

# A Large-System Analysis of the Imperfect-CSIT Gaussian Broadcast Channel with a DPC-based Transmission Strategy

Chinmay S. Vaze and Mahesh K. Varanasi

**Abstract**—The Gaussian broadcast channel (GBC) with  $K$  transmit antennas and  $K$  single-antenna users is considered for the case in which the channel state information is obtained at the transmitter via a finite-rate feedback link of capacity  $r$  bits per user. The throughput (i.e., the sum-rate normalized by  $K$ ) of the GBC is analyzed in the limit as  $K \rightarrow \infty$  with  $\frac{r}{K} \rightarrow \bar{r}$ . Considering the transmission strategy of zeroforcing dirty paper coding (ZFDPC), a closed-form expression for the asymptotic throughput is derived. It is observed that, even under the finite-rate feedback setting, ZFDPC achieves a significantly higher throughput than zeroforcing beamforming. Using the asymptotic throughput expression, the problem of obtaining the number of users to be selected in order to maximize the throughput is solved.

**Index Terms**—broadcast channel, dirty paper coding, inflation factor, zeroforcing beamforming.

## I. INTRODUCTION

THE Gaussian broadcast channel (GBC) has been intensely researched in recent years. It is well-known that dirty paper coding (DPC) [1] achieves the capacity region of the GBC if perfect channel state information (CSI) is available at the transmitter (CSIT) and the receivers (CSIR) [2]. However, even though the assumption of perfect CSIR can be justified, it is unrealistic to assume the same about CSIT. Moreover, the rate achievable over the GBC is quite sensitive to the quality of CSIT as has been demonstrated in [3]–[5] (see also references in [5]). This paper therefore tackles the important problem of achieving high throughputs using DPC over the GBC with imperfect CSIT.

It is well known that under perfect CSIT the DPC based transmission outperforms other known strategies such as zeroforcing beamforming (ZFBF) [3]. Nevertheless, under *imperfect* CSIT, it is the ZFBF strategy that has been intensely researched rather than DPC, mainly because ZFBF is analytical tractable [4], [6] and because of a perception that DPC based schemes are either not feasible without perfect CSIT or, even if feasible, they may be analytically intractable. The main hurdle with DPC is seen to be the difficulty of designing the inflation factor – a parameter that can critically affect its performance [1] – without perfect CSIT; it is even generally believed that the inflation factor cannot be effectively designed without perfect CSIT [4] implying that DPC may be overly

sensitive to the imperfection in CSIT, thereby rendering it less desirable than even ZFBF.

Making progress to this end, we recently developed iterative numerical algorithms for the determination of inflation factor under imperfect CSIT which yield high achievable rates [7], [8]. Some analytical results were also obtained in the high/low SNR (signal-to-noise ratio) regime [9], [10]. However, these results may not always reveal much insight on how DPC works with imperfect CSIT nor does it shed light on the behavior of DPC at moderate values of SNR. Moreover, due to the numerical nature of these algorithms, it is almost impossible to derive analytical results regarding the finite SNR performance of DPC based strategies or on how they compare with other transmission strategies such as ZFBF. Furthermore, the algorithms don't lend themselves to answering important design questions about DPC based schemes – such as optimizing the sum-rate by selecting (and transmitting to) only a subset of users – other than through a tedious and un-insightful exhaustive search. Recall that the strategy of transmitting to a subset of users is known to indeed result in a considerable improvement in the sum-rate under perfect CSIT [3] and for the ZFBF even under imperfect CSIT [6].

To address the above issues, we undertake here a large-system or asymptotic analysis of the GBC with  $K$  transmit antennas and  $K$  single-antenna users (i.e., the GBC of size/dimension  $K$ ) in which the CSIT is obtained via a finite-rate feedback link of capacity  $r$  bits per user per channel realization (or coherence interval). In particular, for the transmission strategy of zeroforcing DPC (ZFDPC) [3], [9], we analyze the normalized sum-rate or the throughput (i.e., the sum-rate divided by  $K$ ) of the GBC in the limit as  $K \rightarrow \infty$  with  $\frac{r}{K} \rightarrow \bar{r}$ . Such a problem has been considered before in the special case of perfect CSIT (i.e.,  $\bar{r} = \infty$ ) in [3] and for the finite-rate feedback GBC with the simpler transmission strategy of ZFBF in [6].

In the large-system limit, the involved random variables converge to their deterministic limits [11]. Therefore, the large-system analysis yields a closed-form expression for the asymptotic throughput (i.e., the throughput in the limit of  $K \rightarrow \infty$ ), the evaluation of which involves a simple easy-to-compute numerical integral. Importantly, unlike many works that deal with high SNR characterizations (c.f., [5] and the references therein) the asymptotic throughput is obtained as a function of SNR, and hence, it can provide insights at any finite SNR. It also serves as a simple semi-analytic tool

This work was supported in part by NSF Grant CCF-0728955. The authors are with the Department of Electrical, Computer, and Energy Engineering, University of Colorado, Boulder, CO 80309-0425 USA (e-mail: vaze, varanasi@colorado.edu).

for the comparison of different transmission strategies. In particular, contrary to popular belief, we show that even under imperfect CSIT ZFDPC does indeed achieve a significantly higher throughput than ZFBF. Furthermore, the asymptotic analysis helps to definitively answer the design question of optimizing over the number of users to be transmitted to. It is also seen that this method when mapped simply to finite dimensions works quite accurately even for the relatively small values of  $K$ . Thus the asymptotic analysis is seen to offer useful insights about finite-dimensional GBCs as well.

