Broadcast Channel: Degrees of Freedom with no CSIR

Umer Salim
Mobile Communications Department
Eurecom Institute
06560 Sophia Antipolis, France
umer.salim@eurecom.fr

Dirk Slock
Mobile Communications Department
Eurecom Institute
06560 Sophia Antipolis, France
dirk.slock@eurecom.fr

***Abstract*— We analyze a broadcast channel with no initial assumption of channel state information neither at the base station (BS) nor at the users' side. For the case when there is no possibility of feedback to the BS and it remains oblivious of the channel state information throughout the transmission, it is shown that the capacity region is bounded by the capacity of a point-to-point MISO link and hence the pre-log of the sum rate is ($1 - 1/T$) for a block fading channel of coherence length T. When the BS is allowed to acquire channel knowledge, operating under time-division duplex (TDD) mode, we give a very simple scheme through which BS and all users get necessary channel state information and the high SNR sum rate shows significant multiplexing gain or degrees of freedom (DOF).**

I. Introduction

In multiple-antenna broadcast channels, capacity or achievable data rates can be excessively increased just by adding multiple antennas at the transmitting end. Thus if a base station (BS) has M transmit antennas and the number of users in the system is K with $K \geq M$, this broadcast channel can support data rates M times larger than a single antenna BS, although all users may have single antenna each in both cases [1], [2], [3]. So under favorable conditions, the sum capacity of the broadcast channel is comparable to the capacity of a point-to-point MIMO channel having the same number of transmit and receive antennas. Apart from this sum capacity aspect, there are two advantages of this broadcast channel. It requires mobile users to have a single antenna each so users' terminals are quite inexpensive and simple. The second advantage is that point-to-point MIMO links are plagued by line-of-sight channel conditions where channel matrices are of reduced rank and they lose their multiplexing abilities. In broadcast channel, naturally users are far apart so the assumption of independent channel for each user holds very well and the channel matrix is of full-rank with probability one and is much well-conditioned as compared to the channel matrix of a point-to-point MIMO link [4].

But these promising advantages of broadcast MIMO don't come for free. To realize these high throughputs, BS has to transmit to multiple users over the same bandwidth. Orthogonal transmission schemes such as time-division multiple access (TDMA), frequency-division multiple access (FDMA) and code-division multiple access (CDMA) are highly sub-optimal as effectively BS will be transmitting to a single user over a particular resource. The other price to pay to achieve these high data rates is that BS must know the forward channel to all users [1]. This point is in sharp contrast to point-to-point MIMO. In point-to-point MIMO, channel state information at the transmitter (CSIT) only affects the power offset of the capacity. The slope of the capacity versus SNR curve, normally termed as the multiplexing gain or the degrees of freedom (DOF), remains unaffected by CSIT [5], [3].

We use the term "non-coherent" to mean that initially there is no assumption of channel knowledge on either side. But we don't prevent any side (transmitter and receivers) to learn/feedback the channel and subsequently use this information for precoding/decoding of data. Most of the initial results on the information theoretic capacity analysis of the broadcast channel came with the assumption of perfect channel state information at the transmitter (CSIT), and each user knows its own channel (CSIR). Inherently all channels are non-coherent and the users (receivers) need to estimate the channels implicitly or explicitly by some kind of training (pilots transmission) to get CSIR. In frequency-division duplex (FDD) mode of operation, downlink (forward) channels are normally different from the uplink (reverse) channels. So the users need to feedback their estimated forward channel information on the reverse link. On the other hand, the acquisition of CSIT gets facilitated when the broadcast channel operates under time-division duplex (TDD) mode. In this case, reciprocity implies that the forward channel matrix is the transpose of the reverse channel matrix [6]. So CSIT can be obtained easily compared to the FDD mode by some kind of pilot transmission from user terminals to the BS.

In section III, we analyze the capacity of a broadcast channel when no feedback is allowed to the BS by any means. When the BS is allowed to have the channel information, we develop a complete transmission strategy starting from non-coherent to fully coherent (although imperfect estimates) data transmission for TDD broadcast channel in section IV. High SNR asymptotics of the achievable sum rate are studied in section V and upper bound to the sum rate is also given.

