On the Achievable Throughput of a Multiantenna Gaussian Broadcast Channel

Giuseppe Caire, *Senior Member, IEEE,* and Shlomo Shamai (Shitz), *Fellow, IEEE*

***Abstract*—A Gaussian broadcast channel (GBC) with r single-antenna receivers and t antennas at the transmitter is considered. Both transmitter and receivers have perfect knowledge of the channel. Despite its apparent simplicity, this model is, in general, a nondegraded broadcast channel (BC), for which the capacity region is not fully known.**

For the two-user case, we find a special case of Marton's region that achieves optimal sum-rate (throughput). In brief, the transmitter decomposes the channel into two interference channels, where interference is caused by the other user signal. Users are successively encoded, such that encoding of the second user is based on the noncausal knowledge of the interference caused by the first user. The crosstalk parameters are optimized such that the overall throughput is maximum and, surprisingly, this is shown to be optimal over all possible strategies (not only with respect to Marton's achievable region).

For the case of $r > 2$ users, we find a somewhat simpler choice of Marton's region based on ordering and successively encoding the users. For each user i in the given ordering, the interference caused by users $j > i$ is eliminated by zero forcing at the transmitter, while interference caused by users $j < i$ is taken into account by coding for noncausally known interference. Under certain mild conditions, this scheme is found to be throughput-wise asymptotically optimal for both high and low signal-to-noise ratio (SNR).

We conclude by providing some numerical results for the ergodic throughput of the simplified zero-forcing scheme in independent Rayleigh fading.

***Index Terms*—Dirty-paper coding, Gaussian vector broadcast channel (BC), multiple-antenna systems.**

I. Introduction

CONSIDER the discrete-time complex baseband multiple-input multiple-output (MIMO) channel with t transmitters and r receivers, defined by

$$\boldsymbol{y}_i = \boldsymbol{H}\boldsymbol{x}_i + \boldsymbol{z}_i, \qquad i = 1, \dots, n \tag{1}$$

where $\boldsymbol{x}_i \in \mathbb{C}^t$ is the transmitted vector at time i, $\boldsymbol{y}_i, \boldsymbol{z}_i \in \mathbb{C}^r$ are the corresponding received and noise vectors, and $\boldsymbol{H} \in \mathbb{C}^{r\times t}$ is the channel matrix, where $h_{k,\ell}$ denotes the complex channel gain from the input (transmitter) ℓ to the output (receiver) k.[1] The noise vector sequence $\{\boldsymbol{z}_i\}$ is independent and identically distributed (i.i.d.) with components $z_{k,i} \sim \mathcal{N}_{\mathbb{C}}(0,1)$ for $k = 1, \dots, r$ and $i = 1, \dots, n$.

Despite its simplicity, the model (1) is extremely rich and describes several situations of interest in data communications, depending on the constraints put on the transmitters and the receivers and on the assumptions about the channel matrix (see the recent tutorials [1], [2]). Just to mention some cases: if both transmitters and receivers are allowed to cooperate, (1) represents a single-user MIMO Gaussian channel, arising in multiple-antenna wireless systems [3]. If only the receivers are allowed to cooperate and the transmitters are constrained to encode their signals independently, (1) represents a vector Gaussian multiple-access channel, arising in code-division multiple access (CDMA) [4], [5]. If only the transmitters are allowed to cooperate and the receivers are constrained to decode their signals independently, (1) represents a vector Gaussian broadcast channel (GBC), arising in the downlink of a wireless system where the base station is equipped with an antenna array [6]–[13]. Finally, if both the transmitters and the receivers are not allowed to cooperate, (1) represents an interference Gaussian channel, arising, for example, in peer-to-peer communication wireless networks [14].

This work focuses on the vector GBC, referred to in the following as the $t\times 1 \cdots r$ GBC (to be read "t times 1 to r Gaussian broadcast channel" in order to stress the fact that the r receivers must process their signals separately, as opposed to a $t \times r$ MIMO single-user channel where the signals at the r receive antennas can be processed jointly).

The input is constrained to satisfy

$$\frac{1}{n}\sum_{i=1}^{n} |\boldsymbol{x}_i|^2 \le A \tag{2}$$

where A is the maximum allowed total transmit energy per channel use. Since the noise has unit variance, A takes on the meaning of total *transmit* signal-to-noise ratio (SNR).

For any block length n, a code $\mathcal{C}_n$ for the input-constrained $t \times 1 \cdots r$ GBC is defined by a codebook of $\exp(n\sum_{k=1}^{r} R_k)$ *codewords* of the form $\boldsymbol{X} = [\boldsymbol{x}_1, \dots, \boldsymbol{x}_n] \in \mathbb{C}^{t\times n}$, such that (2) is satisfied for each codeword, by an encoding function ϕ mapping r-tuples of message indexes $(w_1, \dots, w_r)$ (where $w_k \in$

Manuscript received August 2, 2001; revised November 25, 2002. The work of S. Shamai (Shitz) was supported by the Israeli Academy of Science. The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Washington, DC, June 2001.

G. Caire is with EURECOM, 06904 Sophia-Antipolis, France (e-mail: giuseppe.caire@eurecom.fr).

S. Shamai (Shitz) is with the Department of Electrical Engineering, Technion–Israel Institute of Technology, Technion City, Haifa 32000, Israel (e-mail: sshlomo@ee.technion.ac.il)

Communicated by D. N. C. Tse, Associate Editor for Communications.

Digital Object Identifier 10.1109/TIT.2003.813523

[1]Notation: for a matrix $\boldsymbol{A}$ we indicate its ith row, jth column, and (i, j)th element by $\boldsymbol{a}^i$, $\boldsymbol{a}_j$, and $a_{i,j}$ or, equivalently, by $[\boldsymbol{A}]_{i,j}$, respectively. The submatrix obtained by the rows of $\boldsymbol{A}$ numbered by $i \in \mathcal{S}$, where $\mathcal{S}$ is an ordered index set, is denoted by $\boldsymbol{A}[\mathcal{S}]$.

$\{1, 2, \ldots, \exp(nR_k)\}$) onto the codewords, and by r decoding functions $\psi_1, \ldots, \psi_r$, such that

$$\psi_k: \mathbb{C}^n \to \{1, 2, \ldots, \exp(nR_k)\}.$$

The received signal at the kth receiver antenna is given by the kth row $\boldsymbol{y}^k$ of $\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{X} + \boldsymbol{Z}$. The error probability for the kth user is given by $\epsilon_k = \Pr(\psi_k(\boldsymbol{y}^k) \neq w_k)$. An r-tuple of rates $\boldsymbol{R} = (R_1, \ldots, R_r)$ is achievable if there exists a sequence of codes $\{\mathcal{C}_n: n = 1, 2, \ldots\}$ with rates approaching $\boldsymbol{R}$ and vanishing ϵ_k (for all $k = 1, \ldots, r$). The system *throughput* R, measured in bit per channel use (or bit/s/Hz), is defined as the sum rate

$$R = \sum_{k=1}^{r} R_k. \tag{3}$$

The channel matrix is assumed to be perfectly known to the transmitter and to all receivers, and we consider the following scenarios: 1) $\boldsymbol{H}$ is deterministic and fixed. 2) $\boldsymbol{H}$ is fixed during the transmission of each codeword, but it is randomly and independently selected according to a given probability distribution (composite channel). 3) $\boldsymbol{H}$ is generated by an ergodic matrix random process and varies during the transmission of each codeword so that the channel is *information stable* [15], [16].

In case 2), we may also consider the laxer *long-term* constraint [17], [15]

$$E_{\boldsymbol{H}}\left[\frac{1}{n}\sum_{i=1}^{n}|\boldsymbol{x}_i|^2\right] \leq A \tag{4}$$

and the long-term average throughput $\overline{R} = E_{\boldsymbol{H}}[R]$, where $E_{\boldsymbol{H}}$ denotes expectation with respect to $\boldsymbol{H}$. This is achievable by a variable-rate coding scheme which adjusts its throughput according to the instantaneous realization of $\boldsymbol{H}$. By the law of large numbers, the throughput averaged over a long sequence of channel realizations is given by $\overline{R}$. The same result applies when we remove the block-wise independence on the sequence of channel realizations and we consider a channel where $\boldsymbol{H}$ is constant over blocks of n consecutive channel uses and changes from block to block according to an ergodic matrix random process [15].

With perfect channel knowledge at both transmitter and receivers, it turns out that the information-stable channel (case 3) has the same average throughput of the composite channel (case 2) with long-term power constraint (4) (although coding and decoding strategies and error exponents for these cases are generally different [15]).

The $1 \times 1 \cdots r$ GBC coincides with the classical *degraded* GBC, whose capacity region is well known (see [18] in the deterministic case and [19], [20] in the composite or ergodic cases). However, the $t \times 1 \cdots r$ GBC for $t > 1$ is, in general, a nondegraded BC, for which the capacity region is not fully known [8], [10], [21], and cannot be reduced to an equivalent set of parallel degraded BCs (studied in [22]–[24], [19], [25], [20]).

A. Related Work

The $t \times 1 \cdots r$ GBC was extensively studied in the context of a CDMA downlink where the transmitter was constrained to be a *linear* "joint precoder" (see, for example, [26]–[29]). To the best of our knowledge, the $t \times 1 \cdots r$ GBC problem from the information-theoretic BC viewpoint was considered for the first time in [6]. Independently, [11] considered the problem of a digital subscriber line (DSL) system with coordinated transmission and uncoordinated receivers, and proposed a transmitter precoding scheme based on (generalized) zero-forcing equalization and Tomlinson–Harashima precoding [30], which might be interpreted as a suboptimal implementation of the zero-forcing "dirty-paper" precoding scheme proposed in [6] and studied in Section IV of this work. An improvement of the scheme of [11] was proposed in [12], where Tomlinson–Harashima precoding is replaced by more efficient trellis precoding schemes. See also [24] and references therein.

B. Subsequent Work

The relative merit of our work (initially presented in [6], [7]) is twofold: 1) we introduced the central tool to tackle the $t \times 1 \cdots r$ GBC, namely, coding for known interference [31], [32] (nicknamed *dirty-paper* coding, after Costa's famous title [32]) and the Sato upper bound [33]; 2) we found by *direct calculation* the optimal throughput for the case of two users. We note here that the intimate relationship between Marton's achievable region of general BCs [34] and dirty-paper coding was already noticed in the introduction of Gel'fand and Pinsker paper [31]. The surprising fact here is that, with our choice, Marton's region is optimal at least for the sum rate.

Since the publication of [6], there has been a lot of work producing exciting results around the quest for the capacity region of the $t \times 1 \cdots r$ GBC. Before proceeding further, it is beneficial to give a brief survey of this work, and give credit to the many researchers who contributed to it.

Our direct calculation, reported here in the Appendix for the sake of completeness, could not be generalized to more than two users (although it works in principle, as shown in [35]), or more than one antenna per user. The (absolutely nontrivial) generalization to arbitrary number of users and antenna per user was found in [8], [10], [13], [21], by using *fundamental* tools such as convex duality, channel reciprocity, and uplink–downlink duality. This approach not only yields the general optimum throughput result, but also allows a much better understanding of the problem and provides the complete characterization of the region achievable by dirty-paper coding, which is shown to be the best achievable region by restricting the input to be Gaussian [36], [21].

Variations on the theme of the $t \times 1 \cdots r$ GBC dirty-paper achievable region and a *beamforming* interpretation are provided in [37], and [38], [39] consider the case of a multicell downlink with encoding cooperation (and power sharing) between the base stations.

C. Outline of This Work

The reminder of this paper is organized as follows. Section II recalls the main information-theoretic results used to tackle the $t \times 1 \cdots r$ GBC. Section III states the main result of this work, namely, the closed-form expression for the optimal throughput of the two-user channel (for any number of transmit antennas). Section IV presents a suboptimal but simpler scheme, and provides conditions for throughput-wise asymptotic optimality for both low and high SNR. Finally, Section V shows some numerical results for the ergodic throughput in the case of independent Rayleigh fading and Section VI points out our conclusions and some considerations on the downlink of wireless systems where the base station is equipped with an antenna array.

II. Background

We review the information-theoretic results that will be used in the rest of this paper in order to tackle the $t \times 1 \cdots r$ GBC, namely, Costa's dirty-paper coding [31], [32], Marton's achievable region [34], and Sato's "cooperative" upper bound on the sum capacity of general BCs [33].

A. Dirty-Paper Coding

The capacity of a single-user memoryless channel $P_{Y|X,S}$ with input X, output Y, and interference S, where the interference sequence $\boldsymbol{s}$ (with $s_i \sim S$) is noncausally known by the transmitter and unknown to the receiver, was found in [31] and is given by

$$\sup_{P_{X,U,S}} \{I(U;Y) - I(U;S)\} \tag{5}$$

where the supremum is over all

$$P_{X,U,S}(x|u,s) = 1\{x = f(u,s)\} P_{U|S}(u|s) P_S(s)$$

where P_S is given, and $f(u,s)$ is some deterministic function.

When $P_{Y|X,S}$ is given by the additive noise model $Y = X + S + Z$, where $Z \sim \mathcal{N}(0, N)$ and $S \sim \mathcal{N}(0, Q)$ are independent and the input is constrained by $E[X^2] \le P$, the capacity (5) is the same as if the interference were not present [32], given by $\frac{1}{2}\log(1+P/N)$, and it is obtained by letting $P_{U|S} = \mathcal{N}(\alpha S, P)$ and $f(u,s) = u - \alpha s$, with $\alpha = P/(P+N)$.

