MIMO Linear Precoders with Reduced Complexity

Thomas Ketseoglou, *Senior Member, IEEE,* and Ender Ayanoglu, *Fellow, IEEE*

Abstract—In this paper, the problem of designing a linear precoder for Multiple-Input Multiple-Output (MIMO) systems employing Low-Density Parity-Check (LDPC) codes is addressed under the constraint of independent receiving virtual antenna output groups, thus reducing the relevant transmitter and receiver complexities. Our approach constitutes an interesting generalization of Bit-Interleaved Coded Multiple Beamforming (BICMB) which has shown many benefits in MIMO systems. After examining the properties associated with independent output grouping, we proceed to propose an alternative, practical technique, called Per-Group Precoding (PGP), which groups together multiple input symbol streams and corresponding receiving branches in the "virtual" channel domain (after singular value decomposition of the original MIMO channel), and thus results in independent transmitting/receiving streams between groups. We show with numerical results that PGP offers almost optimal performance, albeit with significant reduction both in the precoder optimization and LDPC EXIT chart based decoding complexities.

I. Introduction

The concept of Multiple-Input Multiple-Output (MIMO) still represents a prevailing research direction in wireless communications due to its ever increasing capability to offer higher rate, more efficient communications, as measured by spectral utilization, and under low transmitting or receiving power. Within MIMO research, BICMB [1]–[3] has shown great potential for practical application, due to its excellent diversity gains and its simplicity. For example, BICMB in conjunction with convolutional coding offers maximum diversity and maximum spatial multiplexing simultaneously [1], thus it represents an optimal technique for this type of Forward Error Correction (FEC). In addition, there are many past works available which investigated with success linear precoding through exploitation of a unitary precoding matrix, mainly from a diversity maximization point of view [4], [5]. On the other hand, LDPC coding is the currently prevailing, near-capacity achieving error-correction technique that operates based on input to output mutual information and extrinsic information transfer (EXIT) charts [6], [7]. The problem of designing an optimal linear precoder toward maximizing the mutual information between the input and output was first considered in [8], [9] where the first optimal power allocation strategies are presented (e.g., Mercury Waterfilling (MWF)), together with general equations for the optimal precoder design. In addition, [10] also considered precoders for mutual information maximization and showed that the left eigenvectors of the optimal precoder can be set equal to the

right eigenvectors of the channel. Finally, in [11], a mutual information maximizing precoder for a parallel layer MIMO detection system is presented reducing the performance gap between maximum likelihood and parallel layer detection.

Recently, globally optimal linear precoding techniques were presented [12], [13] for finite alphabet inputs, capable of achieving mutual information rates much higher than the previously presented MWF [8] techniques, by introducing input symbol correlation through a unitary input transformation matrix in conjunction with channel weight adjustment (power allocation). These mutual information maximizing globally optimal precoders are more appropriate for LDPC codes which are very popular currently, than e.g., Maximal Diversity Precoders (MDP) [12]. However, the gains presented in [12], [14], [15] are achieved at the expense of significantly increased system complexity, even for small modulation constellation size M (e.g., $M = 2,\ 4$). In addition, the interesting design of [12] requires a significant computational complexity increase at the receiver, even in its simplified implementation for M-ary Phase Shift Keying (MPSK) systems [14], due to e.g., the dependence present between receiving branches. This increase could be prohibitive if the receiver is the mobile destination, or if the number of receiving branches is high.

In this paper, we propose linear precoding techniques which offer high mutual information between input and output in a MIMO system with Quadrature Amplitude Modulation (QAM) and also offer independence among the receiving branches, thus highly simplifying the Maximum A Posteriori probability (MAP) detector operation, and hence significantly reducing the receiver complexity. We then proceed to propose a new, interesting technique that groups together "similar" small numbers of multiple streams of input data and receiving branches and then it applies optimized precoding on each group. We carefully look at the group selection strategy, and we show that the best selection is based on selecting input and output groups based on maximum separation of their singular values. The proposed technique is named Per Group Precoding (PGP). PGP offers very good performance with significantly reduced complexity both at the precoder design and receiver levels, due to the independence among different groups and it can be successfully applied even to QAM constellations with $M \geq 16$.

II. Linear Precoder Optimization with Reduced Complexity

The N_t transmit antenna, N_r receive antenna MIMO model (Fig. 1) is described by the following equation

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

where $\mathbf{y}$ is the $N_r \times 1$ received vector, $\mathbf{H}$ is the $N_r \times N_t$ MIMO channel matrix comprising independent complex Gaussian

T. Ketseoglou is with the Electrical and Computer Engineering Department, California State Polytechnic University, Pomona, California (e-mail: tketseoglou@csupomona.edu). E. Ayanoglu is with with the Center for Pervasive Communications and Computing, Department of Electrical Engineering and Computer Science, University of California, Irvine (e-mail: ayanoglu@uci.edu).

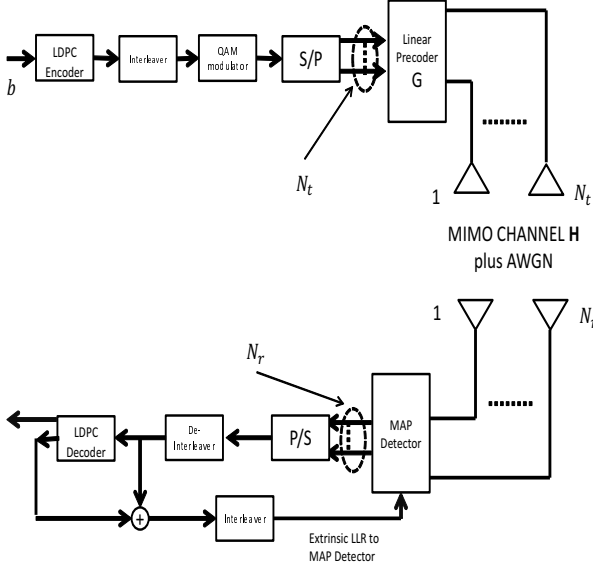components of mean zero and variance one and assumed

