Efficient Scheduling of Multi-User Multi-Antenna Systems

Krishna P Jagannathan†, Sem Borst*, Phil Whiting*, Eytan Modiano†

*Bell Laboratories, Lucent Technologies, P.O. Box 94079, Murray Hill, NJ 07974

†Laboratory for Information and Decision Systems,
Massachusetts Institute of Technology, Cambridge, MA 02139

Abstract

The capacity region of the Gaussian multi-antenna broadcast channel was characterized recently in [20]. It was shown that a scheme based on Dirty Paper Coding [2] achieves the full capacity region, when the transmitter has perfect channel state information. Practically however, this scheme potentially involves considerable amounts of feedback and complex transmission algorithms. In this paper, we study a Gaussian broadcast channel with two transmit antennas and statistically identical users. We obtain an exact asymptotic characterization of the gap between the full sum capacity and the rate that can be achieved by transmitting to a suitably selected pair of users. Specifically, we consider some simple schemes for user-pair selection that take into account the channel norms as well as the phase information. We conclude that one of these schemes, which picks the strongest user and selects a second user to form the best pair, is asymptotically optimal, while also being attractive in terms of feedback and operational complexity. Numerical experiments show that the asymptotic results tend to be remarkably accurate, and that the proposed scheme significantly outperforms a beam-forming strategy for a typical number of users.

I. INTRODUCTION

The multi-antenna broadcast channel (BC) has been the subject of much research interest recently, owing primarily to the impressive capacity benefits that these systems can potentially offer. The ever-increasing pressure to make the best possible use of the available wireless spectrum has been the driving force behind this recent interest. The sum capacity for the Gaussian multi-antenna BC was first obtained for the case of two users with a single receive antenna by Caire & Shamai [1]. Subsequently, Viswanath & Tse [16] and Vishwanath *et al.* [15] extended the result for the sum capacity to an arbitrary number of users and receive antennas by exploiting a powerful duality relation with the multi-access channel, which was further explored in Jindal *et al.* [9]. Recently, Weingarten *et al.* [20] provided a characterization of the entire capacity region. They showed that a scheme based on Dirty Paper Coding (DPC) [2] achieves the full capacity region of the multi-antenna BC.

The above capacity results rely on the assumption that perfect channel state information is available at the transmitter. However, the amount of overhead involved in feeding back the channel information may be prohibitive (especially when the number of users is large), or simply may not be worth the gain in rate. In addition, DPC may be quite difficult to implement in practice.

Motivated by the above issues, extensive efforts have been made to devise practical transmission and coding schemes and find ways to reduce the amount of channel feedback information required. Hochwald *et al.* [3], [4] describe an algorithm based on channel inversion and sphere encoding, and show that it closely approaches the capacity while being simpler to operate than DPC. Jindal [6] considers a multi-antenna BC with limited channel feedback information, and shows that the full capacity gain at high SNR values is achievable as long as the number of feedback bits grows linearly with the SNR (in dB).

In the case of a single receive antenna, it is known that the sum capacity grows with the SNR at rate $\min\{M, K\}$ [5], [7], i.e., the minimum of the number of transmit antennas and the number of users. In other words, multiple transmit antennas provide potentially huge capacity gains, but it is necessary that at least M users are served simultaneously in order to reap the full benefits. Transmitting to less than M users falls short of the maximum capacity as it fails to fully exploit the available degrees of freedom. Transmitting to more than M users may be necessary to achieve the maximum capacity in general, but the results in [5], [7] indicate that transmitting to a suitably selected subset of exactly M near-orthogonal users is close to optimal. When the total number of users to choose from is sufficiently large, such a subset exists with high probability, see [13], [14] for a rigorous asymptotic characterization.

Clearly, the above principle allows for a reduction of the amount of channel feedback and coding complexity. In particular, it suggests the use of beam-forming schemes which construct M (random) orthogonal beams and serve the users with the largest channel gains on each of them with equal power. Transmission schemes along these lines are presented in [19] and [12].