The achievability proof in [32] relies on the fact that both the noise and the interference signal are Gaussian i.i.d. This result has been recently generalized in various ways. In [40], it is shown that the same rate can be achieved for arbitrary noise distribution, provided that the interference is Gaussian i.i.d., or for arbitrary interference distribution provided that the noise is Gaussian (possibly colored). In [41], [42], it is shown that the same result holds for arbitrary interference (arbitrary *interference statistics*, or even arbitrary *interference sequences*, where the transmitter knows the individual sequence but ignores its statistics), provided that the transmitter and the receiver share a common random dither signal.[2]

B. Marton's Achievable Region

The best known achievable region for a general memoryless BC with marginal transition probabilities $P_{Y_1|X}$ and $P_{Y_2|X}$ was found by Marton in [34]. A special case of the Marton region is given by

$$\mathrm{co} \bigcup_{P_{U_1,U_2,X,Y_1,Y_2} \in \mathcal{P}} \left\{ \begin{array}{l} 0 \le R_1 \le I(U_1;Y_1) \\ 0 \le R_2 \le I(U_2;Y_2) \\ R_1 + R_2 \le I(U_1;Y_1) \\ \quad +I(U_2;Y_2) - I(U_1,U_2) \end{array} \right\} \tag{6}$$

("co" denotes convex closure) where $\mathcal{P}$ is the set of all joint probability distributions on U_1, U_2, X, Y_1, Y_2 such that $(U_1,U_2) \to X \to Y_1$, $(U_1,U_2) \to X \to Y_2$, and such that the marginal conditional distributions of Y_1 and Y_2 given X are equal to $P_{Y_1|X}$ and $P_{Y_2|X}$, respectively.

For any joint distribution P_{X,U_1,U_2}, the rate pair

$$R_1 = I(U_1;Y_1) - I(U_1;U_2), \qquad R_2 = I(U_2;Y_2) \tag{7}$$

can be achieved by generating signal $\boldsymbol{u}_2$ i.i.d. $\sim U_2$ for user 2 and $\boldsymbol{u}_1$ for user 1 by treating $\boldsymbol{u}_2$ as the state sequence of a "virtual" single-user channel whose transition probability depends on the *state* (or *interference*) variable U_2. Since $\boldsymbol{u}_2$ is generated by the transmitter itself, the noncausal knowledge of the whole interference sequence can be exploited by the transmitter for generating $\boldsymbol{u}_1$. From (5) it is apparent that for any given P_{X,U_1,U_2} the rates (7) are achievable. In general, the set of achievable rates can be increased by reversing the roles of user 1 and 2, and the region (6) follows [43].[3]

We shall refer to the approach of ordering the users and encoding each user by treating the effect of previous users as noncausally known interference as the *successive encoding* strategy, to stress the parallel with the *successive decoding* strategy achieving the capacity region of degraded broadcast and multiple-access channels [18].

C. The Cooperative Upper Bound

An upper bound to the sum rate of a general BC is obtained in [33] by letting the receivers cooperate and by noticing that the capacity region of the BC depends only on the marginal transition probability distributions $P_{Y_1|X}$ and $P_{Y_2|X}$, and not on the joint distribution $P_{Y_1,Y_2|X}$. By taking the worst case cooperative capacity over all joint transition probabilities with given marginals, we obtain the upper bound

$$R_1 + R_2 \le \inf_{P_{Y_1,Y_2|X} \in \mathcal{U}} \sup_{P_X \in \mathcal{A}} I(X;Y_1,Y_2) \tag{8}$$

where $\mathcal{U}$ is the set of joint transition probabilities with fixed marginals $P_{Y_1|X}$ and $P_{Y_2|X}$ and $\mathcal{A}$ is the set of allowed input distributions (determined by the input constraint).

We use (8) to obtain an upper bound to the throughput R of the $t \times 1 \cdots r$ GBC. In our case, the marginal transition probability density functions (pdfs) are given by

$$p(y_k|\boldsymbol{x}) = \frac{1}{\pi} e^{-|y_k - \boldsymbol{h}^k \boldsymbol{x}|^2}, \qquad k = 1, \ldots, r.$$

Any set of marginal transition pdfs

$$p'(y_k|\boldsymbol{x}) = \frac{1}{\pi \nu_k} e^{-|y_k - \boldsymbol{h}^k \boldsymbol{x}|^2/\nu_k}, \qquad k = 1, \ldots, r$$

with $\nu_k \le 1$ yields a GBC capacity region containing that of the original GBC, since any user k in the new channel can emulate the kth output of the original channel by adding independent Gaussian noise with variance $1 - \nu_k$. This implies that the channels in the family (1) for given $\boldsymbol{H}$ and with $\boldsymbol{z} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{\Sigma}_z)$, where $\boldsymbol{\Sigma}_z$ is any nonnegative definite Hermitian matrix whose diagonal elements are not larger than 1 (we refer to this constraint as the *subunit diagonal constraint*), have all broadcast capacity regions containing the region of the original $t \times 1 \cdots r$

[2]Notice that sharing randomness is common practice in wireless communications. For example, in standard randomly spread CDMA transmitter and receiver share the (pseudo)random spreading code generator.

[3]Note that, in general, the Marton region includes a common rate factor which in certain cases, as the one at hand, happens not to be essential for optimizing sum rates.

GBC (and, therefore, throughput not smaller than R). This fact is summarized in the following.

Lemma 1: For any channel matrix $\boldsymbol{H}$

$$R \leq \min_{\boldsymbol{\Sigma}_z \in \mathcal{U}} \max_{\boldsymbol{\Sigma}_x \in \mathcal{A}} \log \frac{\det\left(\boldsymbol{H}\boldsymbol{\Sigma}_x \boldsymbol{H}^H + \boldsymbol{\Sigma}_z\right)}{\det \boldsymbol{\Sigma}_z} \tag{9}$$

where $\mathcal{A}$ is the set of all input covariance matrices satisfying the input constraint $\mathrm{tr}(\boldsymbol{\Sigma}_x) \leq A$ and $\mathcal{U}$ is the set of all noise covariance matrices $\boldsymbol{\Sigma}_z$ satisfying the subunit diagonal constraint. □

III. The Two-User Case

The inherent problematic nature of Marton's region, when used as a constructive technique, is that one has to "guess" the distribution $P_{U_1, U_2, X}$. For the $t \times 1 \cdots r$ GBC, we propose the following choice. We let $\boldsymbol{x} = \boldsymbol{B}\boldsymbol{u}$, where $\boldsymbol{B} \in \mathbb{C}^{t \times r}$ is a precoding matrix, and the components of $\boldsymbol{u}$ are generated by successive dirty-paper encoding with Gaussian codebooks (without loss of generality, we consider the natural ordering $i = 1, \ldots, r$). The precoding matrix is constrained by $\boldsymbol{B}^H\boldsymbol{B} \in \mathcal{A}$, and the auxiliary input vector $\boldsymbol{u}$ is such that $E[\boldsymbol{u}\boldsymbol{u}^H] = \boldsymbol{I}_r$, so that the input constraint $E[\boldsymbol{x}\boldsymbol{x}^H] \in \mathcal{A}$ is enforced.

The precoded channel yields the set of r interference channels

$$y_i = w_{i,i}u_i + \sum_{j<i} w_{i,j}u_j + \sum_{j>i} w_{i,j}u_j + z_i, \qquad i = 1, \ldots, r \tag{10}$$

where $\boldsymbol{W} = \boldsymbol{H}\boldsymbol{B}$. As stated earlier, the encoder considers the interference signal $\sum_{j<i} w_{i,j}\boldsymbol{u}_j$ caused by users $j < i$ as known noncausally and the ith decoder treats the interference signal $\sum_{j>i} w_{i,j}\boldsymbol{u}_j$ caused by users $j > i$ as additional noise. By applying dirty-paper coding and by using minimum Euclidean distance decoding at each ith receiver, it follows immediately from [41], [44] that the achieved throughput is

$$R^{\mathrm{dp}} = \sum_{i=1}^{r} \log\left(1 + \frac{|w_{i,i}|^2}{1 + \sum_{j>i} |w_{i,j}|^2}\right) \tag{11}$$

where the superscript dp stands for "dirty-paper." This can be further maximized over all precoding matrices $\boldsymbol{B}$ satisfying the trace constraint. Although in this work we are only concerned with the throughput, we stress the fact that the each ith term in the sum (11) is the individual rate of user i. Therefore, we obtain the special case of Marton's region

$$\mathrm{co}\bigcup_{\pi} \bigcup_{\boldsymbol{B}:\boldsymbol{B}^H\boldsymbol{B}\in\mathcal{A}} \left\{ 0 \leq R_{\pi(i)} \leq \log\left(1 + \frac{|w_{\pi(i),\pi(i)}|^2}{1 + \sum_{j>i} |w_{\pi(i),\pi(j)}|^2}\right) : i = 1, \ldots, r \right\} \tag{12}$$

where π runs over all permutations of r elements. This region is fully characterized in [8], [10], [21].

The simplest nontrivial (i.e., nondegraded) GBC with multiple transmit antennas is the two-user case. For this channel, we have the following closed-form result.

Theorem 1: The maximum achievable throughput of the $t \times 1 : 2$ GBC is given by

$$R = \begin{cases} \log(1+|\boldsymbol{h}^1|^2 A), & A \leq A_1 \\ \log \dfrac{\left(A\det(\boldsymbol{H}\boldsymbol{H}^H) + \mathrm{trace}(\boldsymbol{H}\boldsymbol{H}^H)\right)^2 - 4|\boldsymbol{h}^2(\boldsymbol{h}^1)^H|^2}{4\det(\boldsymbol{H}\boldsymbol{H}^H)}, & A > A_1 \end{cases} \tag{13}$$

where, without loss of generality, we assume $|\boldsymbol{h}^1| \geq |\boldsymbol{h}^2|$ and where

$$A_1 = \frac{|\boldsymbol{h}^1|^2 - |\boldsymbol{h}^2|^2}{\det(\boldsymbol{H}\boldsymbol{H}^H)}.$$

Proof: See the Appendix. □

IV. Zero-Forcing Dirty-Paper Coding

In this section, we consider a more intuitive suboptimal choice of the precoding matrix $\boldsymbol{B}$ that applies to general $r > 2$. Let $\boldsymbol{H} = \boldsymbol{G}\boldsymbol{Q}$ be the QR-type decomposition [45] obtained by applying Gram–Schmidt orthogonalization to the rows of $\boldsymbol{H}$. Let $m = \mathrm{rank}(\boldsymbol{H})$, then $\boldsymbol{G} \in \mathbb{C}^{r \times m}$ is lower triangular (i.e., it has zeros above its main diagonal) and $\boldsymbol{Q} \in \mathbb{C}^{m \times t}$ has orthonormal rows. By letting $\boldsymbol{B} = \boldsymbol{Q}^H$, the resulting precoded channel is given by the set of interference channels

$$y_i = g_{i,i}u_i + \sum_{j<i} g_{i,j}u_j + z_i, \qquad i = 1, \ldots, m \tag{14}$$

while no information is sent to users $m+1, \ldots, r$. The input signals $u_1, \ldots, u_m$ are again obtained by successive dirty-paper encoding, where for each i, the noncausally known interference signal is given by $\sum_{j<i} g_{i,j}u_j$. Since the precoding matrix is chosen in order to force to zero the interference caused by users $j > i$ on each user i, we refer to this scheme as the zero-forcing dirty-paper (ZF-DP) coding. As an immediate consequence of [32], [41], we have that the resulting throughput is given by

$$R^{\mathrm{zfdp}} = \sum_{i=1}^{m} [\log(\xi d_i)]_+ \tag{15}$$

where we define $d_i \triangleq |g_{i,i}|^2$ and where ξ is the solution of the waterfilling equation

$$\sum_{i=1}^{m} [\xi - 1/d_i]_+ = A. \tag{16}$$

Obviously, for the composite channel we have $\overline{R}^{\mathrm{zfdp}} = E_{\boldsymbol{H}}[R^{\mathrm{zfdp}}]$, where ξ solves the short-term input constraint (16), or the long-term input constraint $E_{\boldsymbol{H}}[\sum_{i=1}^{m}[\xi - 1/d_i]_+] = A$.