Fig. 1. Original MIMO System model at baseband with Linear Precoder $\mathbf{G}$, N_t transmitting antennas, N_r receiving antennas, and LDPC encoder and decoder. A generic coding rate R is assumed for the LDPC encoder. A similar system can be built for the equivalent channel presented.

quasi-static [1], $\mathbf{G}$ is the precoder matrix of size $N_t \times N_t$, $\mathbf{x}$ is the $N_t \times 1$ data vector with independent, identically distributed components of (normalized) power one (thus with covariance matrix $\mathbf{K}_X = \mathbf{I}_{N_t}$), each of which is in the QAM constellation, and $\mathbf{n}$ represents the circularly symmetric complex Additive White Gaussian Noise (AWGN) of size $N_r \times 1$, with mean zero and covariance matrix $\mathbf{K}_n = \sigma_n^2 \mathbf{I}_{N_r}$, where $\mathbf{I}_{N_r}$ is the $N_r \times N_r$ identity matrix, and $\sigma_n^2 = \frac{1}{SNR}$, SNR being the (coded) symbol signal-to-noise ratio. The precoding matrix $\mathbf{G}$ needs to satisfy the following power constraint

$$\mathrm{tr}(\mathbf{G}\mathbf{G}^h) = N_t, \qquad (2)$$

where $\mathrm{tr}(\mathbf{A})$, $\mathbf{A}^h$ denote the trace, and the Hermitian transpose of matrix $\mathbf{A}$, respectively.

An equivalent model, assuming without loss of generality $N_{tv} = N_{rv} = N_t$ in the virtual domain (after singular value decomposition (SVD)[1], where N_{tv}, N_{rv} represent the number of transmitting and receiving antennas in the virtual domain, respectively, called herein the "virtual" channel can be easily built based on [12] as follows

$$\mathbf{y} = \mathbf{\Sigma}_H \mathbf{\Sigma}_G \mathbf{V}_G^h \mathbf{x} + \mathbf{n}, \qquad (3)$$

where $\mathbf{\Sigma}_H$ and $\mathbf{\Sigma}_G$ are diagonal matrices containing *non-zero* singular values of $\mathbf{H}$, $\mathbf{G}$, respectively, padded with zeroes if necessary for dimension consistency, and $\mathbf{V}_G$ is the matrix of the right singular vectors of $\mathbf{G}$. Thus, all matrices in (3) are of size $N_t \times N_t$, while the vectors are of size $N_t \times 1$. When LDPC coding with sufficient blocklength (see below) is employed in this MIMO system, the overall utilization in $b/s/Hz$ is determined by the mutual information between the

transmitting branches $\mathbf{x}$ and the receiving ones, $\mathbf{y}$ [6], [7]. It is shown [12] that the mutual information between $\mathbf{x}$ and $\mathbf{y}$, $I(\mathbf{x}; \mathbf{y})$, is only a function of $\mathbf{W} = \mathbf{V}_G \mathbf{\Sigma}_H^2 \mathbf{\Sigma}_G^2 \mathbf{V}_G^h$, an important property toward mutual information maximization. The optimal precoder $\mathbf{G}$ is found by solving

$$\begin{aligned} \underset{\mathbf{G}}{\text{maximize}} \quad & I(\mathbf{x}; \mathbf{y}) \\ \text{subject to} \quad & \mathrm{tr}(\mathbf{G}\mathbf{G}^h) = N_t. \end{aligned} \qquad (4)$$

In (4), the constraint present is due to the total (normalized) average MIMO input power which needs to be kept equal to N_t. The average MIMO input power is given as $P_{MIMO} = \mathbb{E}\left(\mathrm{tr}(\mathbf{G}\mathbf{x}\mathbf{x}^h\mathbf{G}^h)\right) = \mathrm{tr}(\mathbf{G}\mathbf{G}^h)$, due to the assumptions made on $\mathbf{x}$. The solution to (4) results in exponential complexity at both transmitter and receiver, as shown in Section II.

When an appropriate LDPC code[2] of sufficient blocklength N_b is employed in the described MIMO system, in conjunction with Gray coding and interleaving employed at the transmitter, followed by MAP detector at the receiver, the system offers very low bit error rate (BER) (e.g., $BER < 10^{-4}$) [6], [7], provided that the coding rate of the LDPC code satisfies the condition

$$R < \frac{I(\mathbf{x}, \mathbf{y})}{N_t \log_2(M)}. \qquad (5)$$

This is due to the fact that LDPC codes are near-capacity achieving codes. For example, based on the published results in [12], a blocklength $N_b \geq 2400$ would suffice toward meeting (5) closely for a 2×2 MIMO system. Thus, designing precoders for high input-output mutual information is more appropriate for LDPC systems than other type of precoders, e.g., MDP. Based on this fact, we focus on this type of precoder designs without special attention on the LDPC code design details.

A. Independent Group Partitions in MIMO Systems

An input-output MIMO Independent Group Partition (IGP) with N_g groups, $\mathcal{P}$, is defined as follows: $\mathcal{P} = \{\mathcal{S}_1, \cdots, \mathcal{S}_{N_g}\}$, where each group in the partition, $\mathcal{S}_i$ with $1 \leq i \leq N_g$, comprises N_{gi} unique, non-intersecting input-output vector pairs, $(\mathbf{x}_{\mathcal{S}_i}, \mathbf{y}_{\mathcal{S}_i})$, which satisfy $\cup_{i=1}^{N_g} \mathcal{N}(\mathbf{x}_{\mathcal{S}_i}) = \cup_{i=1}^{N_g} \mathcal{N}(\mathbf{y}_{\mathcal{S}_i}) = \{1, 2, \cdots, N_t\}$ and $\mathcal{N}(\mathbf{x}_{\mathcal{S}_i}) \cap_{j \neq i} \mathcal{N}(\mathbf{x}_{\mathcal{S}_j}) = \mathcal{N}(\mathbf{y}_{\mathcal{S}_i}) \cap_{j \neq i} \mathcal{N}(\mathbf{y}_{\mathcal{S}_j}) = \emptyset$, where $\mathcal{N}(\cdot)$ represents the index set (nodes) present in the argument set (within the parenthesis). In addition, IGP requires that each input node, n, in a partition set $\mathcal{S}_i$ is connected only to the corresponding outputs in $\mathcal{S}_i$, in other words, the elements of $\mathbf{V}_G$ corresponding to all other outputs are set to zero, i.e., $\mathbf{V}_G(n, j) = 0$, for output nodes j with $j \notin \mathcal{N}(\mathbf{y}_{\mathcal{S}_i})$, and where $\mathbf{V}_G(n, j)$ represents the ith row, jth column entry of matrix $\mathbf{V}_G$ of (3). All the IGP schemes considered herein use square group structure, i.e., with same number of input and output elements in each group. In other words, $|\mathcal{N}(\mathbf{x}_{\mathcal{S}_i})| = |\mathcal{N}(\mathbf{y}_{\mathcal{S}_i})| = N_{vi}$, with $\sum_{i=1}^{N_g} N_{vi} = N_t$.