**Notations:** For a matrix/vector  $A$ ,  $A^*$  is its complex-conjugate transpose.  $\mathcal{CN}(0, 1)$  denotes the circularly symmetric complex normal mean-0 variance-1 random variable (RV), while  $\chi^2_{2K}$  denotes the chi-square RV of mean  $K$ .  $h \sim \mathcal{CN}(K)$  denotes the vector  $h$  of dimension  $K$  consisting of independent  $\mathcal{CN}(0, 1)$  RVs. For any vector  $a$ ,  $\tilde{a}$  denotes its direction, i.e.,  $\tilde{a} = \frac{a}{\|a\|}$ , where  $\|a\|$  denotes the norm of  $a$ . For a vector  $\hat{h}_i$ , the perpendicular space of it and the orthonormal basis vectors spanning the perpendicular space, both, are denoted by  $\hat{p}_i$  (the meaning is to be understood from the context). Almost-sure convergence [12] is denoted by a.s.  $I_K \in \mathbb{C}^{K \times K}$  is an identity matrix. For RVs  $A$  and  $B$ ,  $A \perp B$  denotes independence. All logarithms are to base 2.

## II. SYSTEM MODEL OF GBC

Consider the GBC of size  $K$ . The received signal at the  $i^{th}$  user is given by  $y_i = h_i^* x + z_i$ , where  $h_i^* \in \mathbb{C}^{1 \times K}$  is the channel vector of the  $i^{th}$  user,  $x \in \mathbb{C}^{K \times 1}$  is the signal transmitted under the power constraint of  $P$ , and  $z_i \sim \mathcal{CN}(0, 1)$  is the additive noise. We assume that  $h_i \sim \mathcal{CN}(K)$  are independent. Let  $\omega_i \in \mathbb{C}^{K \times 1}$  denote the BF vector for the  $i^{th}$  user and let  $\|\omega_i\| = 1$ . Let  $u_i$  be the data symbol to be sent to the user  $i$  ( $u_i$ 's are independent). Then the total transmitted signal is given by  $x = \sum_{i=1}^K \omega_i u_i$ . Let  $H = [h_1 \ h_2 \ h_3 \ \dots \ h_K]$ . We assume perfect CSIR. Define  $\text{SNR} = P$ .

We consider the so-called ‘on-off’ power allocation policy which is that the transmitter selects a set, denoted  $\mathcal{A}_{on}$ , of ‘on’ users and transmits with equal power to the selected users. In fact, under perfect CSIT, such a scheme is near optimal. To be precise, the difference between the asymptotic throughput achieved with the optimal waterfilling-type power allocation policy [3] and that obtained using the on-off power policy is negligible. Hence, we consider here only the on-off power policy. Under this scheme, if  $i \in \mathcal{A}_{on}$  then  $u_i \sim \mathcal{CN}(0, \frac{P}{s})$ , where  $s := |\mathcal{A}_{on}|$ , else  $u_i = 0$ . We let  $\frac{s}{K} \rightarrow \bar{s}$  as  $K \rightarrow \infty$ . Thus, the asymptotic throughput is obtained as a function of  $\bar{s}$ , which allows us to answer the design problem of optimization over the fraction of users.

### A. Quantization Scheme

In the limit of large  $K$ , RVs  $\max_{1 \leq i \leq K} \frac{1}{K} \|h_i\|^2$  and  $\min_{1 \leq i \leq K} \frac{1}{K} \|h_i\|^2$ , both, converge to 1 in probability [6, Proposition 1]. Therefore, we find it sufficient, for the present purpose, to feedback only the channel directions, namely the  $\hat{h}_i$ 's. Let  $r$  denote the number of feedback bits per user. Each user has a codebook  $\mathcal{C} = \{q_j\}_{j=1}^{2^r}$  consisting of  $2^r$   $K$ -dimensional unit-norm vectors. The vector  $\hat{h}_i$  is quantized according to the rule:  $\hat{h}_i = \arg \min_{q_j \in \mathcal{C}} \sin^2(\angle(\hat{h}_i, q_j))$ . Denote by  $d_c^2(i)$  (or simply  $d_c^2$ ) the quantization error  $\sin^2(\angle(\hat{h}_i, \hat{h}_i))$ . We further assume that  $\frac{r}{K} \rightarrow \bar{r}$ . We define  $\hat{H} = [\hat{h}_1 \ \hat{h}_2 \ \hat{h}_3 \ \dots \ \hat{h}_K]$ .

In this paper, we assume the quantization cell approximation (called the quantization-cell upper-bound (QUB)) [13]. Under this approximation, we assume ideally that the quantization cell around each vector of the codebook is a spherical cap of area  $2^{-r}$  times the total surface area of the unit sphere [13]. It has been shown that this approximation yields an upper-bound to the performance [13], [14]; and this upper-bound has been observed to be tight [14]. The tightness of the bound is a property that depends mainly on how the quantization error is modeled under the approximation, and not so much on the channel or the transmission scheme for which the approximation is being used. Hence, QUB is a reasonable assumption to make for the present analysis.

**Lemma 1:** If we write  $\tilde{h}_i = \sqrt{1 - d_c^2(i)} \hat{h}_i + \sqrt{d_c^2(i)} \tilde{e}_i$  then, under the QUB model,  $\tilde{e}_i$  is isotropically distributed in  $\hat{p}_i$  and is independent of  $d_c^2(i)$ . Also,  $d_c^2(i) \rightarrow 2^{-\bar{r}} =: \bar{D}$  a.s.