For the sake of comparison, we review here the zero-forcing (ZF) linear beamforming and the cooperative schemes. ZF linear beamforming consists of inverting the channel matrix at the transmitter in order to create orthogonal channels between the transmitter and the receivers without receivers' cooperation. Let $\mathcal{S} \subseteq \{1, \ldots, r\}$ be a subset of cardinality $\leq m$ for which the corresponding submatrix $\boldsymbol{H}[\mathcal{S}]$ is full row rank, i.e., $\mathrm{rank}(H[\mathcal{S}]) = |\mathcal{S}|$. Let

$$\boldsymbol{H}_{\mathcal{S}}^{+} = \boldsymbol{H}[\mathcal{S}]^H \left(\boldsymbol{H}[\mathcal{S}]\boldsymbol{H}[\mathcal{S}]^H\right)^{-1} \tag{17}$$

be the Moore–Penrose pseudoinverse [45] of $\boldsymbol{H}[\mathcal{S}]$. In ZF beamforming, the transmit signal is obtained as $\boldsymbol{x} = \boldsymbol{H}_{\mathcal{S}}^{+}\boldsymbol{u}$ and yields the set of parallel channels $y_i = u_i + z_i$ for $i \in \mathcal{S}$ while no information is sent to users $i \notin \mathcal{S}$. For $\mathcal{S}$ given, the throughput of ZF beamforming is easily found to be

$$R^{\rm zf} = \sum_{i=1}^{|\mathcal{S}|} [\log(\xi b_i)]_+ \tag{18}$$

where we define the coefficients

$$b_i \triangleq \frac{1}{\left[\left(\boldsymbol{H}[\mathcal{S}]\boldsymbol{H}[\mathcal{S}]^H\right)^{-1}\right]_{i,i}} \tag{19}$$

and where ξ is the solution of $\sum_{i=1}^{|\mathcal{S}|}[\xi - 1/b_i]_+ = A$.

The throughput of ZF beamforming can be further optimized with respect to the *active user set* $\mathcal{S}$. In particular, by optimizing over the sets of cardinality one, we obtain maximal ratio combining (MRC) beamforming to the user with highest individual channel capacity, whose throughput is given by

$$R^{\rm mrc} = \log\left(1 + |\boldsymbol{h}^{\max}|^2 A\right) \tag{20}$$

where $\boldsymbol{h}^{\max}$ is the row of $\boldsymbol{H}$ with largest 2-norm. Interestingly, for $t > 1$, this choice does not yield generally the largest throughput, in sharp contrast to the standard degraded BC $(t = 1)$ for which the throughput is maximized by transmitting to the best user only [19].

By letting the receivers cooperate, we obtain a single-user multiple-antenna system with throughput (capacity) given by [3]

$$R^{\rm coop} = \sum_{i=1}^{m} [\log(\xi c_i)]_+ \tag{21}$$

where $(c_1, \ldots, c_m)$ are the nonzero squared singular values [45] of $\boldsymbol{H}$ and where ξ is the solution of $\sum_{i=1}^{m}[\xi - 1/c_i]_+ = A$. The ZF and cooperative throughputs for the composite channel (with short- or long-term power constraints) are immediately obtained from (18) and (21) by taking expectation with respect to $\boldsymbol{H}$.

Remark—On the User Ordering Problem: Since for any unitary matrix $\boldsymbol{Q}$ the matrix $\boldsymbol{QH}$ has the same singular values of $\boldsymbol{H}$, $R^{\rm coop}$ is obviously independent of the user ordering (permutation matrices are unitary). On the contrary, $R^{\rm zf}$ depends on the choice of the *unordered* active user set $\mathcal{S}$, and $R^{\rm zfdp}$ depend on the *ordered* active user set $\mathcal{S}$ (whose rows are considered in order to perform Gram–Schmidt orthogonalization). In order to emphasize this dependence, we introduce the following notation:

$$\begin{aligned} R^{\rm zfdp\text{-}\max} &\triangleq \max_{\mathcal{S}} R^{\rm zfdp} \\ R^{\rm zf\text{-}\max} &\triangleq \max_{\mathcal{S}} R^{\rm zf} \end{aligned} \tag{22}$$

where in the first line $\mathcal{S}$ ranges over the ordered user sets with cardinality m, and in the second line $\mathcal{S}$ ranges over the unordered user sets with cardinality $|\mathcal{S}| \le m$.

If $\mathrm{rank}(\boldsymbol{H}) = m$ then $R^{\rm zfdp\text{-}\max}$ is achieved by an ordered set of m users, for every SNR $A \ge 0$. In fact, suppose that for a given A the maximum of (15) is achieved by an ordered set $\mathcal{S}'$ of cardinality $k < m$, such that $\boldsymbol{H}[\mathcal{S}'] = \boldsymbol{G}'\boldsymbol{Q}'$. Then, there exists an ordered set of users $\mathcal{S} = \mathcal{S}' \cup \{i_1, \ldots, i_{m-k}\}$ such that $\boldsymbol{H}[\mathcal{S}] = \boldsymbol{GQ}$ where $|g_{i,i}|^2 = |g'_{i,i}|^2$ for $i = 1, \ldots, k$ and $|g_{i,i}|^2 > 0$ for $i = k+1, \ldots, m$. Therefore, ZF-DP coding applied to $\boldsymbol{H}[\mathcal{S}']$ and to $\boldsymbol{H}[\mathcal{S}]$ yields the same throughput.

Moreover, the user ordering is irrelevant for ZF-DP coding for asymptotically large SNR if $\boldsymbol{H}$ is full row rank, i.e., if $m = r$. In fact, let $\{d_1, \ldots, d_r\}$ and $\{d'_1, \ldots, d'_r\}$ be the two sets of ordered squared diagonal elements of the matrices $\boldsymbol{G}$ and $\boldsymbol{G}'$ in the QR decompositions $\boldsymbol{H} = \boldsymbol{GQ}$ and $\boldsymbol{\Pi H} = \boldsymbol{G}'\boldsymbol{Q}'$, respectively, where $\boldsymbol{\Pi}$ is an $r \times r$ permutation matrix. We have that

$$\begin{aligned} \prod_{i=1}^{r} d_i &= |\det(\boldsymbol{G})|^2 = \det\left(\boldsymbol{H}\boldsymbol{H}^H\right) \\ &= \det\left(\boldsymbol{\Pi}\boldsymbol{H}\boldsymbol{H}^H\boldsymbol{\Pi}^H\right) = |\det(\boldsymbol{G}')|^2 = \prod_{i=1}^{r} d'_i. \end{aligned}$$

Define the following arithmetic means:

$$M_a \triangleq \frac{1}{r}\sum_{i=1}^{r}\frac{1}{d_i}, \qquad M'_a \triangleq \frac{1}{r}\sum_{i=1}^{r}\frac{1}{d'_i}$$

and the geometric mean

$$M_g \triangleq \left(\prod_{i=1}^{r} 1/d_i\right)^{1/r} = \left(\prod_{i=1}^{r} 1/d'_i\right)^{1/r}.$$

There exist $A_0 < \infty$ such that the equation

$$\sum_{i=1}^{r}[\xi - 1/d_i]_+ = A_0$$

has solution $\xi_0 = A_0/r + M_a$ and the equation

$$\sum_{i=1}^{r}[\xi - 1/d'_i]_+ = A_0$$

has solution $\xi'_0 = A_0/r + M'_a$. Then, for all $A \ge A_0$, the ZF-DP throughputs corresponding to the original and permuted row orders are given by $r\log\frac{A/r+M_a}{M_g}$ and by $r\log\frac{A/r+M'_a}{M_g}$, respectively, implying that their difference vanish as $A \to \infty$.

With ZF beamforming, for a given SNR A the maximum throughput $R^{\rm zf\text{-}\max}$ might be achieved by a user subset $\mathcal{S}$ of cardinality strictly less than $m = \mathrm{rank}(\boldsymbol{H})$. However, it is easy to see from the properties of the waterfilling power allocation in (18) that there exists a finite value A_0 (which depends on $\boldsymbol{H}$) for which, for all $A \ge A_0$, $R^{\rm zf\text{-}\max}$ is achieved by a subset of cardinality m. ◇

From Lemma 1 we have that $R^{\rm coop}$ upperbounds both $R^{\rm zf\text{-}\max}$ and $R^{\rm zfdp\text{-}\max}$. The ZF-DP coding scheme yields generally a larger maximal throughput than ZF beamforming, as stated in the following theorem.

Theorem 2: For any channel matrix $\boldsymbol{H}$,

$$R^{\rm zfdp\text{-}\max} \ge R^{\rm zf\text{-}\max}.$$

Proof: Assume that, after a suitable row permutation, the first k rows of $\boldsymbol{H}$ are linearly independent, choose the user subset $\mathcal{S} = \{1, \ldots, k\}$. The columns $\boldsymbol{v}_i$ of $\boldsymbol{H}_{\mathcal{S}}^{+}$ satisfy

$$\boldsymbol{h}^j\boldsymbol{v}_i = \delta_{i,j}, \qquad j = 1, \ldots, k.$$

Therefore, $\boldsymbol{v}_i^H$ must lie in the orthogonal complement of the subspace $\mathcal{V}_i = \mathrm{span}\{\boldsymbol{h}^j : j = 1, \ldots, k, j \ne i\}$. Let $\boldsymbol{P}_i^{\perp}$ be the orthogonal projector [45] on $\mathcal{V}_i^\perp$. From the above orthonormality condition we get

$$\boldsymbol{v}_i^H = \frac{\boldsymbol{h}^i \boldsymbol{P}_i^\perp}{\boldsymbol{h}^i \boldsymbol{P}_i^\perp (\boldsymbol{h}^i)^H}.$$

The inverse of the ith diagonal element of

$$(\boldsymbol{H}[\mathcal{S}]\boldsymbol{H}[\mathcal{S}]^H)^{-1} = (\boldsymbol{H}_\mathcal{S}^+)^H \boldsymbol{H}_\mathcal{S}^+$$

is given by

$$b_i = \frac{1}{|\boldsymbol{v}_i|^2} = \frac{|\boldsymbol{h}^i \boldsymbol{P}_i^\perp (\boldsymbol{h}^i)^H|^2}{\boldsymbol{h}^i \boldsymbol{P}_i^\perp (\boldsymbol{h}^i)^H} = |\boldsymbol{h}^i \boldsymbol{P}_i^\perp|^2 \tag{23}$$

where we used the fact that orthogonal projectors are idempotents [45]. The rows $\boldsymbol{q}^i$ of $\boldsymbol{Q}$ in the QR decomposition $\boldsymbol{H} = \boldsymbol{G}\boldsymbol{Q}$ are obtained by applying Gram–Schmidt orthogonalization to the ordered rows $\boldsymbol{h}^1, \boldsymbol{h}^2, \ldots, \boldsymbol{h}^k$. We obtain

$$\boldsymbol{h}^i = \sqrt{\boldsymbol{h}^i \tilde{\boldsymbol{P}}_i^\perp (\boldsymbol{h}^i)^H}\, \boldsymbol{q}^i + \sum_{j=1}^{i-1} \boldsymbol{h}^i (\boldsymbol{q}^j)^H \boldsymbol{q}^j$$

where $\tilde{\boldsymbol{P}}_i^\perp$ is the orthogonal projector on the orthogonal complement of $\tilde{\mathcal{V}}_i = \mathrm{span}\{\boldsymbol{h}^1, \ldots, \boldsymbol{h}^{i-1}\}$. From the definition of d_i in (15) and the preceding formula we obtain

$$d_i = \boldsymbol{h}^i \tilde{\boldsymbol{P}}_i^\perp (\boldsymbol{h}^i)^H = |\boldsymbol{h}^i \tilde{\boldsymbol{P}}_i^\perp|^2. \tag{24}$$

Since $\mathcal{V}_i \supset \tilde{\mathcal{V}}_i$, then $b_i \le d_i$ for all $i = 1, \ldots, k$. Finally, since both k and the user ordering were arbitrary, this implies that $R^{\text{zf-max}} \le R^{\text{zfdp-max}}$. □

We conclude that $R^{\text{coop}} \ge R^{\text{zfdp-max}} \ge R^{\text{zf-max}}$ holds for any channel matrix. The next result makes this statement stronger in the limits for high and low SNR.

