}\}$ and each σ_i is chosen from the set $\{0, \frac{1}{K}, \frac{2}{K}, \ldots, \frac{K}{K}\}$. Thus, there are totally $(K+1)^{N_i}$ combinations for all the N_i diagonal elements in $\mathbf{\Sigma_G}$. For each combination of $\{\sigma_1, \ldots, \sigma_{N_i}\}$, we normalize their values as $\tilde{\sigma}_i = \sigma_i / \sum_{i=1}^{N_i} |\sigma_i|^2$ so that the power of the resulting precoding matrix $\mathbf{G}$ is N_i. Each normalized diagonal matrix $\mathbf{\Sigma}_{\mathbf{G}^k}$ along with the fixed $\mathbf{V}_{MD}$ corresponds to a precoding matrix $\mathbf{G}^k = \mathbf{V_H \Sigma}_{\mathbf{G}^k} \mathbf{V}_{mod}$. By computing the lower bound of $\mathcal{I}_L\left(\mathbf{H}, \mathbf{G}^k\right)$, $k = 1, \ldots, (K+1)^{N_i}$, and choosing the $\mathbf{G}^k$ with the maximum lower bound, we find the approximate optimal precoder. We note that the proposed discretization of the power allocation diagonal matrix can be viewed as an extension to the well known on/off power allocation technique.

From the afore mentioned procedure, we note that the proposed simplified precoding is a non-iterative method. Moreover, it avoids the computation of the derivative of $\mathcal{I}_L(\mathbf{H}, \mathbf{G})$ that is more time consuming than computing $\mathcal{I}_L(\mathbf{H}, \mathbf{G})$ itself, especially when the size of the precoding matrix N_i is large. Therefore, the proposed method has a lower complexity than the two-step iterative algorithm in [14].

To verify the complexity and performance merits of the proposed precoders, we consider a constant 2×2 channel and a complex 4×4 channel:

$$\mathbf{H}_1 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \tag{23}$$

Table 2. The CPU time (seconds) of designing precoders for the 4×4 channel $\mathbf{H}_2$

SNR (dB)	QPSK		8PSK	
	Two-step Iterative Algorithm	Proposed Method	Two-step Iterative Algorithm	Proposed Method
-10	177.3064	3.0144	19655	743.6521
-5	163.5079	3.0518	10885	753.4039
5	82.2510	3.0593	11692	755.4079
10	106.9154	3.0773	12707	755.8201
15	178.5296	3.1564	12204	779.5273

and

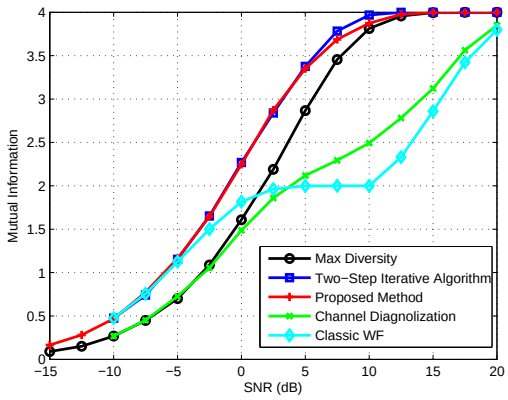
$$\mathbf{H}_2 = \begin{bmatrix} 0.1897 + 0.6602i & 0.5417 + 0.5341i & 0.8600 + 0.5681i & 0.8998 + 0.4449i \\ 0.1934 + 0.3420i & 0.1509 + 0.7271i & 0.8537 + 0.3704i & 0.8216 + 0.6946i \\ 0.6822 + 0.2897i & 0.6979 + 0.3093i & 0.5936 + 0.7027i & 0.6449 + 0.6213i \\ 0.3028 + 0.3412i & 0.3784 + 0.8385i & 0.4966 + 0.5466i & 0.8180 + 0.7948i \end{bmatrix} \tag{24}$$

in which the 2×2 channel $\mathbf{H}_1$ is also used in [4, 11, 14]. The SNR is defined as $SNR_i = \mathrm{Tr}(\mathbf{H}_i \mathbf{H}_i^h)/(N_o \sigma^2)$, for $i = 1, 2$, [11].

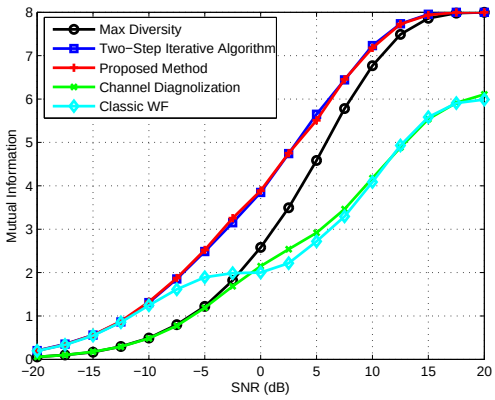
Table 2 lists the CPU time of obtaining precoders using two-step iterative algorithm [14] and the proposed method for the 4×4 channel $\mathbf{H}_2$. The codes for both methods are mostly written in MATLAB except that the part of computing lower bound $\mathcal{I}_L(\mathbf{H}, \mathbf{G})$ is implemented in C++. The simulations are executed on an Intel E8600 3.33 GHz duo core processor. In the proposed method, K is set to be 4. Thus the total number of calculated lower bounds is $5^4 = 625$. The results in Table 2 show that the proposed method consumes much less time than the two-step iterative algorithm for QPSK and 8PSK modulations under various SNRs. For example, when $SNR = 10$ dB, the CPU time of running the proposed method for QPSK and 8PSK modulations is about 3% and 6% that of the two-step iterative algorithm, respectively.

