Precoding with X-codes to Increase Capacity with Discrete Input Alphabets

Saif Khan Mohammed,* Emanuele Viterbo,† Yi Hong,† and Ananthanarayanan Chockalingam*
* Indian Institute of Science, India, saifind2007@yahoo.com, achockal@ece.iisc.ernet.in
† University of Calabria, Italy, yi.winnie.hong@gmail.com, viterbo@deis.unical.it

***Abstract*—We consider Gaussian multiple-input multiple-output (MIMO) channels with discrete input alphabets. We propose a non-diagonal precoder based on X-Codes in [1] to increase the mutual information. The MIMO channel is transformed into a set of parallel subchannels using Singular Value Decomposition (SVD) and X-codes are then used to pair the subchannels. X-Codes are fully characterized by the pairings and the 2×2 real rotation matrices for each pair (parameterized with a single angle). This precoding structure enables to express the total mutual information as a sum of the mutual information of all the pairs. The problem of finding the optimal precoder with the above structure, which maximizes the total mutual information, is equivalent to *i*) optimizing the rotation angle and the power allocation within each pair and *ii*) finding the optimal pairing and power allocation among the pairs. It is shown that the mutual information achieved with the proposed pairing scheme is very close to that achieved with the optimal precoder by Cruz *et al.*, and significantly better than mercury/waterfilling strategy by Lozano *et al.*. Our approach greatly simplifies both the precoder optimization and the detection complexity, making it suitable for practical applications.**

***Index Terms*—Mutual information, MIMO, OFDM, precoding, singular value decomposition, condition number.**

I. Introduction

Many modern communication channels are modeled as a Gaussian multiple-input multiple-output (MIMO) channel. Examples include multi-tone digital subscriber line (DSL), orthogonal frequency division multiplexing (OFDM) and multiple transmit-receive antenna systems with channel state information at transmitter (CSIT). It is known that the capacity of the Gaussian MIMO channel is achieved by beamforming a Gaussian input alphabet along the left and right singular vectors, resulting in a set of parallel Gaussian subchannels. Optimal power allocation between the subchannels is achieved by waterfilling [2]. In practice, the input alphabet is not Gaussian and it is generally chosen from a finite signal set.

We distinguish between two kinds of MIMO channels: *i*) *diagonal* (or parallel) channels and *ii*) *non-diagonal* channels.

For a diagonal MIMO channel with discrete input alphabets, assuming only power allocation on each subchannel (i.e. diagonal precoder), mercury/waterfilling was shown to be optimal in [3]. With discrete input alphabets, Cruz *et al.* ([4]) later showed that the optimal precoder is, however, non-diagonal, i.e., some precoding is performed across all the subchannels.

For a general non-diagonal Gaussian MIMO channel, it was also shown in [4] that the optimal precoder is non-diagonal. Such an optimal precoder is given by a fixed point equation, which requires a high complexity numeric evaluation. Since the precoder jointly codes all the n inputs, joint decoding is also required at the receiver. Thus, the decoding complexity can be very high, specially for large n, as in the case of DSL and OFDM applications. This motivates our quest for a practical low complexity precoding scheme achieving near optimal capacity.

In this paper, we consider a general MIMO channel and a non-diagonal precoder based on X-Codes [1]. The MIMO channel is transformed into a set of parallel subchannels using Singular Value Decomposition (SVD) and X-codes are then used to pair the subchannels. X-Codes are fully characterized by the pairings and the 2×2 real rotation matrices for each pair, where each matrix is parameterized with a single angle. This precoding structure enables to express the total mutual information as a sum of the mutual information of all the pairs.

The problem of finding the optimal precoder with the above structure, which maximizes the total mutual information, can be split into two tractable problems: *i*) optimizing the rotation angle and the power allocation within each pair and *ii*) finding the optimal pairing and power allocation among the pairs. It is shown by simulation that the mutual information achieved with the proposed pairing scheme is very close to that achieved with the optimal precoder in [4], and significantly better than the mercury/waterfilling strategy in [3]. Our approach greatly simplifies both the precoder optimization and the detection complexity, making it suitable for practical applications.

