

# DOWNLINK WITH PARTIALLY COOPERATING BASE STATIONS

Moritz Wiese\*, Holger Boche\*†, Igor Bjelaković\*, and Volker Jungnickel†

\*Technische Universität Berlin  
Heinrich-Hertz-Lehrstuhl für Informationstheorie und Theoretische Informationstechnik  
{moritz.wiese, holger.boche, igor.bjelakovic}@mk.tu-berlin.de  
†Heinrich-Hertz-Institut  
volker.jungnickel@hhi.fhg.de

## ABSTRACT

For a downlink scenario with two base stations and one receiving mobile, where the base stations are connected by a noiseless finite-capacity backbone, we determine the optimal cooperation and coding strategy. The performance is not influenced by the degree of channel state information at the mobile, but it increases with increasing channel state information at the base stations. If the link capacities exceed finite thresholds, the system achieves full-cooperation performance. The optimal cooperation protocol is a simple one-shot strategy.

## 1. INTRODUCTION

Fourth generation wireless systems are designed to support high-rate data applications such as internet browsing and video streaming. The traditional cellular design of wireless systems, where only one base station is responsible for each mobile, proves to be insufficient for achieving such a goal, as inter-cell interference is a great obstacle for reliable transmission. In addition, the rise of multiple-antenna systems (MIMO) has inspired researchers to ask whether several base stations could form a *virtual* MIMO system, thus allowing for the transmission of higher rates without requiring additional power or bandwidth. Both of these issues are addressed in systems such as LTE (Long-Term Evolution) Advanced, which incorporate in their standard the connection of several base stations over a high-speed backbone. Much research has been invested into finding how much of what kind of data, like channel state information (CSI) and the modulation type, must be exchanged between the base stations in order to achieve the goals mentioned above. In many cases, it was assumed that the backbone has infinite capacity. Then, the only question that remains is what kind of preprocessing needs to be done for achieving high rates. For example, [1] consider zero forcing and dirty paper coding, assuming a noiseless infinite capacity link between the base stations and full CSI at the transmitters.

On the information-theoretic side, Willems [4, 5] has investigated the following related problem: one is given a downlink channel with two senders and one receiver and with dis-

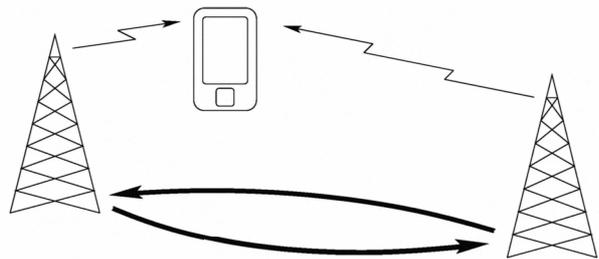


Fig. 1. Two base stations with connecting backbones transmitting to a mobile.

crete alphabets. The channel is known perfectly at all three nodes, and the senders are connected by noiseless links with *finite* capacity. The region of achievable rates can be completely characterized. In this paper, we consider the same problem, but with more realistic channel characteristics. More exactly, our setting has the following features:

- the output of the channel at one time only depends on the corresponding input and not on the input at other times (flat fading);
- the channel remains constant over each transmission block;
- the senders do not have full channel state information (CSIT). Instead, for each sender, there is a partition of the set of possible channel states, and the sender knows which element of his partition the actual channel state belongs to. This includes the case of no CSIT;
- channel state information at the receiver (CSIR) may be arbitrary, anything between full CSIR and no CSIR at all is possible;
- CSI may be asymmetric between all three nodes.

The last point is particularly important in applications. Asymmetric CSIT may occur because the mobile is closer to one base station than to the other.

We obtain a complete characterization of the region of achievable rates. CSIR will prove to be of no influence on

the region. On the other hand, we will see that the amount of CSIT has considerable influence on the capacity region of the channel described above. The exact scaling of the region with CSIT can be read off from the complete characterization in Theorem 1 in Section 2.

We are also able to determine what is the optimal protocol among a class of conferencing protocols defined analogous to [4, 5]. It turns out that this is very simple and that no iterative conferencing is needed – a one-shot strategy is sufficient. Further, we determine how strong the links from one transmitter to the other need to be in order to achieve the full cooperation region. In wireless systems such as LTE Advanced, one will mainly be interested in the sum rate, as usually, the data streams from the two base stations will be the two parts of one aggregate data stream which contains the data intended for the mobile. It only depends on the sum of the capacities of the links between the base stations when the full-cooperation sum rate is achieved, and we will determine this sum as well.