---

[1]Due to SVD, the original $N_r \times N_t$ system described by equation (1) is equivalent to a size $N_t \times N_t$ virtual MIMO system.

[2]Other types of near-capacity achieving channel coding, e.g., Turbo coding could also be employed in our MIMO precoding schemes, as well. However, as LDPC codes represent one form of the currently prevailing channel coding techniques, we decided to focus on this type of channel coding herein.

Thus, the IGP structure clearly defines N_g independent MIMO input-output groups, i.e., N_g independent, smaller dimension MIMO systems. In the sequel we will also need the notion of an Output only Independent Group Partition (OIGP). Let $p(\cdot)$ be the pdf of the quantity within the parenthesis, and let $\Pr(\cdot)$ be the probability of the quantity within the parenthesis. An OIGP MIMO system is defined as a partition of the output vector $\mathbf{y}$ into N_g independent groups, i.e., with $p(\mathbf{y}) = \prod_{l=1}^{N_g} p(\mathbf{y}_{\mathcal{S}_l})$. An interesting fact is that an AWGN MIMO system as defined by (3) is equivalent to an IGP with the same number of groups, N_g, as we show below. Now we present two lemmas that are essential to the overall paper understanding, with their proofs presented in the Appendices.

Lemma 2.1: Any MIMO OIGP with N_g groups is equivalent to a MIMO IGP with the same number of groups.
Based on this lemma, we see that seeking for higher degree of independence in the output of the MIMO system is equivalent to searching for an IGP.

Now, let $\mathcal{S}(l)$ $(1 \leq l \leq N_t)$, be the mapping from the lth input $\mathbf{x}_l$ to its corresponding group, and let $|\mathcal{S}(l)|$ be the number of elements in this group.

Lemma 2.2: Any MIMO OIGP with N_g groups results in an equivalent expression of the symbol MAP probability, $\Pr(\mathbf{x}_i|\mathbf{y})$, as follows

$$\Pr(\mathbf{x}_i|\mathbf{y}) = \Pr(\mathbf{x}_i|\mathbf{y}_{\mathcal{S}(i)})$$

where we have used $\Pr(\cdot)$ since $\mathbf{x}_i$ takes values in a discrete set.

Due to this lemma, the MAP detector with OIGP needs to only look at the MIMO system defined by $\mathcal{S}(i)$ in order to calculate the MAP probability, thus simplifying the overall MAP detector complexity as only $N_{vi} < N_t$ input and output nodes are present in the MAP detector of $\mathbf{x}_i$. Then each group can be separately precoded and decoded at the receiver thus simplifying the overall system complexity. This reduction in the receiver complexity is evaluated in more detail in Section III.

B. The PGP approach

Consider the following optimization problem:

$$\begin{aligned}
\underset{\mathbf{V}_G, \mathbf{\Sigma}_G}{\text{maximize}} \quad & H(\mathbf{y}) \\
\text{subject to} \quad & H(\mathbf{y}) = \sum_{i=1}^{N_g} H_i(\mathbf{y}_{\mathcal{S}_i}) \\
\text{and} \quad & \text{tr}(\mathbf{\Sigma}_G^2) = N_t.
\end{aligned} \quad (6)$$

This way, the previous result can be easily utilized to exploit this type of inter-group independence. This is the generalized PGP problem (G-PGP). A simpler version is obtained if we further specialize the power constraint in (6) to $\text{tr}(\mathbf{\Sigma}_{G_i}^2) = N_{t_i}$, for $i = 1, 2, \cdots, N_g$. The corresponding solution is called Per Group Precoding (PGP) and it is found as follows: For a particular variable selection method, let $\mathbf{x}_{s_i}$, $\mathbf{y}_{s_i}$ be the data variables and the receiving vector variables in the ith selection

Attribute	Computational Complexity Order, $\mathcal{O}$
GE	$2M^{2N_t}$
OE	$2M^{2N_t}$
UE	$8N_t^3$

subset (group), respectively. Let us denote by N_{t_i}, N_{r_i}, N_g, the numbers of spatially multiplexed data streams, spatially multiplexed receiving antennas per group, and of PGP groups, respectively, then, $N_t = \sum_{i=1}^{N_g} N_{t_i}$ and $N_r = \sum_{i=1}^{N_g} N_{r_i}$. PGP solves the following N_g optimization sub-problems, one for each i (group) $(i = 1, 2, \cdots, N_g)$:

$$\begin{aligned}
\underset{\mathbf{W}_{s_i}}{\text{maximize}} \quad & I(\mathbf{x}_{s_i}, \mathbf{y}_{s_i}) \\
\text{subject to} \quad & \mathbf{W}_{s_i}^h = \mathbf{W}_{s_i} \\
\text{and} \quad & \text{tr}(\mathbf{\Sigma}_{G_i}^2) = N_{t_i}.
\end{aligned} \quad (7)$$

Theorem 2.3: The PGP solution is in the feasible region of the original problem (4).
The proof of this theorem can be found in [16].

