

Code Design for MIMO Broadcast Channels

Momin Uppal, *Student Member, IEEE*, Vladimir Stanković, *Member, IEEE*, and Zixiang Xiong *Fellow, IEEE*

Abstract—Recent information-theoretic results show the optimality of dirty-paper coding (DPC) in achieving the full capacity region of the Gaussian multiple-input multiple-output (MIMO) broadcast channel (BC). This paper presents a DPC based code design for BCs. We consider the case in which there is an individual rate/signal-to-interference-plus-noise ratio (SINR) constraint for each user. For a fixed transmitter power, we choose the linear transmit precoding matrix such that the SINRs at users are uniformly maximized, thus ensuring the best bit-error rate performance. We start with Cover's simplest two-user Gaussian BC and present a coding scheme that operates 1.44 dB from the boundary of the capacity region at the rate of one bit per real sample (b/s) for each user. We then extend the coding strategy to a two-user MIMO Gaussian BC with two transmit antennas at the base-station and develop the first limit-approaching code design using nested turbo codes for DPC. At the rate of 1 b/s for each user, our design operates 1.48 dB from the capacity region boundary. We also consider the performance of our scheme over a slow fading BC. For two transmit antennas, simulation results indicate a performance loss of only 1.4 dB, 1.64 dB and 1.99 dB from the theoretical limit in terms of the total transmission power for the two, three and four user case, respectively.

Index Terms—Dirty-paper coding, MIMO broadcast channels, multiple access channel, nested turbo codes, zero-forcing linear beamforming.

I. INTRODUCTION

AN achievable rate region of a degraded Gaussian broadcast channel (BC), where different users receive signals at different signal-to-interference-plus-noise ratios (SINRs), was provided by Cover in 1972 [1]. The scheme is based on superposition coding, where the message for one user is embedded in that for the other. Bergman [2] showed that Cover's rate region is in fact the capacity region by proving the converse. However, the capacity region of the Gaussian multiple-input multiple-output (MIMO) BC, where the channels might not necessarily be degraded, was an open problem until recently — the rate region found in [3] was first shown to achieve the sum-rate capacity in [3]–[6] and then proven to be the actual capacity region in [7].

The core of the capacity-achieving scheme [3] for a Gaussian MIMO BC is a non-linear technique called dirty-paper coding (DPC) [8], which is a coding technique for channels

with encoder side information. According to [8], in a Gaussian interference channel, if the interfering signal is known non-causally at the transmitter, then there is no loss in capacity due to the interference. This scenario is typical in the Gaussian BC, where each user treats messages intended for other users as interference, which is available non-causally at the transmitter.

DPC is the only optimal, i.e., capacity-achieving, technique for the non-degraded Gaussian MIMO BC. It provides significant gains in terms of the achievable rates over suboptimal techniques (e.g., time-division multiple-access and beamforming) for the MIMO BC in many setups [3], [9], especially when the signal-to-noise ratio (SNR) is high and the number of transmit antennas large. Practical DPC involves both source and channel coding [10], and near-capacity code designs have appeared recently [11]–[14]. The time is thus ripe to develop limit approaching DPC-based code designs¹ for the MIMO BC and compare them with others based on suboptimal strategies.

We start with Cover's simplest yet most celebrated two-user degraded Gaussian BC [1] and develop a code design that employs the dirty-paper code of [11] using nested turbo codes. Owing to the powerful nested turbo codes in DPC, our code design for the degraded Gaussian BC is superior to previously reported DPC-based schemes of [17], [18] and loses only 1.44 dB from the boundary of the capacity region when transmitting one bit per real sample (b/s) for each user.

Note that, besides DPC, superposition coding [1] also achieves the capacity of this simple Gaussian BC. The practical scheme of [19] that exploits superposition coding performs only 1.0 dB from the minimum total power at one b/s for each user. However, the drawback of the superposition coding approach of [19] is that it is not clear how it can be extended to handle the MIMO Gaussian BC, where the channels are not necessarily degraded. In contrast to superposition coding, DPC achieves the capacity of both degraded and non-degraded BCs, and our DPC-based design for Cover's Gaussian BC naturally applies to the MIMO BC. Thus, we additionally consider the two-user MIMO Gaussian BC with two transmit antennas at the base-station and develop the first limit-approaching DPC-based code design under the assumption that the transmitter and all the receivers have perfect knowledge of the channel state information (CSI). We point out that in practice, the CSI might not be perfectly available at the transmitters. However, as the first step towards developing DPC-based code design for MIMO broadcast channels, we make such a rather strong assumption in this paper. Extending our design to the case with imperfect/partial CSI at the transmitter is the next step, which we leave to future publications.

Paper approved by I. Lee, the Editor for Wireless Communication Theory of the IEEE Communications Society. Manuscript received July 1, 2007; revised December 4, 2007.

M. Uppal and Z. Xiong are with the Dept of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843 (e-mail: {momin, zx}@ece.tamu.edu).

V. Stanković was with the Dept of Electrical and Computer Engineering, Texas A&M University. He is now with the Dept of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, UK (e-mail: Vladimir.Stankovic@eee.strath.ac.uk).

Work supported by NSF grant CCF-0729149 and Qatar National Research Fund NPRP 28-6-7-28.

Digital Object Identifier 10.1109/TCOMM.2009.04.070300

¹To the best of our knowledge, all existing DPC-based designs [15], [16] for the MIMO BC use scalar forms of DPC such as Tomlinson-Harashima precoding and hence incur a large performance loss.

Most information-theoretic works [3]–[6] on MIMO BCs focus on maximizing the sum-rate, i.e., the sum of the transmission rates for different users under a fixed total transmission power constraint. This allows arbitrary rate allocation to different users depending on the channel coefficients. The sum-rate maximizing scheme in this case might result in an unfair rate allocation (e.g., by assigning a very low rate to one user). However, many applications require the users to operate at a certain minimum rate/SINR. In addition, our DPC code design with nested turbo codes [11] operates at a fixed transmission rate. In light of these two reasons, we fix the rate of the employed code at each user. Thus, our coding scheme has the design objective of uniformly maximizing the SINR for each user, subject to a total transmit power constraint and certain per user rate/SINR constraints. Compared to the sum-rate maximizing scheme, our scheme will obviously achieve a smaller or equal sum-rate. To determine a precoding scheme at the transmitter that uniformly maximizes the individual rate/SINR at each user, one can invoke the duality [4], [5], [20] between the BC and the multiple access channel (MAC); as shown in [21], optimal precoding can be accomplished with reasonable computational complexity by employing an iterative procedure at the encoder.

We employ one conventional channel code for the first user and one dirty-paper code for each of the remaining users for coding over MIMO Gaussian BC. Experiments show that our design for two users and two transmit antennas operates 1.48 dB from the capacity region boundary. We also simulate sub-optimal strategies such as zero-forcing DPC (ZFDPC) and zero-forcing beamforming (ZFBF). However, the performance gap between the optimal strategy and the zero-forcing approaches depend on the exact realization of the channel coefficients. Thus, in order to provide a fair comparison, we also consider the outage performance of our design over a slow Rayleigh fading MIMO BC. Simulations indicate that when the number of transmit antennas is fixed at two, our practical DPC-based design performs 1.4 dB, 1.64 dB and 1.99 dB worse than the theoretical limits for the two, three and four user case, respectively. Preliminary results of this work appeared in [22].

The contributions of this paper are twofold: our DPC-based design for Cover’s celebrated two-user degraded Gaussian BC significantly outperforms existing schemes [17], [18], and is the only scheme in the literature that beats the time-sharing line (see Fig. 4); we also develop the first limit approaching code design for the MIMO Gaussian BC using the powerful turbo-based DPC scheme of [11].

Finally, we point out that recent dirty-paper code designs [12]–[14] based on vector quantizers and LDPC/irregular repeat accumulate codes work very well at the low rate regime (e.g., at 0.25 b/s). However, it is not easy to redesign them (e.g., the best performing one [14] in this class) to operate well at high transmission rates for broadcast applications.

II. CHANNEL CAPACITIES

In this section, we give a brief overview of the capacity regions for the two-user degraded Gaussian BC [1] and the MIMO Gaussian BC [3].

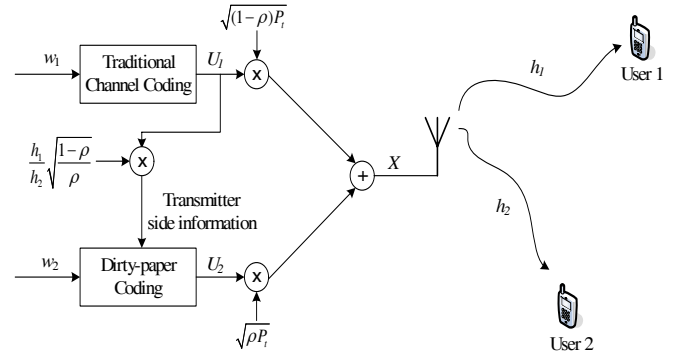


Fig. 1. Encoding setup for the two user degraded Gaussian BC.