In the present paper we focus on the case of two transmit antennas and statistically identical users, and derive an exact asymptotic characterization of the gap between the full sum capacity and the rate that can be achieved by transmitting to a suitably selected pair of users. In particular, we consider a scheme that picks the user with the largest channel gain, and selects a second user from the next $L-1$ strongest ones to form the best pair, taking channel angles into account as well. We prove that the rate gap converges to $1/(L-1)$ when the total number of users K tends to infinity. Allowing L to increase with K, we conclude that the gap asymptotically vanishes, and that the sum capacity is achievable by transmitting to a properly chosen pair of users. Numerical results show that the asymptotics tend to be remarkably accurate, even for a relatively moderate number of users. The fact that the rate gap decays as $1/(L-1)$ also suggests that a modest value of L is adequate for most practical purposes. The above results have significant implications for the design of channel feedback mechanisms and transmission techniques.

The remainder of the paper is organized as follows. In Section II we present a detailed model description and review some known results for the sum rate capacity of the Gaussian multi-antenna BC. We present our main asymptotic results in Section III. In Section IV, we present some numerical experiments to show that our asymptotic results are quite accurate, even for a moderate number of users. In Section V we point out some directions for future work.

II. System Model and Preliminary Results

A. Model

We consider a broadcast channel (BC) with $M > 1$ transmit antennas and K independent receivers each with a single antenna as schematically represented in Figure 1(a). This provides a model for the downlink transmission from a single base station with $M > 1$ transmit antennas to K independent users each with a single receive antenna.

Let $\boldsymbol{x} \in \mathbb{C}^{M\times 1}$ be the transmitted vector signal and let $\boldsymbol{h}_k \in \mathbb{C}^{1\times M}$ be the channel-gain vector of the k-th receiver. Denote by $\boldsymbol{H} = [\boldsymbol{h}_1^\dagger \boldsymbol{h}_2^\dagger \cdots \boldsymbol{h}_K^\dagger]^\dagger$ the concatenated channel matrix of all K receivers. For now, the matrix $\boldsymbol{H}$ is arbitrary but assumed to be fixed. We further assume that the transmitter has perfect channel state information, i.e., exact knowledge of the matrix $\boldsymbol{H}$. The circularly symmetric complex Gaussian noise at the k-th receiver is $n_k \in \mathbb{C}$ where $n_k \sim \mathcal{CN}(0,1)$. Thus the received signal at the k-th receiver

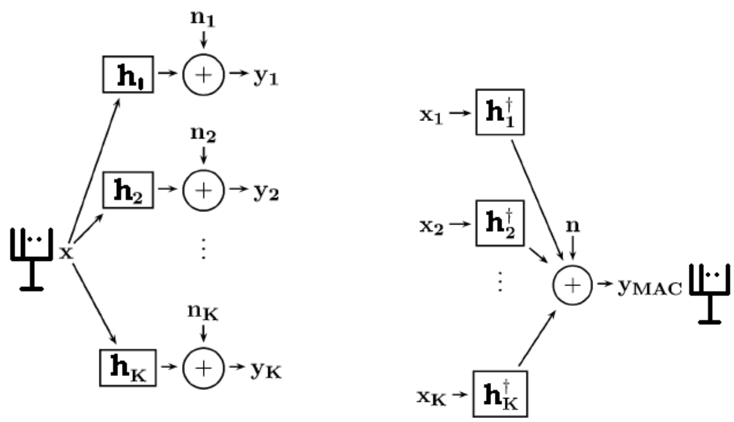

Fig. 1. (a) Multi-antenna broadcast channel (b) Multiple access channel

is $y_k = \boldsymbol{h}_k \boldsymbol{x}_k + n_k$. The covariance matrix of the transmitted signal is $\Sigma_x = \mathbb{E}\left[\boldsymbol{x}\boldsymbol{x}^\dagger\right]$. The transmitter is subject to a power constraint P, which means $Tr(\Sigma_x) \leq P$.

B. Sum capacity

The sum capacity is a key metric of interest for the BC as it measures the maximum achievable system throughput. Since it only considers the aggregate throughput, it does not reflect potential fairness issues that arise when users with widely disparate channel characteristics obtain vastly different throughput portions. In the present paper, however, we focus on the case of statistically identical users, which by symmetry will obtain equal long-term throughput shares.

