Downlink Capacity Evaluation of Cellular Networks With Known-Interference Cancellation

Harish Viswanathan, *Member, IEEE*, Sivarama Venkatesan, *Member, IEEE*, and Howard Huang, *Member, IEEE*

***Abstract*—Recently, the capacity region of a multiple-input multiple-output (MIMO) Gaussian broadcast channel, with Gaussian codebooks and known-interference cancellation through dirty paper coding, was shown to equal the union of the capacity regions of a collection of MIMO multiple-access channels. We use this duality result to evaluate the system capacity achievable in a cellular wireless network with multiple antennas at the base station and multiple antennas at each terminal. Some fundamental properties of the rate region are exhibited and algorithms for determining the optimal weighted rate sum and the optimal covariance matrices for achieving a given rate vector on the boundary of the rate region are presented. These algorithms are then used in a simulation study to determine potential capacity enhancements to a cellular system through known-interference cancellation. We study both the circuit data scenario in which each user requires a constant data rate in every frame and the packet data scenario in which users can be assigned a variable rate in each frame so as to maximize the long-term average throughput. In the case of circuit data, the outage probability as a function of the number of active users served at a given rate is determined through simulations. For the packet data case, long-term average throughputs that can be achieved using the proportionally fair scheduling algorithm are determined. We generalize the zero-forcing beamforming technique to the multiple receive antennas case and use this as the baseline for the packet data throughput evaluation.**

***Index Terms*—Broadcast channel, dirty paper coding (DPC), duality, multiple-input multiple-output (MIMO), packet data, scheduling.**

I. Introduction

THE DEMAND for wireless connectivity is expected to rise significantly over the next decade for both delay-tolerant packet data and delay-sensitive circuit data such as voice and video conferencing. Mobile cellular networks are constantly evolving to support the higher market demands. New innovations in signal processing, coding, resource allocation, network planning, etc., are rapidly being incorporated into cellular wireless systems, to enhance capacity and coverage. Multiple antennas at the base station and possibly at the terminals, are likely to be deployed in future cellular systems. The capacity analysis of multiple-input multiple-output (MIMO) point-to-point systems in [1] and [2] showed significant gains over single antenna systems. We consider the application of MIMO techniques on the downlink of a cellular system, which is a point-to-multipoint or broadcast system. Recently, it has been demonstrated that *known-interference cancellation* techniques—where the base station, through appropriate coding of the signal intended for a given user, can cancel the interference from the signals of previously encoded users—can be used to further enhance the capacity of MIMO broadcast systems, beyond the direct application of point-to-point MIMO signaling techniques. While known-interference cancellation at the base station can clearly enhance capacity, it is unclear how much can be gained in a typical cellular system operating under cochannel interference from neighboring cells. In this paper, we seek to understand, through a simulation study, the potential gains from employing known-interference cancellation at the base stations of a cellular network, while maintaining the current network architecture of independent transmissions of signals from each base station. The optimization algorithms required to obtain these capacity estimates are also studied in this paper.

MIMO broadcast channels belong to the category of non-degraded broadcast channels, for which the capacity region remains unknown. Recently, however, significant advances have been made toward characterizing this capacity region [3], [4], [9], by exploiting a surprising result due to Costa [7] on known-interference cancellation at the transmitter. Briefly, Costa showed that the capacity of the standard scalar single-user additive white Gaussian noise channel is *unchanged* in the presence of an independent additive Gaussian interferer, provided that the interferer's signal is known noncausally to the transmitter. This result effectively shows that, while encoding the desired user's signal, the transmitter can perform a pre-cancellation of the interfering signal *without* a power or rate penalty, a process termed "writing on dirty paper" in [7]. This result can easily be extended to the MIMO single-user Gaussian channel, too [23], [24].

In the context of a multiuser Gaussian broadcast channel, Costa's result suggests the following encoding strategy. Suppose the transmitter encodes the users's data sequentially (in some order), using Gaussian signals. Then, at the time a given user's data is encoded, the signals to be transmitted to all previously encoded users are fully known to the transmitter. The dirty paper coding (DPC) technique then ensures that the given user will not suffer any interference from all the previously encoded users (of course, there will still be interference from users encoded *after* the given user).

Using the above ideas, it has been demonstrated recently that the sum capacity of the Gaussian MIMO broadcast channel is

Manuscript received May 2, 2002; revised November 26, 2002. This paper was presented in part at the DIMACS Workshop on Signal Processing for Wireless Transmission, Rutgers University, NJ, October 2002.

H. Viswanathan is with the Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974 USA (e-mail: harishv@lucent.com).

S. Venkatesan and H. Huang are with the Bell Laboratories, Lucent Technologies, Holmdel, NJ 07733 USA (e-mail: sivarama@lucent.com; hchuang@lucent.com).

Digital Object Identifier 10.1109/JSAC.2003.810346

achieved [4], [8], [9]. Further, in [9], a characterization of the entire rate region achievable with DPC is given, based on a duality between broadcast channels and multiple-access channels (MAC). Using this characterization as a tool, we present optimization algorithms to evaluate capacity gains for the MIMO downlink cellular system employing known-interference cancellation. Specifically, we consider the two scenarios in which the base station has to transmit to all the users in the cell at a *fixed rate* (circuit data case) in each frame and in which the base station *schedules* transmissions to the users at variable rate (packet data case). We evaluate the circuit data capacity as the number of users that can be supported at the required rate subject to a specified outage probability and the packet data capacity in terms of the total throughput that can be provided under a proportionally fair scheduler [16].

Algorithms for determining the optimal weighted rate sum, which is the basic optimization required to evaluate the packet data throughput, are considered in [13]. However, [13] does not consider the problem of achieving a specific set of target rates for the different users, as required for the circuit data capacity. We present an algorithm for determining whether a specified target rate vector is achievable with the available power and also present an alternative algorithm for determining the optimal weighted rate sum. These algorithms are used in our simulation study of the cellular network.

The paper is organized as follows. In Section II, we present the system description and the relevant results from MIMO broadcast channel information theory. In Section III, we present some results that characterize the rate region and provide algorithms for the optimization problems involved in the cellular network capacity evaluation. In Section IV, we present the circuit data capacity evaluation results and in Section V the packet data throughput results. We conclude with a summary and suggestions for future work in Section VI.

II. Known Interference-Cancellation and the MIMO Broadcast Rate Region

Consider a single cell of a cellular system in which the base station has M transmit antennas and each of the K users has N receive antennas. User k's complex-baseband channel output at time j, denoted $\mathbf{y}_k(j)$, is modeled as

$$\mathbf{y}_k(j) = \mathbf{H}_k\mathbf{x}(j) + \mathbf{z}_k(j). \tag{1}$$

Here, $\mathbf{x}(j)$ is the vector signal transmitted by the base station at time j, constrained to have power no greater than P, i.e., $tr(E[\mathbf{x}(j)\mathbf{x}(j)^\dagger]) \le P$; $\mathbf{H}_k$ is a matrix of size $N \times M$, whose entries are assumed to be IID zero-mean complex-Gaussian random variables (independent across users); and $\mathbf{z}_k(j)$ represents the random additive interference that user k experiences due to both receiver noise and the transmissions from all other base stations in the system (for simplicity, we do not allow intercell coordination). We assume that $\mathbf{z}_k(j)$ is complex-Gaussian with zero mean and covariance $\mathbf{I}_N$ (the identity matrix of size N) and that each entry of $\mathbf{H}_k$ has variance $1/\sigma_k^2$. Thus, the average signal-to-interference-plus noise ratio (SINR) of user k's channel is P/σ_k^2.