A. Two-user degraded Gaussian BC

Consider a two-user Gaussian BC with a single-antenna transmitter which sends signal X ; the two users receive $Y_1 = h_1X + Z_1$ and $Y_2 = h_2X + Z_2$, where Z_1 and Z_2 are independent, identically distributed (i.i.d.) zero-mean Gaussian noises with unit variances, independent of X , and h_1, h_2 are constant channel coefficients. Since the channel is degraded, without loss of generality, we assume that $|h_2| > |h_1|$. The transmitted signal X is given by $X = \mathbf{B}[U_1, U_2]^T$, where U_1 and U_2 are the coded messages intended for user 1 and 2, respectively, with $E[|U_1|^2], E[|U_2|^2] \leq 1$; $\mathbf{B} = [\sqrt{(1-\rho)P}, \sqrt{\rho P}]$ is the precoding matrix, with P being the total transmission power constraint, i.e., $E[|X|^2] \leq P$, and ρ ($0 \leq \rho \leq 1$) a parameter that controls the power allocation between the two users. Cover [1] obtained the capacity region for this setup by using superposition coding. However, DPC also achieves all points in the capacity region. Indeed, U_2 can be dirty-paper coded with U_1 as the encoder side information as shown in Fig. 1. This way, user 2 achieves the same rate as if the interference from U_1 were not present. User 1 on the other hand, treats U_2 as interference. Then the achievable rates per real dimension R_1 and R_2 for user 1 and user 2, respectively, are

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{|h_1|^2(1-\rho)P}{|h_1|^2\rho P + 1} \right) \quad (1)$$

and

$$R_2 \leq \frac{1}{2} \log (1 + |h_2|^2\rho P). \quad (2)$$

Since user 1 always gets a degraded version of the signal received by user 2, user 2 can also decode U_1 , provided that it knows the codebook of user 1. Then, the effective achievable rate for user 2 is $R_1 + R_2$.

B. Non-degraded MIMO BC

When the transmitter employs multiple antennas, the channels at the users are no longer degraded. However, the DPC scheme described in the previous subsection can be generalized to MIMO case where a transmitter with M antennas sends messages to K users, each with a single antenna. If h_{ij} is a complex channel gain between user i ($1 \leq i \leq K$) and transmit antenna j ($1 \leq j \leq M$), then $Y_i = \sum_{j=1}^M h_{ij}X_j + Z_i$ is the complex baseband equivalent of the signal received by

user i , X_j is the complex baseband equivalent of the transmitted signal at antenna j , and the Z_i 's are i.i.d. complex zero-mean Gaussian noises with unit variances, independent of the X_j 's. Let $\mathbf{Y} = [Y_1, Y_2, \dots, Y_K]^T$, $\mathbf{X} = [X_1, X_2, \dots, X_M]^T$, and $\mathbf{Z} = [Z_1, Z_2, \dots, Z_K]^T$; then the received vector becomes $\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{Z}$, with h_{ij} corresponding to the element at the i^{th} row and j^{th} column of the $K \times M$ channel matrix \mathbf{H} .

Let w_i be the message intended for user i , then the transmitter sends $\mathbf{X} = \mathbf{B}\mathbf{U}$, where \mathbf{B} is an $M \times K$ precoding matrix, and

$$\begin{aligned} \mathbf{U} &= [U_1, U_2, U_3, \dots, U_K]^T \\ &= [U_1(w_1), U_2(w_2; U_1), \dots, U_K(w_K; U_1, \dots, U_{K-1})]^T \end{aligned}$$

is generated using successive DPC with all K codebooks being uncorrelated and Gaussian with unit power. Here, $U_i = U_i(w_i; U_1, \dots, U_{i-1})$, $2 \leq i \leq K$, indicates that w_i is encoded as the codeword U_i using DPC with the linear combination of U_1, U_2, \dots, U_{i-1} as the encoder side information (i.e., known interference).

Let \mathbf{b}_i be the i^{th} column of the precoding matrix \mathbf{B} , then the $M \times M$ transmitter covariance matrix for each user is a positive semi-definite matrix given by $\mathbf{S}_i = \mathbf{b}_i \mathbf{b}_i^H$, and the transmitter power constraint is $\sum_{i=1}^K \text{tr}(\mathbf{S}_i) \leq P$. We can now write the signal received at user i as

$$Y_i = \mathbf{h}_i \sum_{k=1}^{i-1} \mathbf{b}_k U_k + \mathbf{h}_i \mathbf{b}_i U_i + \mathbf{h}_i \sum_{k=i+1}^K \mathbf{b}_k U_k + Z_i, \quad (3)$$

where \mathbf{h}_i represents the i^{th} row of the channel matrix \mathbf{H} . Due to DPC, user i can cancel out the first term, whereas the second term is the useful signal, and the third is treated as Gaussian interference. Therefore, the achievable rate R_i^{BC} at user i satisfies [6]

$$R_i^{BC} \leq \frac{1}{2} \log \left(1 + \frac{\mathbf{h}_i \mathbf{S}_i \mathbf{h}_i^H}{1 + \mathbf{h}_i \sum_{k=i+1}^K \mathbf{S}_k \mathbf{h}_i^H} \right), \quad i = 1, \dots, K. \quad (4)$$

Note that the rate vector $\mathbf{R}^{BC} = [R_1^{BC}, \dots, R_K^{BC}]$ is achievable under a fixed encoding order, where the message of user i is dirty-paper coded by treating signals for previously encoded users as known interference. One can therefore obtain $K!$ different achievable rate vectors, one for each distinct encoding order. The capacity region of the MIMO BC for a fixed channel matrix \mathbf{H} and a power constraint P is the convex hull of the union of all rate vectors obtained over all possible encoding orders and all covariance matrices \mathbf{S}_i satisfying the power constraint $\sum_{i=1}^K \text{tr}(\mathbf{S}_i) \leq P$.

C. Duality between the Gaussian BC and MAC

The duality between the capacity regions of the Gaussian MIMO BC and MAC was pointed out in [4], [5]. Unlike the achievable rates for the MIMO BC, given by (4), the rates for the MAC are concave functions of the input covariances. Therefore, it is easier to find the boundary of the capacity region of the MAC than that of the BC. By exploiting the duality [4], [5], the achievable rates for the MIMO BC can be derived from those of its dual MAC. Here we briefly review this duality principle which will be used in Section III to determine optimal precoding.

The dual MAC of the MIMO BC described in Section II-B has K users, each with a single antenna transmitting messages simultaneously over Gaussian channels to a single receiver with M receive antennas. If user i transmits signal U_i , the sum of the individual transmission powers satisfies the same power constraint as that of the BC, i.e., $\sum_{i=1}^K \xi_i \leq P$, where $\xi_i = E[|U_i|^2]$. The decoding order of the MAC is the reverse of the encoding order for the dual MIMO BC, i.e., the signals of users $i+1$ ($1 \leq i < K$) through K are treated as perfectly known, and the rest is unknown interference. Then, the achievable rate for user i is [23]

$$R_i^{MAC} \leq \frac{1}{2} \log \left(\frac{|\mathbf{I}_M + \sum_{k=1}^i \xi_k \mathbf{h}_k^H \mathbf{h}_k|}{|\mathbf{I}_M + \sum_{k=1}^{i-1} \xi_k \mathbf{h}_k^H \mathbf{h}_k|} \right). \quad (5)$$

According to [4], a transformation which depends on the channel matrix and the noise statistics can be defined to map the set of MAC powers $\{\xi_i\}$ to the BC transmitter covariance matrices $\{\mathbf{S}_i\}$ and hence $\{\mathbf{b}_i\}$, and vice versa. Moreover, the sets of achievable rates are equal in both domains, i.e., $R_i^{MAC} = R_i^{BC}$.

III. OPTIMAL PRECODING UNDER INDIVIDUAL RATE CONSTRAINTS

A code design for MIMO BC requires finding an appropriate precoding matrix \mathbf{B} . Previous information-theoretic works [3]–[6] have mainly focused on evaluating \mathbf{B} to maximize the achievable sum-rate, in which case the encoder and decoder for each user must be able to operate at an arbitrary code rate. However, as mentioned earlier, such a scheme might result in an unfair rate allocation. Also it is not clear how to design good dirty-paper codes at different rates on the fly. Therefore, we consider a scenario in which each user is assumed to operate at a fixed transmission rate, for which we already have an efficient dirty-paper code design [11]. Note that this strategy, as opposed to the sum-rate maximizing scheme, cannot optimally allocate rates to the user, and thus for the same sum-rate the former will suffer a performance loss compared to the latter, as indicated in Section VI.

Let R_i be the rate of the practical channel code (for $i = 1$) and the dirty-paper code (for $2 \leq i \leq K$). We assume that there is a minimum bit-error rate (BER) requirement at each user, which translates into a minimum requirement on the operating SINR. Let the rate- R_i code perform δ_i dB away from the corresponding Shannon limit of $10 \log_{10}(2^{2R_i} - 1)$ dB at the given minimum BER. Then, in order to satisfy the BER requirement, its operating SINR $\eta_i = \frac{\mathbf{h}_i \mathbf{b}_i \mathbf{b}_i^H \mathbf{h}_i^H}{1 + \mathbf{h}_i \sum_{k=i+1}^K \mathbf{b}_k \mathbf{b}_k^H \mathbf{h}_i^H}$ must satisfy

$$\eta_i \geq \eta_i^o = 10 \log_{10}(2^{2R_i} - 1) + \delta_i \text{ dB}. \quad (6)$$

For example, the dirty-paper code [11] we use in this work performs $\delta_i = 1.53$ dB away from the Shannon limit at $R_i = 1.0$ b/s and a BER of 10^{-5} . Since there is a one-to-one correspondence between the rate R_i of the practical code and its minimum operating SINR η_i^o given by (6), in the sequel, we exclusively speak of the SINR constraints instead of the rate constraints.