The paper is organized as follows: in Section 2, we present the channel model and give the characterization of the capacity region. In Section 3, we derive an outer bound to the capacity region. We show that this outer bound can be achieved by a simple conferencing protocol in Section 4. A numerical example will be presented in Section 5. Finally, in Section 6, we discuss what we have done and conclude the paper.

## 2. CHANNEL MODEL AND RESULTS

We consider a discrete channel. That means that we have a finite alphabet  $A$  for transmitter 1 and a finite alphabet  $B$  for transmitter 2. The outputs of the channel are given in the finite set  $C$ . Such a channel could arise from a channel with additive Gaussian noise by using finite alphabets like M-QAM and by quantizing the output. The possible channel states are contained in an arbitrary set  $S$ . For each state  $s \in S$ , the channel is given by a stochastic matrix  $H_s$ , whose entries are denoted by  $H_s(c|a, b)$ . If transmitter 1 sends a word  $x \in A^n$  of length  $n$  and transmitter 2 a word  $y \in B^n$  of equal length, then the probability that a word  $z \in C^n$  is received if the channel is in state  $s$  equals  $H_s(z_n|x_n, y_n) \cdot \dots \cdot H_s(z_1|x_1, y_1)$ , which we write as  $H_s^n(z|x, y)$ . This reflects that the channel is flat fading and that it remains constant during the transmission of the word.

CSIT for sender  $\nu$  is given by a function  $t_\nu$  from the set of states  $S$  to some finite set  $T_\nu$ ,  $\nu = 1, 2$ . This gives a partition of  $S$  into finitely many blocks. If the channel is in state  $s \in S$ , sender  $\nu$  knows  $t_\nu(s)$ . Further, in this case, the receiver knows  $r(s)$ , where  $r$  is a function from  $S$  to an arbitrary set  $R$ . Note that if  $S$  is infinite, the senders cannot have full CSIT, but full CSIR is always possible by setting  $R = S$  and  $r(s) = s$ .

### 2.1. Cooperation via Conferencing Protocols

The central feature of our communication model is that the transmitters are allowed to communicate over noiseless finite-capacity links. Denote the capacity of the link from transmitter 1 to transmitter 2 by  $C_1$  and by  $C_2$  the capacity of the link in the other direction. To explain what this means, assume that the transmitters send their messages to the receiver in codewords of blocklength  $n$ . To simplify the description, we introduce the notation  $\bar{\nu} = 1$  if  $\nu = 2$  and  $\bar{\nu} = 2$  if  $\nu = 1$ . Before the transmission of the  $m$ -th codeword (and during the transmission of the  $(m - 1)$ -th codeword), the senders have the opportunity of coordinating their transmission to some extent. The total number of possible messages sender  $\nu$  may receive from sender  $\bar{\nu}$  before encoding may not exceed  $2^{nC_{\bar{\nu}}}$ , so per channel use, at most  $2^{C_{\bar{\nu}}}$  messages can be received.

For the exchange of messages between the transmitters, an iterative conferencing protocol is used. The idea is that in the first step, each sender conveys to the other some information about the message he would like to transmit and about his CSI. For the next conferencing step, each sender again sends some information the other, but this time, it may also depend on what he has learned from the other transmitter in the preceding step. And so on. We assume that  $K$  iterations are done. We now formalize this idea. Let  $V_\nu$  be the Cartesian product of finite sets  $V_{\nu,1}, \dots, V_{\nu,K}$  which fulfill

$$\log|V_{\nu,1}| + \dots + \log|V_{\nu,K}| = \log|V_\nu| \leq nC_\nu. \quad (1)$$

For a positive integer  $m$ , we abbreviate  $\{1, \dots, m\} =: [1, m]$ . For each  $k \in [1, K]$ , define a function  $h_{1,k}$  from  $[1, M_\nu] \times V_{\bar{\nu},1} \times \dots \times V_{\bar{\nu},k-1} \times T_\nu$  into  $V_{\nu,k}$ . Assume that sender  $\nu$  would like to transmit message  $\ell_\nu$  and that his channel knowledge is  $\tau_\nu \in T_\nu$ , and further assume that  $\nu$  has received  $(v_{\bar{\nu},1}, \dots, v_{\bar{\nu},k-1}) \in V_{\bar{\nu},1} \times \dots \times V_{\bar{\nu},k-1}$  in the first  $k - 1$  iterations. Then,  $h_{\nu,k}(\ell_\nu, v_{\bar{\nu},1}, \dots, v_{\bar{\nu},k-1}, \tau_\nu)$  is the information sender  $\nu$  sends to sender  $\bar{\nu}$  in the  $k$ -th step of the conference. Note that these conferencing functions are by no means random, but they both depend deterministically on the messages and CSI of *both* of the senders. This will be important in Section 3. We call the protocol just described a Willems conferencing protocol, as the idea of iterative message exchange in this context is taken from [4], [5].