In case of a single transmit antenna, the sum capacity is simply equal to the largest single-user capacity in the system [11], i.e., the aggregate throughput is maximized by transmitting only to the user with the largest channel gain. However, this is not true when there are multiple transmit antennas. In that case, the sum rate capacity is achieved by using DPC techniques to simultaneously transmit to several users. The expression for the sum capacity of the BC (denoted by $\mathcal{C}_{BC}(\boldsymbol{H}, P)$) is derived in [15], [16] as an optimization problem. However, the objective function is an unwieldy non-concave function of the input covariance matrices, and therefore hard to deal with numerically as well as analytically. Fortunately, a duality was shown to exist [15] between the BC and the Gaussian multiple-access channel (MAC) (Figure 1(b)) with a sum-power constraint P. That is, the dual MAC which is formed by reversing the roles of transmitters and receivers has the same capacity region as the BC.

Because of the duality, the sum capacity of the BC can be written in terms of the sum capacity of the dual MAC as

$$\mathcal{C}_{BC}(\boldsymbol{H},P) = \max_{P_k \geq 0, \sum_{k=1}^{K} P_k \leq P} \log \left(\det \left(I_M + \sum_{k=1}^{K} P_k \boldsymbol{h}_k^\dagger \boldsymbol{h}_k \right) \right), \tag{1}$$

where P_k denotes the power allocated to the k-th receiver. Note that the objective function in (1) is indeed concave in the values of the P_k's. Specialized algorithms for calculating the BC sum capacity have been developed in [8]. The next simple yet extremely useful upper bound is established in [7, Theorem 1]:

$$\mathcal{C}_{BC}(\boldsymbol{H},P) \leq M \log \left(1 + \frac{P}{M} ||\boldsymbol{h}_{(1)}||^2 \right). \tag{2}$$

Here $\boldsymbol{h}_{(i)}$ denotes the channel vector of the receiver with the i-th largest norm, i.e., $||\boldsymbol{h}_{(1)}||^2 \geq ||\boldsymbol{h}_{(2)}||^2 \geq \cdots \geq ||\boldsymbol{h}_{(K)}||^2$. In the case of $M = 2$ transmit antennas, we obtain a simple lower bound on the sum capacity by scheduling any two users i and j with equal power:

$$\mathcal{C}_{BC}(\boldsymbol{H},P) \geq C(\boldsymbol{h}_i, \boldsymbol{h}_j, P) := \log \left(\det \left(I_M + \frac{P}{2} (\boldsymbol{h}_i^\dagger \boldsymbol{h}_i + \boldsymbol{h}_j^\dagger \boldsymbol{h}_j) \right) \right). \tag{3}$$

III. Asymptotics

In this section we will rigorously show that, as the number of users grows large, the sum capacity can be closely approached by transmitting to a suitably selected subset of two users and allocating equal power to each of them. So far, we have assumed the channel vectors to be arbitrary but fixed. In order to derive meaningful asymptotic results, we henceforth assume the channel vectors to be random, and primarily consider the *expected* sum capacity. Specifically, we assume that the various components of the channel vector of a user are independent and identically distributed according to $\mathcal{CN}(0,1)$, which corresponds to independent Rayleigh fading. It can further be shown that this assumption is not essential, and with minor modifications the asymptotic results extend to more general distributions of the channel vectors, as long as the phases are independent.

We will consider three heuristic selection schemes. Scheme I picks two arbitrary users among the L strongest ones. Scheme II selects an arbitrary user among the L strongest ones, and a second one from that group to form the best pair with it. Scheme III picks the best pair among the L strongest users. Note that scheme II dominates scheme I and that scheme III in turn dominates scheme II, and that all three schemes coincide when $L = 2$.

The following theorem considers ratio-asymptotics for scheme I; specifically, it shows that the ratio of the rate obtained by using scheme I to the upper bound in (2) converges to unity as K becomes large.

Theorem 3.1: For any fixed value of $L \geq 2$,

$$\lim_{K\to\infty} \frac{\mathbb{E}\left[C(\boldsymbol{h}_{(i)}, \boldsymbol{h}_{(j)}, P)\right]}{\mathbb{E}\left[2\log\left(1+\frac{P}{2}||\boldsymbol{h}_{(1)}||^2\right)\right]} = 1, \tag{4}$$

for all $i, j \leq L$, $i \neq j$.