We consider here an ensemble of codebooks  $\{\bar{U}\mathcal{C}\}$  where  $\bar{U}$  is a Haar distributed unitary matrix and  $\mathcal{C}$  is a given codebook. Then  $\hat{h}_i$  is isotropic. We take the expectation over the codebooks as well although this is not shown explicitly in subsequent formulas.

### B. The Achievable Rate

Assume that the users are encoded according to their natural order. We focus here on ZFDPC scheme [3], [9]. To obtain the BF vectors under this scheme, we perform QR-decomposition [15] of  $H$  when there is perfect CSIT [3]; i.e., let  $H = QR$ . The columns of  $Q$  are respectively the BF vectors of the users. Under finite-rate feedback, we perform the same procedure with  $\hat{H}$  [9]. Note that the BF vectors are orthogonal and (under finite-rate feedback)  $\hat{h}_i^* \omega_j = 0, \forall j > i$ .

We select the auxiliary random variable for the  $i^{th}$  user as  $U_i = u_i + W_i[u_1^* \dots u_{i-1}^*]^*$ , where  $W_i \in \mathbb{C}^{1 \times (i-1)}$  is the inflation factor for the  $i^{th}$  user [1], [8], [9]. Then the achievable rate for the  $i^{th}$  user is given by equation (1) below (see [8]).

In this scheme, the interference (at user  $i$ ) due to the users encoded previously (i.e., users 1 to  $i-1$ ) is treated by DPC

$$R_i = \mathbb{E}_{\hat{H}} \max_{W_i} \mathbb{E}_{H|\hat{H}} \log \frac{\text{nr}}{\text{nr} \cdot (1 + \|W_i\|^2) - \frac{P}{s} \left| h_i^* (\omega_i + [\omega_1 \dots \omega_{i-1}] W_i^*) \right|^2} \text{ where } \text{nr} = 1 + \frac{P}{s} \sum_{j=1}^K |h_i^* \omega_j|^2. \quad (1)$$

$$W_i = \frac{P}{s} \mathbb{E}_{H|\hat{H}} \left( \omega_i^* h_i h_i^* [\omega_1 \dots \omega_{i-1}] \right) M_i^{-1} \text{ where } M_i = \mathbb{E}_{H|\hat{H}} \left( \text{nr} \cdot I_{i-1} - \frac{P}{s} [\omega_1 \dots \omega_{i-1}]^* h_i h_i^* [\omega_1 \dots \omega_{i-1}] \right). \quad (2)$$

whereas that due to the users encoded afterwards (i.e., users  $i+1$  to  $s$ ) is treated by zeroforcing; hence the name ZDFPC.

Our choice of the inflation factor is stated in equation (2), which is obtained from [8] and [9]. It is derived by first moving the conditional expectation in equation (1) inside the logarithm to obtain an upper-bound on the rate and then maximizing this upper-bound over the inflation factor.

### III. EVALUATION OF THE ASYMPTOTIC THROUGHPUT

We want to evaluate the limit,  $\lim_K \frac{1}{K} \sum_i R_i$ , where each  $R_i$  depends on the user index  $i$ , and hence, on the normalized user index  $\frac{i}{s}$ . In the limit, the normalized user index would take a value from the continuum, i.e.,  $\bar{i} := \lim_K \frac{i}{s} \in [0, 1]$ . We thus anticipate that, as  $K \rightarrow \infty$ , the above summation would converge to an integral over  $\bar{i}$  and the integrand of which would be the limit of  $R_i$ .

To compute the limit, we first need to evaluate the conditional expectations in (2). We start below with  $\mathbb{E}_{H|\hat{H}}(\hat{h}_i \tilde{h}_i^*)$ . Then we compute  $\mathbb{E}_{H|\hat{H}} \text{nr}$  and  $\lim_K \text{nr}$ . To this end, note that  $\text{nr} = 1 + \frac{P}{s} |h_i^* \omega_i|^2 + \frac{P}{s} \sum_{j < i} |h_i^* \omega_j|^2 + \frac{P}{s} \sum_{k > i} |h_i^* \omega_k|^2$ ; each of these terms is dealt separately in Subsections III-A to III-C, respectively. Later, in Subsection III-D, we find  $W_i$  in closed form. Finally, we evaluate the limit of the terms involved in the denominator of the argument of the logarithm in (1). We define  $D = \mathbb{E}_{H|\hat{H}} d_c^2(i)$ . Then  $\lim_K D = \bar{D} = 2^{-\bar{r}}$ .

**Lemma 2:**  $\mathbb{E}_{H|\hat{H}}(\tilde{h}_i \tilde{h}_i^*) = (1 - D - \frac{D}{K-1}) \hat{h}_i \hat{h}_i^* + \frac{D}{K-1} I_K$ .

*Proof:* We first prove that  $\mathbb{E} \tilde{e}_i = 0$ . To this end, consider the unitary matrix  $U = [\hat{h}_i \ \hat{p}_i] \begin{bmatrix} 1 & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} \hat{h}_i^* \\ \hat{p}_i^* \end{bmatrix}$ , where  $V \in \mathcal{C}^{(K-1) \times (K-1)}$  is any arbitrary unitary matrix. Now,  $U \hat{h}_i = \hat{h}_i$  and  $U \hat{p}_i = \hat{p}_i V$ . Hence,  $U \tilde{e}_i$  is isotropic in  $\hat{p}_i$ , which implies that  $U \mathbb{E}_{H|\hat{H}}(\tilde{e}_i) = \mathbb{E}_{H|\hat{H}}(\tilde{e}_i) = 0$ .