### 2.2. Cooperative Coding

For a blocklength  $n$  and conferencing capacities  $C_1, C_2$ , we now formally define a code  $(n, R_1, R_2, C_1, C_2)$ . For each  $R_\nu$ , there must be an integer  $M_\nu$  with  $nR_\nu = \log M_\nu$ . Further, there is a Willems conferencing protocol as described above; in particular, there are sets  $V_1, V_2$  fulfilling (1). The encoding procedure is given by functions  $f_1, f_2$ , where  $f_1$  maps the set  $[1, M_1] \times T_1 \times V_2$  into  $A^n$ , and  $f_2$  maps  $[1, M_2] \times T_2 \times V_1$  into  $B^n$ . That means that if transmitter  $\nu$  would like to transmit message  $\ell_\nu$ , if he knows that the channel state is such that

$t_\nu(s) = \tau_\nu$ , and if during the conference, he has received the total information  $v_{\bar{\nu}} \in V_{\bar{\nu}}$  from sender  $\bar{\nu}$ , he uses the codeword  $f_\nu(\ell_\nu, \tau_\nu, v_{\bar{\nu}})$ . As each codeword only depends on the quadruple  $(\ell_1, \ell_2, \tau_1, \tau_2)$ , we can also write  $f_1(\ell_\nu, \tau_\nu, v_{\bar{\nu}}) = x_{\ell_1 \ell_2}^{\tau_1 \tau_2}$  and  $f_2(\ell_\nu, \tau_\nu, v_{\bar{\nu}}) = y_{\ell_1 \ell_2}^{\tau_1 \tau_2}$ . On the decoding side, for each instance  $\rho \in R$  of CSIR, the decoder partitions the output set  $C^n$  into  $M_1 M_2$  disjoint sets  $D_{ij}^\rho$ . If the channel state is  $s \in r^{-1}(\rho)$  and if the channel output is contained in  $D_{ij}^\rho$ , the receiver decides that the message pair  $(i, j)$  has been sent. A code  $(n, R_1, R_2, C_1, C_2)$  is a code  $(n, R_1, R_2, C_1, C_2, \lambda)$  if

$$\frac{1}{M_1 M_2} \sum_{(i,j) \in [1, M_1] \times [1, M_2]} H_s^n(D_{ij}^\rho | x_{ij}^{\tau_1 \tau_2}, y_{ij}^{\tau_1 \tau_2}) \leq \lambda$$

for every possible channel state  $s \in S$ . This means that the average probability of erroneous transmission is smaller than  $\lambda$  no matter which is the actual channel realization, so our encoding-decoding strategy is *universally* good.

A pair  $(R_1, R_2)$  is an achievable rate pair with respect to the CSI functions  $t_1, t_2, r$  if for every  $\varepsilon, \lambda > 0$  and for  $n$  large enough, there is a code  $(n, R'_1, R'_2, C_1, C_2, \lambda)$  with  $R'_\nu \geq R_\nu - \varepsilon$ . Note that only the capacities of the links between the base stations are fixed; in order to achieve a given rate pair, a Willems conferencing protocol may be chosen accordingly. The set of all achievable rate pairs (the capacity region) is denoted  $\mathcal{C}(C_1, C_2)$ .