Notation: $\mathbb{E}$ denotes statistical expectation. Lowercase letters represent scalars, boldface lowercase letters represent vectors, and boldface uppercase letters denote matrices. $\mathbf{A}^\dagger$

denotes the Hermitian of matrix $\mathbf{A}$.

II. System Model

The system we consider consists of one BS having M transmitting antennas and K single-antenna user terminals. In the downlink, the signal received by k-th user can be expressed as

$$y_k = \mathbf{h_k^\dagger x} + n_k, \qquad k = 1, 2, \ldots, K \tag{1}$$

where $\mathbf{h_1}, \mathbf{h_2}, \ldots, \mathbf{h_K}$ are the channel vectors of users 1 through user K with $\mathbf{h_k} \in \mathbb{C}^{M\times 1}$ ($\mathbb{C}^{M\times 1}$ denotes the M-dimensional complex space), $\mathbf{x} \in \mathbb{C}^{M\times 1}$ denotes the M-dimensional signal transmitted by the BS and $n_1, n_2, \ldots, n_K$ are independent complex Gaussian additive noise terms with zero mean and unit variances. We denote the concatenation of the channels by $\mathbf{H}^\dagger = [\mathbf{h_1} \mathbf{h_2} \cdots \mathbf{h_K}]$, so $\mathbf{H}$ is the $K \times M$ forward channel matrix with k-th row equal to the channel of the k-th user ($\mathbf{h_k^\dagger}$). The input must satisfy an average transmit power constraint of P i.e., $\mathbb{E}[||\mathbf{x}||^2] \leq P$.

The channel is assumed to be block fading having coherence length of T symbol intervals where fading remains the same, with independent fading from one block to the next [7]. The entries of the forward channel matrix $\mathbf{H}$ are independent and identically distributed (i.i.d.) complex Gaussian with zero mean and unit variance. We don't impose unrealistic assumptions of the presence of CSIR or CSIT. So initially all receivers and BS transmitter are oblivious of the channel realization in each block.

In normal broadcast scenarios, the number of users (K) will be more than the number of BS transmit antennas (M). It is well-known that with perfect CSIT and CSIR, broadcast channel with M transmit antennas and K single antenna users with $K \geq M$ achieves the multiplexing gain (DOF) of M [8] i.e., the dominant term of the sum-capacity of this broadcast channel is $M \log(P)$. Extra number of users does not contribute to increasing the multiplexing gain of this system although definite power gain can be achieved by scheduling over the users. In this contribution, our main point of concern is the multiplexing gain or the DOF of non-coherent broadcast channel so we focus our attention on the case with $K = M$. We mention explicitly when we are not following this assumption.

III. Broadcast channel with no feedback

For the discussion in this section, we impose the restriction that no uplink transmission is allowed from the user terminals to the BS. This may portray a practical scenario where user terminals are inexpensive devices with only reception capabilities. Hence the transmission mode in this section is half-duplex and conclusions will hold for both FDD and TDD broadcast channel when uplink transmission is not allowed.

For a broadcast channel, if all the users are distributed symmetrically i.e., have the same fading distributions and the transmitter has no instantaneous knowledge of CSI but each receiver knows its own channel perfectly, the sum capacity of this channel is equal to the capacity of the point-to-point channel from the transmitter to any one of the receivers. Thus TDMA is the optimal strategy in this case of no CSIT [9], [10],[11]. Hence the multiplexing gain of such a broadcast channel with CSIR and no CSIT is only one.

In this section, we focus on the broadcast channel where even the users have no channel information (no CSIR case) and all of the M users have symmetrical channel distributions. Because of the symmetry of the fading distributions among users, these channels fall under the category of "bottleneck channels" of Cover [10]. So any code transmitted by the BS, which is decodable at any user i is also decodable at any other user j. It means every user can decode all the information transmitted by the BS for M users. Hence the capacity region for such a broadcast channel is bounded by the capacity of the single user channel from BS to any one of the users. And the maximum sum rate with the restriction of no feedback is given by

$$R_{\text{sum}}^{\text{NO-FB}} = C_{\text{SU}} \tag{2}$$

where C_{SU} is the single-user capacity of MISO link from M-antenna BS to any single antenna user. Although for the case of interest (no CSIT, no CSIR), exact expression for C_{SU} is not known but high SNR asymptotics are available. Using the non-coherent capacity result of block fading channel from [12], we can write

$$R_{\text{sum}}^{\text{NO-FB}} = \left(1 - \frac{1}{T}\right) \log(P) + c \tag{3}$$

where c is a constant that does not depend upon SNR.