Note that for a given transmit power, one might not always be able to satisfy the minimum SINR requirement. Thus,

in order to get the best BER performance, we uniformly maximize the SINR at each user. Specifically, we consider the following optimization problem:

$$\max_{\mathbf{B}} \beta \quad (7)$$

such that

$$\frac{\mathbf{h}_i \mathbf{b}_i \mathbf{b}_i^H \mathbf{h}_i^H}{1 + \mathbf{h}_i \sum_{k=i+1}^K \mathbf{b}_k \mathbf{b}_k^H \mathbf{h}_i^H} \geq \beta \eta_i^0, \quad i = 1, \dots, K$$

and

$$E[\mathbf{u}^H \mathbf{B}^H \mathbf{B} \mathbf{u}] \leq P,$$

where the first K constraints correspond to the SINR requirements and the last one defines the transmitter power constraint. For the sake of simplicity, we describe the solution to this optimization procedure for a fixed encoding order. A search for an encoding order that results in the largest β is then needed. Such an encoding order can be found by using the iterative algorithm of [15]. In the following, we discuss optimal precoding for the case of a degraded Gaussian BC [1] before moving on to the case of MIMO BC.

A. Precoding for two-user degraded Gaussian BC

The problem here is to find a precoding matrix \mathbf{B} that solves the optimization problem of (7). Recall from Section II-A that for a given transmit power P , the precoding matrix is given by $\mathbf{B} = [\sqrt{(1-\rho)P}, \sqrt{\rho P}]$. Thus, the objective is to find the power allocation parameter ρ which maximizes β . It can be shown that β is maximized when the SINR constraints and the power constraint in (7) are all met with equality. Thus, the optimum ρ^* and β^* can be obtained by solving the equations $\eta_1 = \frac{|h_1|^2(1-\rho)P}{|h_1|^2\rho P + 1} = \beta\eta_1^0$ and $\eta_2 = |h_2|^2\rho P = \beta\eta_2^0$, yielding

$$\rho^* = \frac{-\left(\frac{\eta_1^0}{|h_1|^2} + \frac{\eta_2^0}{|h_2|^2}\right) + \sqrt{\left(\frac{\eta_1^0}{|h_1|^2} + \frac{\eta_2^0}{|h_2|^2}\right)^2 + 4\frac{\eta_1^0\eta_2^0P}{|h_2|^2}}}{2\eta_1^0P} \quad (8)$$

and

$$\beta^* = \frac{\rho^*P|h_2|^2}{\eta_2^0}. \quad (9)$$

B. Precoding for MIMO BC

In this subsection, we first present an iterative approach to solving (7), we then discuss some sub-optimal choices of the precoding matrix \mathbf{B} .

1) *Optimal precoding*: The SINR constraints in (7) in the case of MIMO BC are not convex. In order to simplify the problem, we can invoke duality between BC and MAC. As a result, the problem (7) is equivalent to the following MAC domain problem:

$$\max_{\{\xi_i\}} \beta \quad (10)$$

such that

$$\xi_i \mathbf{h}_i (\mathbf{I} + \sum_{j=1}^{i-1} \xi_j \mathbf{h}_j^H \mathbf{h}_j)^{-1} \mathbf{h}_i^H \geq \beta \eta_i^0, \quad i = 1, \dots, K$$

and

$$\sum_{k=1}^K \xi_k \leq P.$$

The optimal choice of $\{\xi_i\}$ results in the SINR constraints and the power constraint being met with equality. Thus, the optimal β^* and $\{\xi_i^*\}$ are solutions to the equations obtained by replacing the $K+1$ inequalities in (10) by equalities. However, obtaining a closed form solution to this problem is not easy. We thus consider an alternative approach which involves solving the converse problem of (10), i.e., the problem of finding the minimum transmitter power such that the minimum SINR constraints of $\beta\eta_i^0$ are satisfied (here β is assumed to be a constant). This converse problem can be stated as

$$\min_{\{\xi_i\}} \sum_{k=1}^K \xi_k \quad (11)$$

such that

$$\xi_i \mathbf{h}_i (\mathbf{I} + \sum_{j=1}^{i-1} \xi_j \mathbf{h}_j^H \mathbf{h}_j)^{-1} \mathbf{h}_i^H \geq \beta \eta_i^0, \quad i = 1, \dots, K.$$

The solution to this converse problem also requires the SINR constraints to be met with equality and can easily be obtained by following the procedure in [15], [21]. It is obvious that if the transmitter power in (10) is equal to the optimum objective function value P^* of (11), the optimum β^* obtained by solving (10) would be the same as the β used in (11). In addition, P^* is a monotonically increasing function of β . Thus, the solution to (10) can be obtained from the following iterative procedure:

- 1: Select initial β_{min} and β_{max} .
- 2: $\beta := (\beta_{min} + \beta_{max})/2$.
- 3: Solve (11) using procedure in [15], [21] to obtain $\{\xi_i^*\}$. Evaluate $P^* = \sum_{k=1}^K \xi_k^*$.
- 4: If $P^* > P$, $\beta_{max} = \beta$ Else If $P^* < P$, $\beta_{min} = \beta$.
- 5: If $|P - P^*| \geq \epsilon$ Goto step 2 Else End.

One can then apply the MAC-to-BC transformation of [4] to obtain the optimal precoding matrix \mathbf{B} .

2) *Zero-forcing*: We briefly mention the two suboptimal approaches considered in [3], namely, zero-forcing DPC (ZFDPC) and zero-forcing beamforming (ZFBF) [24]. As the name zero-forcing suggests, the choice of the precoding matrix forces the interference at each user to be zero, and hence induces K non-interfering channels between the transmitter and the K users.

Let $\mathbf{H} = \mathbf{G}_{K \times K} \mathbf{Q}_{K \times M}$ be the QR decomposition of the channel matrix obtained by Gram-Schmidt orthogonalization, where \mathbf{G} is a lower triangular matrix, i.e., $g_{ij} = 0$ for $j > i$, and \mathbf{Q} satisfies $\mathbf{Q}\mathbf{Q}^H = \mathbf{I}_K$. As before, M denotes the number of transmitter antennas, and K is the number of users each with a single antenna. The precoding matrix is chosen as $\mathbf{B} = \mathbf{Q}^H \mathbf{R}_{K \times K}$, where \mathbf{R} is a *diagonal* matrix. This choice of \mathbf{B} introduces an additional equality constraint in the optimization problem (7) and ensures that at user i the interference from all users $j > i$ is forced to zero. With the introduction of the equality constraint, (7) is equivalent to

$$\max_{\{|r_{ii}|\}} \beta \quad (12)$$

such that

$$|g_{ii} r_{ii}|^2 \geq \beta \eta_i^0, \quad i = 1, \dots, K$$

and

$$\sum_{k=1}^K |r_{ii}|^2 \leq P,$$

the solution to which can be easily obtained as $|r_{ii}^*|^2 = \frac{\beta^* \eta_i^o}{|g_{ii}|^2}$ with $\beta^* = P \sum_{k=1}^K \frac{|g_{ii}|^2}{\eta_i^o}$.

On the other hand, in ZFBF, the precoding matrix is chosen as $\mathbf{B} = \mathbf{H}^\dagger \mathbf{R}$, where \mathbf{H}^\dagger is the pseudo inverse of the channel matrix \mathbf{H} and \mathbf{R} is again a diagonal matrix. The received signal vector in this case is $\mathbf{Y} = \mathbf{H}\mathbf{B}\mathbf{U} + \mathbf{Z} = \mathbf{R}\mathbf{U} + \mathbf{Z}$. Thus, the interfering signals from all users are forced to zero, which simplifies the code design since DPC is no longer required. In this case, one needs to solve the following problem:

$$\max_{\{r_{ii}\}} \beta \quad (13)$$

such that

$$|r_{ii}|^2 \geq \beta \eta_i^o, \quad i = 1, \dots, K$$

and

$$\sum_{j=1}^M \sum_{k=1}^K |h_{jk}^\dagger r_{kk}|^2 \leq P.$$

The solution to the ZFBF problem can be obtained as $|r_{ii}^*|^2 = \beta^* \eta_i^o$ with $\beta^* = \frac{P}{\sum_{j=1}^M \sum_{k=1}^K |h_{jk}^\dagger|^2 \eta_i^o}$ ($1 \leq i \leq K$).

Although zero-forcing is near optimal when the sum-rate is maximized [3], simulations (see Section VI) show that when considering outage performance in a fading environment, it is far from optimum in the setup with individual rate constraints. One disadvantage of zero-forcing is that under individual rate constraints, it fails when there are more users than the total number of transmit antennas, i.e., the solution to the optimization problems (12) and (13) always yields $\beta^* = 0$ when $M < K$. This problem was recently addressed in [25], where different suboptimal solutions based on partial interference cancellation are proposed.