### 2.3. The Main Results

In order to characterize  $\mathcal{C}(C_1, C_2)$ , we need to introduce some more notation. Denote by  $\mathcal{P}_0$  the set of all pairs  $(D, p_0)$ , where  $D$  is any finite set and  $p_0$  is a probability measure on  $D$ . Let  $\mathcal{P}_1$  ( $\mathcal{P}_2$ ) be the set of all stochastic matrices with input alphabet  $D$  and output alphabet  $A$  ( $B$ ). Given a  $(D, p_0) \in \mathcal{P}$  and a pair  $(p_1, p_2) \in \mathcal{P}_1 \times \mathcal{P}_2$ , we define for each channel state  $s \in S$  a probability measure  $p_s$  on  $D \times A \times B \times C$  by setting

$$p_s(d, a, b, c) = H_s(c|a, b) p_2(b|d) p_1(a|d) p_0(d).$$

Then, define a quadruple  $(U, X, Y, Z_s)$  of random variables taking values in  $D \times A \times B \times C$  and having joint probability distribution  $p_s$ . Then, define the set

$$\begin{aligned} \mathcal{R}(p_0, p_1, p_2, C_1, C_2, s) := \{ & (R_1, R_2) \geq 0 : \\ & R_1 \leq I(Z_s; X|Y, U) + C_1, \\ & R_2 \leq I(Z_s; Y|X, U) + C_2, \\ & R_1 + R_2 \leq (I(Z_s; X, Y|U) + C_1 + C_2) \\ & \quad \wedge I(Z_s; X, Y)\}, \end{aligned}$$

where  $a \wedge b := \min\{a, b\}$  and where for random variables  $X, Y, Z$ , the term  $I(Y; X|Z)$  is the mutual information of  $X$  and  $Y$  given  $Z$ . Note that there is nothing random about the sets  $\mathcal{R}(p_0, p_1, p_2, C_1, C_2, s)$ . We can now state the main theorem characterizing the capacity region of the channel.

**Theorem 1.** *For the channel described above, with the CSI functions  $t_1, t_2$ , and  $r$ , using Willems conferencing protocols, one achieves*

$$\mathcal{C}(C_1, C_2) = \bigcup_{(D, p_0) \in \mathcal{P}_0} \bigcap_{(\tau_1, \tau_2) \in T_1 \times T_2} \bigcup_{(p_1, p_2) \in \mathcal{P}_1 \times \mathcal{P}_2} \bigcap_{s \in t_1^{-1}(\tau_1) \cap t_2^{-1}(\tau_2)} \mathcal{R}(p_0, p_1, p_2, C_1, C_2, s). \quad (2)$$

One may assume that  $|D| \leq \max\{|A||B| + 2, |C| + 3\}$ .

Note that the capacity region is not influenced by CSIR. For the special case of  $|S| = 2$ , of no CSIT and of full CSIR, the capacity region was already characterized in [2].

In Section 4, we will see that no iteration is needed in the conferencing between the transmitters. Roughly, the information sent by each transmitter to the other can be described as follows: each transmitter conveys all his channel knowledge to the other. Then, each transmitter partitions his set of messages intended for the receiver into approximately equally-sized blocks. The partition must be so coarse that the index can be transmitted to the other sender without exceeding the remaining link capacity.

The full cooperation region is given by the triangle with corners  $(0, C_\infty)$ ,  $(C_\infty, 0)$ ,  $(0, 0)$ , where

$$C^\infty := \sup_{(D, p_0)} \min_{(\tau_1, \tau_2)} \max_{(p_1, p_2)} \inf_{s \in t_1^{-1}(\tau_1) \cap t_2^{-1}(\tau_2)} I(Z_s; X, Y)$$

and where the sup, min, max range over the same sets as in the characterization of the rate region. This can be seen from the result from [3] cited in more detail in Section 4. Further, if no cooperation is possible, but every transmitter also has the other transmitter's CSI, then the maximal sum rate equals

$$C^0 := \sup_{(D, p_0)} \min_{(\tau_1, \tau_2)} \max_{(p_1, p_2)} \inf_{s \in t_1^{-1}(\tau_1) \cap t_2^{-1}(\tau_2)} I(Z_s; X, Y|U).$$

The full cooperation sum rate then is achieved if

$$C_1 + C_2 \geq C^\infty - C^0.$$

Further, the full cooperation rate region is achieved if  $C_\nu \geq C^\nu$ , where  $C^\nu$  is the capacity of the induced single-user compound channel for transmitter  $\nu$ ,

$$C^\nu = \min_{\tau_\nu \in T_\nu} \max_{p_\nu} \inf_{s \in t_\nu^{-1}(\tau_\nu)} I(Z_\nu; X)(I(Z_\nu; Y)).$$