The mutual information results of different modulations for channel $\mathbf{H}_1$ and $\mathbf{H}_2$ are shown in (a) - (f) of Fig.2, where for comparison, maximum diversity, classic water-filling and channel diagonalization are also provided. For the 2×2 channel with QPSK and 8PSK inputs, the proposed method performs almost the same as the two-step iterative algorithm in the low and medium SNR regions. As the SNR increases, its performance becomes apart from the two-step iterative algorithm but still better than that of the maximum diversity method. On the other hand, it is seen from (c) - (e) in Fig. 2 that the proposed method and the two-step iterative algorithm have nearly the same performances for a large range of SNR region under the 2×2 channel with 16QAM inputs, and the 4×4 channel with QPSK and 8PSK inputs. Due to the extremely high computation complexity, the two-step iterative algorithm has not been simulated for the 4×4 channel with 16QAM inputs. Yet, Fig. 3.2(f) shows that the proposed method achieves higher mutual information than the maximum diversity for the 16QAM. In all cases, the proposed method performs better than the classic water-filling and channel diagonalization methods especially in the medium and high SNR regions.

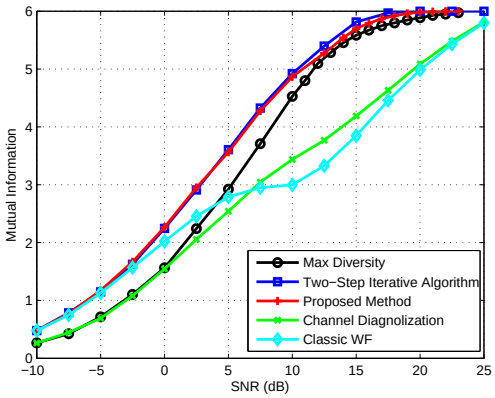
To leverage the benefit of precoding in practical MIMO-OFDM systems, we further propose to apply linear precoding to extended channel matrices that combines $N_r \times N_t$ MIMO and K_G subcarriers. Specifically, we divide the total of N_f subcarriers into N_g groups and perform linear precoding for the extended $(N_r \cdot K_G) \times (N_t \cdot K_G)$ channel in each group. For example, when $N_t = 2$, $N_t = 2$, $K_G = 2$, and $N_f = 2048$, precoders of size 4×4 are designed for 1024 groups. We group the symbols from the two subcarriers and two transmitted data streams into a 4×1 column vector as the equivalent channel input. Each signal vector is then multiplied by the 4×4 linear precoding matrix. By using a larger-size precoder, higher precoding gain can be achieved than that of precoding without subcarrier grouping.



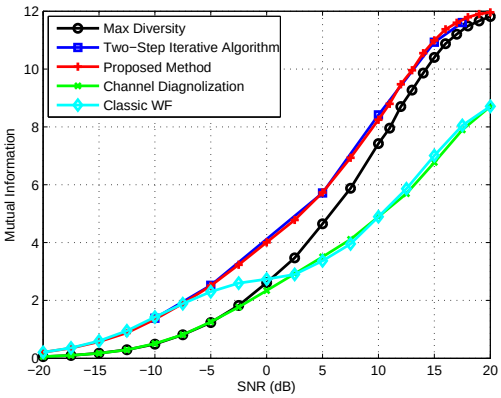
(a) 2×2 channel with QPSK inputs



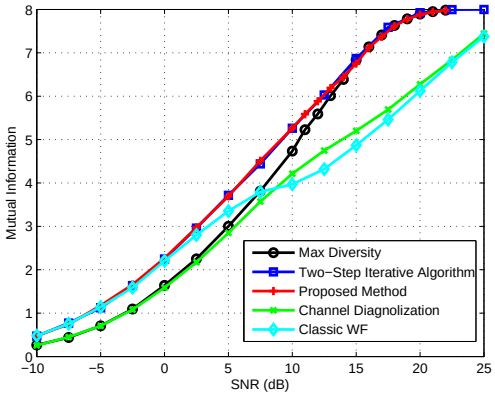
(b) 4×4 channel with QPSK inputs



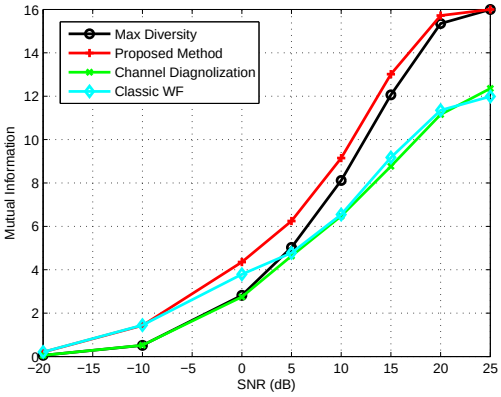
(c) 2×2 channel with 8PSK inputs



(d) 4×4 channel with 8PSK inputs



(e) 2×2 channel with 16QAM inputs



(f) 4×4 channel with 16QAM inputs

Figure 2. Mutual information vs. SNR

In addition, an interleaver is employed for the subcarrier grouping so that each interleaved group contains both statistically strong and weak frequency tones; thus diversity gain is achieved by leveraging both spatial and frequency correlation of the MIMO channels. To illustrated this, we assume that $\mathbf{H}_k$ represents the k-th frequency tone of all the channel matrices, the matrix $\mathbf{H}_k$ has the same distribution as $\sqrt{f_k}\mathbf{\Psi}_{RX}^{\frac{1}{2}}\mathbf{H}_{w,k}\mathbf{\Psi}_{RX}^{\frac{1}{2}}$ [30] for Rayleigh fading, where $\mathbf{\Psi}_{RX}$ and $\mathbf{\Psi}_{TX}$ are the antenna spatial correlation matrices due to angle spreads at the transmitter and receiver, respectively. The matrix $\mathbf{H}_{w,k}$ is an $N_r \times N_t$ random matrix with i.i.d $\sim \mathcal{CN}(0,1)$ elements. The scalar f_k is the channel spectrum function given by:

$$f_k = 1 + 2\sum_{i=1}^{L-1} a_i \cos\left[\frac{2\pi i(k-0.5)}{N_f}\right], k = 1, \cdots, N_f \tag{25}$$

and

$$a_i = \sum_{l=1}^{L-i} \mathbf{\Psi}_{ISI}(l, l+i) \tag{26}$$

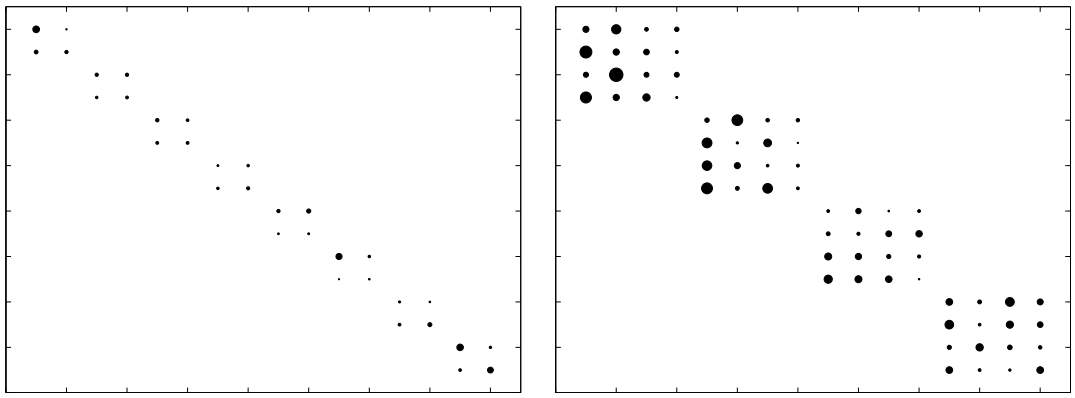
in which $\mathbf{\Psi}_{ISI}$ is the inter-tap correlation matrix [30]. The average power of the k-th frequency tone is

$$\mathrm{Tr}\left[E(\mathbf{H}_k\mathbf{H}_k^h)\right] = f_k \cdot \mathrm{Tr}(\mathbf{\Psi}_{RX}) \cdot \mathrm{Tr}(\mathbf{\Psi}_{TX}). \tag{27}$$

According to (25) and (26), the low frequency tones (when k is small) have stronger average power than the high frequency tones (when k is around $N_f/2$). Thus, the objective of the interleaver is is that the interleaved frequency tones in different groups have similar descending patters on the average power.

Figure 3 gives an example which indicates the effect of subcarrier grouping for a 2×2 MIMO-OFDM system. For comparison purpose, signals without subcarrier grouping are shown in Fig. 3.3(a). In Fig. 3.3(b), each group has signals from two

subcarriers, i.e., $K_G = 2$. It is seen that each group contains strong and weak signals when interleaving is employed in the subcarrier grouping. In this way, both spatial and frequency diversity are utilized.



(a) No subcarrier grouping (b) Precoding with subcarrier grouping

Figure 3. Amplitudes of precoded elements in a 2×2 MIMO-OFDM system. The size of the dots is proportional to the amplitude of precoder coefficients. Only eight subcarriers are shown.

4 APPLICATIONS OF LINEAR PREDING TO MIMO-OFDM

In this section, we present the MIMO-OFDM system that employs the designed 4×4 linear precoding matrices for verification purposes. Specifically, we will discuss the data frame structure, system diagram, and receiver algorithms.

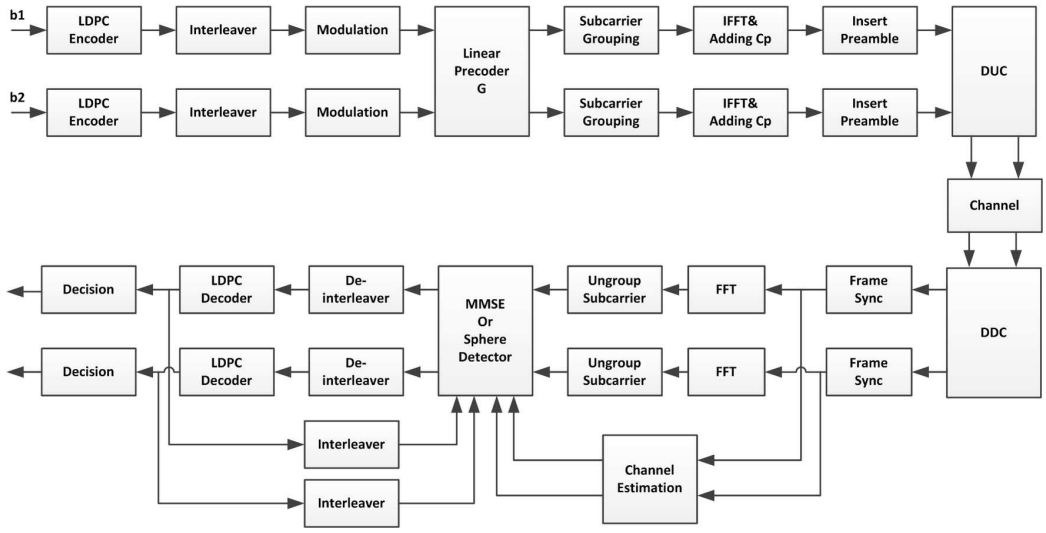


Figure 4. System diagram of 2×2 MIMO-OFDM system.