We consider a block fading model in which each $\mathbf{H}_k$ is constant for the duration of a frame. The frame itself is assumed to be long enough to allow communication at rates close to capacity. The base station is assumed to have perfect knowledge of all the users' channels $\mathbf{H}_k$. The signal transmitted from the base station is the sum of the signals transmitted to each of the users: $\mathbf{x}(j) = \sum_{k=1}^K \mathbf{x}_k(j)$. Since we consider only those coding schemes in which the $\mathbf{x}_k$ are independent, the overall transmit covariance is given by $\sum_{l=1}^K E[\mathbf{x}_l\mathbf{x}_l^\dagger] = \sum_{l=1}^K \mathbf{\Gamma}_l$, where $\mathbf{\Gamma}_l$ is the transmit covariance matrix of user l. The total power transmitted from the base is given by $tr(\sum_{l=1}^K \mathbf{\Gamma}_l)$ and has to satisfy the transmit power constraint $tr(\sum_{l=1}^K \mathbf{\Gamma}_l) \le P$.

With DPC, users can be encoded sequentially in such a way that each user sees no interference from previously encoded users. The rate region achievable by DPC is, therefore, given by [4] and [9]

$$\mathcal{R}_{\text{BC}}^{\text{DPC}} = Co\left(\bigcup_\pi \mathcal{R}_\pi\right) \tag{2}$$

where the union is over all permutations π of $1, 2, \ldots, K$ and

$$\mathcal{R}_\pi = \Bigg\{(R_1,\ldots,R_K) : R_{\pi(k)} = \log \frac{\left|\mathbf{I} + \sum_{l=1}^{k} \mathbf{H}_{\pi(l)}\mathbf{\Gamma}_{\pi(l)}\mathbf{H}_{\pi(l)}^\dagger\right|}{\left|\mathbf{I} + \sum_{l=1}^{k-1} \mathbf{H}_{\pi(l)}\mathbf{\Gamma}_{\pi(l)}\mathbf{H}_{\pi(l)}^\dagger\right|} \cdot \sum_{l=1}^K tr\left(\mathbf{\Gamma}_l\right) \le P\Bigg\}. \tag{3}$$

Here, $Co(\mathcal{A})$ denotes the convex hull of the set $\mathcal{A}$ and $|\mathbf{A}|$ denotes the determinant of a matrix $\mathbf{A}$.

The above rate region achievable with DPC can be characterized in terms of the rate regions of a set of dual MAC channels [9]. In the dual MAC, the users transmit to the base station on channels that are dual to the downlink channels. The dual MAC is given by

$$\mathbf{y}(j) = \sum_{i=1}^K \mathbf{H}_i^\dagger \mathbf{x}_i(j) + \mathbf{w}(j) \tag{4}$$

where $\mathbf{x}_i(j)$ is now the vector signal transmitted by user i and $\mathbf{w}(j)$ is the additive white complex-Gaussian noise in the base station receiver, assumed to have zero mean and a covariance of I_M (the identity matrix of size M). In the dual MAC, there is a separate transmit power constraint for each user, say, P_i for user i.

The MAC rate region corresponding to a given transmit power constraint vector $\mathbf{P} = (P_1,\ldots,P_K)$ is given by [9] and [17]

$$\begin{aligned}\mathcal{R}_{\text{MAC}}(\mathbf{P}) &= \bigcup_{\substack{\mathbf{Q}_1,\ldots,\mathbf{Q}_K \\ tr(\mathbf{Q}_i)\le P_i}} \mathcal{R}_{\text{MAC}}^{cov}\left(\mathbf{Q}_1,\ldots,\mathbf{Q}_K\right) \\ &= \bigcup_{\substack{\mathbf{Q}_1,\ldots,\mathbf{Q}_K \\ tr(\mathbf{Q}_i)\le P_i}} \Bigg\{\mathbf{R} : \sum_{i\in\mathcal{S}} R_i \le \log\left|\mathbf{I} + \sum_{i\in\mathcal{S}} \mathbf{H}_i^\dagger \mathbf{Q}_i \mathbf{H}_i\right| \\ &\qquad \text{for each subset } \mathcal{S}\Bigg\}\end{aligned} \tag{5}$$

where the $\mathbf{Q}_i$ are the MAC transmit covariance matrices. The duality result of [9] states that the DPC rate region defined in (2) equals the union of the MAC rate regions corresponding to all power constraint vectors $\mathbf{P} = (P_1, \ldots, P_K)$ satisfying $\sum_i P_i \le P$.

Lemma 1:

$$\mathcal{R}_{\mathrm{BC}}^{\mathrm{DPC}} = \bigcup_{\mathbf{P}: \sum_i P_i \le P} \mathcal{R}_{\mathrm{MAC}}(\mathbf{P}). \tag{6}$$

III. Optimization Algorithms

Given the characterization of the rate region in (6), one has to determine the optimal operating points on the boundary of the rate region in order to determine the cellular network capacity. For example, to determine if the network can support a given number of users at a given target rate requirement for each user, one has to determine whether the target rate vector is in the interior of the rate region. Similarly, to determine the optimum throughput when using a proportionally fair scheduler (described in Section V), it is necessary to determine the point on the boundary of the rate region that achieves the maximum weighted rate sum for a given set of weights. It may also be necessary to determine the optimal covariance matrices that achieve the target rate vector. We present several results in this section that further characterize the properties of the rate region and provide optimization algorithms for network capacity evaluation.

Lemma 2: The boundary of the rate region $\mathcal{R}_{\mathrm{BC}}^{\mathrm{DPC}}(P)$ is characterized by the set of vectors $\mathbf{R}^*$ that are solutions to the optimization problem

$$\max_{\mathbf{R}} \mu \cdot \mathbf{R} \text{ subject to } \mathbf{R} \in \mathcal{R}_{\mathrm{BC}}^{\mathrm{DPC}}(P) \tag{7}$$

for some choice of the weight vector μ. Furthermore

$$\begin{aligned}\max_{\mathbf{R} \in \mathcal{R}_{\mathrm{BC}}^{\mathrm{DPC}}} \mu \cdot \mathbf{R} = & \max_{\mathbf{Q}_i: \sum_{i=1}^K tr(\mathbf{Q}_i) = P} \sum_{i=1}^{K-1} \left(\mu_{\pi(i)} - \mu_{\pi(i+1)}\right) \\ & \log \left| \mathbf{I} + \sum_{l=1}^{i} H_{\pi(l)}^{\dagger} \mathbf{Q}_{\pi(l)} H_{\pi(l)} \right| \\ & + \mu_{\pi(K)} \log \left| \mathbf{I} + \sum_{l=1}^{K} H_{\pi(l)}^{\dagger} \mathbf{Q}_{\pi(l)} H_{\pi(l)} \right| \end{aligned} \tag{8}$$

where π is the permutation of the indices $\{1, \ldots, K\}$ such that $\mu_{\pi(1)} \ge \mu_{\pi(2)} \ge \cdots \ge \mu_{\pi(K)}$.

Proof: The first statement follows from the fact that $\mathcal{R}_{\mathrm{BC}}^{\mathrm{DPC}}(P)$ is a convex set. The second statement follows essentially from the fact that $\mathcal{R}_{\mathrm{MAC}}(\mathbf{P})$ is a polymatroid [18]. □

The optimization involved in (8) is the maximization of a concave function over a convex set and, thus, any standard convex optimization procedure can be used. Note that once the optimum uplink covariance matrices $\mathbf{Q}_k$ are determined, the equivalent downlink covariance matrices can be obtained through the duality transformations in [9]. We use the following algorithm derived from the general technique in [10] and [11] for determining the maximum weighted rate sum in Section V.

Algorithm 1: Without loss of generality assume that $\mu_1 \ge \mu_2 \ge \cdots \mu_K$.