The above results shows that scheme I (and consequently all the schemes) are asymptotically optimal in the ratio sense. However, it is possible that there is a performance gap between the achieved sum rate and the capacity limit. We now proceed to prove more refined asymptotic results regarding the *difference* between the sum rate and the capacity limit.

Theorem 3.2: For any fixed value of $L \geq 2$, $l \leq L$,

$$\mathbb{E}\left[\mathcal{C}_{BC}(\boldsymbol{H}, P)\right] - \mathbb{E}\left[\max_{k=1,\ldots,L,\ k\neq l} C(\boldsymbol{h}_{(l)}, \boldsymbol{h}_{(k)}, P)\right] \to \frac{1}{L-1}$$

as $K \to \infty$.

The above theorem shows that the asymptotic performance gap of scheme II decays as $1/(L-1)$, which suggests that a relatively moderate value of L may be adequate for most practical purposes.

The next corollaries follow as direct consequences from Theorem 3.2.

Corollary 3.1: For any fixed value l and sequence $L(K)$ with $\lim_{K\to\infty} L(K) = \infty$,

$$\mathbb{E}\left[\mathcal{C}_{BC}(\boldsymbol{H}, P)\right] - \mathbb{E}\left[\max_{k=1,\ldots,L(K),\ k\neq l} C(\boldsymbol{h}_{(l)}, \boldsymbol{h}_{(k)}, P)\right] \to 0$$

as $K \to \infty$.

The above corollary shows that scheme II is asymptotically optimal when sufficiently many users are considered. Because of the dominance relationship, it immediately follows that scheme III is asymptotically optimal as well.

Corollary 3.2:

$$\mathbb{E}\left[\mathcal{C}_{BC}(\boldsymbol{h}_1, \ldots, \boldsymbol{h}_K, P)\right] - \mathbb{E}\left[C(\boldsymbol{h}_{(1)}, \boldsymbol{h}_{(2)}, P)\right] \to 1$$

as $K \to \infty$.

Notice that this corollary corresponds to a special case of scheme I. We see that simply selecting the two strongest users leaves a performance gap of 1.

In conclusion, the above results show that scheme II is asymptotically optimal in the sense that the absolute gap with the sum capacity vanishes to zero provided $L(K) \to \infty$ as $K \to \infty$. Thus, transmitting to a suitably selected pair of users is asymptotically optimal, where one of the users may in fact be chosen arbitrarily. The gain from considering all pairs of users (scheme III) is asymptotically negligible. However, picking an arbitrary pair of users is not optimal (scheme I), not even the two strongest ones ($L = 2$).

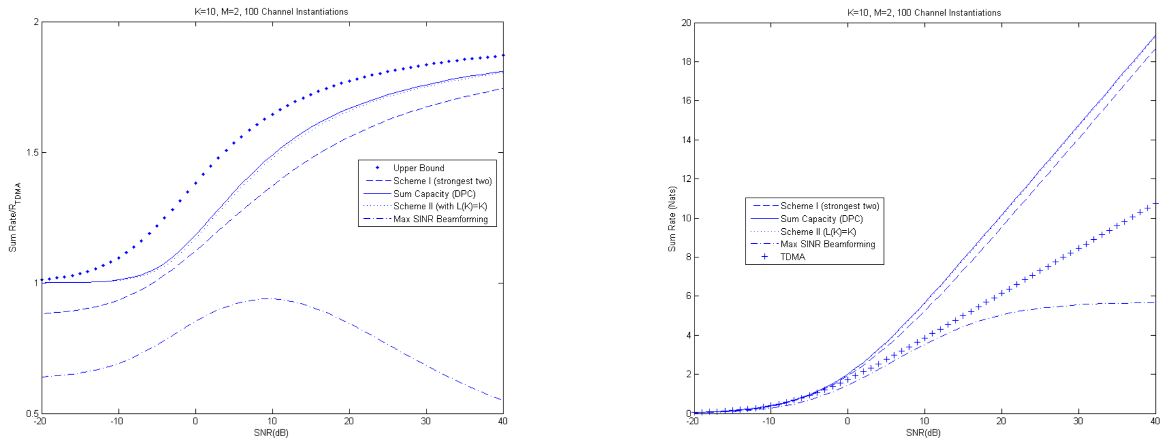

Fig. 2. (a) Comparison of various schemes with TDMA (b) Absolute rates for various schemes