C. Computational Complexity Evaluation and Comparison

The original linear precoder (global) optimization problem as described by (4) is solved in [12]. Assuming $N_r = N_t$, as it is the case for the presented results herein, we can perform a basic calculation of the complexity involved at the transmitter and receiver sites of a MIMO system employing linear precoding techniques in order to increase the input-output mutual information. We are evaluating the computational complexity order $\mathcal{O}$ for the different transmitter and receiver processing stages employed. $\mathcal{O}$ represents the total number of calculations, including the number of real additions and multiplications required for a task, expressed as a constant times the dominating calculation factor involved in the task. The computational complexity order is then this calculation dominating factor of the task.

1) Evaluation of Transmitter Complexity: Table I shows the complexity order for the different precoder stages at the transmitter. These are the gradient evaluations (GE), as required by the two backtracking line search algorithms, the two required evaluations of the objective function I (OE) per backtracking line search iteration, and the calculation of $d\mathbf{\Theta}$ of the differential of the unitary matrix $\mathbf{\Theta}$ (UE). Table I presents the corresponding orders of the computational complexities for these attributes. We see that the transmitter computational complexity order of a global optimal precoder is M^{2N_t}.

2) Evaluation of Receiver Complexity: The major ramification introduced at the receiver due to transmitter precoding is the one regarding the MAP detector [6], [7] which evaluates the channel bit log-likelihood-ratios (LLR). Table II shows the corresponding complexity which is $\mathcal{O}(M^{2N_t} \log_2(M))$.

3) Comparison of PGP to Global Optimization Complexity: Concerning the comparison of PGP complexity versus the globally optimal precoder one for the precoder and the MAP

TABLE II
MAP DETECTOR COMPUTATIONAL COMPLEXITY ORDER OF GLOBALLY
OPTIMAL LINEAR PRECODER.

Attribute	Computational Complexity Order, $\mathcal{O}$
MAP Detector	$M^{2N_t} \log_2(M)$

detector, when receiver branches are partitioned in independent groups (e.g., high values of RIF), each group represents exponentially smaller complexity both at the precoder determination and at the MAP detector, due to intergroup independence and smaller dimensions per group. To see how PGP reduces the MAP detector complexity, consider PGP in a MIMO system with N_g IGP groups. Let us focus on calculating the probability $\Pr(\mathbf{x}_i|\mathbf{y}) = \Pr(\mathbf{x}_i|\mathbf{y}_{\mathcal{S}(i)})$, for input symbol $\mathbf{x}_i$ in the system. Since PGP relies on an IGP to apply, let us use a corresponding IGP, $\mathcal{P} = \{\mathcal{S}_1, \cdots, \mathcal{S}_{N_g}\}$, we have N_g groups of nodes at both input and output domain. Each group in the partition, $\mathcal{S}_i$ with $1 \leq i \leq N_g$, comprises N_{gi} unique, non-intersecting input-output vector pairs, $(\mathbf{x}_{\mathcal{S}_i}, \mathbf{y}_{\mathcal{S}_i})$, as explained earlier, with $\sum_{i=1}^{N_g} N_{vi} = N_t$. Let $n_{in}(i,j)$, $n_o(i,j)$ be the mapping from the partition groups to input and output indices, respectively. In other words, $n_{in}(i,j)$ represents the input index (taking values in $\{1, \cdots, N_t\}$) of the jth ($1 \leq j \leq N_{vi}$) input of the ith group ($1 \leq i \leq N_g$), and similarly for $n_o(i,j)$. Also, let $\mathcal{S}(l)$ ($1 \leq l \leq N_t$), be the mapping from the lth input $\mathbf{x}_l$ to its corresponding group. Finally, let $\mathbf{x}_{\mathcal{S}(i)} = \{\mathbf{x}_{n_{in}(i,1)}, \cdots, \mathbf{x}_i, \cdots, \mathbf{x}_{n_{in}(i,|\mathcal{S}(i)|)}\}$ be the set of input nodes present in the ith group in the IGP. Then, due to Lemma 2.2 and the fact that the MIMO channel in (3) is AWGN, we can write

$$\Pr(\mathbf{x}_i|\mathbf{y}) = \Pr(\mathbf{x}_i|\mathbf{y}_{\mathcal{S}(i)})$$
$$\propto \frac{1}{M^{N_{v\mathcal{S}(i)}}} \cdot \sum_{x_{n_{in}(i,l)}, \, n_{in}(i,l) \neq i} p(\mathbf{y}_{\mathcal{S}(i)}|\mathbf{x}_{\mathcal{S}(i)})$$
$$= \frac{1}{M^{N_{vi}}} \frac{1}{\pi^{N_v} \sigma_n^{2N_v}} \cdot \qquad (8)$$
$$\sum_{x_{n_{in}(i,l)}, \, n_{in}(i,l) \neq i} \exp\left(-\frac{||\mathbf{y}_{\mathcal{S}(i)} - \tilde{\mathbf{H}}_{\mathcal{S}(i)}\mathbf{x}_{\mathcal{S}(i)}||^2}{\sigma_n^2}\right),$$

where $\propto$ means proportional to, i.e., equal except for a multiplication constant independent of $\mathbf{x}_i$, and $\tilde{\mathbf{H}}_{\mathcal{S}(i)}$ is the part of $\tilde{\mathbf{H}} \doteq \Sigma_H \Sigma_G \mathbf{V}_G^h$ that corresponds to $\mathbf{x}_{\mathcal{S}(i)}$. This equation clearly shows that due to the existing IGP, the MAP detector only needs to invoke inputs from the input group corresponding to the particular input under consideration. This also shows that the summations required in each evaluation are of the order $M^{N_{v|\mathcal{S}(i)|}}$, due to PGP, as there are only $N_{v|\mathcal{S}(i)|} - 1$ interfering symbols to $\mathbf{x}_i$, due to inter-group independence as different input groups connect to different antennas due to IGP. Further, as we need $M^{N_{v|\mathcal{S}(i)|}}$ values of this, i.e., over all possible group input combinations, one per different input combination, the total computational complexity order becomes $M^{2N_{v|\mathcal{S}(i)|}}$ per symbol, or $M^{2N_{v|\mathcal{S}(i)|}} \log_2(M)$ per bit in group $\mathcal{S}(i)$.