Tx 1	BPSK Seq 1	Preamble 1		Preamble 2		Data Block 1		Data Block 2	
	144	128CP	256	128CP	256	128CP	2048	128CP	2048

Tx 2	BPSK Seq 2	Preamble 3		Preamble 4		Data Block 3		Data Block 4	
	144	128CP	256	128CP	256	128CP	2048	128CP	2048

Figure 5. Frame structure of MIMO-OFDM signaling

The system diagram of the 2×2 MIMO-OFDM testbed is shown in Fig. 4, where the transmitter employs a signaling frame structure shown in Fig. 5. The 144-bit BPSK sequence at each transmitter antenna is a frame header for frame and symbol synchronization. The two BPSK sequences are two different pseudo-random noise (PN) sequences generated by linear feedback shift registers with different initial seeds. The two preambles at each antenna in Fig. 5 serve as training sequences for channel estimation. Each preamble is consisted of a 256-bit length Zadroff-Chu (ZC) sequence [31] and a 128-bit length cyclic prefix (CP). The CP length is chosen so that it is greater than the length of the channel impulse response (CIR). The preambles have the following structure [32, 33]:

$$\begin{bmatrix} \text{Preamble 1} & \text{Preamble 2} \\ \text{Preamble 3} & \text{Preamble 4} \end{bmatrix} = \begin{bmatrix} S_{chu} & -S_{chu}^{*} \\ S_{chu} & S_{chu}^{*} \end{bmatrix} \tag{28}$$

where S_{chu} is a 256-bit ZC sequence. Two OFDM data blocks follow the preambles, and each data block consists of 2048 data symbols preceded by 128 symbols of CP.

As shown in Fig. 4, the data blocks at the transmitter are generated by two raw source bit streams $\mathbf{b}_1$ and $\mathbf{b}_2$ encoded by a LDPC encoder with 3/4 coding rate and codeword length of 2040 bits. We employ the LDPC channel codes specified by the latest WiMAX standard [34], which has also been realized in the software package in [35]. After LDPC encoding, each codeword is added 8 bits of zeros at the end to form a data block of length 2048, which is then interleaved and modulated to symbols. The supported modulation schemes in our implementation are QPSK, 8PSK, or 16QAM.

After precoding, the symbols of each stream are fed into a subcarrier grouper, which is discussed in Section 3. After the subcarrier grouping, each data block is converted to time domain signals by inverse fast Fourier transform (IFFT), followed by CP insertion and preamble insertion. The digital up convertor (DUC) is used

to up-sample the baseband signals and modulate them into the 17.5 MHz IF. The output of the DUC module is fed to the digital to analog converter (DAC) directly.

At the receiver, the ADC bandpass sampling rate is 125 MHz and the IF is 70 MHz. The digital down convertor (DDC) module down-samples the IF signals to baseband I and Q signals. At each receive antenna, the frame synchronizer captures the received frame based on the time domain correlations of the local PN sequence. After the start of the frame is located, the received preambles are used for channel estimation which uses frequency domain LS method [32, 33]. The received preambles are first converted to frequency domain with a 256-point FFT. For each subcarrier of the preamble, LS estimation is performed to obtain an initial estimation of a 2×2 channel matrix. Then we convert the estimated channel matrices to time domain with a 256-point IFFT and pass them through a rectangular window whose width is larger than the length of the channel impulse response. Finally, we convert the windowed CIR to frequency domain channel response $\hat{\mathbf{H}}_i$ using a 2048-point FFT.

A turbo receiver [36] is employed to iteratively exchange log likelihood ratio (LLR) between the MIMO MMSE-IC soft detector [16, 17] and the LDPC channel decoder for each bit stream. Alternatively, the list fixed sphere decoder (FSD) [19, 18] is also implemented. Since the soft MIMO detector group received signals from two subcarriers to detect symbols, the effective channel matrices $\mathbf{H}_{ei} = \hat{\mathbf{H}}_i\mathbf{G}_i, i = 1, 2, \cdots, 1024$, are used in both MMSE-IC and FSD schemes. The MIMO detector generates extrinsic LLR, based on $\mathbf{y}_i$ and the *a priori* information from the LDPC channel decoders. The extrinsic information of MIMO detector is then interleaved and sent to the LDPC decoder as its *a priori* input. By employing soft channel decoding methods such as log domain sum-product algorithm [37], the LDPC decoder computes its extrinsic information as output, which is fed back to the MIMO detector as the *a priori* information. After two turbo iterations, hard decisions are made at the output of the LDPC decoder.

5 TESTBED, EXPERIMENT SETUP, AND RESULTS

In this section, we present the setup of the MIMO-OFDM testbed and discuss the experiment results.

5.1 TESTBED AND EXPERIMENT SETUP

The testbed was built using the equipment listed in Tabel 3. The baseband processing was implemented on Altera Stratix III FPGA development kits [38]. The FPGA-DSP development kit features Stratix III EP3SL150F1152 high-performance FPGA, and ADC/DAC daughter board. Baseband signal processing was implemented in Altera FPGA and daughter board. The detailed architecture setup of the transmitter and receiver are shown in Fig. 6 and Fig. 7.

Table 3. Description of Key Equipments

Function	**Equipment**
FPGA platform for Tx	Altera Stratix III FPGA Development kit with daughter board
FPGA platform for Rx	Altera Stratix IV FPGA Development kit with daughter board
Receive antennas	Pharad's wearable antennas, BW-800-900-D
RF up-converter	NuWaves, RF2-3000UCV1
Transmitter power amplifiers (optional)	RF Bay, MPA-10-40
Receiver low noise amplifier (optional)	RF Bay, LNA-0915
RF down-converter	NuWaves, RF200-2500RV1
External clock for frequency reference	FS725 Rubidium Frequency Standard

In the transmitter shown in Fig. 6, we implemented the baseband modules and digital up-converter (DUC) in the Stratix III FPGA and utilized the DAC daughter

board to convert two data streams to analog IF signals. At each stream of the transmitter, the IF signals were fed into the low pass filter SLP-50 with 50 MHz bandwidth to reduce the out of band noise. Then the NuWave RF2-3000UCV1 RF up-converter was employed to up-convert the IF signal to 915 MHz RF signals. The MPA-10-40 RF was chosen for power amplification, before the signals being transmitted through wearable antennas. The computer was used to program the FPGA and to configure the RF up-converters. The transmit power can be adjusted manually by changing the attenuation level of the RF up-converter. The FS725 rubidium clock serve as an external reference clock for FPGA board and RF up-converters.