Define the objective function

$$\begin{aligned} f(\mathbf{Q}_1, \ldots, \mathbf{Q}_K) = & \sum_{i=1}^{K-1} (\mu_i - \mu_{i+1}) \log \left| \left\{ \mathbf{I} + \sum_{l=1}^{i} \mathbf{H}_l^{\dagger} \mathbf{Q}_l \mathbf{H}_l \right\} \right| \\ & + \mu_K \log \left| \left\{ \mathbf{I} + \sum_{l=1}^{K} \mathbf{H}_l^{\dagger} \mathbf{Q}_l \mathbf{H}_l \right\} \right|. \end{aligned}$$

The gradient of the objective function with respect to the covariance matrix $\mathbf{Q}_j, 1 \le j \le K$, is then given by

$$\begin{aligned} \nabla f_j(\mathbf{Q}_1, \ldots, \mathbf{Q}_K) = & \sum_{i=j}^{K-1} (\mu_i - \mu_{i+1}) \left[\mathbf{H}_j \left\{ \mathbf{I} + \sum_{l=1}^{i} \mathbf{H}_l^{\dagger} \mathbf{Q}_l \mathbf{H}_l \right\}^{-1} \mathbf{H}_j^{\dagger} \right] \\ & + \mu_K \left[\mathbf{H}_j \left\{ \mathbf{I} + \sum_{l=1}^{K} \mathbf{H}_l^{\dagger} \mathbf{Q}_l \mathbf{H}_l \right\}^{-1} \mathbf{H}_j^{\dagger} \right]. \end{aligned}$$

The algorithm proceeds iteratively as follows. Given the nth iterate $\mathbf{Q}_i(n), 1 \le i \le K$, determine the principal eigenvectors $\mathbf{v}_i$ (of unit norm) and the corresponding principal eigenvalues λ_i of the gradients $\nabla f_i(\mathbf{Q}_1(n), \ldots \mathbf{Q}_K(n))$ for $1 \le i \le K$. Let $j^* = \arg\max(\lambda_1, \ldots, \lambda_K)$.

Then, compute the $(n+1)$st iterate $\mathbf{Q}_i(n+1), 1 \le i \le K$, as follows:

$$\begin{aligned} \mathbf{Q}_{j^*}(n+1) = & t^* \mathbf{Q}_{j^*}(n) + (1 - t^*) P \mathbf{v}_{j^*} \mathbf{v}_{j^*}^{\dagger} \\ \mathbf{Q}_i(n+1) = & t^* \mathbf{Q}_i(n), i \ne j^* \end{aligned}$$

where t^* is the solution to the following one-dimensional optimization that can be solved through bisection.

$$\begin{aligned} t^* = \arg \max_{0 \le t \le 1} f \Big(& t\mathbf{Q}_1(n), \ldots t\mathbf{Q}_{j^*}(n) \\ & + (1-t) P \mathbf{v}_{j^*} \mathbf{v}_{j^*}^{\dagger}, \ldots, t\mathbf{Q}_K(n) \Big) \end{aligned}$$

As $n \to \infty$, the covariance matrices converge to the optimum covariance matrices.

While every point on the boundary of the rate region is a solution to the optimization in (7) for some value of μ, it is not clear that for each value of μ there is a unique $\mathbf{R}^*$ on the boundary that achieves the maximum in (7). Hence, it is not guaranteed that the optimum covariance matrices for every rate vector on the boundary of the rate region can be obtained by the above algorithm for maximizing $\mu \cdot \mathbf{R}$ for some value of μ. Furthermore, it is not clear that it is possible to achieve all points on the boundary of the rate region without time-sharing. In general, time-sharing or rate splitting may be necessary to achieve all points on the boundary of the rate region [19].

The following example explicitly demonstrates the need for time-sharing by computing the rate region $\bigcup_\pi \mathcal{R}_\pi$, where $\mathcal{R}_\pi$ is defined in (3).

Example: Consider the two-user broadcast channel, where the transmitter has four transmit antennas and each receiver has one antenna with the complex channel gain matrices given by the equation at the bottom of the next page. The total transmit power is set to one and the noise variance at the receivers is also set to one. Fig. 1 shows the rate regions obtained for this set up for the two different encoding orders $\pi_1 = (12)$ and $\pi_2 = (21)$. Clearly, the region $\mathcal{R}_{\pi_1} \cup \mathcal{R}_{\pi_2}$ is not convex. Time-

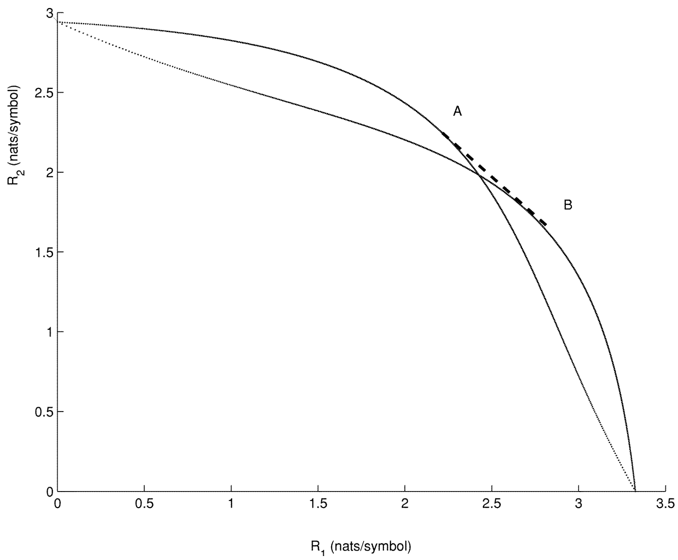

Fig. 1. Rate region for a two-user broadcast channel.

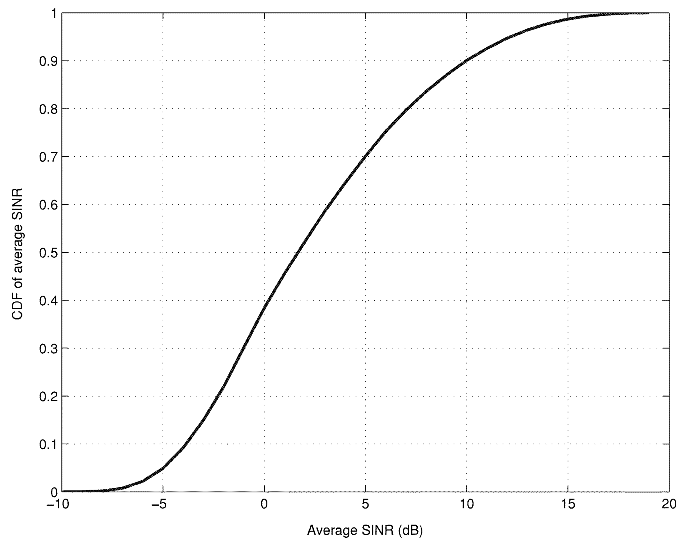

Fig. 2. CDF of average SINR.

sharing between the points A and B is required to achieve all the points in the region $\mathcal{R}_{\mathrm{BC}}^{\mathrm{DPC}}$ which is by definition the convex hull of $\mathcal{R}_{\pi_1} \cup \mathcal{R}_{\pi_2}$. It is clear that for all rate vectors $\mathbf{R}^*$ on the line segment AB in the rate region, the weight vectors $\mu^*(\mathbf{R}^*)$ such that $\mathbf{R}^* = \arg\max_{\mathbf{R}\in\mathcal{R}_{\mathrm{BC}}^{\mathrm{DPC}}} \mu^* \cdot \mathbf{R}$ are identical thereby establishing that the solution to the maximum weighted rate sum is not unique.

In general, as demonstrated by the example above, time-sharing may be required between permutations and, thus, it is not possible to find all points in the rate region through the weighted rate sum maximization for different weight vectors. The following lemma shows how to obtain the μ for a given target rate on the boundary of the rate region. Furthermore, it shows that we can determine whether a given rate vector is achievable with the available power, which we use in the circuit data capacity evaluation study described in Section IV.