II. System model and Precoding

We consider a $n_t \times n_r$ MIMO channel, where the channel state information (CSI) is known perfectly at both transmitter and receiver. Let $\mathbf{x} = (x_1, \ldots, x_{n_t})^T$ be the vector of input symbols to the channel, and let $\mathbf{H} = \{h_{ij}\}$, $i = 1, \ldots, n_r$, $j = 1, \ldots, n_t$, be a full rank $n_r \times n_t$ channel coefficient matrix, with h_{ij} as the complex channel gain between the j-th input symbol and the i-th output symbol. The vector of n_r channel output symbols is given by

$$\mathbf{y} = \sqrt{P_T}\mathbf{H}\mathbf{x} + \mathbf{w} \tag{1}$$

*Saif K. Mohammed is currently visiting DEIS, Università della Calabria, Italy thanks to the Italian Ministry of University and Research (MIUR) collaborative research program: “Bando per borse a favore di giovani ricercatori indiani A.F. 2008”. This work was supported in part by the DRDO-IISc Program on Advanced Research in Mathematical Engineering.

where $\mathbf{w}$ is a uncorrelated Gaussian noise vector, such that $\mathbb{E}[\mathbf{w}\mathbf{w}^\dagger] = \mathbf{I}_{n_r}$, and P_T is the total transmitted power. The power constraint is given by $\mathbb{E}[\|\mathbf{x}\|^2] = 1$.

The maximum multiplexing gain of this channel is $n = \min(n_r, n_t)$. Let $\mathbf{u} = (u_1, \ldots, u_n)^T \in \mathbb{C}^n$ be the vector of n information symbols to be sent through the MIMO channel, with $\mathbb{E}[|u_i|^2] = 1, i = 1, \ldots, n$. Then $\mathbf{u}$ can be precoded using a $n_t \times n$ matrix $\mathbf{T}$, resulting in $\mathbf{x} = \mathbf{T}\mathbf{u}$.

The capacity of the deterministic Gaussian MIMO channel is then achieved by solving

Problem 1:

$$\begin{aligned} C(\mathbf{H}, P_T) &= \max_{\mathbf{K_x} | \mathrm{tr}(\mathbf{K_x})=1} I(\mathbf{x}; \mathbf{y}|\mathbf{H}) \\ &\geq \max_{\mathbf{K_u}, \mathbf{T} \,|\, \mathrm{tr}(\mathbf{T}\mathbf{K_u}\mathbf{T}^\dagger)=1} I(\mathbf{u}; \mathbf{y}|\mathbf{H}) \end{aligned} \tag{2}$$

where $I(\mathbf{x}; \mathbf{y}|\mathbf{H})$ is the mutual information between $\mathbf{x}$ and $\mathbf{y}$. $\mathbf{K_x} \triangleq \mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]$ and $\mathbf{K_u} \triangleq \mathbb{E}[\mathbf{u}\mathbf{u}^\dagger]$ are the covariance matrices of $\mathbf{x}$ and $\mathbf{u}$ respectively. The inequality in (2) follows from the data processing inequality [2].

Let us consider the singular value decomposition (SVD) of the channel $\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}$, where $\mathbf{U} \in \mathbb{C}^{n_r \times n}$, $\mathbf{\Lambda} \in \mathbb{C}^{n \times n}$, $\mathbf{V} \in \mathbb{C}^{n \times n_t}$, and $\mathbf{U}^\dagger\mathbf{U} = \mathbf{V}\mathbf{V}^\dagger = \mathbf{I}_n$, and $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with $\lambda_1 \geq \lambda_2, \cdots, \geq \lambda_n \geq 0$.

In [2] it is shown that Problem 1 is solved when $\mathbf{x}$ is Gaussian distributed and $\mathbf{V}\mathbf{K_x}\mathbf{V}^\dagger$ is diagonal. This can be achieved by using the optimal precoder matrix $\mathbf{T} = \mathbf{V}^\dagger\mathbf{P}$, where $\mathbf{P} \in (\mathbb{R}^+)^n$ is the diagonal power allocation matrix such that $\mathrm{tr}(\mathbf{P}\mathbf{P}^\dagger) = 1$. Further $u_i, i = 1, \ldots, n$ are i.i.d. Gaussian components. Optimal power allocation is achieved through waterfilling between the n parallel channels. This also implies that, the second line of (2) is actually an equality.

III. Optimal precoding with discrete inputs

In practice, discrete input alphabets are used. Subsequently, we assume that the i-th information symbol is given by $u_i \in \mathcal{U}_i$, where $\mathcal{U}_i \subset \mathbb{C}$ is a finite signal set. Let $\mathcal{S} \triangleq \mathcal{U}_1 \times \mathcal{U}_2 \times \cdots \times \mathcal{U}_n$. The capacity of the Gaussian MIMO channel with discrete input alphabet $\mathcal{S}$ is achieved by solving

Problem 2:

$$C_{\mathcal{S}}(\mathbf{H}, P_T) = \max_{\mathbf{T} \,|\, \mathbf{u} \in \mathcal{S}, \|\mathbf{T}\|_F = 1} I(\mathbf{u}; \mathbf{y}|\mathbf{H}) \tag{3}$$