Theorem 3: For any channel matrix $\boldsymbol{H}$ with full row rank

$$\lim_{A\to\infty} \left(R^{\text{coop}} - R^{\text{zfdp-max}}\right) = 0. \tag{25}$$

For any channel matrix $\boldsymbol{H}$

$$\lim_{A\to 0} \frac{R^{\text{zfdp-max}}}{R^{\text{zf-max}}} = 1. \tag{26}$$

Proof: Consider first (25). Let $c_1, \ldots, c_r$ denote the (nonzero) squared singular values of $\boldsymbol{H}$ and $d_1, \ldots, d_r$ the (nonzero) squared diagonal elements of $\boldsymbol{G}$ in the QR decomposition $\boldsymbol{H} = \boldsymbol{G}\boldsymbol{Q}$. Define the following arithmetic means:

$$M_a \triangleq \frac{1}{r}\sum_{i=1}^r \frac{1}{c_i}, \quad \tilde{M}_a \triangleq \frac{1}{r}\sum_{i=1}^r \frac{1}{d_i}$$

and the geometric mean

$$M_g \triangleq \left(\prod_{i=1}^r 1/c_i\right)^{1/r} = \left(\prod_{i=1}^r 1/d_i\right)^{1/r}$$

where the last equality follows from

$$\prod_{i=1}^r c_i = \det\left(\boldsymbol{H}\boldsymbol{H}^H\right) = |\det(\boldsymbol{G})|^2 = \prod_{i=1}^r d_i.$$

It is immediate to see that there exist $A_0 < \infty$ such that the equation

$$\sum_{i=1}^r [\xi - 1/c_i]_+ = A_0$$

has solution $\xi_0 = A_0/r + M_a$ and the equation

$$\sum_{i=1}^r [\xi - 1/d_i]_+ = A_0$$

has solution $\tilde{\xi}_0 = A_0/r + \tilde{M}_a$. Then, for all $A \ge A_0$, the maximum throughputs can be written as

$$R^{\text{coop}} = r\log\frac{A/r + M_a}{M_g}$$

$$R^{\text{zfdp}} = r\log\frac{A/r + \tilde{M}_a}{M_g}. \tag{27}$$

By substituting these expressions in the limit (25) we obtain

$$\lim_{A\to\infty} r\log\frac{1 + rM_a/A}{1 + r\tilde{M}_a/A} = 0.$$

In order to show (26), consider first the case where there is a single row $\boldsymbol{h}^{\max}$ in $\boldsymbol{H}$ of maximum squared Euclidean norm. We notice both ZF-DP and ZF achieve the MRC throughput $R^{\text{mrc}} = \log(1 + |\boldsymbol{h}^{\max}|^2 A)$ by choosing an active user set containing only the user corresponding to the row $\boldsymbol{h}^{\max}$. Let $\mathcal{S}$ denotes an arbitrary user subset of cardinality k for which $\boldsymbol{H}[\mathcal{S}]$ has rank k, let

$$b_i(\mathcal{S}) = \frac{1}{\left[(\boldsymbol{H}[\mathcal{S}]\boldsymbol{H}[\mathcal{S}]^H)^{-1}\right]_{i,i}}$$

and let $R^{\text{zf}}(\mathcal{S})$ denote the maximum of $\sum_{i=1}^k \log(1 + b_i(\mathcal{S})a_i)$ subject to $\sum_{i=1}^k a_i \le A$, $a_i \ge 0$. There exists $A_1(\mathcal{S}) > 0$ such that, for all $A \le A_1(\mathcal{S})$

$$R^{\text{zf}}(\mathcal{S}) = \log\left(1 + \max_i\{b_i(\mathcal{S})\}A\right).$$

Hence, by definition of $\boldsymbol{h}^{\max}$, for all $A \le A_1(\mathcal{S})$ we have $R^{\text{zf}}(\mathcal{S}) \le R^{\text{mrc}}$. By considering all possible subsets $\mathcal{S}$ of cardinality $k = 1, \ldots, m$ we conclude that $R^{\text{zfdp-max}} = R^{\text{mrc}}$ for $0 < A \le \min_\mathcal{S} A_1(\mathcal{S})$.

Similarly, consider an ordered set $\mathcal{S}$ of cardinality m, let $d_1(\mathcal{S}), \ldots, d_m(\mathcal{S})$ denote the squared diagonal elements of $\boldsymbol{G}$ in the QR decomposition $\boldsymbol{H}[\mathcal{S}] = \boldsymbol{G}\boldsymbol{Q}$ and let $R^{\text{zfdp}}(\mathcal{S})$ denote the maximum of $\sum_{i=1}^m \log(1+d_i(\mathcal{S})a_i)$ subject to $\sum_{i=1}^m a_i = A$. There exists $A_2(\mathcal{S}) > 0$ such that, for all $A \le A_2(\mathcal{S})$

$$R^{\text{zfdp}}(\mathcal{S}) = \log\left(1 + \max_i\{d_i(\mathcal{S})\}A\right).$$

Hence, by definition of $\boldsymbol{h}^{\max}$, for all $A \le A_2(\mathcal{S})$ we have $R^{\text{zfdp}}(\mathcal{S}) \le R^{\text{mrc}}$. By considering all possible such subsets $\mathcal{S}$ we conclude that

$$R^{\text{zfdp-max}} = R^{\text{mrc}}, \qquad \text{for } 0 < A \le \min_\mathcal{S} A_2(\mathcal{S}).$$

Then, there exists an $A_3 > 0$ such that for $A \in [0, A_3]$ we have $R^{\text{zfdp-max}} = R^{\text{zf-max}} = R^{\text{mrc}}$.

In the case where $\boldsymbol{H}$ has more than one row with maximum squared Euclidean norm, we have to distinguish the case where there exists a subset of mutually orthogonal rows with maximal norm from the case where any subset of the maximal norm rows is mutually nonorthogonal. In the latter case, the above proof still holds, and the MRC throughput can be obviously achieved by transmitting to anyone of the users corresponding to the maximal norm rows. In the former case, it is not difficult to show that there exists $A_4 > 0$ for which for every $A \in [0, A_4]$ both the ZF and the ZF-DP throughputs are maximized by transmitting with equal power to the users corresponding to the subset of mutually orthogonal rows with maximal norm. This concludes the proof. □

From Theorem 3 and Lemma 1 we can prove the following.

Theorem 4: If $\boldsymbol{H}$ has full row rank, then

$$\lim_{A\to\infty}(R - R^{\text{zfdp-max}}) = 0 \tag{28}$$

and

$$\lim_{A\to 0} R/R^{\text{zfdp-max}} = 1. \tag{29}$$

Proof: It is clear that $R^{\text{zfdp-max}}$ and R^{coop} are a lower and an upper bound on R. The first statement follows directly from the first part of Theorem 3, since

$$R^{\text{zfdp-max}} \le R \le R^{\text{coop}} \text{ and } \lim_{A\to\infty}(R^{\text{coop}} - R^{\text{zfdp-max}}) = 0$$

imply the statement.

In order to prove the second statement, let $\boldsymbol{\Pi}$ be the $r \times r$ permutation matrix which sorts the rows of $\boldsymbol{H}$ such that $|\boldsymbol{h}^1| \ge \cdots \ge |\boldsymbol{h}^r|$, and consider the QR decomposition $\boldsymbol{\Pi H} = \boldsymbol{GQ}$. We apply Lemma 1 by choosing as noise covariance $\boldsymbol{\Sigma}_z = \boldsymbol{GD}^{-2}\boldsymbol{G}^H$, where $\boldsymbol{D} = \text{diag}(|\boldsymbol{h}^1|, \ldots, |\boldsymbol{h}^r|)$. By construction, $\boldsymbol{\Sigma}_z$ is positive definite (recall that $\boldsymbol{H}$ has rank r) and satisfies the subunit diagonal constraint, in fact

$$[\boldsymbol{\Sigma}_z]_{i,i} = \sum_{j=1}^{i} \frac{|g_{i,j}|^2}{|\boldsymbol{h}^j|^2} \le \sum_{j=1}^{i} \frac{|g_{i,j}|^2}{|\boldsymbol{h}^i|^2} = 1.$$

With this choice, the right-hand side (RHS) in (9) becomes

$$\log \frac{\det\left(\boldsymbol{GQ\Sigma}_x\boldsymbol{Q}^H\boldsymbol{G}^H + \boldsymbol{GD}^{-2}\boldsymbol{G}^H\right)}{\det \boldsymbol{GD}^{-2}\boldsymbol{G}^H} = \log \frac{\det\left(\boldsymbol{Q\Sigma}_x\boldsymbol{Q}^H + \boldsymbol{D}^{-2}\right)}{\det \boldsymbol{D}^{-2}}.$$

From Hadamard inequality [45] we obtain the maximizing signal covariance in the form $\boldsymbol{\Sigma}_x = \boldsymbol{Q}^H \text{diag}(a_1, \ldots, a_r)\boldsymbol{Q}$, which yields the bound

$$R \le \sum_{i=1}^{r} [\log(\xi a_i)]_+ \tag{30}$$

where ξ is the solution of

$$\sum_{i=1}^{r} [\xi - 1/|\boldsymbol{h}_i|^2]_+ = A.$$

If $|\boldsymbol{h}^1| > |\boldsymbol{h}^2|$, then there exists a value $0 < A_1 < \infty$ such that, for all $A \le A_1$, the RHS in (30) is equal to $\log(1 + |\boldsymbol{h}^1|^2 A) = R^{\text{mrc}}$. This is clearly achievable by ZF-DP and by ZF, therefore, for $A \in [0, A_1]$ the ZF-DP (and ZF) strategy is optimal (not only asymptotically for $A \to 0$). If there exist $\kappa > 1$ rows with maximal 2-norm $|\boldsymbol{h}^{\max}|$, then there exists a value $0 < A_2 < \infty$ such that, for all $A \le A_2$, the RHS in (30) is equal to $\kappa \log(1 + |\boldsymbol{h}^{\max}|^2 A/\kappa)$. In this case

$$\lim_{A\to 0} \frac{\kappa \log(1 + |\boldsymbol{h}^{\max}|^2 A/\kappa)}{\log(1 + |\boldsymbol{h}^{\max}|^2 A)} = 1$$

and the statement of Theorem 4 still holds (but only in the limit for vanishing A). □

Remark: Downlink Strategies: Theorem 4 shows an interesting feature of the $t \times 1 \cdots r$ GBC and of the ZF-DP coding strategy, which might have a relevant impact on the design of the downlink of wireless communication systems. If the base station is strongly power limited, then the throughput-maximizing strategy consists of MRC beamforming to the best user, which is the same optimal strategy for the standard degraded GBC $(t = 1)$. In this case, the transmit antenna array is used to enhance the received SNR of the best user but does not expand the useful dimensions for transmission. Practical downlink protocols for high-rate packet communications are currently proposed and implemented according to this principle: only one user in each time slot is served according to a channel-driven scheduling allocating the channel to the user enjoying the instantaneous highest individual capacity [46], [47].

On the contrary, if the base station can transmit at large power, the throughput of a single-user multiple-antenna system (as if the receivers were allowed to cooperate) can be approached. In particular, by letting the number of users served at the same time and on the same frequency band equal the number of transmit antennas, under mild conditions on the channel matrix statistics, the slope of the throughput as a function of SNR in decibels is proportional to t. Hence, in the GBC setting, the "capacity boost" typical of multiple-antenna systems depends strongly on the available transmit power. Notice also that the same throughput slopes for low and high SNR are obtained by the conventional ZF beamforming, although this is generally asymptotically suboptimal for high SNR. Beyond its theoretical value, the real advantage of dirty-paper coding on the downlink spectral efficiency of actual wireless systems (including a multicell scenario with intercell interference) is still a matter of debate [48]. ◊

V. Performances in Rayleigh Fading

In this section, we provide some numerical examples for the $t \times 1 \cdots 2$ GBC average throughputs $\overline{R}$ obtained previously, for the composite channel when $\boldsymbol{H}$ has i.i.d. entries $\sim\mathcal{N}_{\mathbb{C}}(0, 1)$ (independent Rayleigh fading). We also provide closed-form results for the general $t \times 1 \cdots r$ GBC with ZF-DP coding (without maximization with respect to the user ordering) for both given t and r and in the large-system limit, i.e., when t and r become large while the ratio r/t users/transmit antenna converges to a given constant α (referred to in the following as *antenna loading*).

Following [49], we define E_b/N_0 for the $t \times 1 \cdots r$ GBC as

$$\frac{E_b}{N_0} \triangleq \frac{tA}{R} \tag{31}$$

where the factor t in the numerator takes into account that, under mild conditions on $\boldsymbol{H}$, the average *received* energy per channel use increases linearly with t for MRC beamforming to any given user. For the cooperative system, since $t \times r$ and $r \times t$ channels yield the same throughput [3], we adopt the definition

$$\frac{E_b}{N_0} \triangleq \frac{\max\{r, t\}A}{R}. \tag{32}$$

We start with the two-user case. Subject to the short-term constraint, we have simply $\overline{R}(A) = E_{\boldsymbol{H}}[R(A)]$ where $R(A)$ is

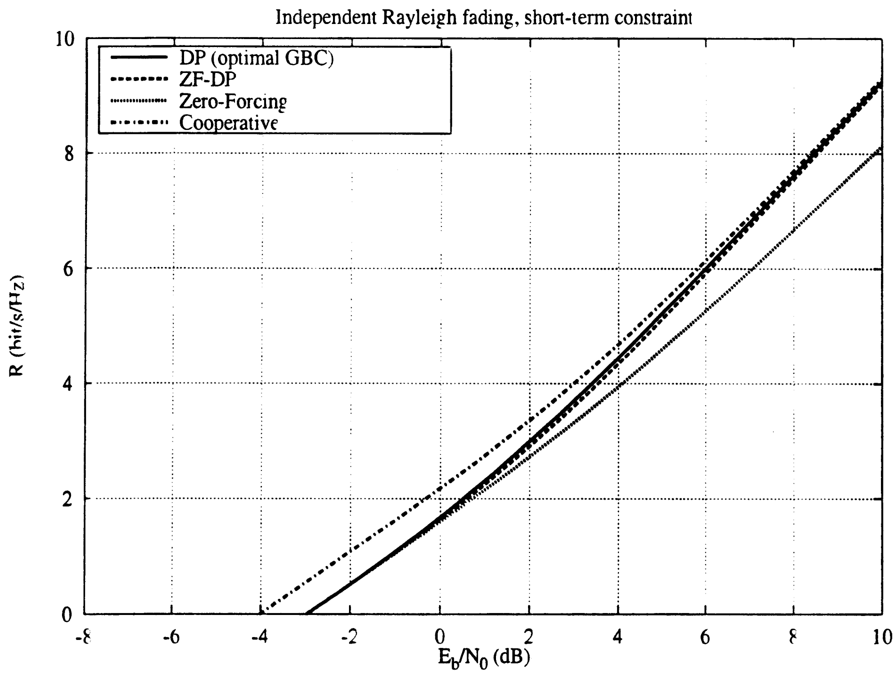

Fig. 1. Throughput of the $2 \times 1:2$ GBC with independent Rayleigh fading and short-term input constraint.