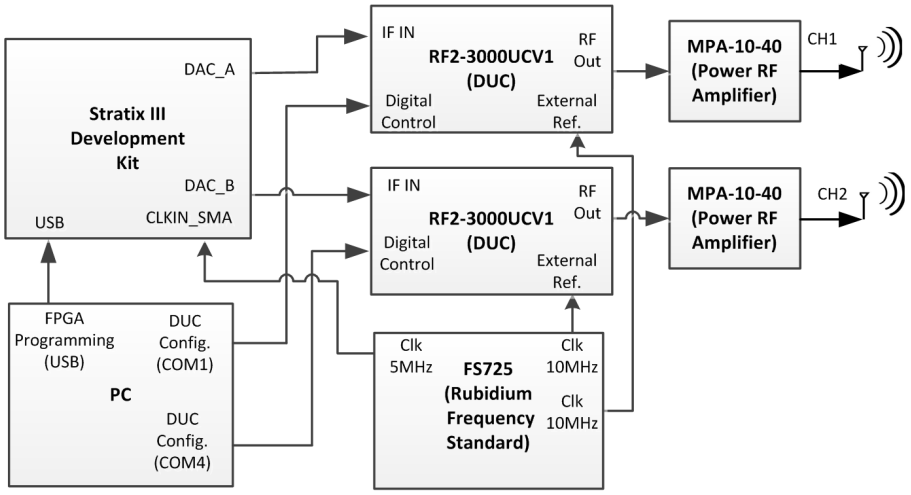


Figure 6. Transmitter setup architecture

At the receiver side, the signals were received by antennas and then down-converted to 70 MHz IF signals by the RF front-end. The 2-channel 14-bit analog to digital converter (ADC) daughter board was employed to convert signals into digital

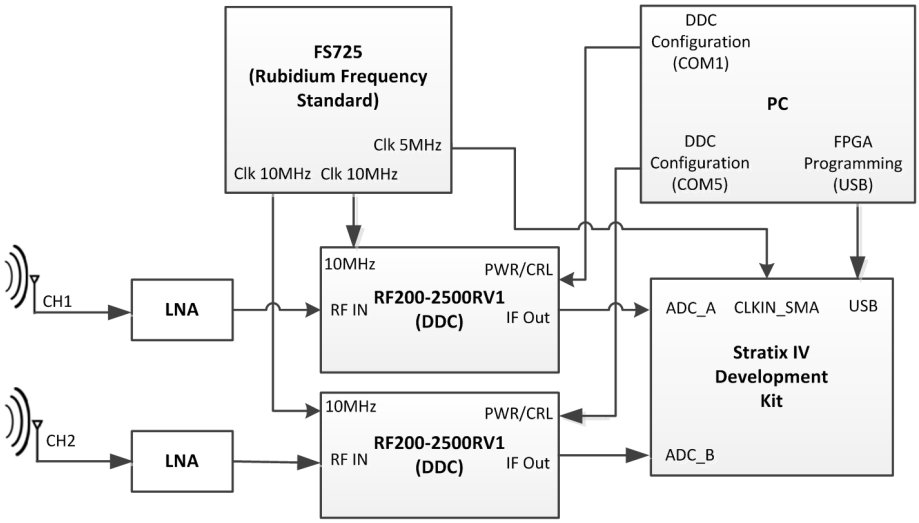


Figure 7. Receiver setup architecture

signals via bandpass sampling. The sampling frequency was also 125 MHz and the folded IF frequency was 55 MHz.

5.2 FIELD EXPERIMENTS

Our experiments were conducted in three steps. Step 1 was channel acquisition without precoding. We transmited packets using identity matrices as precoders, received the signals and estimated the channels. Step 2 was offline precoder design for all subcarriers based on the estimated channels. After the precoding matrices were computed at a fixed SNR, they were quantized and stored in read-only memories (ROMs) of the transmitter FPGA. Step 3 was performance verification with precoding. A particular precoding method was selected and multiple frames were transmitted with the assumption that the channels are static. We change the attenuation level of the RF upconverter at the transmitter to obtain different values of signal transmission power. Under a specific modulation scheme, we use the manual gain controller of the RF downconverter at the receiver side to adjust the spectra of the received signals. Experiment data was logged at the DDC output through

Signal Tap II logic analyzer of Quartus II and baseband receiver was implemented in MATLAB.

We conducted indoor non-line-of-sight (NLOS) experiments in Room 208 of the Emerson Electric Company Hall at the Missouri University of Science and Technology campus. The transmitter and receiver are located in the same room with dimensions of 9.1 m by 7.0 m. In addition, a metal bookshelf is placed between the transmitter and receiver to block the line-of-sight communication path. The transmitter and receiver antennas are located at half of the height of the blocking bookshelf above the floor. The floor plan is shown in Fig. 8, where the ceiling is meters above floor.