Note that there is no maximization over the pdf of $\mathbf{u}$, since we fix $\mathbf{K_u} = \mathbf{I}_n$. The optimal precoder $\mathbf{T}^*$, which solves Problem 2, is given by the following fixed point equation [4]

$$\mathbf{T}^* = \frac{\mathbf{H}^\dagger\mathbf{H}\mathbf{T}^*\mathbf{E}}{\|\mathbf{H}^\dagger\mathbf{H}\mathbf{T}^*\mathbf{E}\|_F} \tag{4}$$

where $\mathbf{E}$ is the minimum mean-square error (MMSE) matrix of $\mathbf{u}$ given by

$$\mathbf{E} = \mathbb{E}[(\mathbf{u} - \mathbb{E}[\mathbf{u}|\mathbf{y}])(\mathbf{u} - \mathbb{E}[\mathbf{u}|\mathbf{y}])^\dagger] \tag{5}$$

It is observed that with discrete input alphabets, it is not optimal to beamform along the column vectors of $\mathbf{V}^\dagger$ and then use waterfilling on the parallel subchannels. Even when $\mathbf{H}$ is diagonal (parallel non-interfering subchannels), the optimal precoder $\mathbf{T}^*$ is *non diagonal*, and can be computed numerically (using a gradient based method) as discussed in [4]. However, the complexity is prohibitive for practical applications. In order to overcome this complexity issue, we propose to use a simple precoder based on the X-Codes [1].

IV. Precoding with X-Codes

X-Codes are based on a pairing of n subchannels $l = \{(i_k, j_k) \in [1, n] \times [1, n], i_k < j_k, k = 1, \ldots, n/2\}$. For a given n, there are $(n-1)(n-3)\cdots 3\,1$ possible pairings. Let $\mathcal{L}$ denote the set of all possible pairings. For example, with $n = 4$, we have $\mathcal{L} = \{\{(1,4),(2,3)\}, \{(1,2),(3,4)\}, \{(1,3),(2,4)\}\}$.

X-Codes are generated by a $n \times n$ real orthogonal matrix, denoted by $\mathbf{G}$. When precoding with X-Codes, the precoder matrix is given by $\mathbf{T} = \mathbf{V}^\dagger\mathbf{P}\mathbf{G}$, where $\mathbf{P} = \mathrm{diag}(p_1, p_2, \cdots, p_n) \in {\mathbb{R}^+}^n$ is the diagonal power allocation matrix such that $\mathrm{tr}(\mathbf{P}\mathbf{P}^\dagger) = 1$. The k-th pair consists of subchannels i_k and j_k. For the k-th pair, the information symbols u_{i_k} and u_{j_k} are jointly coded using a 2×2 real orthogonal matrix $\mathbf{A}_k$ given by

$$\mathbf{A}_k = \begin{bmatrix} \cos(\theta_k) & \sin(\theta_k) \\ -\sin(\theta_k) & \cos(\theta_k) \end{bmatrix} \quad k = 1, \ldots n/2 \tag{6}$$

The angle θ_k can be optimized by maximizing the mutual information for the k-th pair. Each $\mathbf{A}_k$ is a submatrix of the code matrix $\mathbf{G}$ as shown below

$$\begin{array}{ll} g_{i_k,i_k} = \cos(\theta_k) & g_{i_k,j_k} = \sin(\theta_k) \\ g_{j_k,i_k} = -\sin(\theta_k) & g_{j_k,j_k} = \cos(\theta_k) \end{array} \tag{7}$$

where $g_{i,j}$ is the entry of $\mathbf{G}$ in the i-th row and j-th column. It was shown in [1], that for achieving the best diversity gain, an optimal pairing is one in which the k-th subchannel is paired with the $(n-k+1)$-th subchannel. For example, with this pairing and $n = 6$, the X-Code structure is given by

$$\mathbf{G} = \begin{bmatrix} \cos(\theta_1) & & & & & \sin(\theta_1) \\ & \cos(\theta_2) & & & \sin(\theta_2) & \\ & & \cos(\theta_3) & \sin(\theta_3) & & \\ & & -\sin(\theta_3) & \cos(\theta_3) & & \\ & -\sin(\theta_2) & & & \cos(\theta_2) & \\ -\sin(\theta_1) & & & & & \cos(\theta_1) \end{bmatrix}$$