### 3. THE OUTER BOUND TO THE CAPACITY REGION

We prove the outer bound to the capacity region assuming full CSIR. In Section 4, we will see that this outer bound can be achieved with no CSIR, thus proving that CSIR is immaterial for the system performance. We modify the original model in two steps. First, we allow a greater set of conferencing

protocols. The capacity region can only become greater by doing this. For a given blocklength  $n$ , we say that  $(g_1, g_2)$  is an admissible pair of conferencing functions if  $g_\nu$  is a function from  $[1, M_1] \times [1, M_2] \times T_1 \times T_2$  into a set  $V_\nu$  with  $\log|V_\nu| \leq nC_\nu$  and if for fixed  $(\ell_\nu, \tau_\nu)$ , the pairs of values of  $(g_1, g_2)$  (contained in  $V_1 \times V_2$ ) attain at most  $|V_\nu|$  values as  $(\ell_{\bar{\nu}}, \tau_{\bar{\nu}})$  ranges over  $[1, M_{\bar{\nu}}] \times V_{\bar{\nu}}$ . By induction over the number of conferencing iterations, one shows that every Willems conferencing function is admissible in this extension. In the second step, we modify the model thus obtained. We now assume that both transmitters know  $t_1(s)$  and  $t_2(s)$  if the actual channel state is  $s$ , but it is no more allowed to exchange CSI between the senders. The conferencing functions need to fulfill a condition analogous to the one fulfilled by the conferencing functions used in the first step. We call every code obtained in the latter (extended) model having average error smaller than  $\lambda$  a  $\text{code}_{\text{ext}}(n, R_1, R_2, C_1, C_2, \lambda)$ . The following lemma formalizes the intuition that one does not lose anything during this modification.

**Lemma 1.** *For every  $\delta > 0$  there exists a positive integer  $n_0$  such that for every code  $(n, R_1, R_2, C_1, C_2, \lambda)$  with  $n \geq n_0$  there is a  $\text{code}_{\text{ext}}(n, R_1, R_2, C_1 + \delta, C_2 + \delta, \lambda)$ .*

The proof of this lemma makes use of the fact that the amount of channel state information is finite. Now, it suffices to show that all rates achievable with the extended channel model are contained in the set

$$\bigcup_{(D, p_0) \in \mathcal{P}_0} \bigcap_{(\tau_1, \tau_2) \in T_1 \times T_2} \bigcup_{(p_1, p_2) \in \mathcal{P}_1 \times \mathcal{P}_2} \bigcap_{s \in t_1^{-1}(\tau_1) \cap t_2^{-1}(\tau_2)} \mathcal{R}(p_0, p_1, p_2, C_1 + \delta, C_2 + \delta, s),$$

where the sets  $\mathcal{R}$  are the same as those in Theorem 1. The proof can be done using the well-known Fano's inequality. As  $\delta > 0$  is arbitrary, together with Lemma 1 this provides a tight upper bound to the capacity region which coincides with the right half of (2).

#### 4. CAPACITY-ACHIEVING COOPERATION

We assume no CSIR and show that the outer bound given for full CSIR in Section 3 can be achieved.

##### 4.1. The Optimal Conferencing Protocol

We first give an exact definition of the optimal Willems conferencing protocol. In order to do that, we need to define partitions of the message sets. Let positive integers  $M_1, M_2, V_1$ , and  $V_2$  be given with  $|V_\nu| \geq |T_\nu|$ . For  $\nu = 1, 2$ , define numbers

$$\mu_\nu = \left\lfloor \frac{|V_\nu|}{|T_\nu|} \right\rfloor \wedge M_\nu, \quad \xi_\nu = \left\lfloor \frac{M_\nu - 1}{\mu_\nu - 1} \right\rfloor.$$

Then, every  $\ell_\nu \in [1, M_\nu]$  can be represented as

$$\ell_\nu = (v_\nu - 1)\xi_\nu + \ell'_\nu,$$

where  $v_\nu \in [1, \mu_\nu]$  and where

$$\ell'_\nu \in \begin{cases} [1, \xi_\nu] & \text{if } v_\nu \leq \mu_\nu - 1; \\ [1, M_\nu - (\mu_\nu - 1)\xi_\nu] & \text{if } v_\nu = \mu_\nu. \end{cases}$$

This describes a partition of the set  $[1, M_\nu]$  into  $\mu_\nu$  blocks, where the first  $\mu_\nu - 1$  each contain  $\xi_\nu$  elements and where the last block contains the rest. The conferencing functions  $h_\nu$  are functions from  $[1, M_\nu] \times T_\nu$  to  $[1, \mu_\nu] \times T_\nu$ . Each message-channel state pair  $(\ell_\nu, \tau_\nu)$  is mapped by  $h_\nu$  to the pair  $(\ell'_\nu, \tau_\nu)$ , where  $\ell'_\nu$  gives the index of the block containing  $\ell_\nu$ . Later, the  $V_\nu$  will be chosen such that  $\log|V_\nu| \leq nC_\nu$ , so with  $|[1, \mu_\nu] \times T_\nu| \leq |V_\nu|$ , one sees that this is a Willems conferencing protocol. Note that this is a one-shot protocol.