Now,  $\mathbb{E}_{H|\hat{H}}(\tilde{e}_i \tilde{e}_i^*)$  must be of the form  $\hat{p}_i Q \hat{p}_i^*$  where  $Q$  is positive semi-definite matrix. We can prove that  $U \mathbb{E}_{H|\hat{H}}(\tilde{e}_i \tilde{e}_i^*) U^* = \mathbb{E}_{H|\hat{H}}(\tilde{e}_i \tilde{e}_i^*) \Rightarrow V Q V^* = Q, \forall$  unitary  $V$ . This implies that  $Q = k \cdot I_{K-1}$  with  $k$  chosen such that  $\text{tr}(Q) = 1$ . Hence  $\mathbb{E}_{H|\hat{H}}(\tilde{e}_i \tilde{e}_i^*) = \hat{p}_i \frac{1}{K-1} I_{K-1} \hat{p}_i^*$ .

Now using the decomposition of  $\tilde{h}_i$  given in Lemma 1 and noting the fact that  $\hat{p}_i \hat{p}_i^* = I_K - \hat{h}_i \hat{h}_i^*$  we obtain the result. ■

#### A. Analysis of $\frac{1}{s} |h_i^* \omega_i|^2$

$$\begin{aligned} \frac{1}{s} |h_i^* \omega_i|^2 &= \frac{1}{s} \left| h_i^* \begin{bmatrix} \hat{h}_i & \hat{p}_i \end{bmatrix} \begin{bmatrix} \hat{h}_i & \hat{p}_i \end{bmatrix}^* \omega_i \right|^2 \\ &= \frac{1}{s} \left\{ \left| h_i^* \hat{h}_i \hat{h}_i^* \omega_i \right|^2 + \left| h_i^* \hat{p}_i \hat{p}_i^* \omega_i \right|^2 + \text{cross terms} \right\}. \end{aligned}$$

1)  $\frac{1}{s} |h_i^* \hat{h}_i \hat{h}_i^* \omega_i|^2 = \frac{\|h_i\|_s^2}{s} |\hat{h}_i^* \hat{h}_i|^2 |\hat{h}_i \omega_i|^2$ . Therefore,  $\mathbb{E}_{H|\hat{H}} \frac{1}{s} |h_i^* \hat{h}_i \hat{h}_i^* \omega_i|^2 = \frac{K}{s} (1 - D) |\hat{h}_i \omega_i|^2$ .

We know  $|\hat{h}_i \hat{h}_i|^2 \rightarrow 1 - \bar{D}$  a.s. Note that  $\left| \|h_i\| \cdot \hat{h}_i^* \omega_i \right|^2 \sim \chi_{2(K-i+1)}^2$  [11]. Hence,  $\left| \hat{h}_i^* \omega_i \right|^2 = \frac{s}{\|h_i\|^2} \frac{1}{s} \left| \|h_i\| \hat{h}_i^* \omega_i \right|^2 \rightarrow (1 - \bar{i} \bar{s})$  a.s. Therefore  $\frac{1}{s} |h_i^* \hat{h}_i \hat{h}_i^* \omega_i|^2 \rightarrow \frac{1}{s} (1 - \bar{D})(1 - \bar{i} \bar{s})$  a.s.

2) Now  $\hat{p}_i \hat{p}_i^* \tilde{h}_i = e_i$  with  $\mathbb{E}_{H|\hat{H}} \|e_i\|^2 = D$ . Let us consider  $h' \sim \mathcal{CN}(K-1) \perp \tilde{e}_i$ . Let  $a_i = \hat{p}_i \hat{p}_i^* \omega_i$ ;  $\|a_i\|^2 = 1 - |\hat{h}_i^* \omega_i|^2$ . Conditioned on  $\hat{H}$ ,  $\tilde{a}_i \in \hat{p}_i$  is a deterministic direction. Hence,

conditioned on  $\hat{H}$ ,  $\tilde{a}_i^* \cdot \|h'\| \tilde{e}_i \sim \mathcal{CN}(0, 1)$ . Therefore,

$$\begin{aligned} \mathbb{E}_{H|\hat{H}} \left| \tilde{e}_i^* \omega_i \right|^2 &= \frac{\|a_i\|^2}{K-1} \mathbb{E}_{H|\hat{H}} \left| \tilde{a}_i^* \cdot \|h'\| \tilde{e}_i \right|^2 = \frac{1 - |\hat{h}_i^* \omega_i|^2}{K-1} \text{ and} \\ \mathbb{E}_{H|\hat{H}} \frac{1}{s} |h_i^* \hat{p}_i \hat{p}_i^* \omega_i|^2 &= \frac{DK}{s(K-1)} (1 - |\hat{h}_i^* \omega_i|^2). \end{aligned}$$

To compute the limit, note that  $\tilde{a}_i^* \tilde{e}_i$  behaves like  $\frac{\mathcal{CN}(0,1)}{K-1}$ , as far as the limit is concerned. Since the other multiplicative terms remain bounded in limit, this term converges to zero a.s.

3) One of the cross terms is  $\|h_i\|^2 \hat{h}_i^* \hat{h}_i \hat{h}_i^* \omega_i \omega_i^* \hat{p}_i \hat{p}_i^* \tilde{h}_i$ . Now,  $\tilde{h}_i^* \hat{h}_i = \sqrt{1 - d_c^2(i)}$ ;  $\hat{p}_i \hat{p}_i^* \tilde{h}_i = \sqrt{d_c^2(i)} \tilde{e}_i$ . Since  $d_c(i) \perp \tilde{e}_i$  and  $\mathbb{E}_{H|\hat{H}} \tilde{e}_i = 0$ , the conditional expectation of the cross terms is zero. Their limit can also be shown to be zero a.s.