Lemma 3: Define

$$g(\mu) \stackrel{\triangle}{=} \max_{\mathbf{R}\in\mathcal{R}_{\mathrm{BC}}^{\mathrm{DPC}}(P)} \mu \cdot \mathbf{R} \tag{9}$$

and let

$$\mu^* = \arg \min_{\mu\geq 0, \sum_i \mu_i = 1} g(\mu) - \mu \cdot \mathbf{R}^*. \tag{10}$$

The rate vector $\mathbf{R}^*$ is achievable with power P if and only if

$$g(\mu^*) - \mu^* \cdot \mathbf{R}^* \geq 0. \tag{11}$$

Furthermore, equality in (11) is achieved if and only if $\mathbf{R}^*$ is on the boundary of the rate region.

Proof: The proof follows from the application of some basic results about convex sets and supporting hyperplanes to the convex set $\mathcal{R}_{\mathrm{BC}}^{\mathrm{DPC}}(P)$ [15]. □

Note that the function $g(\mu)$ is a convex function and, thus, has a global minimum. Furthermore, (10) is also a convex optimization and, thus, can be solved through standard techniques. The main advantage in using the above result lies in the fact that one can quickly determine rate vectors that lie outside the rate region since it is sufficient to find some μ for which $g(\mu) - \mu \cdot \mathbf{R}^* < 0$.

We now use the above characterization of the rate region and the optimization algorithm to numerically evaluate the gains from using known-interference cancellation in a cellular system. We first present the results for the fixed rate circuit data scenario and then for the packet data scenario.

IV. Circuit Data Capacity Evaluation

In this section, we present numerical results to quantify the capacity gains that are possible through the use of DPC, in a system of *fixed-rate*, or "circuit data," users. More precisely, we focus on a single cell in a cellular network and assume that users within this cell desire service at a fixed rate R from the base station in the cell. The base station is assumed to have perfect knowledge of the users' channels.

We follow the channel model in (1) with total power P set equal to one and the noise variance σ_k^2 to $1/\rho_k$, where ρ_k represents the average SINR at user k's location when all of the base station power is assigned to this user. In the simulations, this average SINR is assigned independently to all the users according to the distribution in Fig. 2.

The baseline for our capacity comparisons is a conventional time-division multiple-access (TDMA) system, in which the base station serves only one user at any time, using all the power available to it. This system is compared with one in which the base station uses DPC to serve *simultaneously* as many users in the cell as can be supported under the transmit power constraint. These two systems are described in further

---

$$\mathbf{H}_1 = [-1.5 - 0.53i \quad 2.91 - 2.81i \quad -1.84 + 0.44i \quad 0.62 - 1.99i],$$
$$\mathbf{H}_2 = [-0.61 - 0.75i \quad 0.26 - 2.15i \quad 0.92 - 0.11i \quad -1.16 - 3.18i].$$

detail in the following subsections. In both cases, the number of users that can be supported depends on the users' locations within the cell (captured through their average SINRs) and the realizations of the users' channel matrices and is therefore random. So, we represent system capacity through the cumulative density function (CDF) of the maximum number of users that can be supported.

A. Conventional TDMA System

In a conventional TDMA system, the base station serves only one user at a time, using the entire transmit power of 1 available to it. When serving user k, the base station chooses the distribution of its transmit signal $\mathbf{x}$ to achieve the capacity of user k's channel, thus allowing it to serve each user at the desired rate R as quickly as possible.

Let the capacity of user k's channel when the base station serves it exclusively, using all the transmit power available, be C_k. The value of C_k can be obtained by a standard "waterfilling" procedure over the eigenmodes of the matrix $\mathbf{H}_k^\dagger \mathbf{H}_k$. The base station must then dedicate the channel to user k for a fraction of time R/C_k. The maximum number of users that can be supported is, therefore, given by

$$N_{\text{TDMA}} = \max\left\{N : \sum_{k=1}^{N} \frac{R}{C_k} \leq 1\right\}. \tag{12}$$

B. System Based on DPC With Arbitrary Ordering

In the DPC system, the base station serves all users simultaneously, using DPC techniques. This requires imposing an encoding order on users, determining which users experience interference from which other users. We assume that users are simply ordered according to their indices (which might represent, for example, their order of arrival into the system as suggested in [14]). With this "first-come first-served" (FCFS) ordering, user k experiences interference only from users $1, 2, \ldots, k-1$, but not from users indexed $k+1$ or higher. The transmitted vector signal $\mathbf{x}$ in (1) can then be expressed as $\mathbf{x}(j) = \sum_k \mathbf{x}_k(j)$, where $\mathbf{x}_k(j)$ is the signal intended for user k at time j.

Note that, since user k is oblivious to the presence of users indexed $k+1$ or higher, its channel can effectively be modeled by

$$\mathbf{y}_k(j) = \mathbf{H}_k \mathbf{x}_k(j) + \mathbf{H}_k \sum_{i<k} \mathbf{x}_i(j) + \mathbf{z}_k(j). \tag{13}$$

We restrict the transmitted signals $\mathbf{x}_k(j)$ to have stationary complex-Gaussian distributions. It is easy to see that there is no loss of optimality in further restricting $\mathbf{x}_k(j)$ to be circularly symmetric and to have zero mean. Therefore, the distribution of $\mathbf{x}_k(j)$ is completely characterized by the covariance $\mathbf{Q}_k = E[\mathbf{x}_k(j)\mathbf{x}_k(j)^\dagger]$.

The base station chooses the transmit covariance $\mathbf{Q}_k$ for user k to minimize the power required to achieve the required rate of R, given user k's channel matrix $\mathbf{H}_k$ and the transmit covariances $\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_{k-1}$ already assigned to users $1, 2, \ldots, k-1$. This optimization can be accomplished through a standard "waterfilling" procedure over the eigenmodes of $\tilde{\mathbf{H}}_k^\dagger \tilde{\mathbf{H}}_k$, where

$$\tilde{\mathbf{H}}_k = \left[I + \mathbf{H}_k \left(\sum_{i<k} \mathbf{Q}_i\right) \mathbf{H}_k^\dagger\right]^{-1/2} \mathbf{H}_k. \tag{14}$$

With the above scheme, the base station serves as many users as the assumed transmit power constraint of one allows. The maximum number of users that can be supported is, therefore, given by

$$N_{\text{DPC}} = \max\left\{N : \sum_{k=1}^{N} tr\left(\mathbf{Q}_k\right) \leq 1\right\}. \tag{15}$$

C. Other Encoding Orders for DPC System

So far, we have assumed that the encoding order of users for DPC is determined simply by their indices, which are essentially assigned arbitrarily. The transmit covariances for the users are then determined in a greedy fashion, i.e., each user's transmit covariance is chosen optimally for the noise and interference from all previously encoded users, by a waterfilling procedure. While this scheme has the virtue of being simple, it is possible that a more sophisticated encoding order and method of choosing transmit covariances could result in a significantly higher circuit-data capacity (measured, as before, by the number of fixed-rate users that can be served by the base station at a specified outage probability). In this section, we consider two such advanced schemes and compare their capacities with the simple greedy FCFS scheme.

The first scheme encodes users in increasing order of the power that they would require to achieve their target rate, in the absence of any interference from other users in the same cell (other-cell interference, whose strength is represented by the user's average SINR, is of course assumed to be present for this computation). This encoding order will henceforth be referred to as "minimum power first." Once the encoding order is determined, the transmit covariances are still chosen in a greedy manner, as in the FCFS case.

The second scheme uses both an optimal encoding order and an optimal choice of transmit covariances, in that it determines whether or not it is possible to accommodate a given number of users with *some* choice of encoding order and transmit covariances. Lemma 3 provides an efficient method for making this decision. Thus, this scheme also illustrates the use of Lemma 3.