The achievability of this high SNR asymptotic of sum rate is straightforward. BS activates any one of its M transmit antennas and we also focus on a single user. So the broadcast channel reduces to a point-to-point SISO channel. In each coherence block of length T, first symbol is dedicated to training when the selected user estimates the only channel coefficient present. On rest of $T-1$ symbol intervals, user decodes the data based upon this channel knowledge, so extracting $T-1$ DOF out of each T symbol interval, matching the rate of equation (3).

C:1 For a broadcast channel (having no initial assumption of channel information) with M transmit antennas and K single antenna receivers, with no feedback to the BS throughout the transmission, the capacity region is bounded by the capacity of a ($M \times 1$) MISO link and the high SNR sum capacity behaves as the capacity of a point-to-point MISO or SISO channel with no CSIR (as pre-log is same for both).

C:2 This sum capacity is achievable by using one transmitting antenna at the BS, imposing unit length training and then transmitting to any single user or doing TDMA in the data phase of ($T-1$) symbol intervals. Thus increasing the number of antennas at the BS is not always beneficial as argued in [6], [13], in particular at high SNR.

IV. Broadcast channel with feedback

For a broadcast channel having a transmitter equipped with M transmit antennas and $K = M$ single antenna receivers with perfect CSIT and CSIR, the first order term of the sum capacity is $M \log(SNR)$ [8]. If we compare this to the capacity of the same broadcast channel with only CSIR available where the first order term of the sum capacity is only $\log(SNR)$, it clearly gives the strong motivation of having a learned transmitter BS. Thus if there is possibility of making the channel state information known at the BS, the difference in the sum capacity of broadcast channel with and without CSIT forcefully dictates that this is the right thing to do.

For our block fading channel with coherence length of T symbol intervals, we divide this interval in three phases, 1) uplink training, 2) downlink training and 3) coherent data transmission. The first phase is the uplink training phase

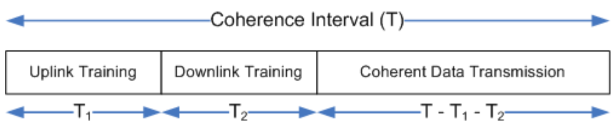

Fig. 1. Coherence interval Divided in Three Transmission Phases

where users train the BS about the forward channel and thus BS makes an estimate of the forward channel matrix comprising of the channel vectors of all users. So this phase is equivalent to feeding the BS about CSI. Based upon this channel information, BS may choose some transmission strategy which could be a simple linear beamforming strategy like zero forcing (ZF), some non-linear strategies like vector perturbation or the optimal dirty paper coding (DPC). The second phase is the downlink training phase where the BS transmits pilots so that users estimate their corresponding effective channels. When this second phase ends, both sides of the broadcast channel have necessary channel state information albeit imperfect. Thus starting from a broadcast channel with no CSIT and no CSIR, reaching up to the third data phase, we have a broadcast channel with imperfect CSIT and CSIR and hence in this data phase, BS may choose good transmission strategies and users can decode data coherently. The data rates obtained and their scaling with SNR show that these training phases are beneficial.

Below we give a detailed analysis of the three transmission phases mentioned above.