IV. Numerical results

In Figure 2(a), we plot the sum rate obtained by the various schemes, as relative to the TDMA sum rate versus the SNR (in dB), i.e., the value of P. The absolute sum rate (in Nats) is graphed as a function of SNR in Figure 2(b). We consider a system with two transmit antennas, $K = 10$ users, and average over 100 channel realizations. The solid line corresponds to the optimal DPC scheme. The dotted line just underneath the solid line corresponds to a special case of scheme II. Specifically, we schedule the user with the largest channel norm, and the second user to form the best possible pair with it (i.e., $L(K) = K$). It is clear that even for this moderate value of K, scheme II performs very well, in addition to being asymptotically optimal. The broken line corresponds to a special case of scheme I, where the two strongest users are scheduled with equal power. It is clear that scheme II dominates scheme I quite significantly. It is also interesting to note that the upper bound in (2) (shown in the figure with diamonds), although asymptotically tight, is quite loose for practical values of K and SNR. We finally observe that TDMA is optimal in the very low SNR regime.

We also make a comparison with a beam-forming (BF) scheme along the lines described in [19] and [12]. Specifically the BF scheme selects two users $k_m^* := \arg\max_{k=1,\ldots,K} SNR_{k,m}$, $m = 1, 2$, with

$$SNR_{k,m} := \frac{|< \boldsymbol{h}_k, \phi_m >|^2}{2/P + |< \boldsymbol{h}_k, \phi_{3-m} >|^2},$$

and ϕ_1, ϕ_2 two orthonormal vectors. The expected rate thus obtained is

$$R_{BF} := \mathbb{E}\left[\log(1 + SNR_{k_1^*,1}) + \log(1 + SNR_{k_2^*,2})\right]$$

(ignoring potential complications when $k_1^* = k_2^*$), and is plotted as the lower curve in Figure 2(a). The lower bound $C(\boldsymbol{h}_{k_1^*}, \boldsymbol{h}_{k_2^*})$ (not shown in Figure 2(a)) indicates that the BF scheme does well in terms of selecting a pair of users. However, the lower curve shows that transmitting along the pre-determined beams without using actual phase information performs poorly. In fact, it does consistently worse than a simple TDMA strategy for this particular case with 10 users, even though it is known to be asymptotically optimal in the limit of a large number of users.

Note that as $P \downarrow 0$, we have

$$R_{BF} \approx \frac{P}{2}\mathbb{E}\left[|<\boldsymbol{h}_{k_1^*}, \phi_1>|^2 + |<\boldsymbol{h}_{k_2^*}, \phi_2>|^2\right] = P\mathbb{E}\left[|<\boldsymbol{h}_{k_1^*}, \phi_1>|^2\right] \leq P\mathbb{E}\left[||\boldsymbol{h}_{(1)}||^2\right] \approx R_{TDMA}.$$

Also, as $P \to \infty$, we find that R_{BF} approaches

$$\mathbb{E}\left[\log\left(1 + \frac{|<\boldsymbol{h}_{k_1^*}, \phi_1>|^2}{|<\boldsymbol{h}_{k_1^*}, \phi_2>|^2}\right) + \log\left(1 + \frac{|<\boldsymbol{h}_{k_2^*}, \phi_2>|^2}{|<\boldsymbol{h}_{k_2^*}, \phi_1>|^2}\right)\right] = 2\mathbb{E}\left[\log\left(1 + \frac{|<\boldsymbol{h}_{k_1^*}, \phi_1>|^2}{|<\boldsymbol{h}_{k_1^*}, \phi_2>|^2}\right)\right]$$

i.e., it saturates at a finite value, whereas R_{TDMA} grows without bound, albeit slowly.

V. Future Work

Several natural topics for further investigation present themselves. First of all, the above results have evident implications for the design of channel feedback mechanisms and transmission techniques. It would be interesting to address these aspects in more detail. A second major avenue that would be worth pursuing is to generalize the results to an arbitrary number of transmit antennas, and possibly several receive antennas. A further challenging subject that is under ongoing investigation, concerns the extension to a scenario with heterogeneous users and maximizing a *weighted* sum rate or achieving an optimal fair operating point of the capacity region. Some interesting results along the latter lines may be found in [10], [18].