Note that since the method of determining covariance matrices in the FCFS and MPF schemes is based on sequential waterfilling, these schemes are suboptimal even when the optimal encoding order is used. While waterfilling to the channel and interference so far may minimize the transmit power for one user, doing so might hurt users who have to deal with this interference. Thus, these schemes are suboptimal not only because of the suboptimal ordering but also because the transmit covariance matrices may be suboptimal. Thus, the optimum scheme based on the rate region described above is advantageous not simply in the encoding order but also in the transmit covariance matrices. However, both the MPF and the optimum schemes require the encoding order and transmit covariances to be recomputed each time a new user enters or exits the system. For this reason, they are much harder to implement in practice.

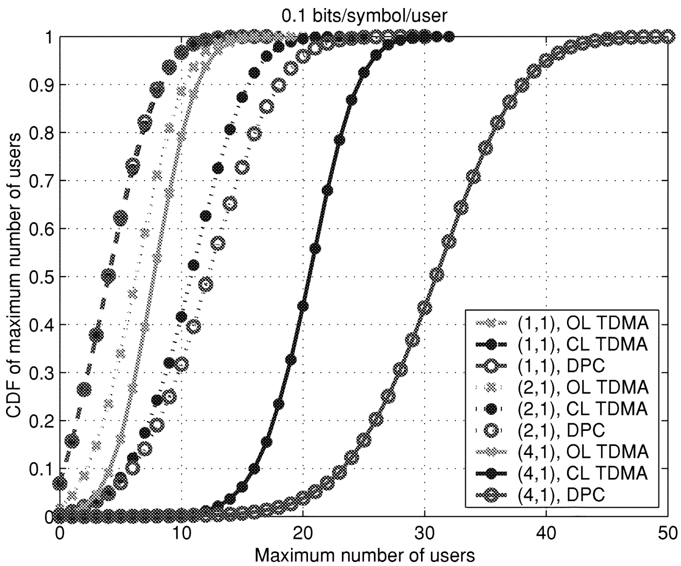

Fig. 3. Outage capacity with single receive antenna.

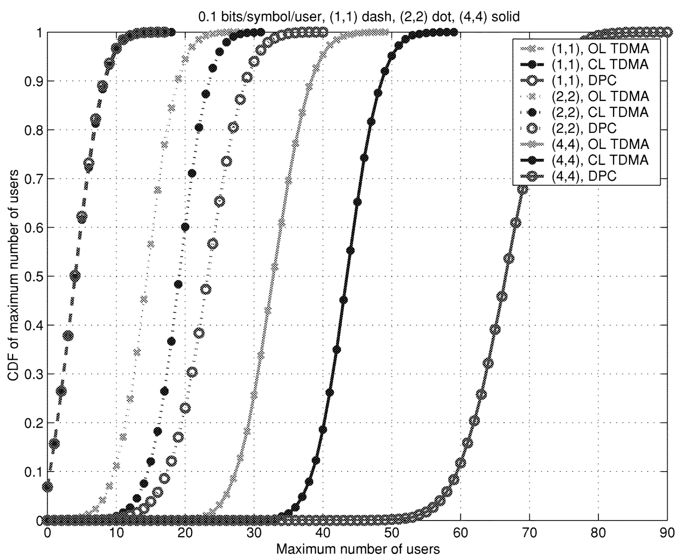

Fig. 4. Outage capacity with multiple receive antennas.

D. Simulation Results

Figs. 3 and 4 show the CDFs of the maximum number of users that can be supported by the TDMA system and DPC system (indicated by "CL TDMA" and "DPC," respectively) with FCFS ordering for various antenna configurations, assuming a desired rate per user of 0.1 bits/symbol. These CDFs were obtained by running 10 000 trials, in each of which users were added to the system one by one, till it was impossible to accommodate another user. The users were independently assigned average SINRs and channel matrices, as described earlier.

For comparison, we have also shown the CDFs for an open-loop TDMA system (indicated by "OL TDMA"), in which the base station does not know the users' channels and must, therefore, use a transmit covariance that is a multiple of the identity matrix. In this case, the capacity of user k's channel is $\log\left|I_N + (\rho_k/M)\mathbf{H}_k\mathbf{H}_k^\dagger\right|$ bits/symbol.

If we use the number of users that can be supported with at least 90% probability as the measure of capacity, we find that in the (4,1) and (4,4) cases, DPC provides about a 50% capacity improvement over closed-loop TDMA. The improvement is somewhat smaller in the (2,2) case and is almost negligible in the (1,1) and (2,1) cases. In general, we can expect the capacity advantage of DPC over TDMA to increase as the number of antennas deployed at one or both ends of the communication link increases. The higher dimensionality results in less interference between the users, on average, which makes it more advantageous to transmit to multiple users at the same time.

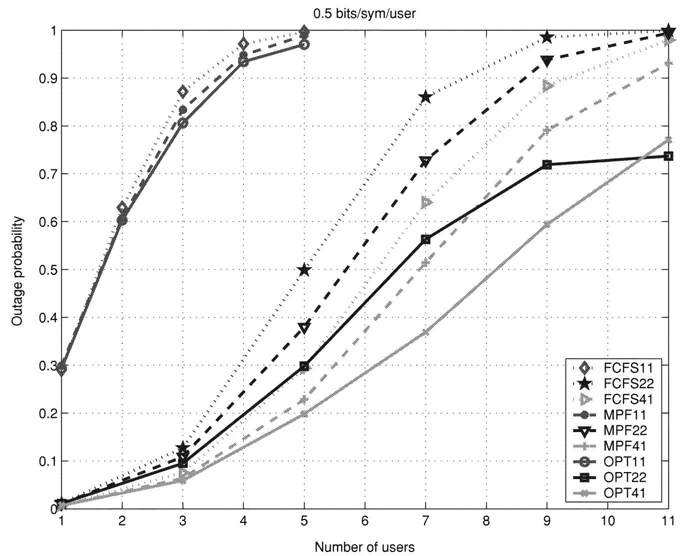

Fig. 5. Outage capacity for optimum and minimum power first orderings.

Fig. 5 shows the outage probability as a function of the number of the users desiring fixed-rate service, for the FCFS, minimum-power-first and optimal-ordering schemes, for three antenna configurations: (1,1), (2,2), and (4,1). In all cases, the target rate for each user is assumed to be 0.5 bits/symbol. These results show that at low outage probabilities, say 10%, the simple FCFS ordering actually has a negligible shortfall with respect to the more advanced schemes. However, at higher outages we do see a benefit from using the optimum order. A possible explanation for this is as follows. At low outages an outage event is a rare event and most likely occurs because there is one user in the system with a very bad channel condition and, hence, cannot meet the target rate requirement even when the optimal ordering is used. On the other hand at higher outages, where the outage scenario is more common and there are a larger number of users in system, it possible that some users even with reasonable channel conditions do not meet the target rate with the FCFS scheme whereas using the optimum ordering makes it possible to transmit at the target rate to all the users. Thus, we see a larger gap between the optimum and suboptimum ordering at higher outages.

V. Packet Data Throughput Evaluation

Consider a packet data cellular system with multiple antennas at the base station and one or more antennas at the terminal. In delay-tolerant packet data systems, it is possible, with the aid of channel condition feedback from the users, to schedule transmission to users when their channel conditions are favorable, thereby achieving multiuser diversity. Several scheduling algorithms have been proposed for packet data scheduling on the downlink of a cellular system [12], [20], [21]. These scheduling algorithms are designed primarily for single antenna base stations and do not consider cancellation of known interference through DPC. In this section, we evaluate the throughput gains from DPC through simulation of a particular scheduling scheme which is derived from the *proportionally fair* scheduling algorithm described in [12] and [21].