##### 4.2. Optimal Coding

Now, we show the achievability of an arbitrary pair  $(R_1, R_2)$  contained in the set on the right-hand side of (2). To do so, we use a result proved in [3]. It describes the capacity region of a downlink channel where both transmitters have their own message for the receiver, but where there is an additional common message that both transmitters would like to transmit. Assume that this channel can take on the same states as our channel, but that each sender knows  $t_1(s)$  and  $t_2(s)$ . CSIR may be arbitrary. There is no cooperation between the transmitters. Using the average error criterion in the definition of achievability again, the capacity region  $\mathcal{C}_{\text{cm}}$  of this channel is

$$\mathcal{C}_{\text{cm}} = \bigcup_{(D, p_0) \in \mathcal{P}_0} \bigcap_{(\tau_1, \tau_2) \in T_1 \times T_2} \bigcup_{(p_1, p_2) \in \mathcal{P}_1 \times \mathcal{P}_2} \bigcap_{s \in t_1^{-1}(\tau_1) \cap t_2^{-1}(\tau_2)} \mathcal{R}_{\text{cm}}(p_0, p_1, p_2, s),$$

where

$$\begin{aligned} \mathcal{R}_{\text{cm}}(p_0, p_1, p_2, s) := \{ & (R'_0, R'_1, R'_2) \geq 0 : \\ & R'_1 \leq I(Z_s; X|Y, U), \\ & R'_2 \leq I(Z_s; Y|X, U), \\ & R'_1 + R'_2 \leq I(Z_s; X, Y|U) \\ & R'_0 + R'_1 + R'_2 \leq I(Z_s; X, Y) \}. \end{aligned}$$

Here,  $R'_0$  denotes the rate at which the common message is transmitted. We obtain a particular  $(R'_0, R'_1, R'_2) \in \mathcal{C}_{\text{cm}}$  from the pair  $(R_1, R_2)$  by setting  $\tilde{R}_\nu := R_\nu \wedge C_\nu$  and then defining

$$R'_0 := \tilde{R}_1 + \tilde{R}_2, \quad R'_\nu := R_\nu - \tilde{R}_\nu \text{ for } \nu = 1, 2.$$

As  $(R'_0, R'_1, R'_2)$  is achievable, we know that for large enough blocklength  $n$ , there is a  $\text{code}_{\text{cm}}(n, R'_0, R'_1, R'_2, \lambda)$  for the downlink channel with common message with rate triple  $(R''_0, R''_1, R''_2)$  fulfilling  $R'_\nu \geq R''_\nu \geq R'_\nu - \varepsilon$  for  $\nu = 0, 1, 2$  and with average error smaller than  $\lambda$ .

Back in our original situation, suppose that the message pair  $(\ell_1, \ell_2)$  is to be transmitted and that sender  $\nu$  knows that

the channel state is contained in the subset  $t_\nu^{-1}(\tau_\nu)$  of the set of all possible channel states. Then through the second component of the received information  $h_{\bar{\nu}}(\ell_{\bar{\nu}}, \tau_{\bar{\nu}})$  (using the conferencing protocol defined in Subsection 4.1), sender  $\nu$  obtains the channel information available to sender  $\bar{\nu}$ . Thus, transmission can be done using CSIT =  $(t_1^{-1}(\tau_1), t_2^{-1}(\tau_2))$  at each sender. Further, the pair of first components of the  $h_\nu(\ell_\nu, \tau_\nu)$  can be interpreted as common message. For each message pair  $(\ell_1, \ell_2)$ , it contains the information which cell in  $[1, M_\nu]$  contains  $\ell_\nu$ , for both  $\nu = 1, 2$ .