Acknowledgment

The authors wish to express their gratitude to Gerhard Kramer and Hanan Weingarten for many helpful suggestions and interesting discussions. The second author also gratefully acknowledges insightful comments of Nihar Jindal and Harish Viswanathan.

References

[1] G. Caire, S. Shamai (2003). On the achievable throughput of a multiantenna Gaussian broadcast channel. *IEEE Trans. Inf. Theory* **49** (7), 1691–1706.

[2] M. Costa (1983). Writing on dirty paper. *IEEE Trans. Inf. Theory* **29** (3), 439–441.

[3] B.M. Hochwald, C.B. Peel, A.L. Swindlehurst (2005). A vector-perturbation technique for near-capacity multiantenna multiuser communication – part I: channel inversion and regularization. *IEEE Trans. Commun.* **53** (1), 195–202.

[4] B.M. Hochwald, C.B. Peel, A.L. Swindlehurst (2005). A vector-perturbation technique for near-capacity multiantenna multiuser communication – part II: perturbation. *IEEE Trans. Commun.* **53** (3), 537–544.

[5] N. Jindal (2005). A high SNR analysis of MIMO broadcast channels. Preprint.

[6] N. Jindal (2005). MIMO broadcast channels with finite rate feedback. Preprint.

[7] N. Jindal, A. Goldsmith (2005). Dirty-paper coding versus TDMA for MIMO broadcast channels. *IEEE Trans. Inf. Theory* **51** (5), 1783–1794.

[8] N. Jindal, W. Rhee, S. Vishwanath, S.A. Jafar, A. Goldsmith (2005). Sum power iterative water-filling for multi-antenna Gaussian broadcast channels. *IEEE Trans. Inf. Theory* **51** (4), 1570–1580.

[9] N. Jindal, S. Vishwanath, A. Goldsmith (2004). On the duality of Gaussian multiple-access and broadcast channels. *IEEE Trans. Inf. Theory* **50** (5), 768–783.

[10] J. Lee, N. Jindal (2005). Symmetric capacity of MIMO downlink channels. Preprint.

[11] L. Li, A. Goldsmith (2001). Capacity and optimal resource allocation for fading broadcast channels – Part I: Ergodic capacity. *IEEE Trans. Inf. Theory* **47** (3), 1083–1102.

[12] M. Sharif, B. Hassibi (2005). On the capacity of MIMO broadcast channels with partial side information. *IEEE Trans. Inf. Theory* **51** (2), 506–522.

[13] C. Swannack, E. Uysal-Biyikoglu, G. Wornell (2004). Low-complexity multi-user scheduling for maximizing throughput in the MIMO broadcast channel. In: *Proc. 42nd Annual Allerton Conf. Commun. Control, Comput.*

[14] C. Swannack, E. Uysal-Biyikoglu, G. Wornell (2004). Finding NEMO: near-orthogonal sets for multiplexing and opportunistic scheduling in MIMO broadcast. Preprint.

[15] S. Vishwanath, N. Jindal, A. Goldsmith (2003). Duality, achievable rates and sum-rate capacity of MIMO broadcast channels. *IEEE Trans. Inf. Theory* **49** (10), 2658–2668.

[16] P. Viswanath, D.N.C. Tse (2003). Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality. *IEEE Trans. Inf. Theory* **49** (8), 1912–1921.

[17] H. Viswanathan, S. Venketesan (2005). Asymptotics of sum rate for dirty paper coding and beamforming in multiple-antenna broadcast channels. Preprint.

[18] H. Viswanathan, S. Venketesan, H. Huang (2005). Downlink capacity evaluation of cellular networks with known interference cancellation. *IEEE J. Sel. Areas Commun.* **21** (5), 802–811.

[19] P. Viswanath, D.N.C. Tse, R. Laroia (2002). Opportunistic beamforming using dumb antennas. *IEEE Trans. Inf. Theory* **48** (6), 1277–1294.

[20] H. Weingarten, Y. Steinberg, S. Shamai (2005). The capacity region of the Gaussian MIMO broadcast channel. *IEEE Trans. Inf. Theory*, to appear.