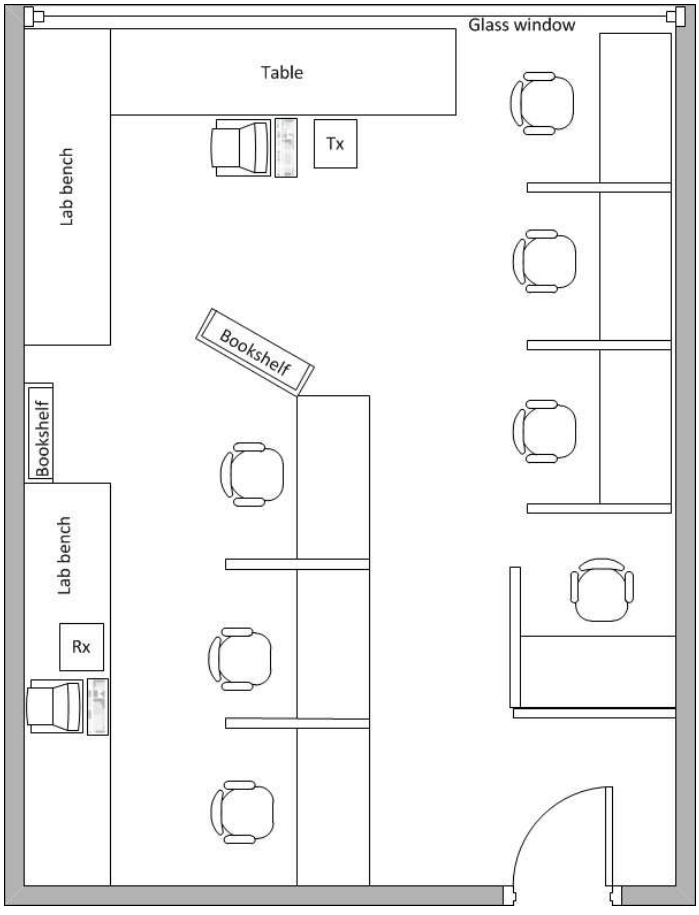


Figure 8. Floor plan of the indoor NLOS environment (9.1 m × 7.0 m)

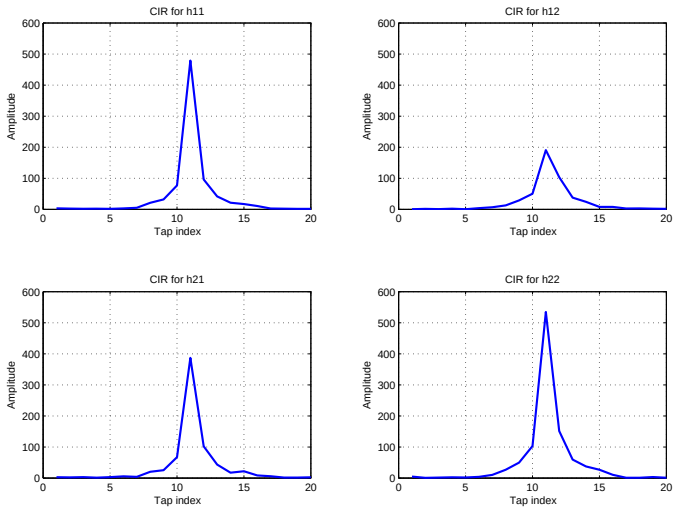


Figure 9. Estimated channel impulse responses of indoor same room NLOS environment

5.3 EXPERIMENT RESULTS

Figure 9 plots a snapshot of the estimated channel impulse responses for the indoor same room NLOS environment, where h_{ij} represents the channel from j-th transmit antenna to i-th receive antenna. It is observed that the length of all the channels are less than 20 taps. Our extensive experiments have shown similar characteristics of the channel impulse responses. Therefore, the width of the time domain window for channel estimation are set to 20.

Table 4 lists BER results of using list FSD method for data processing. The experiments include different modulations for precoding schemes 1, 2, and 3 discussed in Section 2, and the proposed precoding method in Section 3. From the results of QPSK and 8PSK shown in Table 4, it is seen that the proposed method achieved lowest BER among all the implemented precoding method in our system. In addition,

Table 4. BER results using List FSD for indoor same room NLOS experiments

Mod	**Transmitter attenuation (dB)**	**Classic WF**	**Channel diagonalization**	**Maximum diversity**	**Proposed method**
QPSK	18	0.3896	0.3776	0.07614	0.00784
	16	0.3806	0.1003	0	0
	14	0.3578	0.05833	0	0
	10	0.3858	0	0	0
8PSK	18	0.4149	0.4419	0.2015	0.121
	16	0.4031	0.3425	0.08562	0.0488
	14	0.4116	0.2091	0.05403	0.03475
	10	0.3700	0.2170	0.009695	0.002505

the maximum diversity outperformed the classic water-filling and channel diagonalization. The performance gains of the the proposed method and the maximum diversity over the other two schemes are because of the fact that these two approaches take into account the specific structure of finite-alphabet inputs and offer higher mutual information. Such observations from the experiments are similar to the simulation results in [14], which includes the optimal precoding and the maximum diversity methods with MAP detection. For the list FSD in our receiver, the node distribution [19] for QPSK was $\mathbf{n}_s = [1, 1, 4, 4]$, and the number of survival candidates to calculate LLR was $N_L = 16$. For 8PSK, the distribution was set to $\mathbf{n}_s = [1, 1, 8, 8]$ and 16 candidates with the minimum distances were used for LLR calculation.

6 CONCLUSION

In this paper, a practical linear precoding method for MIMO-OFDM system has been proposed that simplifies the two-step algorithm maximizing the lower bound of the mutual information with finite-alphabet inputs. The proposed algorithm achieves similar performance as the original two-step algorithm but requires only 3-6% of the computational complexity, thus making the precoding design feasible to practical systems.

The precoding design has been applied to a 2×2 MIMO-OFDM system with 2048 subcarriers by designing 1024 precoding matrices of size 4×4 based on the extended channel matrices that combine the 2×2 MIMO with 2 subcarriers. Experimental results of indoor NLOS experiments have showed that the proposed precoding method has achieved the lowest BER in comparison to maximum diversity, classic water-filling and channel diagonalization methods.