The special case with $\theta_k = 0, k = 1, 2, \cdots, n/2$ results in no coding across subchannels. Given the encoder matrix $\mathbf{G}$, the subchannel gains $\mathbf{\Lambda}$, and the power allocation matrix $\mathbf{P}$, the mutual information between $\mathbf{u}$ and $\mathbf{y}$ is given by

$$\begin{aligned} I_{\mathcal{S}}(\mathbf{u}; \mathbf{y}|\mathbf{\Lambda}, \mathbf{P}, \mathbf{G}) &= h(\mathbf{y}|\mathbf{\Lambda}, \mathbf{P}, \mathbf{G}) - h(\mathbf{w}) \\ &= -\int_{\mathbf{y} \in \mathbb{C}^{n_r}} p(\mathbf{y}|\mathbf{\Lambda}, \mathbf{P}, \mathbf{G}) \log_2(p(\mathbf{y}|\mathbf{\Lambda}, \mathbf{P}, \mathbf{G})) d\mathbf{y} - n\log_2(\pi e) \end{aligned} \tag{8}$$

where the received vector pdf is given by

$$p(\mathbf{y}|\mathbf{\Lambda}, \mathbf{P}, \mathbf{G}) = \frac{1}{|\mathcal{S}|\pi^n} \sum_{\mathbf{u} \in \mathcal{S}} e^{-\|\mathbf{y} - \mathbf{U}\mathbf{\Lambda}\mathbf{P}\mathbf{G}\mathbf{u}\|^2} \tag{9}$$

and when $n = n_r$ (i.e., $n_r \leq n_t$), it is equivalently given by

$$p(\mathbf{y}|\mathbf{\Lambda}, \mathbf{P}, \mathbf{G}) = \frac{1}{|\mathcal{S}|\pi^n} \sum_{\mathbf{u} \in \mathcal{S}} e^{-\|\mathbf{r} - \mathbf{\Lambda}\mathbf{P}\mathbf{G}\mathbf{u}\|^2} \tag{10}$$

where $\mathbf{r} = (r_1, r_2, \cdots, r_n)^T \triangleq \mathbf{U}^\dagger \mathbf{y}$.

We next define the capacity of the MIMO Gaussian channel when precoding with X-Codes and further receiver processing[2] is done with the vector $\mathbf{r}$. Towards this, we need the following definitions. For a given pairing l, let $\mathbf{r}_k \triangleq (r_{i_k}, r_{j_k})^T$, $\mathbf{u}_k \triangleq (u_{i_k}, u_{j_k})^T$, $\mathbf{\Lambda}_k \triangleq \text{diag}(\lambda_{i_k}, \lambda_{j_k})$, $\mathbf{P}_k \triangleq \text{diag}(p_{i_k}, p_{j_k})$ and $\mathcal{S}_k \triangleq \mathcal{U}_{i_k} \times \mathcal{U}_{j_k}$. Due to the pairing structure of $\mathbf{G}$, the mutual information $I_\mathcal{S}(\mathbf{u};\mathbf{r}|\mathbf{\Lambda},\mathbf{P},\mathbf{G})$ can be expressed as the sum of mutual information of all the $n/2$ pairs as follows:

$$I_\mathcal{S}(\mathbf{u};\mathbf{r}|\mathbf{\Lambda},\mathbf{P},\mathbf{G}) \;=\; \sum_{k=1}^{n/2} I_{\mathcal{S}_k}(\mathbf{u}_k;\mathbf{r}_k|\mathbf{\Lambda}_k,\mathbf{P}_k,\theta_k) \quad (11)$$

Having fixed the precoder structure to $\mathbf{T} = \mathbf{V}^\dagger\mathbf{P}\mathbf{G}$, we can then formulate the following

Problem 3:

$$C_X(\mathbf{H}, P_T) = \max_{\mathbf{G},\mathbf{P}\,|\,\mathbf{u}\in\mathcal{S},\text{tr}(\mathbf{P}\mathbf{P}^\dagger)=1} I_\mathcal{S}(\mathbf{u};\mathbf{r}|\mathbf{\Lambda},\mathbf{P},\mathbf{G}) \quad (12)$$

It is clear that the solution of the above problem is still a formidable task, although apparently simpler than Problem 2. In fact, instead of the $n\times n$ variables of $\mathbf{T}$, we now deal with n variables for power allocation in $\mathbf{P}$, $n/2$ variables for the angles defining $\mathbf{A}_k$, and the pairing selection $l \in \mathcal{L}$. In the following, we will show how to efficiently solve Problem 3 by splitting it into two simpler problems.