IV. DPC DESIGN BASED ON NESTED TURBO CODES

DPC [8] can be implemented using nested lattice codes [10], where the source code is nested within a channel code. Chou *et al.* [26], [27] reported a turbo-coded trellis-based construction for DPC by nesting a trellis coded quantization (TCQ) code inside a turbo-trellis coded modulation (TTCM). However, owing to the structural dissimilarity between TCQ and TTCM, the actual performance of TCQ is severely degraded and thus it affects the overall performance of the scheme, especially at low transmission rates. An improved design which alleviates the effect of this structural dissimilarity was proposed in [11] where a stronger source code referred to as the *turbo-like* TCQ was employed. In order to provide a description of how the improved scheme of [11] works, we first briefly review the TCQ/TTCM scheme of [26], [27].

The trellis structure in the TCQ/TTCM scheme of [26], [27] is constructed via a rate- $k/n/m$ concatenated code (denoted by C_1+C_2 , with C_1 being the rate- $\frac{k}{n}$ convolutional code and C_2 being the rate- $\frac{n}{m}$ convolutional code) as shown in the encoder block diagram in Fig. 2 (a). TCQ essentially relies on the trellis Γ_1 formed by C_1+C_2 ; the TTCM code consists of a parallel concatenated code with C_2 in both branches. C_2 in the

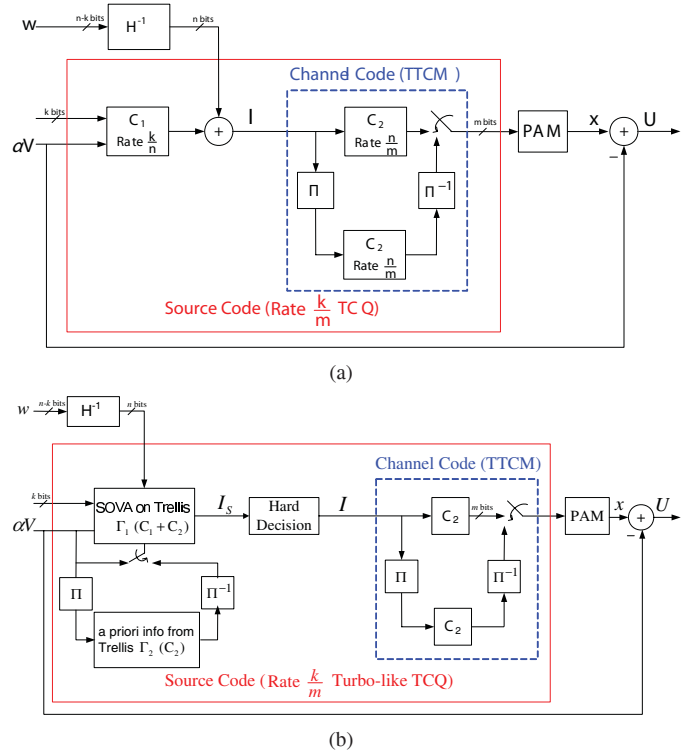


Fig. 2. Block diagram of (a) the TCQ/TTCM encoder of [26], [27] and (b) the turbo-like TCQ/TTCM encoder of [11].

bottom branch is preceded by an n -bit symbol interleaver and followed by an m -bit symbol deinterleaver. The two branches are multiplexed by taking even/odd-indexed symbols from the top/bottom branch before PAM.

At the encoder, every $(n-k)$ -bit segment of the message w is mapped to an n -bit symbol by the pseudo inverse of the parity-check matrix H of C_1 before being added to an output n -bit symbol of C_1 . This way, the codewords of C_1+C_2 are shifted by a fixed amount as determined by the message w . Consequently, one coset of TTCM codewords is selected by w to be used for TCQ, which uses the Viterbi algorithm to search for its input sequence of k -bit symbols so that the scaled side information sequence αV is quantized to an output sequence x . The quantization error sequence $U = x - \alpha V$ is finally transmitted. At the decoder, the n -bit input symbols to TTCM (or codewords of C_1) are recovered by an iterative BCJR decoder. Finally, the transmitted message w is reconstructed by calculating the syndromes of the recovered codewords of C_1 .

Whereas the presence of an interleaver in Fig. 2 (a) greatly boosts the performance of the TTCM code over TCM, the TCQ source code suffers because the interleaver significantly increases the number of paths that need to be searched, making the Viterbi algorithm no longer a viable solution to finding the closest codeword x to αV . In [26], [27], the bottom branch of TTCM is simply ignored during TCQ. But the actual average quantization error $E[U^2]$ includes contributions from both even-indexed symbols from the top branch and odd-indexed symbols from the bottom branch. This leads to an *extra* quantization error in $E[U^2]$ that is responsible for the degradation of the source code performance in TCQ/TTCM.

In order to reduce this extra quantization error, a nested turbo code design was proposed in [11] to take into account the bottom branch of TTCM in source coding. The encoder block diagram of that code design is depicted in Fig. 2 (b), where the obvious difference from Fig. 2 (a) is the employment of a new turbo-like TCQ in place of TCQ. The main difference between turbo-like TCQ and TCQ lies in the computation of the input sequences of symbols $\mathbf{I} = [I(0), \dots, I(L-1)]$ to the TTCM encoder. This sequence is evaluated using a soft-output Viterbi algorithm, which adopts a composite distortion metric to take both branches into account. This distortion metric is the same as the TCQ metric for the even positions on the top branch of TTCM, whereas the distortion for the odd positions on the bottom branch of TTCM is provided by the trellis Γ_2 in the form of *a-priori* information.

The source coding component of Fig. 2 (b) is referred to as turbo-like TCQ because it has a parallel concatenated structure with interleavers Π and Π^{-1} , and more importantly, it essentially implements the first iteration of turbo TCQ, which takes advantage of the fact [28] that turbo TCQ generally improves upon TCQ at the first iteration before losing ground at subsequent iterations. Without iterative quantization (or source encoding), the distortion from the bottom branch of TTCM can only be included in the form of *a priori* information. This limits the improvement of turbo-like TCQ/TTCM over TCQ/TTCM in terms of source coding performance. To alleviate the impact of the inaccurate *a-priori* information from the bottom branch, another novelty in the code design of [11] lies in decreasing the percentage of samples processed by that branch and picking the optimal percentage for best DPC performance.

Results in [11] indicate that nested turbo coding performs significantly better than the scheme of [26], [27] at all transmission rates. For example, at 1.0 b/s, nested turbo coding performs 1.53 dB away from capacity. In contrast, the TCQ/TTCM scheme of [26], [27] shows a gap of 2.07 dB from capacity.

V. OVERALL CODE DESIGN FOR MIMO BC

In this section, we describe the practical coding strategy for the MIMO BC using the capacity approaching DPC design of Section IV. According to the capacity achieving scheme described in Section II-B, user 1 does not have any side information. Thus, user 1 should employ a traditional channel code. For our code design we use conventional TTCM and a PAM constellation for user 1. The remaining users exploit the turbo-like TCQ/TTCM scheme for DPC. We thus require one conventional channel code and $K - 1$ dirty-paper codes. Our overall DPC-based code design is schematically shown in Fig. 3. This design is applicable to both the degraded Gaussian BC and the MIMO Gaussian BC. In the following, we discuss several issues in applying the turbo-like TCQ/TTCM scheme to code design for the MIMO BC.

1. *Normalization at the receiver*: The complex baseband equivalent of the received signal at user i is given by (3). In order for user i to decode its message correctly, the user needs to work with the same constellation as was used at the encoder for X_i . For this purpose, as indicated in Fig. 3 we

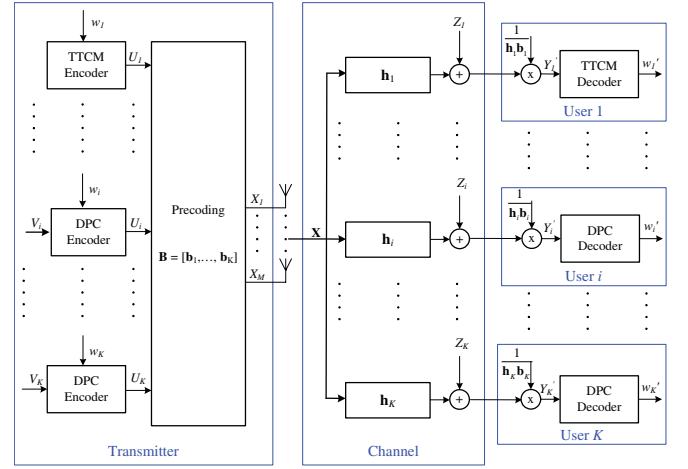


Fig. 3. Block diagram of our proposed overall DPC-based coding scheme for MIMO BC.

normalize the received signal at user i by $\mathbf{h}_i \mathbf{b}_i$. It is apparent that this normalization does not affect the received SINR. The resulting signal can be written as

$$Y_i' = \underbrace{U_i(w_i; U_1, \dots, U_{i-1})}_{\text{Useful signal}} + \underbrace{\frac{\mathbf{h}_i \sum_{j=1}^{i-1} \mathbf{b}_j U_j}{\mathbf{h}_i \mathbf{b}_i}}_{\text{Encoder side information}} + \underbrace{\frac{\mathbf{h}_i \sum_{j=i+1}^K \mathbf{b}_j U_j}{\mathbf{h}_i \mathbf{b}_i}}_{\text{Unknown interference } V_i} + \underbrace{\frac{Z_i}{\mathbf{h}_i \mathbf{b}_i}}_{\text{Gaussian noise}}. \quad (14)$$

Hence, the side information term in Fig. 3 for the i^{th} ($i > 1$) user is given by $V_i = \frac{\mathbf{h}_i \sum_{j=1}^{i-1} \mathbf{b}_j U_j}{\mathbf{h}_i \mathbf{b}_i}$.