For N_g groups, and assuming for simplicity equal number of N_t/N_g beams per group, the corresponding complexity associated with PGP is found from Table I, II, after substituting

N_t/N_g for N_t, and multiplying by N_g since PGP needs N_g smaller global optimal precoders. This means that PGP reduces the corresponding computational complexities at both the transmitter and receiver by $M^{2N_t(1-1/N_g)}$ times, thus offering a significant computational complexity reduction at both the transmitter and receiver. For example, by using a 4×4 MIMO system with QPSK modulation, PGP needs 128 times smaller computational complexity at both the transmitter and receiver. As the number of antennas and the modulation constellation size grow, this PGP complexity reduction becomes more significant.

III. NUMERICAL RESULTS

The results presented herein employ PGP as described above, as well as globally optimal precoders for comparison. Since PGP performs a number of N_g globally optimal precoder determinations, albeit of smaller size, it suffices to describe the globally optimal recoder implementation. The globally optimal precoder implementation methodology is performed by employing two backtracking line searches, one for $\mathbf{W}$, and another one for Σ_G^2, at each iteration, in a fashion similar to [12], but introducing some improvements. Similarly to [12], we follow a block coordinate gradient ascent maximization method to find the solution to the optimization problem described in (4), employing the virtual model of (3). For most cases presented, it is worth mentioning that only a few iterations (e.g., typically < 20) are required to converge to the PGP solution results as presented herein, even for higher size MIMO configurations, e.g., 5×5 MIMO systems. All the results consider Quadrature Phase Shift Keying (QPSK) modulation on narrowband independently fading channels. For a 3×3 MIMO system with

$$\mathbf{H} = \begin{bmatrix} 1 & 0.5j & 0.3 \\ -0.5j & 1.5 & -0.1j \\ 0.3 & 0.1j & 0.5 \end{bmatrix},$$

Fig. 3 presents a plot of the mutual information achieved by PGP, globally optimal precoding, no precoding, plain beamforming ($\mathbf{V}_G = \Sigma_G = \mathbf{I}$ in the model presented in (3), which is known to be optimal in low SNR values [14]), and the MDP design of [4], as a function of the symbol SNR in dB. We observe very close agreement between the performance of the PGP and the globally optimal precoder. We also see that PGP significantly outperforms both the plain beamforming and no precoding case in the "medium" to "high" SNR range, by more than $1.5 \ b/s/Hz$ and $0.7 \ b/s/Hz$, respectively, as expected (e.g., [12]), however with significantly reduced receiver complexity. In addition, PGP still outperforms MDP by about $0.5 \ b/s/Hz$ in the medium SNR range. However, this gain of PGP comes simultaneously with a much lower receiver complexity, due to the output branch independence present with PGP. The reason plain beamforming may appear to be performing worse than the no precoding case is as follows. Plain beamforming is known to be the optimal precoding technique for low SNR, but generally it lacks in performance in higher SNR. Since many authors consider it for comparison purposes, e.g., [14], we decided to include it in our results, as well. The rationale behind this low performance is that
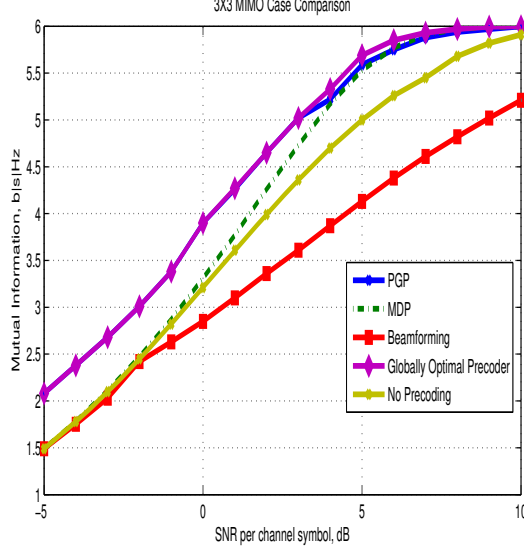
Fig. 3. Results for PGP, globally optimal precoding, no precoding, plain beamforming, and MDP for a 3×3 MIMO system and QPSK modulation.
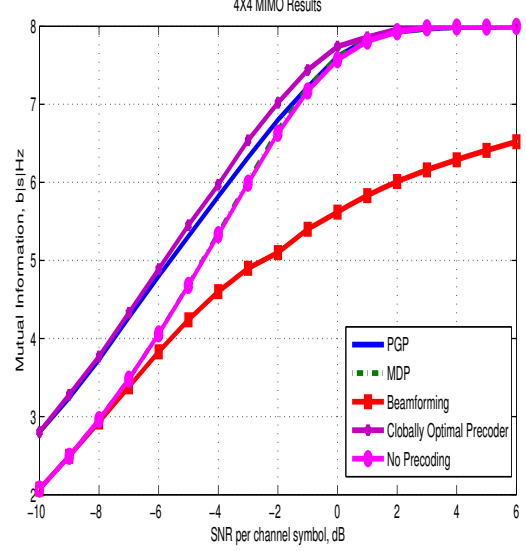

Fig. 4. Results for PGP, globally optimal precoding, no precoding, plain beamforming, and MDP for a 4×4 MIMO system and QPSK modulation.

although the plain beamformer has full independence, it fails to achieve high information transfer. Based on the model of (3), not all precoders can achieve better information transfer than the original channel ($\mathbf{H}$) one.