A. Scheduling Algorithm

The scheduling algorithm used attempts to provide some fairness in the sense that users with low average SINR are not denied service. On any scheduling interval n (duration over which channel is assumed to be constant), the base transmits to a subset of users that maximizes the weighted rate-sum criterion

$$\max \sum_{i=1}^{K} \mu_i(n) R_i(n) \tag{16}$$

where the maximization is over all rate vectors in the feasible rate region, which depends on the coding technique. When DPC is applied the solution is obtained using Algorithm 1. For the zero-forcing (ZF) beamforming technique that we use as the baseline for comparison, the optimization is described in the next section. The weights for the kth user during the nth frame or scheduling interval is

$$\mu_k(n) = \frac{1}{\bar{R}_k(n)}$$

where $\bar{R}_k(n)$ is the average throughput achieved by user k up to time n, updated according to the following:

$$\bar{R}_k(n) = (1-\delta)\bar{R}_k(n-1) + \delta R_k(n-1) \tag{17}$$

where $\delta = 0.99$ is a "forgetting factor," and $R_k(n-1)$ is the rate achieved by the kth user during transmission interval $n-1$. weighted Note that under the restriction of transmission to a single user in each scheduling interval, the above scheme becomes identical to the scheduling algorithm in [12] and [21]. In fact, even when the base station may transmit to multiple users in a single scheduling interval, the scheduling algorithm above can be shown to satisfy the proportionally fair optimality criterion, i.e., $\sum_i (\bar{T}_i^* - \bar{T}_i)/\bar{T}_i^* \geq 0$ where $\bar{T}_i^*$ and $\bar{T}_i$ are, respectively, the long-term throughputs achieved by this and any other scheduling algorithm.

B. Optimization for Conventionally Coded System

As a baseline comparison for a system with DPC, we consider a system with conventional single user coding without known-interference cancellation where, for a given user, the transmissions to all other users can cause interference. The achievable rate region under conventional coding is shown in (18), at the bottom of the page. Because the extreme points of this general capacity region are difficult to determine, we restrict our transmission so that $\mathbf{H}_k \mathbf{Q}_j \mathbf{H}_k^\dagger$ is the zero matrix for all $j \neq k$, where the covariance of user k is $\mathbf{Q}_k = E[\mathbf{x}_k \mathbf{x}_k^\dagger]$. This can be achieved by assuming a *ZF beamforming* constraint. This technique has been discussed in a number contexts including uplink space-division multiple-access systems [5] and multiuser detection [6].

We first consider the case with M transmit antennas and K users, each with a single receive antenna. Let $\mathcal{S}$ be the subset of K users we transmit to under the ZF constraint and let K_S be the cardinality of $\mathcal{S}$. Let $\mathbf{H}_{(1)}, \mathbf{H}_{(2)}, \ldots, \mathbf{H}_{(K_\mathcal{S})}$ denote the channels of these users. We assume that $\mathcal{S}$ is chosen so that the $K_\mathcal{S} \times M$ matrix $\mathbf{H}(\mathcal{S}) = [\mathbf{H}_{(1)}^T, \mathbf{H}_{(2)}^T, \ldots, \mathbf{H}_{(K_\mathcal{S})}^T]^T$ is full rank. (A necessary condition is that $K_\mathcal{S} \leq M$.) Assume that the encoded data stream for the kth user is weighted by the M-vector $\mathbf{w}_k$, so that the transmitted signal for this user is of the form $\mathbf{x}_k = \mathbf{w}_k u_k$. If we let $\mathbf{w}_k$ be the kth column of the $M \times K_\mathcal{S}$ matrix $\mathbf{H}^\dagger(\mathcal{S})(\mathbf{H}(\mathcal{S})\mathbf{H}^\dagger(\mathcal{S}))^{-1}$ then the received signal by the kth user is simply u_k corrupted by additive Gaussian noise but with no interference from the other users' transmissions. Therefore, the throughput for each user is maximized by letting u_k be independent Gaussian signals. If we define $E[|u_k|^2] = \alpha_k^2$, then

$$\mathbf{H}_k \mathbf{Q}_j \mathbf{H}_k^\dagger = \begin{cases} 0, & j \neq k \\ \alpha_k^2, & j = k \end{cases}$$

For a given subset of users $\mathcal{S}$, the rate region in (18) can then be written as

$$\mathcal{R}_{\mathrm{BC}}^{\mathrm{CC-ZF}}(\mathcal{S}) = \bigcup_{\sum_{k \in S} \alpha_k^2 \beta_k \leq P} \left\{ \mathbf{R} : R_k \leq \log\det\left[1 + \alpha_k^2\right] \right\} \tag{19}$$

where β_k is the kth diagonal element of $(\mathbf{H}(\mathcal{S})\mathbf{H}^\dagger(\mathcal{S}))^{-1}$. We wish to maximize the weighted rate sum according to the weights in (16)

$$\max_{R_k \in \mathcal{R}_{\mathrm{BC}}^{\mathrm{CC-ZF}}(\mathcal{S})} \sum_{k \in \mathcal{S}} \mu_k R_k$$

where μ_k is the kth user's weight and the rates $R_k, k \in \mathcal{S}$ belong to the restricted rate region $\mathcal{R}_{\mathrm{BC}}^{\mathrm{CC-ZF}}$. The solution to this optimization problem is the weighted waterfilling solution given by

$$\sum_{k \in \mathcal{S}} \mu_k \log\left(\frac{\mu_k V P}{\beta_k}\right)_+$$

where V is the solution to $\sum_{k \in \mathcal{S}} (\mu_k V - \beta_k/P)_+ = 1$ and where $[x]_+ = \max\{0, x\}$. For a given transmission interval, we wish to maximize the weighted rate sum in (16) over all possible subsets $\mathcal{S}$. A brute-force search would consider all subsets of the

$$\mathcal{R}_{\mathrm{BC}}^{\mathrm{CC}} = \bigcup_{\substack{\mathbf{Q}_1, \ldots, \mathbf{Q}_K \\ \sum_{i=1}^{K} tr(\mathbf{Q}_i) \leq P}} \left\{ (R_1, \ldots, R_K) : R_k \leq \log \frac{\left| \mathbf{I} + \sum_{j=1}^{K} \mathbf{H}_k \mathbf{Q}_j \mathbf{H}_k^\dagger \right|}{\left| \mathbf{I} + \sum_{j=1, j\neq k}^{K} \mathbf{H}_k \mathbf{Q}_j \mathbf{H}_k^\dagger \right|} \right\} \tag{18}$$

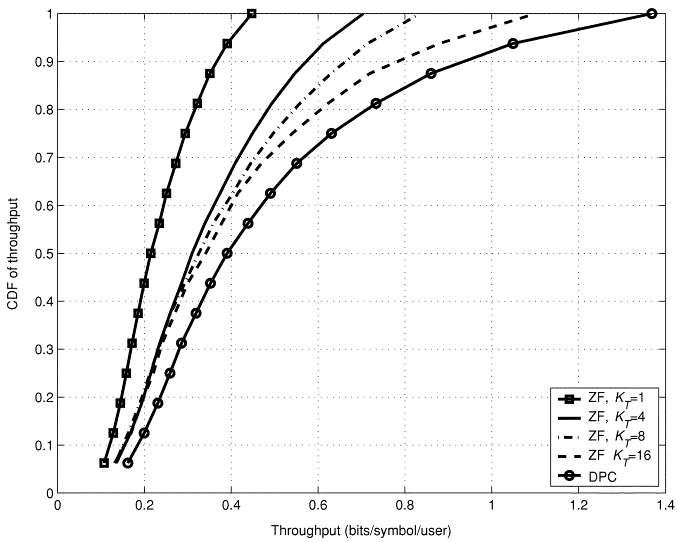

Fig. 6. CDF of throughput for $N = 1$.