2. *Extension to complex baseband*: The theoretical background in the earlier sections assumes the baseband equivalent of the coded messages U_i to be complex numbers. In practice this can be realized by using a two dimensional constellation such as QAM. However, note that the coded message in the DPC scheme of Fig. 2 (b) is mapped to a PAM constellation, indicating that the baseband equivalent of the signals are real. In order to get a complex output, we combine the outputs of two independent nested turbo codes (denoted by U_I and U_Q), in which the phase of U_Q is shifted by 90 degrees via multiplication by $j = \sqrt{-1}$. If V is a complex side information at the encoder, the side information inputs to the two encoders are $V_I = \text{Re}\{V\}$ and $V_Q = \text{Im}\{V\}$. At the decoder, the real part of the received signal Y_I can be tied to the input of one DPC decoder, while the imaginary part Y_Q to another independent decoder. This way we effectively convert the PAM constellation of our DPC scheme to a QAM constellation. Note that the same principle can also be applied to the TTCM code of user 1.

3. *CSI at the transmitter*: Note that the calculation of precoding matrices in various schemes in Section III requires perfect CSI to be available at the transmitter. Moreover, DPC also implicitly entails the requirement of perfect CSI at the transmitter. While obtaining this CSI at the receivers is usually simple, acquiring it perfectly at the transmitter is often unrealistic in practice. In order to obtain the CSI, the transmitter should rely on a feedback mechanism from

the individual users. Even with this feedback, the CSI is not guaranteed to be available perfectly at the transmitter. However, as mentioned earlier, for our designs we assume that the CSI is available at the encoder perfectly before the transmission begins. We will leave the case of imperfect CSI at the transmitter as a possible extension to this work.

4. *Non-Gaussian interference*: As indicated in Section IV, the codewords of the convolutional code C_2 are mapped to a PAM constellation. This results in an interference at the decoders which is not Gaussian. The practical decoders should thus be designed to exploit the non-Gaussian statistics of the interference. This would require them computing the channel likelihood values based on the actual noise plus interference distribution. However, this computation becomes cumbersome, especially for a large number of interferers and/or a large constellation size. For the sake of simplicity, our implemented decoders assume that the noise plus interference is Gaussian with variance equal to that of the actual distribution. Because of this assumption, our decoders suffer a performance loss from the case which exploits the non-Gaussian statistics. Fortunately, simulations (not included in this paper) indicate that this performance loss is not significant. Similarly, the side information V_i in (14) will not be Gaussian, but Costa's capacity result [8] holds also for arbitrary side information [29]. Our simulations with the dirty-paper code construction of Section IV verify this.

VI. SIMULATION RESULTS

In our code design, we use a 16-state, rate- $\frac{1}{2}$ systematic convolutional code for TTCM. The code polynomial is chosen as the constraint-length four Ungerboeck code for the PAM constellation (suboptimally to maximize the average Euclidean distance between TCM codewords). Specifically, the parity check polynomials for this code are $h_0(D) = 23$ and $h_1(D) = 10$ in octal notation. For the practical DPC scheme of Section IV, we choose C_1 as a 16-state, rate- $\frac{1}{2}$, non-systematic convolutional code with generator polynomials $g_0(D) = 23$ and $g_1(D) = 10$. Code C_2 , on the other hand, is a 16-state, rate- $\frac{2}{3}$, systematic convolutional code with parity check polynomials $h_0(D) = 23$, $h_1(D) = 10$, and $h_2(D) = 0$. The block length for both TTCM and dirty-paper code is fixed at 10,000 samples. In the following, we first present our simulation results for the two-user degraded Gaussian BC followed by results for MIMO Gaussian BC. The results for MIMO Gaussian BC include not only the performance of our nested turbo scheme with optimal precoding, but also with ZFDPC and ZFBF. However, as mentioned earlier, the performance gap between the three approaches depends greatly on what the exact realization of the channel coefficients is. Thus, in order to provide a fair comparison, we additionally consider the performance of these strategies over an ensemble of channel coefficients. Specifically, we analyze the performance based on an outage measure for the case of slow Rayleigh fading. All simulations are carried out at equal rates of 1 b/s for each user. This does not mean that our design only works for equal rate constraints, since the precoding strategies described in Section III are valid for any set of (including unequal) rate constraints. In addition, our DPC scheme based on nested turbo codes has been designed to work at rates of 0.5, 1.0 and 2.0 b/s in [11],

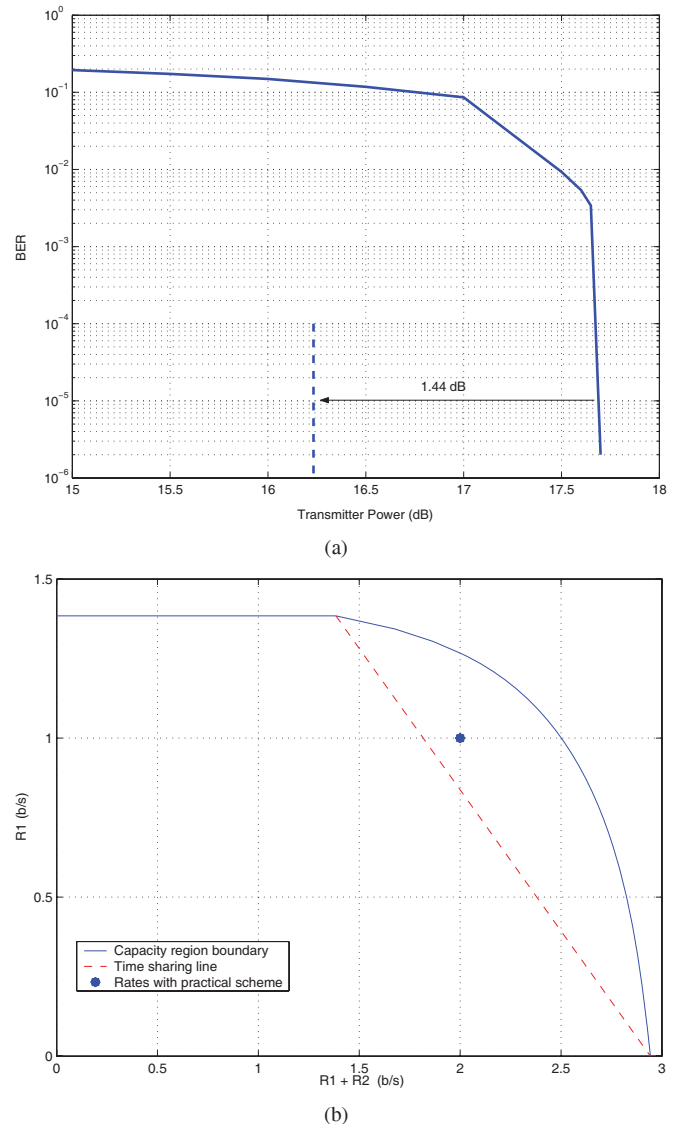


Fig. 4. (a) BER vs. the transmission power P for the degraded Gaussian BC, with $R_1 = R_2 = 1$ b/s, $h_2 = 1$, $h_1 = \sqrt{0.1}$. The dashed line represents the capacity. (b) The capacity region for the degraded Gaussian BC with transmission power $P = 17.65$ dB, $h_2 = 1$, and $h_1 = \sqrt{0.1}$.

and can be modified to work at higher rates as well. The TTCM code can similarly be modified. In short, our schemes are readily applicable to the case with unequal rate constraints on the users.

A. Two-user degraded Gaussian BC

We simulate our DPC-based design for Cover's two-user degraded Gaussian BC with channel coefficients $h_1 = \sqrt{0.1}$ and $h_2 = 1$. Simulations indicate that at a rate of 1.0 b/s and a BER of 10^{-5} , the TTCM code for user 1 performs $\delta_1 = 0.98$ dB, and the dirty-paper code at user 2 performs $\delta_2 = 1.53$ dB away from the Shannon limit. For a given transmit power, we evaluate the optimal ρ^* from (8) and thus the optimal \mathbf{B} , and plot the resulting BER averaged over the two users in Fig. 4 (a).