For the randomly generated 4×4 MIMO system with channel

$$\mathbf{H} = \begin{bmatrix} 1.0919 + 1.0036i & -0.7507 - 0.5688i & -0.1361 + 0.1525i & -1.0058 + 0.5107i \\ 0.0608 + 0.2062i & 1.6620 + 0.4926i & 0.6283 - 0.8244i & 0.3452 + 1.8282i \\ -1.0547 + 0.1399i & -0.4353 + 0.5905i & -0.5408 - 0.8117i & -0.1254 - 0.4716i \\ -0.5249 + 1.1227i & 0.5290 - 0.1723i & -0.9916 + 0.0742i & -0.1386 + 0.1325i \end{bmatrix},$$

and with $N_g = 2$, 2×2 subgroups, for a MIMO system with $N_t = N_r = 4$ and QPSK modulation, we get the results shown in Fig. 4. Clearly, PGP still outperforms MDP by about $0.7 \ b/s/Hz$ in the medium SNR range, while PGP outperforms the plain beamforming case by more than $1.5 \ b/s/Hz$ and MDP in the medium to high SNR range. At the same time PGP offers performance almost equal to the globally optimal precoder one in the low SNR range, while it is still very close to the globally optimal precoder at higher SNR values.

IV. CONCLUSIONS

In this paper, the problem of designing a linear precoder for MIMO systems employing LDPC codes is addressed under the constraint of independent virtual receiving ouput groups, thus reducing the relevant transmitter and receiver complexities. This approach sees the overall precoding problem in an LDPC-coded system from a brand new angle allowing for a more practical deployment of higher dimension MIMO systems and higher QAM constellation sizes with very good performance over a wide SNR range.

We then target a generalization of BICMB and show that this offers a solution to a very meaningful precoding optimization problem that allows for inter-group independence between different transmitting-receiving antenna pairs in the virtual channel domain. We call the new precoding solution

PGP and we show, based on numerical results, that PGP offers indeed excellent performance, while its computational complexity is significantly reduced compared to the original solution. For the results presented, PGP is shown to attain mutual information very close to the globally optimal precoder and higher than MDP by $0.7 - 0.8 \ b/s/Hz$ in medium SNR ranges. Thus, based on presented evidence, PGP is a very good candidate for almost optimal precoding performance in LDPC-coded systems with relatively low system complexity at both the transmitter and receiver. Our future work will look at generalizing the presented approach to higher size modulation constellations, e.g., QAM with $M \geq 16$.

APPENDIX A
PROOF OF LEMMA 2.1

For the *if* part, which is easier, let $\mathbf{x}_S$ be the restriction of the input vector $\mathbf{x}$ to the set of indices described by S (which is a subset of $\{1, 2, \cdots, N_t\}$, using the notation of (3)). We need to prove that if an IGP exists in a MIMO system with output groups $\{\mathbf{y}_{S_{y_i}}\}_{i=1}^{N_g}$, then the output $\mathbf{y}$ pdf, $p(\mathbf{y})$, can be written as $\prod_{m=1}^{N_g} p(\mathbf{y}_{S_m})$, i.e., the output groups are independent. Recalling that an IGP with N_g groups, $\mathcal{P}$, is defined as follows: $\mathcal{P} = \{S_1, \cdots, S_{N_g}\}$, where each group in the partition, S_i with $1 \leq i \leq N_{vi}$, comprises N_{vi} unique, non-intersecting input-output vector pairs (with $\sum_{i=1}^{N_g} N_{vi} = N_t$), $(\mathbf{x}_{S_i}, \mathbf{y}_{S_i})$, which satisfy $\cup_{i=1}^{N_g} \mathcal{N}(\mathbf{x}_{S_i}) = \cup_{i=1}^{N_g} \mathcal{N}(\mathbf{y}_{S_i}) = \{1, 2, \cdots, N_t\}$ and $\mathcal{N}(\mathbf{x}_{S_i}) \cap_{j \neq i} \mathcal{N}(\mathbf{x}_{S_j}) = \mathcal{N}(\mathbf{y}_{S_i}) \cap_{j \neq i} \mathcal{N}(\mathbf{y}_{S_j}) = \emptyset$, where $\mathcal{N}(\cdot)$ represents the index set (nodes) present in the argument set (within the parenthesis). In addition, IGP requires that each input node, n, in a partition set S_i is connected only to the corresponding outputs in S_i, in other words, the elements of $\mathbf{V}_G$ corresponding to all other outputs are set to zero, i.e., $\mathbf{V}_G(n, j) = 0$, for output nodes j with $j \notin \mathcal{N}(\mathbf{y}_{S_i})$, and where $\mathbf{V}_G(n, j)$ represents the ith row, jth column entry of matrix $\mathbf{V}_G$ of (3). All the IGP schemes

considered herein use square group structure, i.e., with same number of input and output elements in each group. In other words, $|\mathcal{N}(\mathbf{x}_{\mathcal{S}_i})| = |\mathcal{N}(\mathbf{y}_{\mathcal{S}_i})| = N_{vi}$, with $\sum_{i=1}^{N_g} N_{vi} = N_t$. We have

$$
\begin{aligned}
p(\mathbf{y}) &= p(\mathbf{y}_1, \cdots, \mathbf{y}_{N_g}) \\
&= \sum_{\mathbf{x}_{\mathcal{S}_1}} \cdots \sum_{\mathbf{x}_{\mathcal{S}_{N_g}}} p(\mathbf{y}_{\mathcal{S}_1}, \cdots, \mathbf{y}_{\mathcal{S}_{N_g}} | \mathbf{x}_{\mathcal{S}_1}, \cdots, \mathbf{x}_{\mathcal{S}_{N_g}}) \cdot \\
&\quad \Pr(\mathbf{x}_{\mathcal{S}_1}, \cdots, \mathbf{x}_{\mathcal{S}_{N_g}}) \\
&= \sum_{\mathbf{x}_{\mathcal{S}_1}} \cdots \sum_{\mathbf{x}_{\mathcal{S}_{N_g}}} \prod_{m=1}^{N_g} p(\mathbf{y}_{\mathcal{S}_m} | \mathbf{x}_{\mathcal{S}_m}) \Pr(\mathbf{x}_{\mathcal{S}_m}) \\
&= \prod_{m=1}^{N_g} p(\mathbf{y}_{\mathcal{S}_m}),
\end{aligned}
\tag{9}
$$

where we have used the conditional independence of outputs from inputs which do not enter their antennas, since they belong to different input groups.