given by (13) and where we put in evidence its dependence on the input constraint A. The maximum average throughput subject to a long-term constraint is obtained by solving

$$\begin{cases} \max & E_{\boldsymbol{H}}[R(a)] \\ \text{subject to} & E_{\boldsymbol{H}}[a] = A, \quad a \geq 0. \end{cases} \tag{33}$$

By using the Lagrange–Kuhn–Tucker conditions [50], after some algebra, we obtain the optimal transmit power allocation in the form

$$a(\boldsymbol{H}) = \begin{cases} [\xi - 1/|\boldsymbol{h}^1|^2]_+, & \text{for } \frac{|\boldsymbol{h}^1|^4 - |\boldsymbol{h}^2(\boldsymbol{h}^1)^H|^2}{|\boldsymbol{h}^1|^2 \det(\boldsymbol{H}\boldsymbol{H}^H)} \geq \xi \\ \xi + \sqrt{\xi^2 + \frac{4|\boldsymbol{h}^2(\boldsymbol{h}^1)^H|^2}{\det(\boldsymbol{H}\boldsymbol{H}^H)^2}} & \\ \quad - \frac{\operatorname{trace}(\boldsymbol{H}\boldsymbol{H}^H)}{\det(\boldsymbol{H}\boldsymbol{H}^H)}, & \text{otherwise} \end{cases} \tag{34}$$

where ξ satisfies $E_{\boldsymbol{H}}[a(\boldsymbol{H})] = A$. The condition

$$\frac{|\boldsymbol{h}^1|^4 - |\boldsymbol{h}^2(\boldsymbol{h}^1)^H|^2}{|\boldsymbol{h}^1|^2 \det(\boldsymbol{H}\boldsymbol{H}^H)} \geq \xi$$

is equivalent to $a(\boldsymbol{H}) \leq A_1$, where A_1 is given in Theorem 5. Hence, by substituting (34) in the place of A in (13), we obtain explicitly the optimal throughput subject to the long-term power constraint as $\overline{R}(A) = E_{\boldsymbol{H}}[f_\xi(\boldsymbol{H})]$ where

$$f_\xi(\boldsymbol{H}) = \begin{cases} [\log(\xi|\boldsymbol{h}^1|^2)]_+, & \text{for } \frac{|\boldsymbol{h}^1|^4 - |\boldsymbol{h}^2(\boldsymbol{h}^1)^H|^2}{|\boldsymbol{h}^1|^2 \det(\boldsymbol{H}\boldsymbol{H}^H)} \geq \xi \\ \log\left[\det\left(\boldsymbol{H}\boldsymbol{H}^H\right)\xi\left(\xi + \sqrt{\xi^2 + \frac{4|\boldsymbol{h}^2(\boldsymbol{h}^1)^H|^2}{\det(\boldsymbol{H}\boldsymbol{H}^H)^2}}\right)\right] - \log 2, & \\ & \text{otherwise.} \end{cases} \tag{35}$$

Figs. 1 and 2 show $\overline{R}$ in the case $t = r = 2$ for the short and the long-term constraints, respectively, versus E_b/N_0. For the sake of comparison, we show also the 2×2 cooperative, ZF, and ZF-DP throughputs.

For the short-term constraint, there exists a minimum $(E_b/N_0)_{\min} > 0$ below which $\overline{R}$ is zero.[4] This can be calculated by letting $A \downarrow 0$ in (31) and in (32). For the $2 \times 1:2$ GBC we obtain

$$\left(\frac{E_b}{N_0}\right)_{\min} = \frac{2\log 2}{E[|\boldsymbol{h}^1|^2]} = \frac{2\log 2}{(11/4)} = -2.97 \text{ dB}$$

where we used the fact that $|\boldsymbol{h}^1|^2$ is distributed as the maximum of two i.i.d. central Chi-squared random variables with four degrees of freedom. For the 2×2 cooperative system we have

$$\left(\frac{E_b}{N_0}\right)_{\min} = \frac{2\log 2}{E[c_1]} = \frac{2\log 2}{(7/2)} = -\,4.02 \text{ dB}$$

where we used the fact that c_1 is the maximum eigenvalue of the 2×2 Wishart matrix [3] $\boldsymbol{H}\boldsymbol{H}^H$.

From Figs. 1 and 2 it is clearly visible that the simple ZF-DP scheme is optimal both for $E_b/N_0 \downarrow (E_b/N_0)_{\min}$ and for $E_b/N_0 \to \infty$. For small SNR, it is (asymptotically) equivalent to ZF beamforming and both strategies reduce to simple MRC beamforming to the best user. For large SNR, it is (asymptotically) equivalent to the cooperative single-user multiple-antenna capacity. This is a consequence of the fact that for independent Rayleigh fading, the channel matrix $\boldsymbol{H}$ has full rank almost surely, therefore Theorem 3 applies to almost all realizations of the channel.

Next, we examine the average throughput of the ZF-DP coding scheme in independent Rayleigh fading under a long-term power constraint for general r and t. As stated

[4] For the long-term constraint $(E_b/N_0)_{\min} = 0$ since $|\boldsymbol{h}^1|^2$ has a distribution with unbounded support [5].

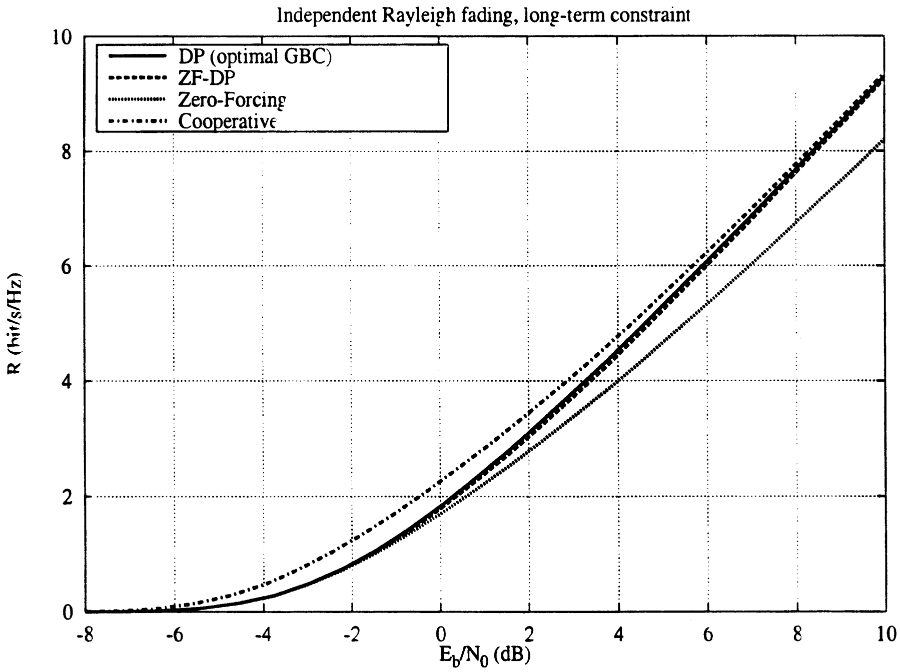

Fig. 2. Throughput of the $2 \times 1 : 2$ GBC with independent Rayleigh fading and long-term input constraint.

earlier, we assume that no effort is made to optimize the user ordering. In other words, the receiver works with the natural ordering of the rows of $\boldsymbol{H}$ and considers the coefficients $\{d_i \colon i = 1, \ldots, m\}$ corresponding to the QR decomposition $\boldsymbol{H} = \boldsymbol{GQ}$. It is interesting to notice that, since the composite channel is symmetric with respect to any user, by time-sharing with uniform probability over all possible user subsets and orderings, every user in the system achieves the same average per-user rate $\rho \triangleq \overline{R}/r$. We make use of the following well-known results [51], [52], [4].

Lemma 2: Let $\boldsymbol{H} \in \mathbb{C}^{r \times t}$ have i.i.d. entries $\sim \mathcal{N}_{\mathbb{C}}(0, 1)$, and let $g_{i,i}$ be the ith diagonal element of $\boldsymbol{G}$ in the QR decomposition $\boldsymbol{H} = \boldsymbol{GQ}$. Then, the random variables $d_i = |g_{i,i}|^2$ are statistically independent and $d_i \sim \chi^2_{2(t-i+1)}$, where χ^2_{2k} denotes the central Chi-squared distribution with $2k$ degrees of freedom, whose pdf is $f(z) = z^{k-1}e^{-z}/(k-1)!$. □

Lemma 3: Let $\nu = i/r \in [0, 1]$ denote the normalized user index. If $\boldsymbol{H} \in \mathbb{C}^{r \times t}$ (for $r/t = \alpha$) has i.i.d. circularly symmetric elements with mean 0, variance 1, and bounded fourth moment, then

$$\lim_{r \to \infty} \frac{1}{t} d_i = [1 - \alpha\nu]_+ \tag{36}$$

with probability 1. □

In the case of finite r, t, the long-term water-filling equation is $\sum_{i=1}^m E[[\xi - \frac{1}{d_i}]_+] = A$ where, from Lemma 2, we have

$$\begin{aligned}\sum_{i=1}^m E\left[\left[\xi - \frac{1}{d_i}\right]_+\right] &= \sum_{i=1}^m \int_{1/\xi}^\infty \left(\xi - \frac{1}{z}\right) \frac{z^{t-i}e^{-z}}{(t-i)!}\, dz \\ &= \sum_{i=1}^m \frac{1}{(t-i)!} [\xi\Gamma(t-i+1, 1/\xi) \\ &\quad -\Gamma(t-i, 1/\xi)]\end{aligned} \tag{37}$$

where

$$\Gamma(n, x) \triangleq \int_x^\infty z^{n-1}e^{-z}\, dz$$

and where $\Gamma(0, x) = \mathrm{E_i}(1, x)$ and, for integer $n \ge 1$

$$\Gamma(n, x) = (n-1)! e^{-x} \sum_{j=0}^{n-1} x^j/j!.$$

The resulting ZF-DP average throughput is given by

$$\begin{aligned}\overline{R}^{\text{zfdp}} &= \sum_{i=1}^m E\left[[\log(\xi d_i)]_+\right] \\ &= \sum_{i=1}^m \frac{1}{\xi^{t-i+1}} \int_1^\infty \log(z) \frac{z^{t-i}e^{-z/\xi}}{(t-i)!}\, dz \\ &= \sum_{i=1}^m \mathcal{J}_{t-i+1}(\xi)\end{aligned} \tag{38}$$

where we let

$$\begin{aligned}\mathcal{J}_k(a) &\triangleq a^{-k} \int_1^\infty \log(z) \frac{z^{k-1}e^{-z/a}}{(k-1)!}\, dz \\ &= \mathrm{E_i}(1, 1/a) + \sum_{\ell=1}^{k-1} \frac{1}{\ell} e^{-1/a} \sum_{j=0}^{\ell-1} \frac{a^{-j}}{j!}.\end{aligned} \tag{39}$$

The average throughput for ZF beamforming (optimized over the size of the active user set) is given by

$$\overline{R}^{\text{zf}} = \max_{k=1,\ldots,m} k\mathcal{J}_{t-k+1}(\xi_k) \tag{40}$$

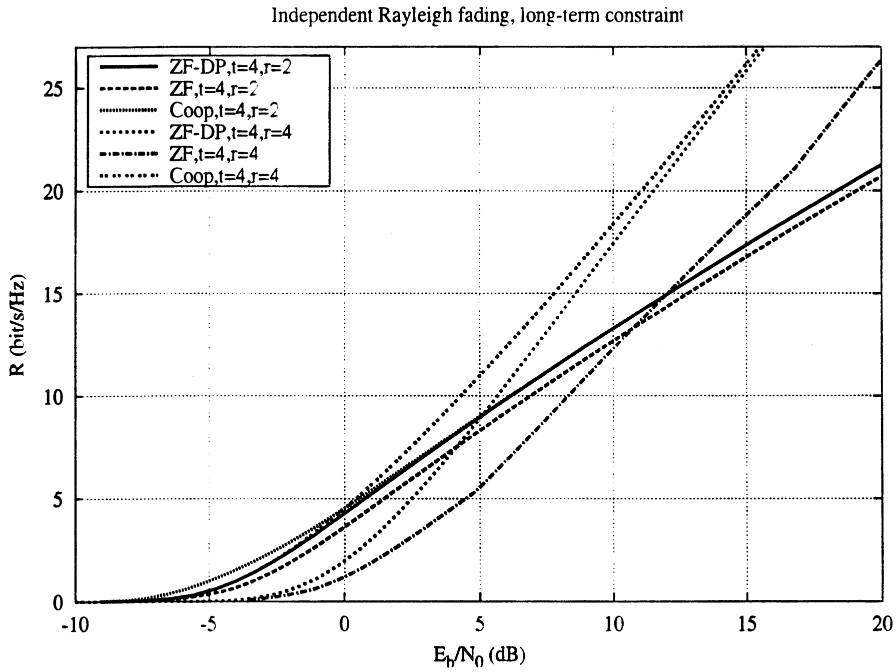

Fig. 3. Throughput of ZF-DP, ZF, and cooperative schemes in the case $t = 4,\ r = 2$ and $t = 4,\ r = 4$, with independent Rayleigh fading.

where ξ_k is the solution of the water-filling equation

$$\frac{k}{(t-k)!}\left[\xi\Gamma(t-k+1,\,1/\xi) - \Gamma(t-k,\,1/\xi)\right] = A$$

analogous to (37). For comparison, the throughput of the cooperative system is given in [3].