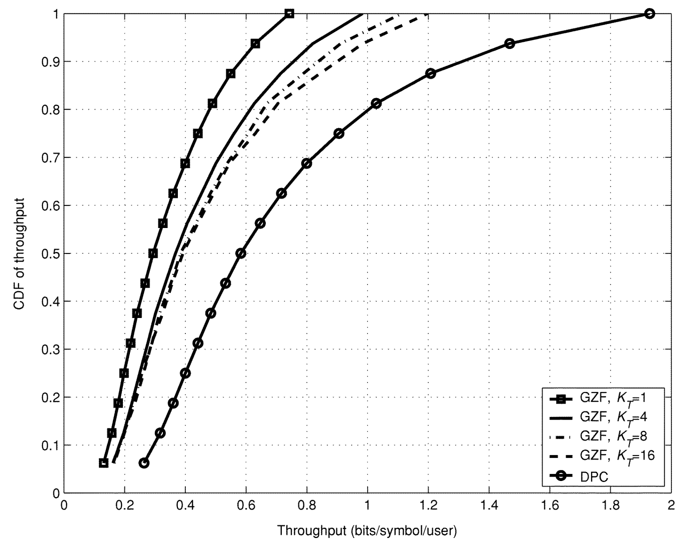

Fig. 7. CDF of throughput for $N = 2$.

K users with cardinality equal or less than M. Even for moderate K and M, the size of the search may be quite large. To reduce the number of subsets for consideration, we can restrict our search to a set of K_T users who achieve the highest single-user weighted rate. In other words, for a given user, we determine its weighted rate $\mu_k R_k$ as if the base were transmitting to this user only. We let $\mathcal{T}$ be the restricted set which contains those K_T users with the highest single-user weighted rate. The loss in optimality in limiting our search to this set $\mathcal{T}$ depends on the size of and the correlation of channels among those users in $\mathcal{T}$.

For multiple receive antennas, we use a generalization of ZF beamforming based on group multiuser detection [22]. This technique is described in the Appendix.

C. Simulation Results

We evaluate the system throughput for a scheduled downlink packet system using either DPC or conventional coding for $M = 4$ transmitters and $K = 16$ users, each with $N = 1$ or two antennas.

We follow the channel model in (1) with total power P set equal to one and the noise variance $\sigma_k = 1/\rho_k$ where ρ_k represents the average SINR at user k's location when all of the base station power is assigned to this user. In the simulations, this average SINR is assigned independently to all the users according to the distribution in Fig. 2 as in the circuit data capacity evaluation in Section IV. Independent realizations of the matrices $\mathbf{H}_k,\ 1 \leq k \leq K$ are generated for each scheduling interval and we assume that there is no delay in obtaining perfect knowledge of $\mathbf{H}_k$ at the transmitter. 1 bps/Hz. We run each simulation over 10 000 transmission intervals. On each transmission interval, an independent realization of the channel matrix $\mathbf{H}_k$ is generated for each of the users. At the end of the simulation, the average rate $\bar{R}_l$ for each user is noted and these values are plotted as a cumulative distribution in Figs. 6 and 7.

Fig. 6 shows the cumulative distribution function (CDF) of the throughput for a system with $M = 4$ transmitters and $K = 16$ users, each with $N = 1$ antennas. We compare the performance of a system with DPC with a conventional coding system under the ZF constraint. We let the size of the restricted set $\mathcal{T}$ be $K_T = 1$, 4, 8,16. When $K_T = 1$, the base transmits to the single user with the highest weighted rate on each transmission interval. Under the assumption that the weight is the reciprocal of the average rate, this case is equivalent to a conventional packet data system but with multiple antennas. The performance improves as K_T increases, showing the benefits of transmitting to more than one user at a time. For $K_T = 4$, the base transmits to $K_S = 4$ users simultaneously. When $K_T = 8$ or 16, $K_S = 4$ since $K_S \leq M$. The performance improvement in going from $K_T = 4$ to 8 occurs because the best set S when $K_T = 8$ does not necessarily correspond to the best four users ranked according to weighted single-user throughput. When $K_T = 16$, the search for the best set S occurs over all users, regardless of their weighted single-user throughput. DPC increases the throughput by about a factor of two to three over the conventional system. A system with conventional coding and ZF provides a significant fraction of the dirty paper gains. An initial investigation on the practical implementation of dirty paper encoding and decoding is given in [25]. In practice, transmission can be accomplished using sign-bit shaping [26], precoding, and vector quantization, however, the implementation of the detection and decoding will be very challenging because of the inherent rate loss induced by the vector quantizer. Therefore, with four transmit antennas and single antenna receivers, conventional coding is a competitive alternative to DPC if the implementation complexity is considered.

Fig. 7 shows the CDF of throughput for group ZF when $N = 2$ antennas are used at each mobile. Compared with $N = 1$, the overall throughputs increase by 50%. The gains of DPC over a conventionally coded system with $K_T = 1$ are still about a factor of two to three. However, the improvement with larger K_T is not as significant as in the single receive antenna case.

VI. Summary

We evaluated the capacity of the downlink of a cellular system without intercell coordination when using known-interference cancellation. Optimization algorithms required for the capacity evaluation were presented. These algorithms can be used to

$$\mathbf{R}_{\mathrm{BC}}^{\mathrm{CC-ZF}}(\mathcal{S}) = \bigcup_{\alpha_k : \sum_{k\in S} \alpha_k^2 tr[\mathbf{V}_k^\dagger \mathbf{G}_k \mathbf{V}_k] \le P} \left\{ \mathcal{R} : R_k \le \log\left(\det\left[1 + \alpha_k^2 \mathbf{S}_k^2\right]\right)\right\}$$

determine the optimal signal covariance matrices for achieving the capacity. Known-interference cancellation using DPC showed gains of about 50% for the circuit data case and a factor of two to three for the packet data cases considered when compared with orthogonal transmission schemes such as TDMA. However, when space-division multiplexing is employed using ZF beamforming along with TDMA to transmit to multiple users in each slot the DPC gains are reduced as demonstrated in Figs. 6 and 7. Further, we found that the 10% outage level the FCFS order was nearly as good as the optimum order among the users for cancelling known interference. Practical coding schemes that incorporate known-interference cancellation have to be developed to realize these gains in practice.

Intercell coordination schemes can be considered where multiple base stations jointly transmit to the users in the coverage area. The optimization algorithms presented in this paper can be applied directly to this case provided a sum power constraint across the base stations is considered. These algorithms can also be adapted to the case when per base power constraints are considered along the lines in [14]. Intercell coordination, however, requires time and phase synchronization of signals transmitted from the multiple base stations which may be difficult to achieve in practice.

Appendix
Group Zero-Forcing Beamforming for Multiple Receive Antennas

Group zero-forcing (GZF) beamforming is a generalization of ZF beamforming based on group multiuser detection [22]. Under GZF, each user has $N > 1$ antennas and receives one data stream per antenna. The GZF constraint weights the transmissions so that the N streams for a given user are orthogonal to the other users' transmissions, however those N streams are not constrained to be mutually orthogonal. Under this GZF restriction, the transmission of each user can be weighted to achieve the multiple-antenna capacity with known channel at the transmitter. As before, let $\mathcal{S}$ be the subset of K users we transmit to and let $K_{\mathcal{S}}$ be the cardinality of $\mathcal{S}$. Let $\mathbf{H}(\mathcal{S})$ denote the $K_{\mathcal{S}}N$-by- M channel matrix obtained by stacking the $N \times M$ channel matrices of the $K_{\mathcal{S}}$ users. Assume $\mathcal{S}$ is chosen so that $\mathbf{H}(\mathcal{S})\mathbf{H}^\dagger(\mathcal{S})$ is full rank. Define $\mathbf{G}_k,\ k = 1\dots K$, to be the matrix inverse of the kth block diagonal of size $N \times N$ of the matrix $(\mathbf{H}(\mathcal{S})\mathbf{H}^\dagger(\mathcal{S}))^{-1}$. Let the singular value decomposition of $\mathbf{G}_k$ be $\mathbf{G}_k = \mathbf{U}_k\mathbf{S}_k\mathbf{V}_k^\dagger$, where $\mathbf{U}_k$ and $\mathbf{V}_k$ are $N\times N$ unitary matrices and $\mathbf{S}_k$ is the $N \times N$ diagonal matrix of eigenvalues. Let the transmitted signal for the kth user be $\mathbf{x}_k = \mathbf{W}_k\mathbf{u}_k$ where $\mathbf{W}_k$ is a $M \times N$ matrix of weights applied to the N-dimensional vector Gaussian signal $\mathbf{u}_k$ with independent components and covariance $E[\mathbf{u}_k\mathbf{u}_k^\dagger] = \mathbf{I}\alpha_k^2$. Define $\tilde{\mathbf{H}}_k$ to be the kth group of N columns of $\mathbf{H}^\dagger(\mathcal{S})(\mathbf{H}(\mathcal{S})\mathbf{H}^\dagger(\mathcal{S}))^{-1}$ so that $\mathbf{H}^\dagger(\mathcal{S})(\mathbf{H}(\mathcal{S})\mathbf{H}^\dagger(\mathcal{S}))^{-1} = [\tilde{\mathbf{H}}_1\dots\tilde{\mathbf{H}}_K]$. The GZF weighting is $\mathbf{W}_k = \tilde{\mathbf{H}}_k\mathbf{G}_k\mathbf{V}_k$ and it follows that