It remains to choose  $n, M_1, M_2, V_1, V_2$  such that

$$\log \mu_\nu \leq n\tilde{R}_\nu \quad (3)$$

and

$$nR_0'' - \frac{1}{n} \leq \log \mu_1 \mu_2 = nR_0''; \quad \log \xi_\nu = nR_\nu'' \text{ for } \nu = 1, 2. \quad (4)$$

Thus  $|V_\nu| \geq |T_\nu|$ , as required in the definition of the  $h_\nu$ . Further, (3) ensures  $\log |V_\nu| \leq nC_\nu$ . That the triple of messages obtained through the conferencing protocol can be encoded into the code $_{\text{cm}}(n, R_0'', R_1'', R_2'', \lambda)$  is ensured by (4), and by (3) and (4),  $\log M_\nu \geq n(R_\nu - \varepsilon)$ , as required in the definition of achievability. This proves that the outer bound from Section 3 is achievable using the simple Willems protocol from Subsection 4.1.

## 5. A NUMERICAL EXAMPLE

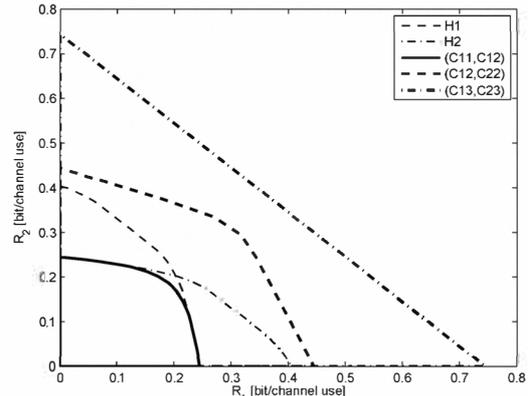
Assume  $A = B = C = \{0, 1\}$ . We set  $S = \{1, 2\}$ . At the different states, the channel is given by the stochastic matrices

$$H_1 = \begin{pmatrix} .9 & .1 \\ .4 & .6 \\ .6 & .4 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad H_2 = \begin{pmatrix} .9 & .1 \\ .6 & .4 \\ .4 & .6 \\ 0 & 1 \end{pmatrix}.$$

Here, the output distribution corresponding to the input combination  $(a, b)$  is written in row  $2a+b+1$ . The capacity region in the case of no CSI is depicted in Fig. 2 for three different choices of  $C_1, C_2$ .

## 6. DISCUSSION

In this paper, we have derived the optimal coding and cooperation protocols in a downlink channel with flat fading remaining constant over the transmission blocks, where the transmitters can exchange information about their messages and their CSI prior to transmission over noiseless finite-capacity links. For CSI, it was assumed that each sender only knows which of finitely many partitioning sets the channel state is contained in, whereas the receiver was allowed to have any amount of CSI. This models a situation where two base stations would like to transmit messages to a mobile and where there is a backbone connecting them. Future wireless systems will rely



**Fig. 2.** The capacity region for the conferencing capacity pairs  $(C_{11}, C_{21}) = (0, 0)$ ,  $(C_{12}, C_{22}) = (.2, .2)$ , and  $(C_{13}, C_{23}) = (.5, .5)$ . Additionally, the capacity regions of the channels  $H_1$  and  $H_2$  are shown.

on cooperation strategies to achieve the high data rates they are designed for. In this work, it was determined how much cooperation is needed. In particular, it was found that finite capacity backbones are sufficient to achieve full cooperation performance, and the required link strength was determined.

## 7. REFERENCES

- [1] G.J. Foschini, K. Karakayali, and R.A. Valenzuela, "Coordinating multiple antenna cellular networks to achieve enormous spectral efficiency", *IEE Proc.-Commun.*, vol. 153, no. 4, pp. 548–555, 2006.
- [2] I. Marić, R.D. Yates, and G. Kramer, "Capacity of Interference Channels With Partial Transmitter Cooperation", *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3536–3548, 2007.
- [3] M. Wiese, H. Boche, I. Bjelaković, and V. Jungnickel, "The Compound Multiple Access Channel with Partially Cooperating Encoders", submitted to *IEEE Trans. Inf. Theory*, 2010.
- [4] F.M.J. Willems, "Information-Theoretical Results for the Discrete Memoryless Multiple Access Channel," Ph.D. dissertation, Katholieke Universiteit Leuven, Belgium, 1982.
- [5] F.M.J. Willems, "The Discrete Memoryless Multiple Access Channel with Partially Cooperating Encoders," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 441–445, 1983.