#### B. Analysis of $\frac{1}{s} \sum_{j < i} |h_i^* \omega_j|^2$

The conditional expectation can be computed using the techniques developed in Subsection III-A. We directly state the main result. Note that  $|\hat{h}_i^* \omega_i|^2 = 1 - \sum_{j < i} |\hat{h}_i^* \omega_j|^2$  and  $|h_i^* \omega_j|^2 = |h_i^* \hat{h}_i \hat{h}_i^* \omega_j|^2 + |h_i^* \hat{p}_i \hat{p}_i^* \omega_j|^2 + \text{cross terms}$ . Then  $\mathbb{E}_{H|\hat{H}} \sum_{j < i} |h_i^* \hat{h}_i \hat{h}_i^* \omega_j|^2 = K(1 - D)(1 - |\hat{h}_i^* \omega_i|^2)$ .

$$\mathbb{E}_{H|\hat{H}} \sum_{j < i} |h_i^* \hat{p}_i \hat{p}_i^* \omega_j|^2 = \frac{KD}{(K-1)} \left( i - 2 + |\hat{h}_i^* \omega_i|^2 \right).$$

To compute the limit, we have  $\omega_j \perp h_i \ \forall j < i$ . Thus,  $h_i^* \omega_j$  are independent  $\sim \mathcal{CN}(0, 1)$  random variables. Hence  $\frac{1}{s} \sum_{j < i} |h_i^* \omega_j|^2 \rightarrow \bar{i}$  a.s.

#### C. Analysis of $\frac{1}{s} \sum_{k > i} |h_i^* \omega_k|^2$

To compute the conditional expectation, the techniques developed in Subsection III-A are used. We omit the details and state the main result:  $\frac{1}{s} \mathbb{E}_{H|\hat{H}} \sum_{k > i} |h_i^* \omega_k|^2 = \frac{K}{K-1} D \frac{s-i}{s}$ .

Since  $\omega_k \in \hat{p}_i$ , we get  $|h_i^* \omega_k|^2 = \|h_i\|^2 \cdot \|e_i\|^2 \cdot |\omega_k^* \tilde{e}_i|^2$ . As in Subsection III-A, introduce  $h' \sim \mathcal{CN}(K-1) \perp \tilde{e}_i$ . Then conditioned on  $\hat{H}$ ,  $\sum_{k > i} |\omega_k^* \cdot \|h'\| \tilde{e}_i|^2 \sim \chi_{2(s-i)}^2$ ; hence unconditionally also, it has the same distribution. This gives  $\frac{1}{s} \sum_{k > i} |h_i^* \omega_k|^2 \rightarrow \bar{D}(1 - \bar{i})$  a.s.

#### D. Computation of the inflation factor

Let us define  $l_i^* = \hat{h}_i^* [\omega_1 \ \omega_2 \cdots \omega_{i-1}]$ . Then we have:

$$\begin{aligned} \mathbb{E}_{H|\hat{H}} (\omega_i^* \tilde{h}_i \tilde{h}_i^* [\omega_1 \cdots \omega_{i-1}]) &= (1 - D - \frac{D}{K-1}) (\omega_i^* \hat{h}_i) l_i^*, \\ \mathbb{E}_{H|\hat{H}} \text{nr} &= \bar{\text{nr}} = 1 + \frac{PK}{s} (1 - D) + PD \frac{K}{K-1} \frac{s-1}{s}, \text{ and} \\ M_i &= \left( \bar{\text{nr}} - \frac{PK}{s} \right) l_{i-1} - \frac{PK}{s} \left( 1 - D - \frac{D}{K-1} \right) l_i l_i^*. \end{aligned}$$

We now need  $l_i^* M_i^{-1}$ . Note that  $M_i$  is positive definite and one of the eigenvectors of  $M_i$  or  $M_i^{-1}$  is  $\tilde{l}_i$  while the rest are orthogonal to  $\tilde{l}_i$ . Hence,  $l_i^* M_i^{-1}$  is equal to  $l_i^*$  times the eigenvalue of  $M_i^{-1}$  corresponding to  $\tilde{l}_i$  as the eigenvector. Therefore, we get

$$W_i = \frac{\frac{PK}{s} \left( 1 - D - \frac{D}{K-1} \right) (\omega_i^* \hat{h}_i) l_i^*}{1 + \frac{PK}{s} \left( 1 - D - \frac{D}{K-1} \right) |\hat{h}_i^* \omega_i|^2 + PD \frac{K}{K-1}}. \quad (4)$$

#### E. Analysis of $(1 + \|W_i\|^2)$

We have  $\|l_i\|^2 = \sum_{j < i} |\hat{h}_i^* \omega_j|^2 = 1 - |\hat{h}_i^* \omega_i|^2$ . Further  $|\hat{h}_i^* \omega_i|^2 \rightarrow (1 - \bar{i} \bar{s})$  a.s. Let  $x^\infty(\bar{i})$  denote the limit of  $(1 + \|W_i\|^2)$ . Then we can easily obtain the following:

$$x^\infty(\bar{i}) = 1 + \frac{\left( \frac{P}{s} \right)^2 (1 - \bar{D})^2 (1 - \bar{i} \bar{s}) (\bar{i} \bar{s})}{(1 + P \bar{D} + \frac{P}{s} (1 - \bar{D}) (1 - \bar{i} \bar{s}))^2} \text{ a.s.}$$

$$F. \text{ Analysis of } f_i := \frac{1}{\bar{s}} \left| h_i^* (\omega_i + [\omega_1 \cdots \omega_{i-1}] W_i^*) \right|^2$$