Fig. 3 shows the ZF-DP, ZF, and cooperative throughputs for $t = 4,\ r = 2$ and $t = 4,\ r = 4$ cases. The throughput gain of the ZF-DP coding over ZF beamforming is very significant for $t = 4,\ r = 4$, and less significant for $t = 4,\ r = 2$.

Next, we study the normalized throughput ρ in the *large-system* regime. From Lemma 3, we have immediately that $\rho^{\rm zfdp}$ is given as the solution of

$$\begin{cases} \max & \displaystyle\int_0^\mu \log\left(1 + [1-\alpha\nu]_+ a(\nu)\right)\, d\nu \\ \text{subject to} & \displaystyle\int_0^\mu a(\nu)\, d\nu = A/\alpha, \quad a(\nu) \ge 0 \end{cases} \tag{41}$$

where $a(\nu)$ is the transmit SNR of the $\lfloor \nu r \rfloor$th signal and $\mu \triangleq m/r = \min\{1,\, 1/\alpha\}$. The optimal $a(\nu)$ is given by

$$a(\nu) = \left[\xi - \frac{1}{1-\alpha\nu}\right]_+ \tag{42}$$

where ξ is the solution of the waterfilling equation

$$\int_0^\mu [\xi - 1/(1-\alpha\nu)]_+\, d\nu = A/\alpha$$

and, obviously, $\xi \ge 1$. After some algebra, we obtain the throughput in the form

$$\rho^{\rm zfdp} = \begin{cases} \log\left[\frac{1}{\alpha}\left(A - \log(1-\alpha)\right)\right] & \\ \quad -(1/\alpha - 1)\log(1-\alpha) - 1 & \text{(case 1)} \\ \frac{1}{\alpha}\left(\log\xi + \frac{1}{\xi} - 1\right) & \text{(case 2)} \end{cases} \tag{43}$$

where "case 1" corresponds to the condition $\alpha < 1$ and $A \ge \frac{\alpha}{1-\alpha} + \log(1-\alpha)$ and "case 2" corresponds to the complement condition, with ξ implicitly given by the unique solution of $\xi - \log\xi - 1 = A$ in $[1, +\infty)$.

For the sake of comparison, we calculate also the normalized throughput with ZF beamforming (optimized with respect to the size of the active user set) and with cooperative receivers in the large-system regime. In the case of ZF, it follows from [4], [52] that

$$\rho^{\rm zf} = \max_{\kappa\in[0,\mu]} \kappa\log(1 + [1-\kappa\alpha]A/(\kappa\alpha)). \tag{44}$$

The cooperative throughput is given (implicitly) by

$$\rho^{\rm coop} = \mu\int [\log(\xi z)]_+ f_\lambda(z)\, dz \tag{45}$$

where ξ is the solution of

$$\int \left[\xi - \frac{1}{z}\right]_+ f_\lambda(z)\, dz = A \tag{46}$$

and where $f_\lambda(z)$ is the limit of the empirical density of the nonzero eigenvalues of the normalized Wishart matrix $\frac{1}{m}\boldsymbol{H}\boldsymbol{H}^H$ (see [4] and references therein).

Fig. 4 shows the normalized throughputs of the ZF-DP, ZF, and cooperative schemes for $\alpha = 0.5,\ 1.0$, and 2.0. For $\alpha \le 1$, ZF-DP coding is asymptotically optimal for large SNR since the channel matrix has rank r with probability 1. On the contrary, we cannot invoke Theorem 4 in the case $\alpha > 1$, since in this case the channel matrix has rank $t < r$ with probability 1.

VI. Conclusion and Discussion

We investigated the achievable throughput (sum rate) of a generally nondegraded broadcast Gaussian channel where the transmitter has t antennas and the r receivers have one antenna each, subject to the assumption that the channel is perfectly

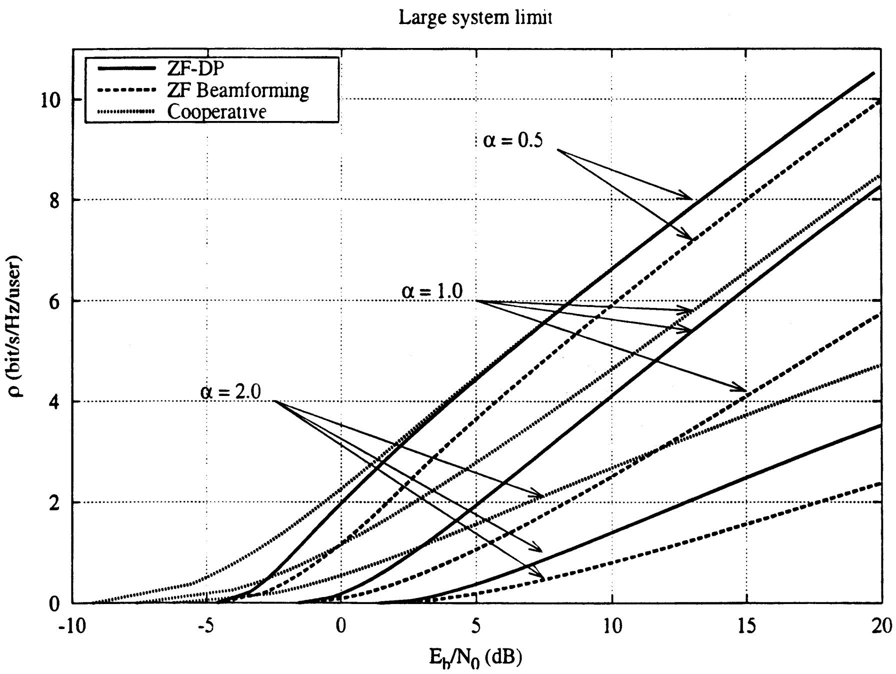

Fig. 4. Normalized throughput in the large-system limit of ZF-DP, ZF, and cooperative schemes for $\alpha = 0.5$, 1.0, and 2.0.

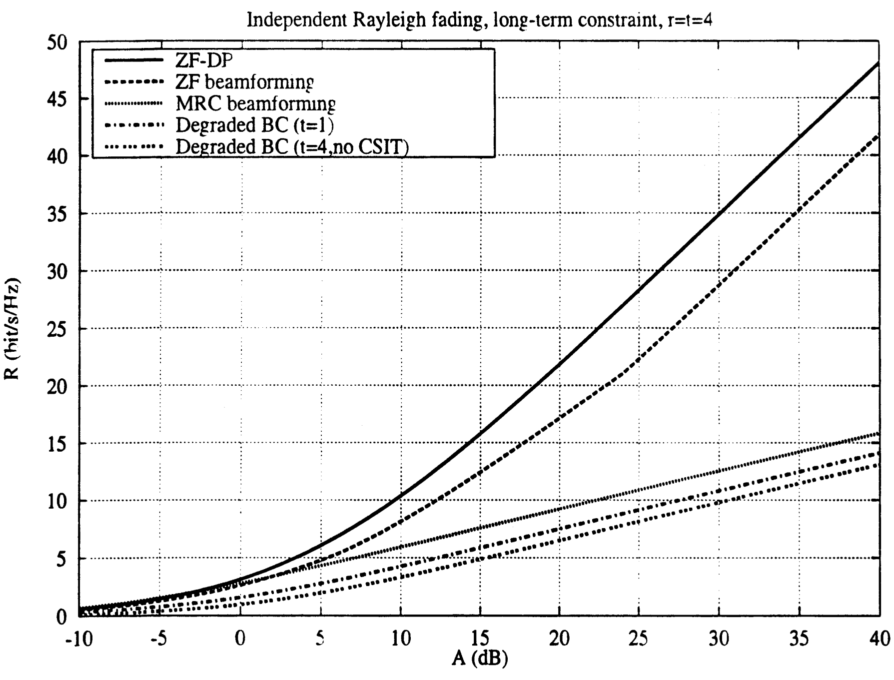

Fig. 5. Throughput versus SNR comparison for a system with independent Rayleigh fading and long-term input constraint, ZF-DP, ZF, MRC with $t = 4$, $r = 4$, the degraded GBC with $t = 1$, $r = 4$, and the degraded GBC without channel state information at the transmitter (no CSIT), $t = 4$ and arbitrary r.

known to all terminals. For this model, we proposed a choice of the Marton's region parameters such that the transmit signal is given by $\boldsymbol{x} = \boldsymbol{B}\boldsymbol{u}$, and the components u_i of the auxiliary input $\boldsymbol{u}$ are obtained by successive dirty-paper encoding, by treating the previous inputs $j < i$ as noncausally known interference. In the two-user case, we proved that, by optimizing $\boldsymbol{B}$, this choice achieves optimal throughput. For the general case, we examined a simple suboptimal choice of $\boldsymbol{B}$ that forces to zero the interference caused by the inputs $j > i$ on each user i. The suboptimal ZF-DP coding strategy is shown to achieve asymptotically optimal throughput for high SNR if the channel matrix has full row rank, while, for vanishing SNR, it reduces to simple MRC beamforming to the best user, which is shown to be also optimal in general, for low SNR.

Perhaps, the relevance of our results lies in the tremendous amount of subsequent work and in a number of new exciting information-theoretic results by other authors that this work generated [8], [10], [13], [21], [29], [37]. Moreover, driven by these results, the study of coding techniques for dirty-paper coding recently experienced a renewed interest [53]–[56].

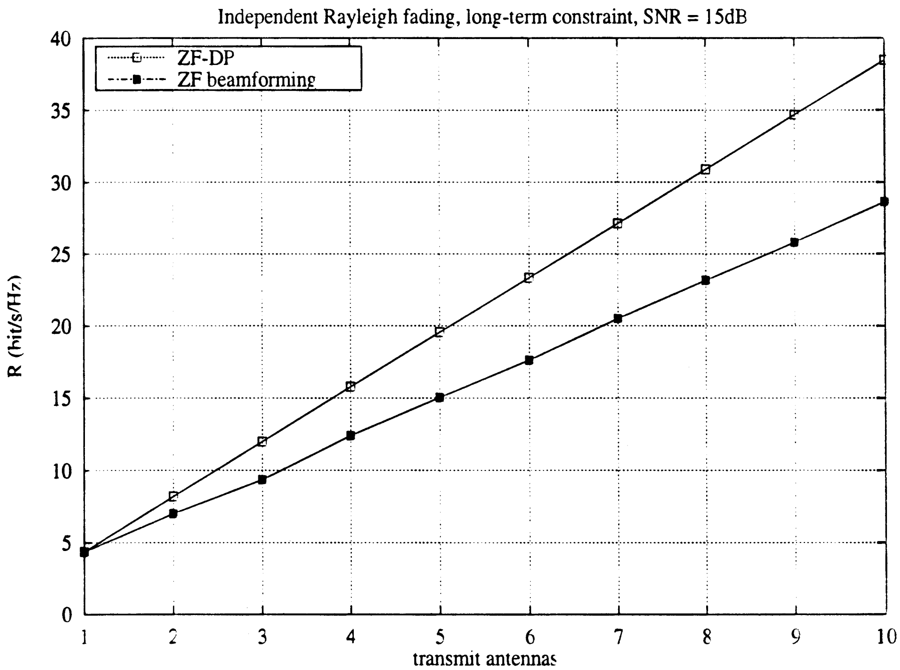

Fig. 6. Throughput versus the number of transmit antennas t for transmit SNR $A = 15$ dB in a system with independent Rayleigh fading and $r > t$ users, for ZF-DP and ZF beamforming.

We would like to conclude by pointing out some considerations for the downlink throughput optimization in a (single-cell) wireless communication system. As an example, consider Fig. 5, showing the throughput of a system with $r = 4$ users and $t = 4$ transmit antennas, independent Rayleigh fading, and long-term power constraint, for ZF-DP coding, ZF beamforming, and the MRC beamforming to the best user only. The optimal throughput for the standard degraded GBC with $t = 1$, also obtained by transmitting to the best user only, and the optimal throughput of the degraded GBC with $t = 4$ but no transmitter channel state information are shown for comparison.

For relatively large SNR, the throughput gain due to $t = 4$ over $t = 1$ antennas at the transmitter is modest if the system is constrained to serve a single user per slot. On the contrary, the throughput gain provided by the asymptotically optimal ZF-DP strategy can be very large even for moderate SNR, and increases with SNR. ZF beamforming yields the same optimal throughput slope for high SNR, but it pays a fairly large throughput penalty with respect to ZF-DP. Moreover, this penalty increases with $\min\{r, t\}$, as illustrated in Fig. 6.

If the transmitter has no channel state information, the throughput slope is independent of t, i.e., the channel "degrees of freedom" depend critically on the availability of the channel knowledge at the transmitter. This is a rare example of a Gaussian channel where channel knowledge has an impact not only on SNR enhancement, but also on the pre-$\log(\text{SNR})$ factor (i.e., on the number of degrees of freedom). For a system with a large number of users $(r \gg 1)$ and fixed (large) transmit SNR, a virtually arbitrarily large downlink throughput can be achieved by simply increasing the number of transmit antennas t and serving t users simultaneously. This, of course, depends on the ability of estimating reliably the channel matrix at the transmitter. In this respect, systems exploiting time-division duplexing might be preferable, since the channel can be estimated from the uplink signals (see, for example, [57] and references therein).