For the *only if* part, we will use mathematical induction on N_g. We also need to invoke the fact that the noise in the MIMO channel in (3) is AWGN and thus the conditional pdf of the output $\mathbf{y}$ given the input vector $\mathbf{x}$ is circular complex Gaussian. For $N_g = 2$ the statement becomes as follows, after setting $\tilde{\mathbf{H}} \doteq \boldsymbol{\Sigma}_H \boldsymbol{\Sigma}_G \mathbf{V}_G^h$, employing the model of (3): If

$$
p(\mathbf{y}) = \prod_{l=1}^{2} p(\mathbf{y}_{\mathcal{S}_l})
\tag{10}
$$

where $\mathbf{y}_{\mathcal{S}_1}, \mathbf{y}_{\mathcal{S}_2}$ represent a partition of the output vector $\mathbf{y}$, then an IGP with $N_g = 2$ needs to exist in the MIMO system. For the model in (3) we can write (here $N_v = \min\{N_t, N_r\}$, where $\min$ stands for the minimum of a set of elements)

$$
\begin{aligned}
p(\mathbf{y}) &= \sum_{\mathbf{x}} p(\mathbf{y}|\mathbf{x}) \Pr(\mathbf{x}) \\
&= \frac{1}{M^{N_v}} \frac{1}{\pi^{N_v} \sigma_n^{2N_v}} \sum_{\mathbf{x}} \exp\left(-\frac{||\mathbf{y} - \tilde{\mathbf{H}}\mathbf{x}||^2}{\sigma_n^2} \right).
\end{aligned}
\tag{11}
$$

Let us separate the $\mathbf{x}$ vector into three, in general, parts: $\mathbf{x}_{\mathcal{S}_1}$, $\mathbf{x}_{\mathcal{S}_2}$, $\mathbf{x}_{\mathcal{S}_c}$, representing inputs entering only outputs in $\mathbf{x}_{\mathcal{S}_1}$, $\mathbf{x}_{\mathcal{S}_2}$, and inputs entering both, respectively. Then,

$$
\begin{aligned}
p(\mathbf{y}) &= \frac{1}{M^{N_v}} \frac{1}{\pi^{N_v} \sigma_n^{2N_v}} \sum_{\mathbf{x}_{\mathcal{S}_1}} \sum_{\mathbf{x}_{\mathcal{S}_2}} \sum_{\mathbf{x}_{\mathcal{S}_c}} \exp\left(-\frac{||\mathbf{y}_{\mathcal{S}_1} - \tilde{\mathbf{H}}_{\mathcal{S}_1}\mathbf{x}||^2}{\sigma_n^2} \right) \cdot \\
&\quad \exp\left(-\frac{||\mathbf{y}_{\mathcal{S}_2} - \tilde{\mathbf{H}}_{\mathcal{S}_2}\mathbf{x}||^2}{\sigma_n^2} \right),
\end{aligned}
\tag{12}
$$

where $\tilde{\mathbf{H}}_{\mathcal{S}_l}$, $l = 1,\ 2$ represents the rows of $\tilde{\mathbf{H}}$ that correspond to the OIGP l. By further breaking down $\tilde{\mathbf{H}}_{\mathcal{S}_l}\mathbf{x}$, we can see that the two output groups cannot be independent, except if $\mathcal{S}_c = \emptyset$, in which case we have an IGP. This proves our assertion.

Assume that the statement is true for $N_g = k$: If

$$
p(\mathbf{y}) = \prod_{l=1}^{k} p(\mathbf{y}_{\mathcal{S}_l})
\tag{13}
$$

where $\mathbf{y}_{\mathcal{S}_1}, \cdots, \mathbf{y}_{\mathcal{S}_k}$ represent a partition of the output vector $\mathbf{y}$, then an IGP with $N_g = k$ needs to exist in the MIMO system. It is rather easy by invoking the already proved $N_g = 2$ case and the assumption for $N_g = k$ that the statement is

also true for $N_g = k + 1$. This proves our assertion.

APPENDIX B
PROOF OF LEMMA 2.2

For an input-output MIMO IGP, $\mathcal{P} = \{\mathcal{S}_1, \cdots, \mathcal{S}_{N_g}\}$, we have N_g groups of nodes at both input and output domain. Each group in the partition, $\mathcal{S}_i$ with $1 \leq i \leq N_g$, comprises N_{gi} unique, non-intersecting input-output vector pairs, $(\mathbf{x}_{\mathcal{S}_i}, \mathbf{y}_{\mathcal{S}_i})$, as explained in the main text, with $\sum_{i=1}^{N_g} N_{vi} = N_t$. Let $n_{in}(i, j)$, $n_o(i, j)$ be the mapping from the partition groups to input and output indices, respectively. In other words, $n_{in}(i, j)$ represents the input index (taking values in $\{1, \cdots, N_v\}$) of the jth ($1 \leq j \leq N_{vi}$) input of the ith group ($1 \leq i \leq N_g$), and similarly for $n_o(i, j)$. Also, let $\mathcal{S}(l)$ ($1 \leq l \leq N_t$), be the mapping from the lth input $\mathbf{x}_l$ to its group number. Finally, let $\mathbf{x}_{\mathcal{S}(i)} = \{\mathbf{x}_{n_{in}(i,1)}, \cdots, \mathbf{x}_i, \cdots, \mathbf{x}_{n_{in}(i,|\mathcal{S}(i)|)}\}$ be the set of input nodes present in the ith group of the IGP. We can then write, employing $p(\cdot)$ for the pdf of the variables in the parenthesis,