From equation (4), we can write  $W_i^* = (\hat{h}_i^* \omega_i) c_i l_i$  for an appropriately chosen scalar  $c_i$ . Then

$$\begin{aligned} f_i &= \frac{1}{\bar{s}} \left| h_i^* [\omega_1 \cdots \omega_{i-1}] (\hat{h}_i^* \omega_i) c_i l_i + \hat{h}_i^* \omega_i \right|^2 \\ &= \frac{\|h_i\|^2}{s |\hat{h}_i^* \omega_i|^2} \left| \tilde{h}_i^* [\omega_1 \cdots \omega_{i-1}] l_i c_i |\hat{h}_i^* \omega_i|^2 + \tilde{h}_i^* \omega_i \omega_i^* \hat{h}_i \right|^2. \end{aligned}$$

Let us first analyze the terms  $A = \tilde{h}_i^* [\omega_1 \cdots \omega_{i-1}] l_i$  and  $B = \tilde{h}_i^* \omega_i \omega_i^* \hat{h}_i$  in the above equation. To this end, we have

$$\begin{aligned} A &= \tilde{h}_i^* [\hat{h}_i \hat{p}_i] [\hat{h}_i \hat{p}_i]^* [\omega_1 \cdots \omega_{i-1}] l_i \\ &= \tilde{h}_i^* \hat{h}_i \hat{h}_i^* [\omega_1 \cdots \omega_{i-1}] l_i + \tilde{h}_i^* \hat{p}_i \hat{p}_i^* [\omega_1 \cdots \omega_{i-1}] l_i, \end{aligned}$$

and  $B = \tilde{h}_i^* \hat{h}_i \hat{h}_i^* \omega_i \omega_i^* \hat{h}_i + \tilde{h}_i^* \hat{p}_i \hat{p}_i^* \omega_i \omega_i^* \hat{h}_i$ .

Using the arguments developed in Subsection III-A Part 2), it can be proved that the terms  $\tilde{h}_i^* \hat{p}_i \hat{p}_i^* [\omega_1 \cdots \omega_{i-1}] l_i$  and  $\tilde{h}_i^* \hat{p}_i \hat{p}_i^* \omega_i \omega_i^* \hat{h}_i$  converge to zero a.s. Hence, we get

$$\begin{aligned} \lim_K \left| A c_i |\hat{h}_i^* \omega_i|^2 + B \right|^2 &= \lim_K \left| \tilde{h}_i^* \hat{h}_i \hat{h}_i^* [\omega_1 \cdots \omega_{i-1}] l_i c_i |\hat{h}_i^* \omega_i|^2 + \tilde{h}_i^* \hat{h}_i \hat{h}_i^* \omega_i \omega_i^* \hat{h}_i \right|^2 \\ &= \lim_K |\tilde{h}_i^* \hat{h}_i|^2 \cdot |\hat{h}_i^* \omega_i|^4 \cdot |c_i| |l_i|^2 + 1 \Big|^2. \end{aligned}$$

Now,  $|\tilde{h}_i^* \hat{h}_i|^2 \rightarrow (1 - \bar{D})$  a.s.,  $|\hat{h}_i^* \omega_i|^4 \rightarrow (1 - \bar{i}\bar{s})^2$  a.s., and  $|l_i|^2 \rightarrow (\bar{i}\bar{s})$ . The limiting value of  $c_i$ , denoted by  $c^\infty(\bar{i})$ , can be easily obtained from the results of the previous subsection (not shown here explicitly). Putting together these results, we obtain  $f_i \rightarrow \frac{1}{\bar{s}} (1 - \bar{D}) (1 - \bar{i}\bar{s}) (c^\infty(\bar{i}) \cdot \bar{i}\bar{s} + 1)^2$  a.s.

Using all the above results, we obtain the asymptotic throughput  $\rho(P, \bar{s}, \bar{r})$  as given by equation (3).

In the simpler case of perfect CSIT (i.e., as  $\bar{r} \rightarrow \infty$ ), the expression for the asymptotic throughput obtained here reduces to  $\rho(P, \bar{s}, \infty) = \bar{s} \int_0^1 \log \left\{ 1 + \frac{1}{\bar{s}} P (1 - \bar{i}\bar{s}) \right\} d\bar{i}$ , which is same as what one would obtain by specializing the result of [3] to the ‘on-off’-type power policy.

#### IV. NUMERICAL RESULTS

Let the optimal value of  $\bar{s}$  maximizing  $\rho$  for the given value of  $P$  and  $\bar{r}$  be  $\bar{s}_{opt}$ . Let  $\rho_{opt}(P, \bar{r}) := \rho(P, \bar{s}_{opt}, \bar{r})$ . The asymptotic throughput for ZFBF is obtained from [6].