Power allocation can be divided into power allocation between pairs, followed by power allocation between the 2 subchannels of each pair. Let $\bar{\mathbf{P}} = \text{diag}(\bar{p}_1, \bar{p}_2, \cdots, \bar{p}_{n/2})$ be a diagonal matrix, where $\bar{p}_k \triangleq \sqrt{p_{i_k}^2 + p_{j_k}^2}$ with $\bar{p}_k^2$ being the power allocated to the k-th pair. The power allocation within each pair can be simply expressed in terms of the fraction $f_k \triangleq p_{i_k}^2/\bar{p}_k^2$ of the power assigned to the first subchannel of the pair. The mutual information achieved by the k-th pair is then given by

$$\begin{aligned} I_{\mathcal{S}_k}(\mathbf{u}_k;\mathbf{r}_k|\mathbf{\Lambda}_k,\mathbf{P}_k,\theta_k) &= I_{\mathcal{S}_k}(\mathbf{u}_k;\mathbf{r}_k|\mathbf{\Lambda}_k,\bar{p}_k,f_k,\theta_k) \quad (13)\\ &= -\int_{\mathbf{r}_k\in\mathbb{C}^2} p(\mathbf{r}_k)\log_2 p(\mathbf{r}_k)\, d\mathbf{r}_k - 2\log_2(\pi e) \end{aligned}$$

where $p(\mathbf{r}_k)$ is given by

$$p(\mathbf{r}_k) = \frac{1}{|\mathcal{S}_k|\pi^2} \sum_{\mathbf{u}_k\in\mathcal{S}_k} e^{-\|\mathbf{r}_k - \bar{p}_k\mathbf{\Lambda}_k\mathbf{F}_k\mathbf{A}_k\mathbf{u}_k\|^2} \quad (14)$$

where $\mathbf{F}_k \triangleq \text{diag}(\sqrt{f_k}, \sqrt{1-f_k})$ and $\mathbf{A}_k$ is given by (6).

The capacity of the discrete input MIMO Gaussian channel when precoding with X-Codes can be expressed as

Problem 4:

$$C_X(\mathbf{H}, P_T) = \max_{l\in\mathcal{L},\bar{\mathbf{P}}|\text{tr}(\bar{\mathbf{P}}\bar{\mathbf{P}}^\dagger)=1} \sum_{k=1}^{n/2} C_{\mathcal{S}_k}(k, l, \bar{p}_k) \quad (15)$$

[2]For $n_r \le n_t$, $I(\mathbf{u};\mathbf{y}|\mathbf{H}) = I(\mathbf{u};\mathbf{r}|\mathbf{H})$. However, when $n_r > n_t$, receiver processing with $\mathbf{r}$ becomes information lossy, and $I(\mathbf{u};\mathbf{y}|\mathbf{H}) > I(\mathbf{u};\mathbf{r}|\mathbf{H})$.

where $C_{\mathcal{S}_k}(k, l, \bar{p}_k)$, the capacity of the k-th pair, is achieved by solving

Problem 5:

$$C_{\mathcal{S}_k}(k, l, \bar{p}_k) = \max_{\theta_k, f_k} I_{\mathcal{S}_k}(\mathbf{u}_k;\mathbf{r}_k|\mathbf{\Lambda}_k,\bar{p}_k,f_k,\theta_k) \quad (16)$$

In other words, we have split Problem 3 into two simpler problems. Firstly, given a pairing l and power allocation between pairs $\bar{\mathbf{P}}$, we can solve Problem 5 for each $k = 1, 2, \cdots, n/2$. Problem 4 uses the solution to Problem 5 to find the optimal pairing l^* and the optimal power allocation $\bar{\mathbf{P}}^*$ between the $n/2$ pairs. For small n, the optimal pairing and power allocation between pairs can always be computed by brute force enumeration of all possible pairings. This is, however, prohibitively complex for large n, and we shall discuss heuristic approaches in Section VI. We will show in the following that, although suboptimal, precoding with X-Codes will perform close to the optimal capacity in [4]. The additional benefit of our scheme is that the detection complexity at the receiver is highly reduced, since there is coupling only between pairs of channels, as compared to the case of full-coupling for the optimal precoder in [4].

In the next section, we solve Problem 5. Since this problem is the same for each pair, it is equivalent to finding the optimal rotation angle and power allocation for a Gaussian MIMO channel with only $n = 2$ subchannels.

V. Gaussian MIMO channels with $n = 2$

With $n = 2$, there is only one pair and only one possible pairing. Therefore we drop the subscript k in Problem 5 and we find $C_X(\mathbf{H}, P_T)$ in Problem 3. The processed received vector $\mathbf{r} \in \mathbb{C}^2$ is given by

$$\mathbf{r} = \sqrt{P_T}\mathbf{\Lambda}\mathbf{F}\mathbf{A}\mathbf{u} + \mathbf{z} \quad (17)$$

where $\mathbf{z} = \mathbf{U}^\dagger\mathbf{w}$ is the equivalent noise vector with the same statistics as $\mathbf{w}$. Let $\alpha \triangleq \lambda_1^2 + \lambda_2^2$ be the overall channel power gain and $\beta \triangleq \lambda_1/\lambda_2$ be the *condition number* of the channel. Then (17) can be re-written as