APPENDIX
PROOF OF THEOREM 1

Assume $|\boldsymbol{h}^1| \geq |\boldsymbol{h}^2|$. The case of $\text{rank}(\boldsymbol{H}) = 1$ is trivial, since in this case the two rows of $\boldsymbol{H}$ are linearly dependent, then, the $t \times 1 : 2$ GBC reduces to a standard degraded GBC with input $x = \boldsymbol{h}^1\boldsymbol{x}$ and outputs

$$y_1 = x + z_1, \quad y_2 = (|\boldsymbol{h}^2|/|\boldsymbol{h}^1|)x + z_2.$$

The throughput is clearly maximized by transmitting to the best user only [19], i.e., to user 1, and is given by $R = \log(1 + |\boldsymbol{h}^1|^2 A)$, which coincides with the first line in (13) since, in this case, $A_1 = +\infty$ and only the first line in (13) is relevant, therefore, Theorem 5 holds in the rank 1 case.

Next, we consider the case of $\boldsymbol{H}$ of rank 2. We use Lemma 1 to find an upperbound and successive dirty-paper coding to find a lower bound in the form (11). Both bounds are shown to coincide with (13), thus proving the theorem.

Cooperative Throughput Upper Bound: We consider first the maximization of mutual information in (9) with respect to $\boldsymbol{\Sigma}_x$, for a given $\boldsymbol{\Sigma}_z$. By letting $\boldsymbol{\Sigma}_z = \boldsymbol{U}\boldsymbol{\Lambda}_z\boldsymbol{U}^H$, with $\boldsymbol{U}$ unitary and $\boldsymbol{\Lambda}_z$ diagonal, we obtain the equivalent problem

$$\max_{\boldsymbol{\Sigma}_x \in \mathcal{A}} \log\det\left(\boldsymbol{H}_z\boldsymbol{\Sigma}_x\boldsymbol{H}_z^H + \boldsymbol{I}\right)$$

where $\boldsymbol{H}_z = \boldsymbol{\Lambda}_z^{-1/2}\boldsymbol{U}^H\boldsymbol{H}$. More explicitly, since

$$\boldsymbol{\Sigma}_z = \begin{bmatrix} 1 & \rho \\ \rho^* & 1 \end{bmatrix}$$

we obtain

$$\boldsymbol{U} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ \theta & -\theta \end{bmatrix}, \qquad \boldsymbol{\Lambda}_z = \begin{bmatrix} 1+r & 0 \\ 0 & 1-r \end{bmatrix}$$

where we let $r = |\rho|$ and $\theta = \exp(-j\angle\rho)$.

Let λ_1 and λ_2 denote the eigenvalues of $\boldsymbol{H}_z\boldsymbol{H}_z^H$. Then, the maximization with respect to $\boldsymbol{\Sigma}_x \in \mathcal{A}$ yields the function of r, θ and A

$$f(r, \theta, A) = \sum_{i=1}^{r} [\log(\xi\lambda_i)]_+ \tag{47}$$

where ξ solves the *water-filling* [50] equation

$$\sum_{i=1}^{2}\left[\xi - \frac{1}{\lambda_i}\right]_+ = A. \tag{48}$$

Explicitly, we have

$$\boldsymbol{H}_z\boldsymbol{H}_z^H = \frac{1}{2}\begin{bmatrix} \dfrac{h^+ + 2\sigma}{1+r} & \dfrac{h^- + j2\omega}{\sqrt{1-r^2}} \\ \dfrac{h^- - j2\omega}{\sqrt{1-r^2}} & \dfrac{h^+ - 2\sigma}{1-r} \end{bmatrix}$$

where we define $h^+ = |\boldsymbol{h}^1|^2 + |\boldsymbol{h}^2|^2$, $h^- = |\boldsymbol{h}^1|^2 - |\boldsymbol{h}^2|^2$ and $\theta\boldsymbol{h}^2(\boldsymbol{h}^1)^H = \sigma + j\omega$. The eigenvalues $\lambda_{1,2}$ are given by

$$\lambda_{1,2} = \frac{1}{2}\left(T \pm \sqrt{T^2 - 4D}\right)$$

with $T = \text{tr}(\boldsymbol{H}_z\boldsymbol{H}_z^H)$ and $D = \det(\boldsymbol{H}_z\boldsymbol{H}_z^H)$. It is immediate to check that D is independent of θ, while T is a decreasing function of σ. We conclude that θ minimizing capacity is given by $\theta^\star = \exp(-j\angle\boldsymbol{h}^2(\boldsymbol{h}^1)^H)$, so that $\sigma = |\boldsymbol{h}^2(\boldsymbol{h}^1)^H|$ is maximum. With this choice, we have

$$T = \frac{h^+ - 2r|\boldsymbol{h}^2(\boldsymbol{h}^1)^H|}{1-r^2}$$

$$D = \frac{|\boldsymbol{h}^1|^2|\boldsymbol{h}^2|^2 - |\boldsymbol{h}^2(\boldsymbol{h}^1)^H|^2}{1-r^2}.$$

In order to obtain the eigenvalues in a convenient form, it is useful to represent the rows $\boldsymbol{h}^1$ and $\boldsymbol{h}^2$ in an orthonormal basis. Applying Gram–Schmidt orthogonalization [45], we can write

$$\boldsymbol{H} = \boldsymbol{G}\boldsymbol{Q} \tag{49}$$

where $\boldsymbol{G}$ is lower triangular and $\boldsymbol{Q}$ is unitary, and we obtain $\boldsymbol{h}^1 = g_{1,1}\boldsymbol{q}^1$, $\boldsymbol{h}^2 = g_{2,1}\boldsymbol{q}^1 + g_{2,2}\boldsymbol{q}^2$. The explicit expression of the eigenvalues is now given by

$$\begin{aligned}\lambda_{1,2} = \frac{1}{2(1-r^2)}\Big[&|g_{1,1}|^2 + |g_{2,1}|^2 + |g_{2,2}|^2 - 2r\sigma \\ &\pm\big((|g_{1,1}|^2 - |g_{2,1}|^2 - |g_{2,2}|^2)^2 \\ &-4r\sigma(|g_{1,1}|^2 + |g_{2,1}|^2 + |g_{2,2}|^2) \\ &+4r^2|g_{1,1}|^2(|g_{2,1}|^2 + |g_{2,2}|^2) + 4\sigma^2\big)^{1/2}\Big].\end{aligned} \tag{50}$$

Next, we have to minimize the maximum mutual information defined by (47) and by (48) with respect to the noise correlation parameter $r \in [0, 1)$ (recall that minimization with respect to θ is already achieved). We partition the SNR range $[0, \infty)$ into two intervals, called in following the *low-SNR* and the *high-SNR* regions, and defined by the range of A for which $\xi \le 1/\lambda_2$ or $\xi > 1/\lambda_2$, respectively (notice that $\lambda_1 \ge \lambda_2$ holds for any channel matrix and value of r). Then, we consider separately the minimization of the upper bound on the two regions.

Low-SNR Region: Let $r = r^\star = |g_{2,1}/g_{1,1}| = \sigma/|g_{1,1}|^2$. Then, we obtain

$$\lambda_1 = |g_{1,1}|^2, \quad \lambda_2 = \frac{|g_{1,1}g_{2,2}|^2}{|g_{1,1}|^2 - |g_{2,1}|^2}.$$

For

$$\begin{aligned}0 \le A \le A_1 &= \frac{|g_{1,1}|^2 - |g_{2,1}|^2 - |g_{2,2}|^2}{|g_{1,1}g_{2,2}|^2} \\ &= \frac{|\boldsymbol{h}^1|^2 - |\boldsymbol{h}^2|^2}{\det(\boldsymbol{H}\boldsymbol{H}^H)}\end{aligned} \tag{51}$$

the resulting mutual information is $\log(1+|g_{1,1}|^2 A)$, which coincides with the first line of (13). This is achievable under the BC constraint (noncooperative receivers) by MRC at the transmitter (beamforming [3]) to user 1 only (the best user), and, therefore, it is clearly a tight upper bound. For $|\boldsymbol{h}^1|^2 = |\boldsymbol{h}^2|^2$, we have $A_1 = 0$ and under this condition the low-SNR case is irrelevant.

High-SNR Region: In this case, (47) and (48) become

$$\begin{aligned} f(r, \theta^\star, A) &= \log(\xi\lambda_1) + \log(\xi\lambda_2) \\ \xi &= \frac{1}{2}\left(A + \frac{1}{\lambda_1} + \frac{1}{\lambda_2}\right). \end{aligned} \tag{52}$$

By substituting we obtain

$$\begin{aligned} f(r, \theta^\star, A) &= 2\log\left(\sqrt{\lambda_1\lambda_2}\left(A + \frac{\lambda_1 + \lambda_2}{\lambda_1\lambda_2}\right)\right) - 2\log 2 \\ &= 2\log\left(\sqrt{D}\,(A + T/D)\right) - 2\log 2. \end{aligned} \tag{53}$$

Trace and determinant T and D are written in terms of the $g_{i,j}$'s as

$$T = \frac{|g_{1,1}|^2 + |g_{2,1}|^2 + |g_{2,2}|^2 - 2r\sigma}{1-r^2}$$

$$D = \frac{|g_{1,1}g_{2,2}|^2}{1-r^2}.$$

By direct substitution of the preceding expressions into (53) and by differentiating with respect to r we obtain a stationary point in

$$\begin{aligned} r^\star &= \frac{2|g_{1,1}g_{2,1}|}{A|g_{1,1}g_{2,2}|^2 + |g_{1,1}|^2 + |g_{2,1}|^2 + |g_{2,2}|^2} \\ &= \frac{2|\boldsymbol{h}^2(\boldsymbol{h}^1)^H|}{A\det\left(\boldsymbol{H}\boldsymbol{H}^H\right) + \text{trace}\left(\boldsymbol{H}\boldsymbol{H}^H\right)}. \end{aligned}$$

Notice that $r^\star$ is a decreasing function of A, and $\lim_{A\to\infty} r^\star = 0$. Therefore, the worst case noise in the large-SNR case is white. Also, for $A = A_1$ we obtain $r^\star = |g_{2,1}/g_{1,1}| \le 1$. Then, the solution $r^\star$ is compatible with the positive definiteness constraint on $\boldsymbol{\Sigma}_z$ for all finite A. Finally, we observe that the worst case noise correlation $r^\star$ is continuous in A for all $A \geq 0$, since the limits of $r^\star$ for $A \to A_1$ from the left and from the right are the same. Eventually, by substituting $r^\star$ found above into (52) we obtain the second line of (13).

Dirty-Paper Lower Bound: The throughput achievable by successive dirty-paper encoding (11) is given explicitly by

$$R^{\rm dp} = \log\left(1 + \frac{|g_{1,1}r_{1,1}|^2 a_1}{1 + |g_{1,1}r_{1,2}|^2 a_2}\right) + \log\left(1 + |g_{2,1}r_{1,2} + g_{2,2}r_{2,2}|^2 a_2\right) \quad (54)$$

where we use again the Gram–Schmidt orthogonalization $\boldsymbol{H} = \boldsymbol{GQ}$ and where, without loss of generality, we let $\boldsymbol{B} = \boldsymbol{Q}^H\boldsymbol{R}$, with $\boldsymbol{R}$ upper triangular, parameterized by

$$\boldsymbol{R} = \begin{bmatrix} \sqrt{a_1}r_{1,1} & \sqrt{a_2}r_{1,2} \\ 0 & \sqrt{a_2}r_{2,2} \end{bmatrix}.$$

The constraint $\boldsymbol{B}^H\boldsymbol{B} \in \mathcal{A}$ is written explicitly as

$$|r_{1,1}|^2 a_1 + \left(|r_{1,2}|^2 + |r_{2,2}|^2\right) a_2 = A.$$

The rate (54) must be maximized with respect to a_1, a_2 and the coefficients $r_{1,1}$, $r_{1,2}$, $r_{2,2}$. To this purpose, we reparameterize the problem by letting $b = |g_{2,1}/g_{2,2}|$, $z = |r_{1,2}/r_{2,2}|$, $q = z^2/(1+z^2)$ and $p = (bz+1)^2/(1+z^2)$, $X_1 = |r_{1,1}|^2 a_1$, and $X_2 = (|r_{1,2}|^2 + |r_{2,2}|^2)a_2$. Then, we obtain

$$R^{\rm dp} = \log\left(1 + \frac{|g_{1,1}|^2 X_1}{1 + |g_{1,1}|^2 q X_2}\right) + \log\left(1 + |g_{2,2}|^2 p X_2\right) \quad (55)$$

with the constraint $X_1 + X_2 = A$. For $z = 0$, $X_1 = A$, and $X_2 = 0$, (55) coincides with (13) in the low-SNR region $0 \leq A \leq A_1$. Therefore, we shall consider only the high-SNR region $A \geq A_1$.