In Fig. 1, we plot  $\bar{s}_{opt}$  as function of  $P$  for  $\bar{r} = 1, 5$ . It can be seen that for lower values of  $P$ ,  $\bar{s}_{opt}$  increases with  $P$ . This behavior can be understood by noting that if the increased power is allocated to only a few users, then there would be a ‘logarithmic’ increase in the sum-rate; however,

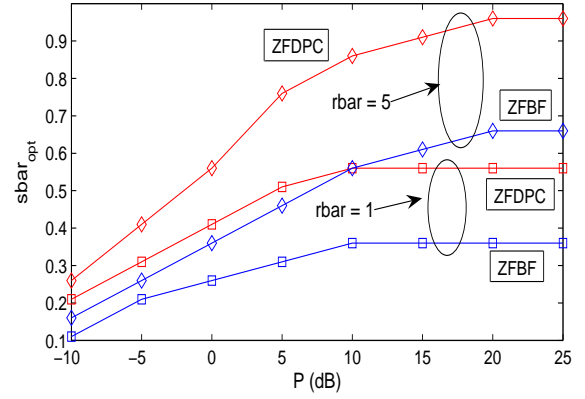


Fig. 1.  $\bar{s}_{opt}$  vs.  $P$  for  $\bar{r} = 1$  and  $5$ .

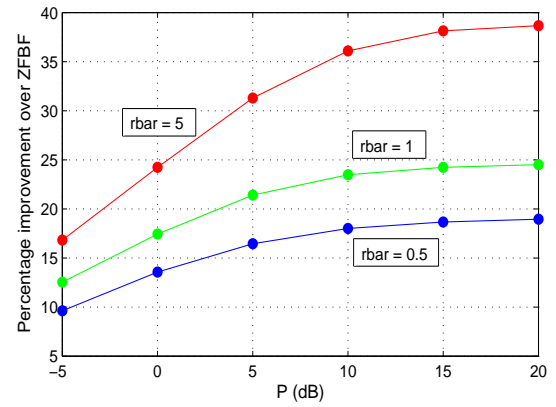


Fig. 2. Percentage improvement in the asymptotic throughput achieved with ZFDPC over ZFBF.

if the power is distributed across more users then the ‘pre-log’ factor gets improved (i.e., more users contribute to the sum-rate). However, increasing  $\bar{s}$  also increases the inter-user interference. Hence, at higher values of  $P$ ,  $\bar{s}_{opt}$  becomes constant. Next,  $\bar{s}_{opt}$  increases with  $\bar{r}$  because the inter-user interference reduces with increasing  $\bar{r}$ . Finally, note that  $\bar{s}_{opt}$  for ZFDPC is higher than that for ZFBF. As we would discuss below, ZFDPC manages the channel gain of the useful signal and the inter-user interference more efficiently than ZFBF and hence,  $\bar{s}_{opt}$  corresponding to it turns out to be higher.

In Fig. 2, we compare the asymptotic throughput obtained using ZFDPC and ZFBF. The numerical results in this figure pertain to  $\rho_{opt}$ . Here we plot the percentage improvement achieved using ZFDPC over ZFBF against  $P$  for three values of  $\bar{r}$ . We can see that ZFDPC achieves a considerably higher

$$\rho(P, \bar{s}, \bar{r}) = \bar{s} \left\{ \log \text{NR} - \int_0^1 \log \left( \text{NR} \cdot x^\infty(\bar{i}) - \frac{P}{\bar{s}} (1 - 2^{-\bar{r}}) (1 - \bar{i}\bar{s}) \left( \frac{\frac{P}{\bar{s}} (1 - 2^{-\bar{r}}) (\bar{i}\bar{s})}{1 + \frac{P}{\bar{s}} (1 - 2^{-\bar{r}}) (1 - \bar{i}\bar{s}) + P 2^{-\bar{r}}} + 1 \right)^2 \right) d\bar{i} \right\}, \quad (3)$$

$$\text{where } \text{NR} := \lim_{K \rightarrow \infty} \text{nr} = 1 + \frac{P}{\bar{s}} (1 - 2^{-\bar{r}}) + P 2^{-\bar{r}} \text{ is independent of } \bar{i}; \text{ and } x^\infty(\bar{i}) = 1 + \frac{(\frac{P}{\bar{s}})^2 (1 - 2^{-\bar{r}})^2 (1 - \bar{i}\bar{s}) (\bar{i}\bar{s})}{(1 + P 2^{-\bar{r}} + \frac{P}{\bar{s}} (1 - 2^{-\bar{r}}) (1 - \bar{i}\bar{s}))^2}.$$

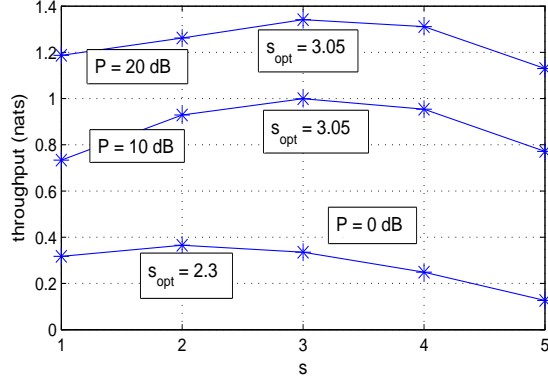


Fig. 3. Throughput achieved using ZFDPC vs.  $P$  for  $K = 5$  and  $r = 10$ .

throughput than ZFBF at all values of  $P$  and  $\bar{r}$ . Note that even for  $\bar{r}$  as low as 0.5, the percentage improvement is in the range of 10% to 20%, which is significant.