$$\mathbf{r} = \sqrt{\tilde{P}_T}\tilde{\mathbf{\Lambda}}\mathbf{F}\mathbf{A}\mathbf{u} + \mathbf{z} \quad (18)$$

where $\tilde{P}_T \triangleq P_T\alpha$ and $\tilde{\mathbf{\Lambda}} \triangleq \mathbf{\Lambda}/\sqrt{\alpha} = \text{diag}(\beta/\sqrt{1+\beta^2}, 1/\sqrt{1+\beta^2})$. The equivalent channel $\tilde{\mathbf{\Lambda}}$ now has a gain of 1, and its channel gains are dependent only upon β. Our goal is, therefore, to find the optimal rotation angle θ^* and the fractional power allocation f^*, which maximize the mutual information of the equivalent channel with condition number β and gain equal to 1. The total available transmit power is now $\tilde{P}_T$.

It is difficult to get analytic expressions for the optimal θ^* and f^*, and therefore we can use numerical techniques to evaluate them and store them in lookup tables. For a given application scenario, given the distribution of β, we decide upon a few discrete values of β which are representative of the actual values observed in real channels. For each such quantized value of β and a given discrete input alphabet, we

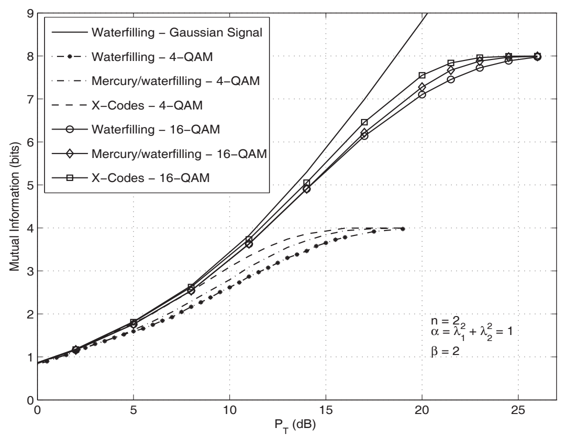

Fig. 1. Mutual information versus P_T for $n = 2$ parallel channels with $\beta = 2$ and $\alpha = 1$ for 4- and 16-QAM.

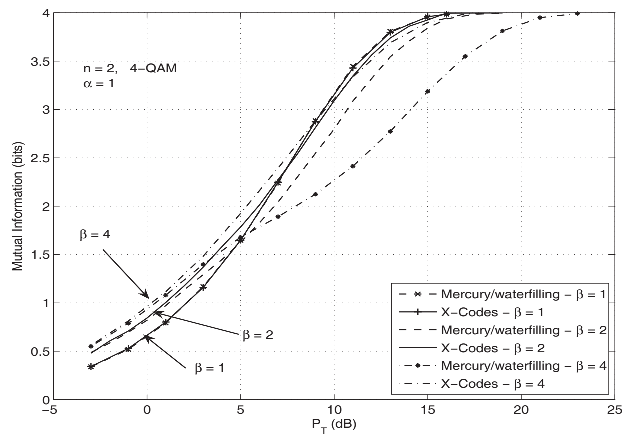

Fig. 2. Mutual information versus P_T for $n = 2$ parallel channels with varying $\beta = 1, 2, 4$. $\alpha = 1$. Input alphabet is 4-QAM.

numerically compute a table of the optimal values f^* and θ^* as a function of $\tilde{P}_T$. These tables are constructed offline. During the process of communication, the transmitter knows the value of α and β from channel measurements. It then finds the lookup table corresponding to the measured β. The optimal values f^* and θ^* are then found by indexing the appropriate entry in the table with $\tilde{P}_T = P_T\alpha$.

We next present some simulation results to show that indeed our simple precoding scheme can significantly increase the mutual information, compared to mercury/waterfilling (i.e., without precoding across subchannels). For the sake of comparison, we also present the mutual information achieved by the waterfilling scheme with discrete input alphabets.

We restrict the discrete input alphabets $\mathcal{U}_i, i = 1, 2$, to be square M-QAM alphabets consisting of two $\sqrt{M}$-PAM alphabets in quadrature. Mutual information is evaluated by solving Problem 5, i.e., numerically maximizing w.r.t. the rotation angle and power allocation.