At a BER of 10^{-5} , it is seen that the transmission power needed to achieve $R_1 = R_2 = 1.0$ b/s is 17.65 dB, which is 1.44 dB away from the minimum required power if ideal

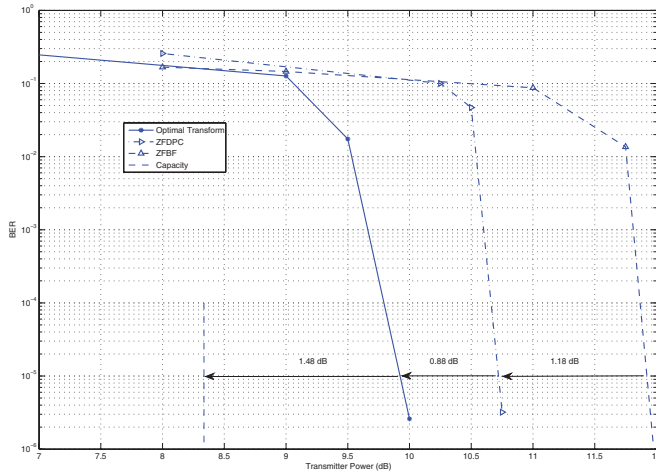


Fig. 5. BER vs. the transmission power for the two-user MIMO Gaussian BC with two transmit antennas, $\mathbf{h}_1 = [1/\sqrt{2}e^{2.37j}, \sqrt{3}/2e^{2.14j}]$, $\mathbf{h}_2 = [e^{2.23j}, 0.5e^{0.87j}]$ and $R_1 = R_2 = 1$ b/s.

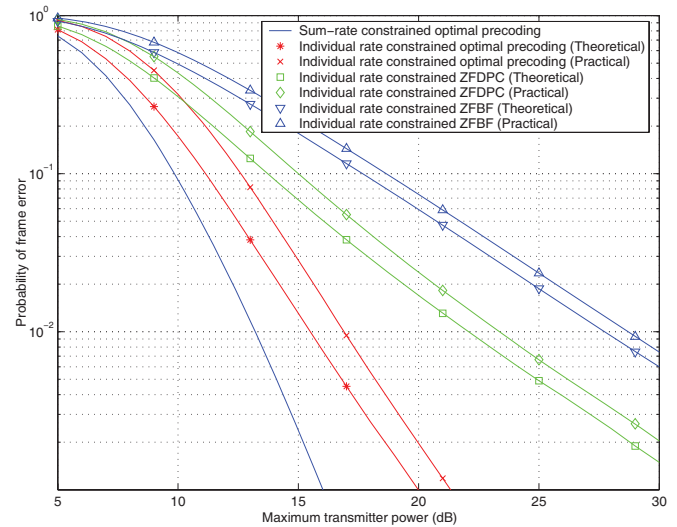
codes were used. This result is 1.8 dB better than that reported in [17]. Fig. 4 (b) depicts the capacity region for $P = 17.65$ dB, which is the required total power for our code design to operate at $R_1 = R_2 = 1.0$ b/s. Our operating point is significantly above the time-sharing line.

B. Non-degraded MIMO Gaussian BC

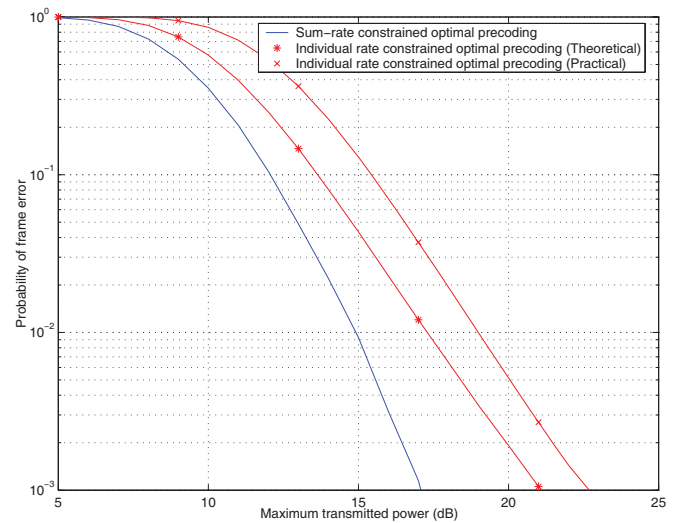
We also simulate our design for a two-user MIMO Gaussian BC with two transmit antennas at the base-station by fixing the channel coefficients as $\mathbf{h}_1 = [1/\sqrt{2}e^{2.37j}, \sqrt{3}/2e^{2.14j}]$ and $\mathbf{h}_2 = [e^{2.23j}, 0.5e^{0.87j}]$. For a given transmit power, we evaluate the optimum precoding matrix using the procedure outlined in Section III-B1. We plot the BER (averaged over the two users) versus the transmitter power curves in Fig. 5. It is seen that our practical code with optimal precoding performs only 1.48 dB away from the theoretical limit. Practical coding with ZFDPC performs 0.88 dB worse than with optimal precoding, while ZFBF loses an additional 1.18 dB.

C. Non-degraded MIMO fading BC

We assume the channels undergo independent Rayleigh slow flat fading, i.e., each element of the matrix \mathbf{H} is i.i.d., circularly symmetric, zero-mean, complex Gaussian with unit variance, and \mathbf{H} is frame-wise constant. For a given realization of the channel coefficients \mathbf{H} and a transmitter power constraint P , we compute the maximum $\beta^*(\mathbf{H}, P)$ (the arguments indicate that β^* is a function of \mathbf{H} and P) as described in Section III-B. Thus, the maximum achievable SINR at user i is $\beta^* \eta_i^o$. The actual BERs at each user will satisfy the minimum BER requirement if $\beta^* \geq 1$. In order to analyse the system performance, we consider the probability of frame error-like measure $P_{fe}(P) = Pr(\beta^*(\mathbf{H}, P) < 1)$, where the probability is calculated by averaging over the ensemble of the channel matrix \mathbf{H} . This probability can be thought of as the outage probability, where an outage event occurs whenever $\beta^*(\mathbf{H}, P) < 1$. P_{fe} is equivalent to the frame error probability if the frames at all users are received in error if and only if an outage event occurs. Note that this might not always be true,



(a)



(b)

Fig. 6. (a) Probability of frame error vs. maximum transmission power P for $K = 2$ and $M = 2$. (b) Probability of frame error vs. maximum transmission power P for $K = 3$ and $M = 2$.

nevertheless we still call this measure as the *probability of frame error*.

One can also evaluate P_{fe} by using an alternative method which involves calculating the minimum required transmitter power $P^*(\mathbf{H})$ such that the set of SINR requirements for $\{\eta_i\}$ in (6) are satisfied ($P^*(\mathbf{H})$ can be obtained by solving (11)). Note that $\beta^* < 1$ if and only if $P^*(\mathbf{H}) > P$, and thus the probability of frame error can also be evaluated as $P_{fe}(P) = Pr(P^*(\mathbf{H}) > P)$. We use this method of evaluating the probability of frame error when discussing our results. In the following, we evaluate the probability of frame error versus the transmitter power for the cases when the number of users or transmit antennas is up to four.

1) *Simulations for the two-user case:* Fig. 6 (a) compares code designs based on optimal DPC, ZFDPC, and ZFBF in terms of the probability of frame error vs. transmission power for two antennas at the transmitter. At a frame error rate of 1%, compared to the sum-rate capacity ($R_1 + R_2 = 2.0$ b/s) curve [3], our practical DPC-based code design loses 3.70 dB

in performance. 2.30 dB of this loss is due to the individual rate constraints $R_1 = 1.0$ and $R_2 = 1.0$ b/s. Practical coding accounts for the remaining 1.40 dB loss. Compared to optimal DPC-based design, ZFDPC is approximately 6.50 dB worse while ZFBF loses an additional 5.50 dB.

2) *Simulations for the three-user case:* The results for three users and two transmit antennas are provided in Fig. 6 (b). Since in this case the number of transmit antennas is less than the number of users, zero-forcing (both ZFDPC and ZFBF) does not work. The sum-rate capacity curve is obtained by using the iterative waterfilling algorithm of [30]. Compared to this sum-rate capacity curve, our practical DPC-based code design loses 4.07 dB in performance. About 2.38 dB of this loss is due to the individual rates being constrained, while the remaining loss of 1.64 dB is due to practical coding.

With three transmit antennas, the loss due to the constraints on the individual rates reduces to 1.48 dB, while the practical coding loss is 1.58 dB. Simulation results show that our practical DPC-based scheme with optimal precoding outperforms theoretical ZFDPC and ZFBF by 6.50 dB and 15.50 dB, respectively.

3) *Simulations for the four-user case:* The frame error rate (FER) versus the maximum transmission power curves for four users and two transmit antennas is provided in Fig. 7 (a); the loss due to individual rates being constrained is 2.25 dB while the practical coding loss is 1.99 dB.

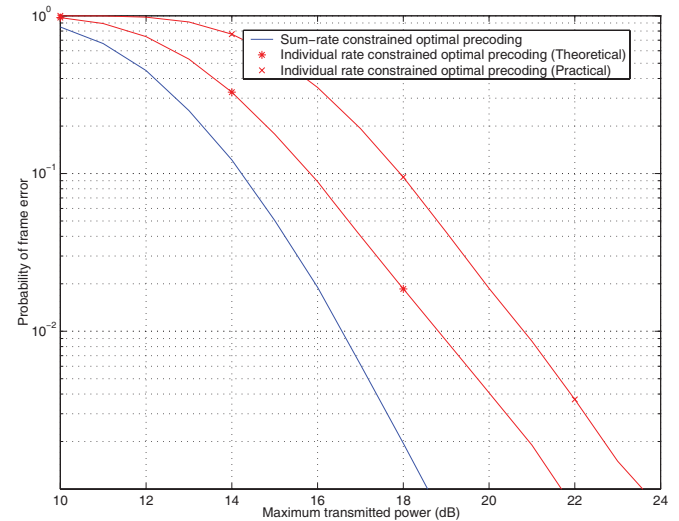
In the case with three transmit antennas, the loss due to the constraints on individual rates is 1.28 dB and the practical coding loss is 1.69 dB.