Let us now examine the differences between ZFDPC and ZFBF. Consider the  $i^{th}$  user. Under ZFDPC, the interference due to users 1 to  $i - 1$  is canceled by DPC. This can be accomplished for any given choice of the BF vectors of users 1 to  $i$ , as long as the inflation factor  $W_i$  is chosen in the appropriate manner. The BF vectors are chosen under ZFDPC in such a way that the interference due to users  $i + 1$  to  $s$  gets zeroforced at the  $i^{th}$  user. In contrast to this, under ZFBF, the BF vectors are selected so as to zeroforce the interference due to all other users. With this background, let us now analyze the channel gain of the useful signal, i.e., the term  $\frac{1}{s} |h_i^* \omega_i|^2$ . Under ZFDPC, it is proportional to  $\frac{1}{s} \chi_{2(K-i+1)}^2$  RV (recall that the other additive terms converge to zero in limit), whereas under ZFBF, it is proportional to  $\frac{1}{s} \chi_{2(K-s+1)}^2$  RV  $\forall i$  [6]. Thus, except for the user  $i = s$ , every other user experiences a stronger channel under ZFDPC. In other words, DPC (or ZFDPC) manages the channel gain of the useful signal and the interference together more effectively than ZFBF. It was known that, due to these differences, ZFDPC outperforms ZFBF under perfect CSIT [3]. In the light of the results obtained here, we conclude that the same is true even under imperfect CSIT as well. Lastly, it must be noted that with ZFBF, the asymptotic throughput is zero when  $\bar{s} = 1$ , i.e.,  $\rho(P, 1, \bar{r}) = 0 \forall P, \bar{r}$ . Note that for ZFDPC,  $\rho(P, 1, \bar{r})$  is comparable to  $\rho_{opt}(P, \bar{r})$ . This behavior can be easily understood by noting the distribution of the channel gain of the useful signal under two transmission schemes.

Consider now Fig. 3. Here, we plot the throughput (i.e., the sum-rate normalized by  $K$ ) for the GBC with  $K = 5$  and  $r = 10$ . We see from the figure that by optimizing over the number of users, an improvement of about 0.2 nats can be obtained (at all values of  $P$ ), over the simple solution of transmitting to all  $K$  users. This improvement is quite significant, especially at  $P = 0$  dB and  $P = 10$  dB. The question now is how to determine the optimal number of ‘on’ users in the  $K$ -dimensional GBC for a given  $P$  and  $r$ . Instead of performing an exhaustive search over  $s$ , we propose

a simple and computationally efficient approach. First find  $\bar{s}_{opt}$  for the given  $P$  and  $\bar{r} = \frac{r}{K}$ . Then we suggest that for the  $K$ -dimensional GBC, select  $s_{opt} := \bar{s}_{opt} K$  (rounded to the nearest integer) number of users. In Fig. 3, we see that the maximum value of the normalized throughput is indeed attained at  $s_{opt}$  (rounded to the nearest integer). We have observed this simple method to work quite accurately, even for the relatively small values of  $K$  ( $K = 5$  here). Note that the method suggested for the ZFBF in [6] using their large system analysis for selecting the number of users is more complicated than the one suggested here but seems to provide no particular benefit over this simple approach.

## V. CONCLUSION

We provide a large-system analysis of the GBC with finite-rate feedback and derive a closed-form expression for the asymptotic throughput achievable using ZFDPC. Using this result, we show that the DPC-based scheme achieves a significantly higher throughput than ZFBF. For the first time, DPC is shown to have a better performance, under imperfect CSIT. Also, using the asymptotic throughput expression, we address the problem of optimizing over the number of ‘on’ users.

## REFERENCES

- [1] M. Costa, “Writing on dirty paper,” *IEEE Trans. Inform. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.
- [2] H. Weingarten, Y. Steinberg, and S. Shamai, “The capacity region of multiple-input multiple-output broadcast channels,” *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.
- [3] G. Caire and S. Shamai, “On the achievable throughput of a multiantenna gaussian broadcast channel,” *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.
- [4] N. Jindal, “MIMO broadcast channels with finite rate feedback,” *IEEE Trans. Inform. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.
- [5] C. S. Vaze and M. K. Varanasi, “The degrees of freedom regions of MIMO broadcast, interference, and cognitive radio channels with no CSIT,” Sep. 2009, Available Online: <http://arxiv.org/abs/0909.5424>.
- [6] W. Dai, Y. Liu, B. Rider, and W. Gao, “How many users should be turned on in a multi-antenna broadcast channel?” *IEEE Journal on Sel. Areas of Comm.*, vol. 26, Issue 8, pp. 1526–1535, Oct. 2008.
- [7] C. S. Vaze and M. K. Varanasi, “Dirty paper coding for fading channels with partial transmitter side information,” in *Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, USA, Oct. 2008, pp. 341–345.
- [8] —, “On the achievable rate of the fading dirty paper channel with imperfect CSIT,” in *43rd Annual Conference on Information Sciences and Systems*, John Hopkins University, Baltimore, USA, Mar. 2009, pp. 346–351.
- [9] —, “Dirty paper coding for the MIMO gaussian broadcast channels with imperfect CSIT,” to be submitted to *IEEE Trans. Inform. Theory*, 2010, under preparation.
- [10] —, “On the scaling of feedback bits to achieve the full multiplexing gain over the gaussian broadcast channel using DPC,” in *submitted to IEEE Intern. Symp. Inform. Theory*, Jun. 2010.
- [11] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*. Foundations and Trends in Communications and Information Theory, Vol. 1, Issue 1, NOW publishers.
- [12] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*. McGraw-Hill Publishing Company, 2002.
- [13] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, “On beamforming with finite rate feedback in multiple-antenna systems,” *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2562–2579, Oct. 2003.
- [14] T. Yoo, N. Jindal, and A. Goldsmith, “Multi-antenna downlink channels with limited feedback and user selection,” *IEEE Journal on Selected Areas in Communication*, vol. 25, no. 7, pp. 1478–1491, Sep. 2007.
- [15] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge Univ. Press, 1985.