In Fig. 1, we plot the maximal mutual information versus P_T, for a system with two subchannels, $\beta = 2$ and $\alpha = 1$. Mutual information is plotted for 4- and 16-QAM signal sets. It is observed that *for a given achievable mutual information, coding across subchannels is more power efficient*. With 4-QAM and an achievable mutual information of 3 bits, X-Codes require only 0.8 dB more transmit power when compared to the ideal Gaussian signalling with waterfilling. This gap increases to 1.9 dB for mercury/waterfilling and 2.8 dB for the waterfilling scheme with 4-QAM as the input alphabet. A similar trend is observed with 16-QAM as the input alphabet. The proposed precoder clearly performs better, since the mutual information is optimized w.r.t. the rotation angle θ and power allocation, while mercury/waterfilling, as a special case of X-Code, only optimizes power allocation and fixes $\theta = 0$.

In Fig. 2, we compare the mutual information achieved by X-Codes and the mercury/waterfilling strategy for $\alpha = 1$ and $\beta = 1, 2, 4$. Input alphabet is 4-QAM. It is observed that both the schemes have the same mutual information when $\beta = 1$. However with increasing β, the mutual information of mercury/waterfilling strategy is observed to degrade significantly at high values of P_T, whereas the performance of X-Codes does not vary as much. For example, for a target mutual information of 3 bits, with $\beta = 4$ the mercury/waterfilling strategy requires 4.6 dB more transmit power as compared to when $\beta = 1$. On the other hand, X-Codes require only 0.1 dB extra transmit power when β increases from 1 to 4. The degradation of mutual information for the mercury/waterfilling strategy is explained as follows. For the mercury/waterfilling strategy, with increasing β, all the available power is allocated to the stronger channel till a certain transmit power threshold. However, since finite signal sets are used, mutual information is bounded from above until the transmit power exceeds this threshold. This also explains the reason for the intermediate change of slope in the mutual information curve with $\beta = 4$ (see the rightmost dash-dot curve in Fig. 2). On the other hand, this problem does not arise when coding across subchannels with X-Codes. Therefore, in terms of achievable mutual information, rotation coding is observed to be more robust to ill-conditioned channels.

For low values of P_T, mutual information of both the schemes are similar, and improves with increasing β. This is due to the fact that, at low P_T, mutual information increases linearly with P_T, and therefore all power is assigned to the stronger channel. With increasing β, the stronger channel has an increasing fraction of the total channel gain, which results in increased mutual information.

For a given α and β, we observed that mutual information is sensitive to the rotation angle except at very low P_T. We also observed that the optimal angle θ^* does not vary significantly with P_T.

VI. Gaussian MIMO channels with $n > 2$

We now consider the problem of finding the optimal pairing and power allocation between pairs for Gaussian MIMO channels with even $n > 2$. We also compare the performance achieved by X-Codes with that of the optimal precoder in [4]. We first observe that mutual information is indeed sensitive

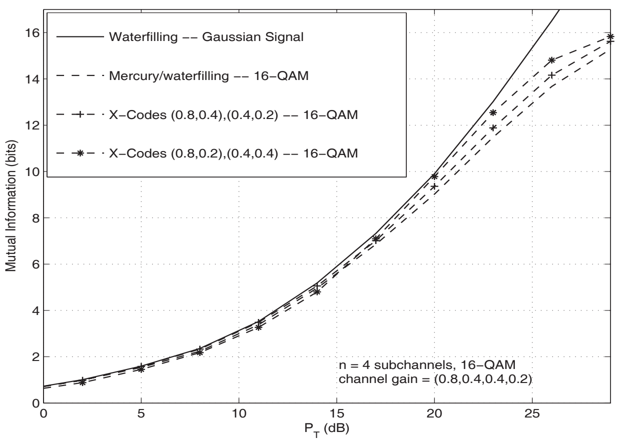

Fig. 3. Mutual information versus P_T for X-Codes with two different pairings. $n = 4$ and 16-QAM signal set.

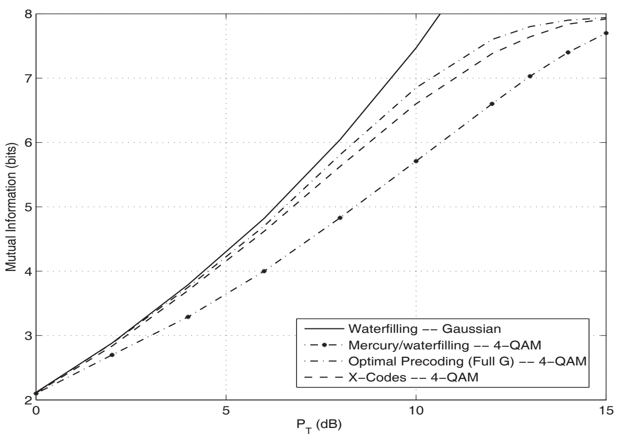

Fig. 4. Mutual information versus P_T for the Gigabit DSL channel given by (42) in [4].

to the chosen pairing, and this therefore justifies the criticality of computing the optimal pairing. This is illustrated through Fig. 3, for $n = 4$, $\mathbf{\Lambda} = \text{diag}(0.8, 0.4, 0.4, 0.2)$ and 16-QAM. Optimal power allocation between the two pairs is computed numerically. It is observed that the pairing $\{(1, 4), (2, 3)\}$ performs significantly better than the pairing $\{(1, 3), (2, 4)\}$.