Fig. 7 (b) shows similar results for the case of four transmit antennas. Compared to the sum-rate capacity curve, our practical DPC scheme loses 2.54 dB. This loss can be broken down into a 0.92 dB loss due to the constraints on the individual rates, and 1.62 dB because of practical coding. In addition, it can be observed that ZFPDC is 7.20 dB worse than the practical DPC scheme with optimal precoding while ZFBF loses an additional 11.50 dB.

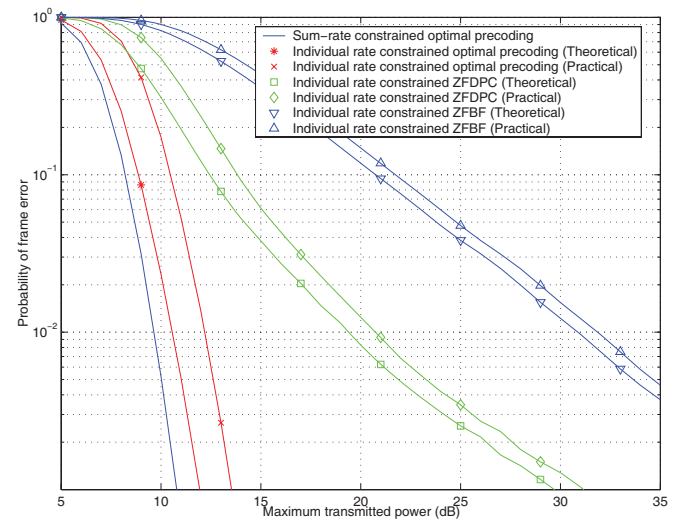
Discussions:

Note that in the results presented above, compared to the optimal sum-rate maximizing scheme [30], our DPC based scheme suffers two types of performance losses: one due to the constraints on individual rates and another from practical channel coding (for the first user) and DPC (for the remaining $K - 1$ users). In the following we make a few intuitive comments about these two types of losses and how they change with the number of transmit antennas M and the number of users K .

Loss due to constraints on individual rates: Our simulations indicate that, for a fixed number of users and increasing number of transmit antennas, the performance gap between the outage curve for the sum-rate capacity and that of the individual rate constrained scheme decreases. This is because the rate allocation to the individual users becomes fairer with increased number of transmit antennas, i.e., the disparity between different rate allocations reduces. For example, with a single transmit antenna, the sum-rate maximizing scheme always allocates all the rate to the user with the best channel while no rate is allocated to the remaining users. On the other hand, with two transmit antennas, the optimal strategy



(a)



(b)

Fig. 7. (a) Probability of frame error vs. maximum transmission power P for $K = 4$ and $M = 2$. (b) Probability of frame error vs. maximum transmission power P for $K = 4$ and $M = 4$

might not always require allocating zero rate to any one of the users. Thus, with increased number of antennas, it becomes more likely that the rate allocations of the sum-rate maximizing scheme also satisfy the minimum rate/SINR requirement of the individual rate constraint scheme, leading to reduced performance loss.

When the number of transmit antennas is fixed, increasing the number of users affects the loss due to individual rate constraints in two ways. For a fixed encoding order, increasing the number of users increases the number of individual rate constraints, hence, the rate allocation which achieves the sum-rate capacity would be *less* likely to satisfy all the individual rate constraints as well. On the other hand, searching over all possible encoding orders has an opposite effect, i.e., it *increases* the likelihood that the rate allocation of the sum-rate maximizing scheme also satisfies the individual rate constraints. This is because the sum-rate capacity is independent of the encoding order, but the rate allocation is not. In other words, different encoding orders result in the same sum-rate

capacity, but different rate allocations. Thus, the effect on the overall loss due to individual rate constraints depends on which one of these two factors outweighs the other. Since the number of encoding orders grows as $K!$, for a large number of users, we expect the effect of the second factor to dominate the first. Indeed, this is verified by our simulations where for two transmit antennas, the loss due to individual rate constraints increases when the number of users is increased from two to three, but decreases when the number of users is increased from three to four.

Loss due to practical coding: Our experiments also indicate that under a fixed number of transmit antennas, increasing the number of users will lead to a higher practical coding loss. This is because each additional user incurs a loss due to practical DPC and thus the overall coding loss increases.

On the other hand, when the number of users is fixed, but the number of transmit antennas increases, the overall coding loss decreases. To explain this, we examine the practical coding loss for a fixed channel matrix \mathbf{H} and the fixed encoding order of Section II-B. For this we use the definition of probability of frame error which involves the minimum possible transmitter power, i.e., $P_{fe}(P) = Pr(P^*(\mathbf{H}) > P)$. Thus, as a first step we attempt to find an expression for the minimum transmit power in terms of the individual coding losses δ_i 's. Let the normalized inner product between the i^{th} and j^{th} row of the \mathbf{H} matrix be denoted by $\rho_{ij} = \frac{\mathbf{h}_i \mathbf{h}_j^H}{\|\mathbf{h}_i\| \|\mathbf{h}_j\|}$. Following the iterative procedure for evaluating $P^*(\mathbf{H})$ as described in [15], [21], we can write the dual MAC covariance of the i^{th} user as a function of the performance gap of the individual decoders from the Shannon's limit ($\delta_1, \dots, \delta_i$), the SINR requirements (η_1, \dots, η_i), and the channel correlations (given by ρ_{ij}). The covariances for the first three users are $\xi_1(\delta_1, \eta_1) = \frac{\delta_1 \eta_1}{\|\mathbf{h}_1\|^2}$, $\xi_2(\delta_1, \delta_2, \eta_1, \eta_2) = \frac{\delta_2 \eta_2}{\|\mathbf{h}_2\|^2} (1 - \frac{\delta_1 \eta_1}{1 + \delta_1 \eta_1} |\rho_{12}|^2)^{-1}$, and $\xi_3(\delta_1, \delta_2, \delta_3, \eta_1, \eta_2, \eta_3) = \frac{\delta_3 \eta_3}{\|\mathbf{h}_3\|^2} \left\{ 1 - \frac{\delta_1 \eta_1}{1 + \delta_1 \eta_1} |\rho_{13}|^2 - \frac{\delta_2 \eta_2}{(1 + \delta_2 \eta_2)(1 - \frac{\delta_1 \eta_1}{1 + \delta_1 \eta_1} |\rho_{12}|^2)} \left| \rho_{23} - \frac{\delta_1 \eta_1}{1 + \delta_1 \eta_1} \rho_{12} \rho_{31} \right|^2 \right\}^{-1}$. The covariances for the remaining users can be written in a similar manner. We define the overall multiplicative coding loss as the ratio of total transmit power required by practical coding over the power required by ideal coding. It is therefore given by $L = \frac{\sum_{i=1}^K \xi_i(\delta_1, \dots, \delta_i, \eta_1, \dots, \eta_i)}{\sum_{i=1}^K \xi_i(1, \dots, 1, \eta_1, \dots, \eta_i)}$. Note that L will heavily depend on the cross correlations ρ_{ij} of the channel matrix rows. To gain an insight into how the correlation affects L , we define an individual multiplicative loss as $l_i = \frac{\xi_i(\delta_1, \dots, \delta_i, \eta_1, \dots, \eta_i)}{\xi_i(1, \dots, 1, \eta_1, \dots, \eta_i)}$. If $p_i = \frac{\xi_i(1, \dots, 1, \eta_1, \dots, \eta_i)}{\sum_{i=1}^K \xi_i(1, \dots, 1, \eta_1, \dots, \eta_i)}$, then L can be written as a weighted average of the individual multiplicative losses, i.e., $L = \sum_{i=1}^K l_i p_i$. With decreasing magnitudes of the cross correlations ρ_{ij} , $i = 1, \dots, K$; $j \neq i$, l_i will decrease. Thus, we can bound it by $l_i \geq \delta_i$, with equality when all the cross correlation terms are zero, i.e., when the corresponding rows of the channel matrix are orthogonal. The overall loss can be bounded as $L \geq \min(\delta_1, \dots, \delta_K)$. On the other hand, we have $l_i \leq \delta_i \frac{\prod_{j=1}^{i-1} (1 + \delta_j \eta_j)}{\prod_{j=1}^{i-1} (1 + \eta_j)}$, with equality when all the cross correlation magnitudes are one, i.e., when the corresponding rows of the channel matrix lie in the same direction. Thus, $L \leq \max(\delta_1, \dots, \delta_K) \frac{\prod_{j=1}^K (1 + \delta_j \eta_j)}{\prod_{j=1}^K (1 + \eta_j)}$.