$$
\begin{aligned}
\Pr(\mathbf{x}_i|\mathbf{y}) &= \frac{p(\mathbf{y}, \mathbf{x}_i)}{p(\mathbf{y})} \\
&= \frac{\sum_{l \neq i, l \in \mathcal{S}(i)} \sum_{m=1, m \neq i}^{N_g} p(\mathbf{y}|\mathbf{x}_{\mathcal{S}_1}, \cdots, \mathbf{x}_{\mathcal{S}(i)}, \cdots, \mathbf{x}_{\mathcal{S}_{N_g}}) \Pr(\mathbf{x})}{p(\mathbf{y})} \\
&= \frac{1}{\prod_{m=1}^{N_g} p(\mathbf{y}_{\mathcal{S}_m})} \cdot \sum_{l \neq i, l \in \mathcal{S}(i)} \sum_{m=1, m \neq i}^{N_g} \\
&\quad \prod_{m=1, m \neq i}^{N_g} p(\mathbf{y}_{\mathcal{S}_m}|\mathbf{x}_{\mathcal{S}_m}) \Pr(\mathbf{x}_{\mathcal{S}_m}) p(\mathbf{y}_{\mathcal{S}_i}|\mathbf{x}_{\mathcal{S}_i}) \Pr(\mathbf{x}_{\mathcal{S}_i}) \\
&= \frac{1}{\prod_{m=1}^{N_g} p(\mathbf{y}_{\mathcal{S}_m})} \sum_{l \neq i, l \in \mathcal{S}(i)} \sum_{m=1, m \neq i}^{N_g} \\
&\quad \prod_{m=1, m \neq i}^{N_g} p(\mathbf{y}_{\mathcal{S}_m}|\mathbf{x}_{\mathcal{S}_m}) \Pr(\mathbf{x}_{\mathcal{S}_m}) p(\mathbf{y}_{\mathcal{S}_i}|\mathbf{x}_{\mathcal{S}_i}) \Pr(\mathbf{x}_{\mathcal{S}_i}) \\
&= \frac{p(\mathbf{y}_{\mathcal{S}_i}, \mathbf{x}_i)}{p(\mathbf{y}_{\mathcal{S}(i)})} = \Pr(\mathbf{x}_i|\mathbf{y}_{\mathcal{S}(i)}),
\end{aligned}
\tag{14}
$$

where the last four equations come from the conditional independence of outputs given inputs of the other groups and the fact that the inputs are independent, the independence of the outputs of different groups, the total probability law, and the definition of conditional probability, respectively. Note that $\Pr(\mathbf{x})$ is the probability of the overall input vector (N_t components), which due to independent input assumption and equally likely inputs, can be written as $\Pr(\mathbf{x}) = \prod_{l=1}^{N_t} \Pr(\mathbf{x}_i) = \prod_{k=1}^{N_g} \Pr(\mathbf{x}_{\mathcal{S}_k}) = \frac{1}{M^{N_t}}$.

This proves our assertion.

REFERENCES

[1] E. Akay, E. Sengul, and E. Ayanoglu, "Bit Interleaved Coded Multiple Beamforming," *IEEE Transactions on Communications*, vol. 55, pp. 1802–1810, September 2007.

[2] B. Li, H. J. Park, and E. Ayanoglu, "Constellation Precoded Multiple Beamforming," *IEEE Transactions on Communications*, vol. 59, pp. 1275–1286, May 2011.

[3] B. Li and E. Ayanoglu, "Diversity Analysis of Bit-Interleaved Coded Multiple Beamforming with Orthogonal Frequency Division Multiplexing," *IEEE Transactions on Communications*, vol. 61, pp. 3794–3805, September 2013.

[4] Y. Xin, Z. Wang, and G. Giannakis, "Space-Time Diversity Systems Based on Linear Constellation Precoding," *IEEE Transactions on Wireless Communications*, vol. 2, pp. 294–309, March 2003.

[5] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-Time Codes for High Data Rate Communications: Performance Criterion and Code Construction," *IEEE Transactions on Information Theory*, vol. 44, pp. 744–765, March 1998.

[6] S. ten Brink, G. Kramer, and A. Ashihkmin, "Design of Low-Density Parity-Check Codes for Modulation and Detection," *IEEE Transactions on Communications*, vol. 52, pp. 670–678, April 2004.

[7] Y. Jian, A. Ashihkmin, and N. Sharma, "LDPC Codes for Flat Rayleigh Fading Channels with Channel Side Information," *IEEE Transactions on Communications*, vol. 56, pp. 1207–1213, August 2008.

[8] F. Perez-Cruz, M. Rodriguez, and S. Verdu, "MIMO Gaussian Channels with Arbitrary Inputs: Optimal Precoding and Power Allocation," *IEEE Transactions on Information Theory*, vol. 56, pp. 1070–1086, March 2010.

[9] A. Lozano, A. Tulino, and S. Verdu, "Optimal Power Allocation for Parallel Gaussian Channels With Arbitrary Input Distributions," *IEEE Transactions on Information Theory*, vol. 52, pp. 3024–3051, July 2006.

[10] M. Payaro and D. Palomar, "On Optimal Precoding in Linear Vector Gaussian Channels with Arbitrary Input Distribution," in *Proceedings IEEE International Symposium on Information Theory*, 2009.

[11] E. Ohlmer, U. Wachsmann, and G. Fettweis, "Mutual Information Maximizing Linear Precoding for Parallel Layer MIMO Detection," in *Proceedings 12th International Workshop on Signal Processing Advances in Wireless Communications*, 2011.

[12] G. Xiao, Y. Zheng, and Z. Ding, "Globally Optimal Linear Precoders for Finite Alphabet Signals Over Complex Vector Gaussian Channels," *IEEE Transactions on Signal Processing*, vol. 59, pp. 3301–3314, July 2011.

[13] M. Lamarca, "Linear Precoding for Mutual Information Maximization in MIMO Systems," in *Proceedings International Symposium of Wireless Communication Systems 2009*.

[14] W. Zeng, G. Xiao, and J. Lu, "A Low Complexity Design of Linear Precoding for MIMO Channels with Finite Alphabet Inputs," *IEEE Wireless Communications Letters*, vol. 1, pp. 38–42, February 2012.

[15] W. Zeng, G. Xiao, W. Mingxi, and J. Lu, "Linear Precoding for Finite Alphabet Inputs Over MIMO Fading Channels with Statistical CSI," *IEEE Transactions on Signal Processing*, vol. 60, pp. 3134–3148, February 2012.

[16] T. Ketseoglou and E. Ayanoglu, "Linear precoding for MIMO with LDPC Coding and Reduced Complexity," in *Proc. Asilomar Conference on Signals, Systems, and Computers*, 2013.