$$\mathbf{U}_j^\dagger\mathbf{H}_j\mathbf{Q}_k\mathbf{H}_j^\dagger\mathbf{U}_j = \begin{cases} \mathbf{0}, & j \ne k \\ \alpha_k^2\mathbf{S}_k^2, & j = k \end{cases}.$$

The transmitted power for the kth user is $tr(\mathbf{Q}_k) = \alpha_k^2 tr\left[\mathbf{V}_k^\dagger\mathbf{G}_k\mathbf{V}_k\right]$. Then, from (18) the constrained rate region under GZF can be written as shown in the equation at the top of the page, where we have used the fact that $\det(\mathbf{I} + \mathbf{AB}) = \det(\mathbf{I} + \mathbf{BA})$. For a given subset $\mathcal{S}$, we wish to maximize the weighted rate sum $\max_{R_k\in\mathcal{R}_{\mathrm{BC}}^{\mathrm{CC-GZF}}(S,P)} \sum_{k\in S}\mu_k R_k$, where μ_k is the kth user's weight. As in the single receive antenna case, the solution to this optimization problem is the weighted waterfilling solution and the same strategies in picking a restricted set of K_T users applies.

Acknowledgment

The authors would like to acknowledge numerous helpful discussions with their colleagues at Bell Labs and thank the anonymous referees for their insightful comments and suggestions.

References

[1] G. J. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, pp. 311–335, 1998.
[2] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun. Related Technol.*, vol. 10, no. 6, pp. 585–596, Nov. 1999.
[3] G. Caire and S. Shamai, "Achievable rates in multi-antenna broadcast downlink," presented at the Annu. Allerton Conf. Communications and Control, Allerton, IL, Oct. 2000.
[4] ——, "On the multiple antenna broadcast channel," in *Proc. Asilomar Conf.*, Pacific Grove, CA, Nov. 2001, pp. 1188–1193.
[5] B. Suard, G. Xu, H. Liu, and T. Kailath, "Uplink channel capacity of space-division-multiple-access schemes," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1468–1476, July 1998.
[6] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
[7] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 24, pp. 374–377, May 1978.
[8] P. Viswanath and D. Tse, "Sum capacity of the multiple antenna broadcast channel," in *Proc. IEEE Int. Symp. Information Theory, ISIT 2002*, Lusanne, Switzerland, July 2002, p. 497.
[9] S. Vishwanath, N. Jindal, and A. Goldsmith, "On the capacity of multiple input multiple output broadcast channel," in *Proc. ICC*, May 2002, pp. 1444–1450.
[10] D. G. Luenberger, *Introduction to Linear and Non-Linear Programming*. Reading, MA: Addison-Wesley, 1973.
[11] K. Kumaran, private communication.
[12] A. Jalali, R. Padovani, and R. Pankaj, "Data throughput of CDMA-HDR: A high efficiency, high data rate personal wireless system," in *Proc. IEEE Vehicular Technology Conf.*, Tokyo, Japan, May 2000, pp. 1854–1858.
[13] S. A. Jafar and A. Goldsmith, "Optimal power allocation for multiuser multicellular multiple antenna systems," in *Proc. 2002 IEEE Int. Symp. Information Theory*, Lausanne, Switzerland, 2002, p. 50.
[14] S. A. Jafar and G. J. Foschini, Optimal multiuser multicellular multiple antenna systems, preprint, Aug. 2001.

[15] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.

[16] P. Bender *et al.*, "CDMA/HDR: A bandwidth-efficient high-speed wireless data service for nomadic users," *IEEE Commun. Mag.*, vol. 38, pp. 70–77, July 2000.

[17] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[18] D. Tse and S. Hanly, "Multiaccess fading channels-Part I: Polymatriod structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2796–2815, Nov. 1998.

[19] B. Rimoldi and R. Urbanke, "A rate splitting approach to the Gaussian multiple-access channel," *IEEE Trans. Inform. Theory*, vol. 42, pp. 364–396, Mar. 1996.

[20] M. A. Andrews *et al.*, "Providing quality of service over a shared wireless link," *IEEE Commun. Mag.*, vol. 39, no. 2, pp. 150–154, Feb. 2001.

[21] D. Tse, "Forward-link multi-user diversity through rate adaptation and scheduling," Bell Labs Presentation, Murray Hill, NJ, 1999.

[22] M. Varanasi, "Group detection for synchronous CDMA systems," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1083–1096, July 1995.

[23] W. Yu, A. Sutivong, D. Julian, T. Cover, and M. Chiang, "Writing on colored paper," in *Proc. Int. Symp. Information Theory*, Washington, June 2001, p. 302.

[24] A. Cohen and A. Lapidoth, "The Gaussian watermarking game," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1639–1667, June 2002.

[25] W. Yu and J. Cioffi, "Trellis precoding for the broadcast channel," in *Proc. IEEE Global Telecommunications Conf.*, San Antonio, TX, Nov. 2001, pp. 1344–1348.

[26] G. D. Forney, Jr., "Trellis shaping," *IEEE Trans. Inform. Theory*, vol. 38, pp. 281–300, Mar. 1992.

Harish Viswanathan (S'93–M'97) was born in Trichy, India, in 1971. He received the B.Tech. degree from the Department of Electrical Engineering, Indian Institute of Technology, Chennai, in 1992 and the M.S. and Ph.D. degrees from the School of Electrical Engineering, Cornell University, Ithaca, NY, in 1995 and 1997, respectively.

He is currently with Lucent Technologies, Bell Labs, Murray Hill, NJ. His research interests include information theory, communication theory, wireless networks, and signal processing.

Sivarama Venkatesan (S'95–M'97) received the B.Tech. degree in electronics and communication engineering from the Indian Institute of Technology, Chennai, in 1991, and the M.S. and Ph.D. degrees in electrical engineering from Cornell University, Ithaca, NY, in 1994 and 1998, respectively.

He is currently a Member of Technical Staff in the Wireless Communications Research Department of Bell Laboratories, Lucent Technologies, Holmdel, NJ. His research interests include information theory and multiple-antenna wireless communication systems.

Howard Huang (S'90–M'95) was born in Houston, TX, in 1969. He received the B.S.E.E. degree from Rice University, Houston, in 1991 and the Ph.D. degree from Princeton University, Princeton, NJ, in 1996.

He is currently a Member of Technical Staff in the Wireless Research Lab at Bell Labs, Lucent Technologies, Holmdel, NJ. His interests include multiuser detection, multiple-antenna technologies and their application in third-generation mobile cellular systems.