7 ACKNOWLEDGMENTS

This work was supported in part by Office of Naval Research under Grant N00014-09-1-0011 and by National Science Foundation under Grant ECCS-0846486 and CCF-0915846. The authors would like to thank Mr. Weiliang Zeng for providing MATLAB codes of iterative two-step algorithm. We also acknowledge Mr. Bing Han and Drs. Fei Ren and Huang Lou for their help with the FPGA transceiver implementation and field experiments.

8 REFERENCES

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley, 2006.

[2] Y. Xin, Z. Wang, and G. B. Giannakis, "Space-time diversity systems based on linear constellation precoding," *IEEE Trans. Wireless Commun.*, vol. 2, pp. 294-309, Mar. 2003.

[3] D. P. Palomar, J. Cioffi and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, pp.2381-2401, Sep 2003.

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

[5] A. Lozano, A. M. Tulino, and S. Verdú, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Trans. Inform. Theory*, vol.52, pp.3033-3051, July 2006.

[6] 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 (Globecom)*, New Orleans, USA, Nov. 30 - Dec. 4, 2008.

[7] C. Xiao and Y. R. Zheng, "Transmit precoding for MIMO systems with partical CSI and discrete-constellation inputs," in *Proc. IEEE International Conference on Communications (ICC)*, Dresden, Germany, June 14-18, 2009.

[8] M. Payaró and D. P. Palomar, "On optimal precoding in linear vector Gaussian channels with arbitrary input distribution," in *Proc. IEEE International Symposium on Information Theory*, Seoul, Korea, Jun. 28 - Jul. 3, 2009.

[9] M. Lamarca, "Linear precoding for mutual information maximization in MIMO systems," in *Proc. of 6th International Symposium on Wireless Communication Systems*, Siena, Italy, 2009.

[10] F. Pérez-Cruz, M. R. D. Rodrigues, and S. Verdú, "MIMO Gaussian channels with arbitrary inputs: optimal precoding and power allocation," *IEEE Trans. Inform. Theory*, vol.56, pp. 1070-1084, Mar. 2010.

[11] 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.

[12] M. Wang, C. Xiao, and W. Zeng, "Linear precoding for MIMO multiple access channels with finite discrete inputs," *IEEE Trans. Wireless Commun.*, vol. 10, pp 3934-3942, Nov. 2011.

[13] W. Zeng, Y. R. Zheng, M. Wang, and J. Lu, "Linear precoding for relay networks: a perspective on finite-alphabet inputs," *IEEE Trans. on Wireless Commun.*, vol. 11, pp. 1146-1157, Mar 2012.

[14] W. Zeng, C. Xiao, M. Wang, and J. Lu, "Linear precoding for finite-alphabet inputs over MIMO fading channels with CSI," *IEEE Trans. Signal Process.*, vol. 60, July 2012 (scheduled to appear).

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

[16] L. Boher, R. Rabineau, and M. Helard, "An efficient MMSE equalizer implementation for 4×4 MIMO-OFDM systems in frequency selective fast varing channels," in *IEEE Proc. Personal, Indoor and Mobile Radio Communications (PIMRC)*, Athens, Greece, Sept 3-7, 2007.

[17] L. Boher, R. Rodrigue and M. Helard, "FPGA implementation of an iterative receiver for MIMO-OFDM Systems," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 857-866, Aug. 2008.

[18] L. G. Barbero and J. S. Thompson, "Fixing the complexity of the sphere decoder for MIMO detection," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 2131-2142, Jun. 2008.

[19] L. G. Barbero and J. S. Thompson, "Extending a fixed-complexity sphere decoder to obtain likelihood information for Turbo-MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 57, pp. 2804-2814, Sep. 2008.

[20] H. Sampath, S. Talwar, J. Tellado, V. Erceg, and A. Paulraj, "A fourth-generation MIMO-OFDM broadband wireless system: design, performance, and field trial results," *IEEE Commun. Mag.*, vol. 40, pp. 143-149, Sep. 2002.

[21] C. Dubuc, D. Starks, T. Creasy, and H. Yong, "A MIMO-OFDM prototype for next-generation wireless WANs," *IEEE Commun. Mag.*, vol. 42, pp. 82-87, Dec. 2004.

[22] A. van Zelst and T. C. W. Schenk, "Implementation of a MIMO OFDM based-wireless LAN system," *IEEE Trans. Signal Process.*, vol. 52, pp. 483-494, Feb. 2004.

[23] H. Yu, M. Kim, E. Choi, T. Jeion, and S. Lee, "Design and prototype development of MIMO-OFDM for next generation wireless LAN," *IEEE Trans. Consumer Electron.*, vol. 51, pp. 1134-1142, Nov. 2005.

[24] T. Haustein, A. Forck, H. Gäbler, V. Jungnickel, and S. Schiffermüller, "Real-time signal processing for multiantenna systems: algorithms, optimization, and implementation on an experimental test-bed," *EURASIP Journal on Applied Signal Processing*, 2006, Article ID 27 573.

[25] S. Haene, D. Perels, and A. Burg, "A real-time 4-stream MIMO-OFDM transceiver: system design, FPGA implementation, and characterization," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 877-889, Aug. 2008.

[26] C. Gómez-Calero, L. C. Navarrete, L. Haro, and R. Martiínez, "A 2×2 MIMO DVB-T2 system: design, new channel estimation scheme and measurements With polarization diversity," *IEEE Trans. on Broadcast.*, vol. 57, pp. 195-203, June 2011.

[27] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecomm. ETT*, vol. 10, pp. 585-596, Nov. 1999.

[28] A. Goldsmith, "Wireless communications," Cambridge University Press, 2005. pp.301

[29] W. Zeng, C. Xiao, and J. Lu, "A low-complexity design of linear precoding for MIMO channels with finite-alphabet inputs," accepted by *IEEE Wireless Commun. Lett.*, Dec. 2011.