As seen from the above, even for the three user case, the expressions for the covariances are coupled in a very complicated manner. One can expect the expressions to get even more complex with increased number of users. The performance gap of the outage curves therefore becomes particularly hard to analyze in a Rayleigh fading environment. However, based on the analysis given above, one can gain an intuitive understanding of the practical loss behavior: with the same number of users but increased number of antennas, the coding loss decreases because the probability for the magnitudes of the correlation coefficients to be small increases. As two extreme examples, consider first the case when $M = 1$, for which the magnitudes of the correlations ρ_{ij} are always one; on the other hand, as $M \rightarrow \infty$, the law of large numbers kicks in and the correlations converge in probability to zero.

Finally, as opposed to the case with optimal precoding, the practical coding loss with ZFDPC is much easier to analyze primarily because there is no coupling between the users. Recall that the transmission power for practical ZFDPC is given by $P_{pr} = \sum_{i=1}^K \frac{\delta_i \eta_i}{|g_{ii}|^2}$, then $P_{pr} \leq \max(\delta_1, \dots, \delta_K) \sum_{i=1}^K \frac{\eta_i}{|g_{ii}|^2} = \max(\delta_1, \dots, \delta_K) P_{th}$, where P_{th} is the required transmission power when all the decoders are operating at Shannon's limit. Thus, regardless of the number of users, the overall practical coding loss of the ZFDPC scheme is upper bounded by $\max(\delta_1, \dots, \delta_K)$, which in our case is 1.53 dB. This is verified by results in Figs. 6 and 7.

VII. CONCLUSIONS

We have presented capacity-approaching code designs for the degraded Gaussian BC and for the MIMO Gaussian BC. We also additionally consider the performance of our designs over a MIMO Rayleigh fading BC. The main component of our code designs is a practical DPC scheme based on nested turbo codes. Simulation results for the two-user MIMO Gaussian BC with two transmit antennas indicate a performance loss of only 1.48 dB from the theoretical limit. For the fading case, our schemes exhibit a practical coding loss of 1.4, 1.64, and 1.99 dB for two transmit antennas and two, three, and four users, respectively. Moreover, our results show a significant performance gain of optimal DPC over other suboptimal strategies (e.g., time sharing and zero-forcing linear beamforming).

There are many possible directions for future work. For example, one research direction is to study adaptive modulation and coding in our nested turbo DPC design in order to get a simple variable rate scheme. Another interesting topic is to design practical schemes for MIMO BCs when the CSI is not perfectly available at the transmitter.

REFERENCES

- [1] T. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory*, vol. 18, pp. 2-14, Jan. 1972.
- [2] P. Bergman, "A simple converse for broadcast channels with additive white Gaussian noise," *IEEE Trans. Inform. Theory*, vol. 20, pp. 279-280, Mar. 1974.
- [3] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1691-1706, July 2003.

- [4] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of MIMO broadcast channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2658-2668, Oct. 2003.
- [5] P. Viswanath and D. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1912-1921, Aug. 2003.
- [6] W. Yu and J. Cioffi, "Sum capacity of Gaussian vector broadcast channel," *IEEE Trans. Inform. Theory*, vol. 50, pp. 1875-1892, Sept. 2004.
- [7] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inform. Theory*, vol. 52, pp. 3936-3964, Sept. 2006.
- [8] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, pp. 439-441, May 1983.
- [9] N. Jindal and A. Goldsmith, "Dirty paper coding vs. TDMA for MIMO broadcast channels," *IEEE Trans. Inform. Theory*, vol. 51, pp. 1783-1794, May 2005.
- [10] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1250-1276, June 2002.
- [11] Y. Sun, M. Uppal, A. Liveris, S. Cheng, V. Stanković, and Z. Xiong, "Nested turbo codes for the Costa problem," *IEEE Trans. Commun.*, vol. 56, pp. 388-399, Mar. 2008.
- [12] U. Erez and S. ten Brink, "A close-to-capacity dirty paper coding scheme," *IEEE Trans. Inform. Theory*, vol. 51, pp. 3417-3432, Oct. 2005.
- [13] A. Bennatan, D. Burshtein, G. Caire, and S. Shamai, "Superposition coding for side-information channels," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1872-1889, May 2006.
- [14] Y. Sun, A. D. Liveris, V. Stanković, and Z. Xiong, "Near-capacity dirty-paper code designs based on TCQ and IRA codes," in *Proc. ISIT'05*, Adelaide, Australia, Sept. 2005.
- [15] C. Fung, W. Yu, and T. Lim, "Precoding for the multi-antenna downlink: multiuser SNR gap and optimal user ordering," *IEEE Trans. Commun.*, vol. 55, pp. 188-197, Jan. 2007.
- [16] M. Stojnic, H. Vikalo, and B. Hassabi, "Asymptotic analysis of the Gaussian broadcast channel with perturbation preprocessing," in *Proc. ICASSP'06*, Toulouse, France, May 2006.
- [17] J. Kasuma and K. Ramachandran, "Communicating by cosets and applications to broadcast," in *Proc. CISS'02*, Princeton, NJ, Mar. 2002.
- [18] W. Yu and J. Cioffi, "Trellis precoding for broadcast channel," *Proc. Globecom'01*, San Antonio, TX, Nov. 2001.
- [19] T. Sun, R. Wesel, M. Shane, and K. Jarett, "Superposition turbo TCM for multirate broadcast," *IEEE Trans. Commun.*, vol. 52, pp. 368-371, Mar. 2004.
- [20] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of Gaussian multiple-access and broadcast channels," *IEEE Trans. Inform. Theory*, vol. 50, pp. 768-783, May 2004.
- [21] M. Schubert and H. Boche, "Iterative multiuser uplink and downlink beamforming under SINR constraints," *IEEE Trans. Signal Processing*, vol. 53, pp. 2324-2334, July 2005.
- [22] M. Uppal, V. Stanković, and Z. Xiong, "Code designs for MIMO broadcast channels," in *Proc. ISIT'06*, Seattle, WA, July 2006.
- [23] W. Yu, W. Rhee, S. Boyd, and J. Cioffi, "Iterative water-filling for vector multiple access channels," *IEEE Trans. Inform. Theory*, vol. 50, pp. 145-152, Jan. 2004.
- [24] T. Yoo and A. Goldsmith, "On the optimality of multi-antenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Select. Areas Commun.*, vol. 24, pp. 528-541, Mar. 2006.
- [25] A. Dabagh and D. Love, "Precoding for multiuser broadcast channels with successive zero-forcing," *IEEE Trans. Signal Processing*, vol. 55, pp. 3837-3850, July 2007.
- [26] J. Chou, S. Pradhan, and K. Ramchandran, "Turbo coded trellis-based constructions for data embedding: channel coding with side information," in *Proc. 35th Asilomar Conf. Signals, Systems Computers*, Pacific Grove, CA, Nov. 2001.
- [27] J. Chou, "Channel coding with side information: theory, practice and applications," Ph.D. dissertation, University of California at Berkeley, Berkeley, CA, 2002.
- [28] V. Chappelier, C. Guillemot, and S. Marinkovic, "Turbo trellis-coded quantization," in *Proc. 3th Intl. Symp. Turbo Codes*, Brest, France, Sept. 2003.
- [29] A. Cohen and A. Lapidoth, "The Gaussian watermarking game," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1639-1667, June 2002.
- [30] N. Jindal, W. Rhee, S. Vishwanath, S. A. Jafar, and A. Goldsmith, "Sum power iterative waterfilling for multi-antenna Gaussian broadcast channels," *IEEE Trans. Inform. Theory*, vol. 51, pp. 1570-1580, Apr. 2005.



Momin Uppal received the B.S. degree in Electronic Engineering with highest distinction from GIK Institute of Engineering Sciences and Technology, Pakistan, in 2002. He received the M.S. degree in Electrical Engineering from Texas A&M University in 2006, where he is currently pursuing the Ph.D. degree. His research interests include dirty-paper coding, broadcast channels, and cooperative communications.



Vladimir Stanković (M'03) received the Dr.-Ing. degree in 2003 from the University of Leipzig, Germany. From 2003 to 2006, he was with the Department of Electrical and Computer Engineering at Texas A&M University, College Station, first as Postdoctoral Research Associate and then as a Research Assistant Professor. He is currently a Lecturer in the Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, UK. His research interests include multimedia networking, network information theory, and wireless

communications.



Zixiang Xiong (S'91-M'96-SM'02-F'07) received the Ph.D. degree in Electrical Engineering in 1996 from the University of Illinois at Urbana-Champaign.

From 1995 to 1997, he was with Princeton University, first as a visiting student, then as a research associate. From 1997 to 1999, he was with the University of Hawaii. Since 1999, he has been with the Department of Electrical and Computer Engineering at Texas A&M University, where he is a professor. He spent the summers of 1998 and 1999 at Microsoft Research, Redmond, WA. His research interests are network information theory, code designs and applications, networked multimedia and genomic signal processing.

Dr. Xiong received a National Science Foundation Career Award in 1999, an Army Research Office Young Investigator Award in 2000, an Office of Naval Research Young Investigator Award in 2001. He also received the 2006 IEEE Signal Processing Magazine best paper award. He served as associate editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY (1999-2005), the IEEE TRANSACTIONS ON IMAGE PROCESSING (2002-2005), and the IEEE TRANSACTIONS ON SIGNAL PROCESSING (2002-2006). He is currently an associate editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS (PART B). He is the publications chair of GENSIPS'06 and ICASSP'07 and the technical program committee co-chair of ITW'07.