By dividing all elements of $\boldsymbol{G}$ by $|g_{1,1}|$ and by replacing the input constraint A by $a \triangleq A|g_{1,1}|^2$, we obtain the equivalent problem

$$R^{\rm dp} = \log\left(1 + \frac{x_1}{1 + q x_2}\right) + \log(1 + p x_2) \quad (56)$$

where $x_1 + x_2 = a$ and where, by letting $z = \sqrt{q/(1-q)}$ and $b = \sqrt{\alpha/\beta}$ with $\alpha = |g_{2,1}/g_{1,1}|^2$ and $\beta = |g_{2,2}/g_{1,1}|^2$, we have the relation

$$p = \left(\sqrt{\alpha q} + \sqrt{\beta(1-q)}\right)^2.$$

The high-SNR condition $A \geq A_1$ translates into the condition $a \geq (1 - \alpha - \beta)/\beta$.

For any fixed $q \in [0, 1]$, we let $x_1 = a - x$ and $x_2 = x$ in (56) and maximize the resulting expression for $x \in [0, a]$. By letting $\frac{\partial}{\partial x} R^{\rm dp} = 0$ and making the substitution $y = 1 + qx$, we obtain the solution

$$y = \sqrt{\frac{(1+aq)(p-q)}{p(1-q)}} \quad (57)$$

which is valid if $p \geq q$ and $y \geq 1$. The first condition yields the inequality

$$\left(\sqrt{\alpha q} + \sqrt{\beta(1-q)}\right)^2 \geq q$$

which implies

$$0 \leq q \leq q_{\max} = \frac{\beta}{(1-\sqrt{\alpha})^2 + \beta} \leq 1.$$

The second condition yields the inequality

$$\left(\sqrt{\alpha q} + \sqrt{\beta(1-q)}\right)^2 \geq \frac{1+aq}{1+a}.$$

By letting $q = z^2/(1+z^2)$, this is turned into the second-order inequality

$$(\alpha - 1)z^2 + 2\sqrt{\alpha\beta}\, z + \beta - 1/(1+a) \geq 0$$

which implies $z \in [z_1, z_2]$, where

$$z_{1,2} = \frac{\sqrt{\alpha\beta} \pm \sqrt{\beta - \frac{1-\alpha}{1+a}}}{1-\alpha}.$$

For $a \geq (1 - \alpha - \beta)/\beta$, it is easy to check that $\beta - (1-\alpha)/(1+a) \geq 0$, therefore, the above solution always exists in the high-SNR region. It is easy to check that

$$0 \leq q_1 \triangleq \frac{z_1^2}{1+z_1^2} \leq q_2 \triangleq \frac{z_2^2}{1+z_2^2} \leq q_{\max}.$$

Therefore, the solution (57) is valid in the interval $q \in [q_1, q_2]$ and the maximum throughput can be obtained by first substituting (57) into (56) and then maximizing with respect to $q \in [q_1, q_2]$. After substitution, the function to be maximized is

$$2\log\left[\frac{1}{q}\left(\sqrt{\left(\sqrt{\alpha q} + \sqrt{\beta(1-q)}\right)^2 (1+aq)} - \sqrt{\left(\left(\sqrt{\alpha q} + \sqrt{\beta(1-q)}\right)^2 - q\right)(1-q)}\right)\right]. \quad (58)$$

By substituting again $q = z^2/(1+z^2)$ into (58) and letting the derivative with respect to z equal zero, we obtain a fifth-order equation, whose roots can be given (quite fortuitously!) in closed form as z_1, z_2 given above, $z_{3,4} = -\sqrt{\beta/\alpha}$, and

$$z_5 = \frac{2\sqrt{\alpha\beta}}{\beta(a+1) + 1 - \alpha}$$

It can be checked that $z_5 \in [z_1, z_2]$, and, therefore, it is the sought maximum.

Finally, by substituting the resulting

$$q^\star \triangleq \frac{z_5^2}{1+z_5^2} = \frac{4\alpha\beta}{(\beta(a+1)+1-\alpha)^2 + 4\alpha\beta}$$

into (58), we obtain the maximum throughput as

$$R^{\rm dp} = \log\frac{(a\beta + 1 + \alpha + \beta)^2 - 4\alpha}{4\beta} \quad (59)$$

which coincides with the second line of (13). This concludes the proof.

REFERENCES

[1] A. Goldsmith, S. Ali Jafar, N. Jindal, and S. Vishwanath, "Fundamental capacity of MIMO channels," *IEEE J. Select. Aareas Commun.*, vol. 21, pp. 000–000, June 2003.
[2] S. Diggavi, N. Al-Dahir, A. Stamoulis, and R. Calderbank, "Great expectations: The value of spatial diversity in wireless networks," *Proc. IEEE*, to be published.
[3] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun., ETT*, vol. 10, no. 6, pp. 585–596, Nov. 1999.
[4] S. Verdú and S. Shamai (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 622–640, Mar. 1999.
[5] ——, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1302–1327, May 2001.
[6] G. Caire and S. Shamai, "On achievable rates in a multi-antenna broadcast downlink," in *Proc. 38th Annu. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Oct. 2000.
[7] ——, "On achievable rates in a multi-antenna Gaussian broadcast channel," in *Proc. IEEE Int. Symp. Information Theory, ISIT 2001*, Washington, DC, June 2001, p. 147.
[8] P. Viswanath and D. N. C. Tse, "Sum capacity of the multiple antenna Gaussian broadcast channel," in *Proc. IEEE Int. Symp. Information Theory, ISIT 2002*, Lausanne, Switzerland, June/July 2002, p. 497. See also "Sum capacity of the vector Gaussian broadcast channel and uplink–downlink duality," *IEEE Trans. Inform. Theory*, vol. 49, Aug. 2003.
[9] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of multiple-access and broadcast channels," in *Proc IEEE Int. Symp. Information Theory*, Lausanne, Switzerland, July 2002, p. 500. To be published in *IEEE Trans. Inform. Theory*.
[10] S. Vishwanath, N. Jindal, and A. Goldsmith, "On the duality of Gaussian multiple-access and broadcast channels," *IEEE Trans. Information Theory*, to be published.
[11] G. Ginis and J. Cioffi, "A multi-user precoding scheme achieving crosstalk cancellation with application to DSL systems," in *Proc. 34th Asilomar Conf. Signals, Systems and Computers*, Pacific Groove, CA, Nov. 2000.
[12] W. Yu and J. Cioffi, "Trellis precoding for the broadcast channel," in *Proc. GLOBECOM 2001*, San Antonio, TX, Nov. 2001.
[13] ——, "Sum capacity of a Gaussian vector broadcast channel," in *IEEE Int. Symp. Information Theory, ISIT 2002*, Lausanne, Switzerland, June/July 2002, p. 498.
[14] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 46, pp. 388–404, Mar. 2000.
[15] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2619–2692, Oct. 1998.
[16] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1147–1157, July 1994.
[17] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channel," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1468–1489, July 1999.
[18] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[19] D. Tse. Optimal power allocation over parallel Gaussian broadcast channels. [Online]. Available: http://degas.eecs.berkeley.edu/dtse/pub.html.
[20] L. Li and A. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels. I. Ergodic capacity," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1083–1102, Mar. 2001.
[21] G. Kramer, S. Vishwanath, S. Shamai, and A. Goldsmith, "Capacity bounds for Gaussian vector channels," in *Proc. Workshop Signal Processing for Wireless Communication (DIMACS 2002)*, Piscataway, NJ, Oct. 2002.
[22] G. Poltyrev, "Capacity for a sum of certain broadcast channels," *Probl. Pered. Inform.*, vol. 15, no. 2, pp. 40–44, 1979.
[23] A. El Gamal, "Capacity for the product and sum of two unmatched broadcast channels," *Probl. Pered. Inform.*, vol. 16, no. 1, pp. 3–23, 1980.
[24] D. Hughes-Hartogs, "The capacity of a degraded spectral Gaussian broadcast channel," Ph.D. dissertation, Stanford Univ., Stanford, CA, July 1975.
[25] A. Goldsmith and M. Effros, "The capacity region of broadcast channels with intersymbol interference and colored Gaussian noise," *IEEE Trans. Inform. Theory*, vol. 47, pp. 219–240, Jan. 2001.
[26] B. Vojcic and W. Mee Jang, "Transmitter precoding in synchronous multiuser communications," *IEEE Trans. Commun.*, vol. 46, pp. 1346–1355, Oct. 1998.
[27] P. Baier, M. Meurer, T. Weber, and H. Troeger, "Joint transmission (JT), an alternative rationale for the downlink of time division CDMA using multi-element transmit antennas," in *Proc. IEEE 6th Int. Symp. Spread-Spectrum Technology and Applications (ISSSTA)*, NJIT, Newark, NJ, Sept. 2000, pp. 1–5.
[28] F. Rashid-Farrokhi, L. Tassiulas, and K. Liu, "Joint optimal power control and beamforming in wireless networks using antenna arrays," *IEEE Trans. Commun.*, vol. 46, pp. 1313–1323, Oct. 1998.
[29] M. Torlak, G. Xu, B. Evans, and H. Liu, "Fast estimation of weight vectors to optimize multi-transmitter broadcast channel capacity," *IEEE Trans. Signal Processing*, vol. 46, pp. 243–246, Jan. 1998.
[30] D. Forney and V. Eyuboglu, "Combined equalization and coding using precoding," *IEEE Comm. Mag.*, vol. 29, pp. 24–34, Dec. 1991.
[31] S. Gelfand and M. Pinsker, "Coding for channel with random parameters," *Probl, Control Inform. Theory*, vol. 9, no. 1, pp. 19–31, Jan. 1980.
[32] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 439–441, May 1983.
[33] H. Sato, "An outer bound on the capacity region of broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 374–377, May 1978.
[34] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 306–311, May 1979.
[35] G. Caire and S. Shamai, "On the multiple-antenna broadcast channel," in *Proc. 36th Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2001.
[36] P. Viswanath and D. N. C. Tse, "Sum capacity of the multi-antenna Gaussian broadcast channel," in *Proc. Workshop on Signal Processing for Wireless Communication (DIMACS 2002)*, Piscataway, NJ, Oct. 2002.
[37] M. Schubert and H. Boche, "Joint 'dirty-paper' precoding and downlink beamforming," in *Proc. Int. Symp. Information Theory and Applications*, Prague, Czech Republic, Sept. 2002.
[38] S. Shamai and B. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitter end," in *IEEE Semiannu. Vehicular Technology Conf. (VTC) Spring*, Rhodes, Greece, May 2001, pp. 1745–1749.
[39] S. Ali Jafar and A. Goldsmith, "Transmitter optimization for multiple antenna cellular systems," in *Proc. Int. Symp. Information Theory, ISIT 2002*, Lausanne, Switzerland, June/July 2002, p. 50.
[40] A. Cohen and A. Lapidoth, "The Gaussian watermarking game," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1639–1667, June 2002.
[41] U. Erez, S. Shamai (Shitz), and R. Zamir, "Capacity and lattice-strategies for cancelling known interference," *IEEE Trans. Inform. Theory*. Presented at the 2000 Cornell Summer Workshop on Information Theory, Cornell University, Ithaca, NY, Aug. 18–19, 2000, submitted for publication.
[42] R. Zamir, S. Shamai (Shitz), and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1250–1276, June 2002.
[43] A. El Gamal and E. van der Meulen, "A proof of Marton's coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 120–122, Jan. 1981.
[44] A. Lapidoth, "Nearest neighbor decoding for additive non-Gaussian noise channels," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1520–1529, Sept. 1996.
[45] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge Univ. Press, 1985.
[46] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushayana, and A. Viterbi, "CDMA/HDR: A bandwidth-efficient high-speed wireless data service for nomadic users," *IEEE Commun. Mag.*, vol. 38, pp. 70–77, July 2000.
[47] P. Viswanath, D.N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1277–1294, June 2002.
[48] H. Huang, S. Venkatesan, and H. Viswanathan, "Downlink capacity evaluation of cellular networks with known interference cancellation," in *Proc. Workshop on Signal Processing for Wireless Communication (DIMACS 2002)*, Piscataway, NJ, Oct. 2002.
[49] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1319–1343, June 2002.
[50] R. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.
[51] A. Edelman, "Eigenvalues and condition numbers of random matrices," Ph.D. dissertation, MIT, Cambridge, MA, May 1989.

[52] D. N. C. Tse and S. V. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and capacity," *IEEE Trans. Inform. Theory*, vol. 45, pp. 641–675, Mar. 1999.

[53] U. Erez and R. Zamir, "Lattice decoded nested codes achieve the Poltyrev exponent," in *Proc. IEEE Int. Symp. Information Theory (ISIT 2002)*, Lausanne, Switzerland, June/July 2002, p. 395.

[54] S. Pradhan, J. Chou, and K. Ramchandran, "Duality between source coding and channel coding with side information," Univ. Calif., Berkeley, UCB/ERL Tech. Memo. M01/34, Dec. 2001.

[55] J. Eggers, R. Bauml, R. Tzschoppe, and B. Girod, "Scalar Costa scheme for information embedding," *IEEE Trans. Signal Processing*, vol. 51, pp. 1003–1019, Apr. 2003.

[56] D. Burshtein, G. Caire, and S. Shamai, "LDPC coding for interference mitigation at the transmitter," in *Proc. 40th Annu. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Oct. 2002.

[57] R. Knopp and G. Caire, "Power control and beamforming for systems with multiple transmit and receive antennas," *IEEE Trans. Wireless. Commun.*, vol. 1, pp. 638–648, Oct. 2002.