In Fig. 4, we compare the mutual information achieved with optimal precoding [4], to that achieved by the proposed precoder with 4-QAM input alphabet. The 4×4 channel is given by (42) in [4]. For X-Codes, the optimal pairing is $\{(1, 4), (2, 3)\}$ and the optimal power allocation between the pairs is computed numerically. It is observed that X-Codes perform very close to the optimal precoding scheme. Specifically, for an achievable mutual information of 6 bits, compared to the optimal precoder [4], X-Codes need only 0.4 dB extra power, whereas 2.3 dB extra power is required with mercury/waterfilling.

In OFDM applications, n is large and Problem 4 becomes too complex to solve, since we can no more find the optimal pairing by enumeration. It was observed in section V, that for $n = 2$, a larger value of the condition number β leads to a

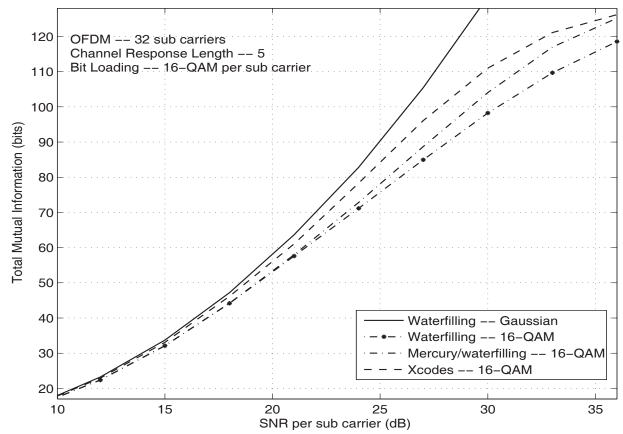

Fig. 5. Mutual information versus per subcarrier SNR for an OFDM system with 32 carriers. Channel impulse response = $[-0.454 + \text{j}0.145, -0.258 + \text{j}0.198, 0.0783 + \text{j}0.069, -0.408 - \text{j}0.396, -0.532 - \text{j}0.224]$

higher mutual information for low and medium values of P_T. Therefore, we conjecture that pairing the k-th subchannel with the $(n/2 + k)$-th subchannel could have mutual information very close to optimal, since this pairing scheme attempts to maximize the minimum β among all pairs. Given a pairing of subchannels, it is also difficult to compute the optimal power allocation between pairs. However, it was observed that for channels with large n, even waterfilling power allocation among the pairs results in good performance. We illustrate this in Fig. 5, for an OFDM system with $n = 32$ subchannels and 16-QAM, where the proposed precoder is used with the conjectured pairing scheme and waterfilling power allocation between pairs. The proposed precoder performs within 1.3 dB of the Gaussian signalling scheme for an achievable total mutual information of 96 bits. The proposed precoder performs about 1.5 dB better than the mercury/waterfilling scheme.

VII. Conclusions

In this paper, we proposed a *low complexity* precoding scheme based on X-Codes, which achieves near optimal capacity for Gaussian MIMO channels with discrete inputs. For large n, typical in OFDM applications, we also presented a heuristic approach for optimizing the pairing of subchannels. The proposed precoder was shown to perform better than the mercury/waterfilling strategy. Future work will focus on finding close to optimal pairings, and close to optimal power allocation strategies between pairs.

References

[1] S.K. Mohammed, E. Viterbo, Y. Hong, and A. Chockalingam, "MIMO Precoding with X- and Y-Codes," *submitted to IEEE Trans. on Information Theory*, Nov 2009 (available at http://arxiv.org/abs/0912.1909v1).

[2] T.M. Cover and Joy A. Thomas, *Elements of information theory*, John Wiley and Sons, 2nd Ed., July 2006.

[3] A. Lozano, A.M. Tulino, and S. Verdu, "Optimum Power Allocation for Parallel Gaussian Channels With Arbitrary Input Distributions," *IEEE. Trans. on Information Theory*, pp. 3033–3051, vol. 52, no. 7, July 2006.

[4] F.P. Cruz, M.R.D. Rodrigues and S. Verdu, "MIMO Gaussian Channels with Arbitrary Inputs: Optimal Precoding and Power Allocation," accepted in *IEEE Trans. on Information Theory*.