[30] C. Xiao and Y. R. Zheng, "On the ergodic capacity of MIMO triply selective Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol.7, pp.2272-2279, June 2008.

[31] D. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Trans. Inform. Theory*, vol.18, pp. 531-532, July 1972.

[32] A. N. Mody and G. L. Stuber, "Receiver implementation for a MIMO OFDM system," in *Proc. Global Communications Conference (Globecom)*, Taipei, Taiwan, Nov. 17-21, 2002.

[33] G. L. Stuber, J. Barry, S. McLaughlin, Y. (G.) Li, M. A. Ingram, and T. Pratt, "Broadband MIMO-OFDM wireless communications," *Proceedings of IEEE*, vol. 92, pp. 271-294, Feb. 2004.

[34] *Part 16:Air Interface for Broadband Wireless Access Systems*, IEEE Std 802.16$^{\text{TM}}$-2009, May. 2009.

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

[36] 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.

[37] 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 (Globecom)*, San Antonio, USA, Nov. 25 - 29, 2001.

[38] Altera. *Stratix III 3SL150 development kit reference manual* [Online]. Available: http://www.altera.com/products/devkits/altera/kit-siii-host.html

SECTION

2 CONCLUSIONS

This dissertation proposed a new approach to maximizing data rate/throughput of practical communication system/networks through linear precoding and power allocation. By maximizing the mutual information with finite-alphabet inputs such as PSK, PAM and QAM signals, this approach improves performance when the designed precoders are applied to practical systems and networks.

Several numerical optimization methods were developed for MIMO multiple access channels, dual-hop relay networks, and point-to-point MIMO systems. In MIMO multiple access channels, an iterative weighted sum rate maximization algorithm was proposed which utilized an alternating optimization strategy and gradient descent update. In dual-hop relay networks, the structure of the optimal precoder was exploited to develop a two-step iterative algorithm based on convex optimization and optimization on the Stiefel manifold. The gradient descent method was also used to obtain the optimal power allocation scheme in dual-hop relay networks. For point-to-point MIMO systems, a low complexity precoding design method was proposed, with discretized power allocation vector in a non-iterative fashion, thus reducing complexity.

Performances of the proposed power allocation and linear precoding schemes were evaluated in terms of both mutual information and BER. Numerical results showed that at the same target mutual information or sum rate, the proposed approaches achieved 3-10 dB gains compared to the existing methods in the medium signal-to-noise ratio region. Such significant gains were also indicated in the coded BER systems.

The contributions of my PhD research work are summarized in five journal papers and seven conference papers, among which, three journal papers and one conference paper are included in this dissertation.

3 PUBLICATIONS

[1] M. Wang, Y. R. Zheng, and C. Xiao, "Practical linear precoder design for finite alphabet MIMO-OFDM with experiment validation," to be submitted to *IEEE Trans. Veh. Technol.*.

[2] Y. Wu, M. Wang, C. Xiao, Z. Ding, and X. Gao, "Linear precoding for finite alphabet signals in multi-antenna broadcast channels,", accepted by *IEEE Int. Conf. Commun. (ICC)*, Ottawa, Canada, Jun., 2012.

[3] W. Zeng, Y. R. Zheng, M. Wang, and J. Lu, "Linear precoding for relay networks: a perspective on finite-alphabet inputs," *IEEE Trans. Wireless Commun.*, vol. 11, pp 1146 - 1157, Mar. 2012.

[4] W. Zeng, C. Xiao, M. Wang, and J. Lu, "Linear precoding for finite-alphabet inputs over MIMO fading channels with statistical CSI", accepted by *IEEE Trans. Signal Process.*, Feb. 2012.

[5] M. Wang, W. Zeng, and C. Xiao, "Linear precoding for MIMO multiple access channels with discrete-constellation inputs," accepted by *IEEE Global Commun. Conf. (Globecom)*, Houston, TX, USA, Dec. 2011.

[6] W. Zeng, C. Xiao, M. Wang, and J. Lu, "On the linear precoder design for MIMO channels with finite-alphabet inputs and statistical CSI," accepted by *IEEE GLOBECOM*, Houston, TX, USA, Dec. 2011.

[7] M. Wang, C. Xiao, and W. Zeng, "Linear precoding for MIMO multiple access channels with finite discrete inputs," *IEEE Trans. Wireless Commun.*, vol. 10, pp 3934 - 3942, Nov. 2011.

[8] Y.Wu, M. Wang, C. Xiao, Z. Ding, and X. Gao, "Linear precoding for MIMO broadcast channels with finite-alphabet constraints," submitted to *IEEE Trans. Wireless Commun.*, Nov. 2011.

[9] S. Subedi, H. Lou, F. Ren, M. Wang, and Y. R. Zheng, "Validation of the triply selective fading channel model through a MIMO test bed and experimental results," in *Proc. IEEE Military Communi. Conf. (MILCOM)*, Nov. 2011.

[10] W. Zeng, C. Xiao, M. Wang, and J. Lu, "Linear precoding for relay networks with finite-alphabet constraints," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kyoto, Japan, Jun. 2011.

[11] W. Zeng, M. Wang, C. Xiao, and J. Lu, "On the power allocation for relay networks with finite-alphabet constraints," in *Proc. IEEE Global Commun. Conf. (Globecom)*, Miami, FL, USA, Dec. 2010.

VITA

Mingxi Wang received his B.S degree in Telecommunication Engineering from Beihang University (formerly: Beijing University of Aeronautics and Astronautics), China, in 2006, and the M.S degree in Signal and Information Processing from the same university in 2009. He began his Ph.D study in August 2009 at the 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. He received his Ph.D. degree in Electrical Engineering from Missouri University of Science and Technology in Aug. 2012.