A. Uplink Training Phase

In this training phase, users transmit pilot signals which are known at the BS. As there are $K = M$ users, so the length of this uplink training interval is $T_1 \geq M$. Here we suppose that the average power constraint of each user is P_u. For this uplink training, the use of orthogonal training sequences by all users is very attractive because in that case all users can transmit simultaneously to the BS with their full power without interfering with each other. Thus pilot signal matrix (combined from all users) is $\sqrt{T_1}\mathbf{A}$ where $\mathbf{A}$ is a $K \times T_1$ unitary matrix hence $\mathbf{A}\mathbf{A}^\dagger = \mathbf{I_K}$ where $\mathbf{I_K}$ denotes a $K \times K$ identity matrix. If $\mathbf{Y_u}$ denotes the $M \times T_1$ matrix of the received signal by M antennas of the BS in this training interval of length T_1, the system equation for this uplink training phase becomes

$$\mathbf{Y_u} = \sqrt{P_u T_1}\mathbf{G}\mathbf{A} + \mathbf{Z_u} \tag{4}$$

where $\mathbf{Z_u}$ is a $M \times T_1$ matrix having i.i.d. zero mean unit variance complex Gaussian noise entries and $\mathbf{G}$ denotes the $M \times K$ uplink channel matrix. As pilot signal matrix $\mathbf{A}$ is known at the BS, it can formulate an MMSE estimate of the uplink channel matrix $\mathbf{G}$ which is given by

$$\hat{\mathbf{G}} = \frac{\sqrt{P_u T_1}}{P_u T_1 + 1}\mathbf{Y_u}\mathbf{A}^\dagger \tag{5}$$

Because this broadcast channel is operating under TDD mode of operation and we have assumed that perfect reciprocity holds between uplink and downlink channels so downlink (forward) channel matrix is just the transpose of the uplink channel matrix hence $\mathbf{H} = \mathbf{G}^\mathbf{T}$. The channel vector for user k can be expressed as $\mathbf{h_k} = \hat{\mathbf{h}}_\mathbf{k} + \tilde{\mathbf{h}}_\mathbf{k}$ where $\tilde{\mathbf{h}}_\mathbf{k}$ is the estimation error vector with i.i.d. Gaussian entries. All entries in the channel matrix are independent hence estimation error variance for any channel entry denoted by σ_1^2 is given by

$$\sigma_1^2 = \mathbb{E}[|\mathbf{H}_{ij} - \hat{\mathbf{H}}_{ij}|^2] = \frac{1}{P_u T_1 + 1} \tag{6}$$

C:3 The length of the uplink training phase T_1 depends solely upon the number of users K. It will remain same even if there is only one antenna employed at the BS.

C:4 The estimation error variance for each channel entry goes inversely proportional to the training length T_1 and the power constraint of the user terminals P_u.

B. BS Transmission Strategy: ZF Precoding

It is known that the dirty paper coding (DPC) is the capacity achieving transmission scheme for MIMO broadcast channel and achieves the full capacity region [14] but this scheme is complex and its implementation is quite tedious. So a lot of research has been carried out to analyze the performance of simpler linear precoding schemes. Zero forcing precoding, one of the simplest linear precoding strategy, has been shown to behave quite optimally at asymptotically high values of SNR and achieves the full DOF of a coherent broadcast channel [8]. It means that the first order term of the sum capacity of the broadcast channel remains the same whether one employs DPC or ZF precoding at the BS. In this contribution we are mainly interested in analyzing the DOF obtainable with some simple transmission scheme hence BS uses ZF precoding based upon the knowledge of the forward channel matrix obtained through explicit training.

In ZF precoding, beamforming vector for user k (denoted as $\bar{\mathbf{v}}_\mathbf{k}$), is selected such that it is orthogonal to the channel vectors of all other users. ZF beamforming vectors are the normalized columns of the inverse of the channel matrix $\mathbf{H}$. Hence with perfect CSIT, each user will receive only the beam directed to it and no multi-user interference will be experienced. For the case in hand, where the BS has imperfect estimate of the channel matrix, there will be some residual interference. If we represent ZF beamforming matrix by $\bar{\mathbf{V}} = [\bar{\mathbf{v}}_\mathbf{1}\bar{\mathbf{v}}_\mathbf{2}\cdots\bar{\mathbf{v}}_\mathbf{K}]$, the transmitted signal $\mathbf{x}$ becomes

$\mathbf{x} = \bar{\mathbf{V}}\mathbf{u}$ and the signal received by user k (1) can be expressed as

$$\begin{aligned} y_k &= \mathbf{h}_\mathbf{k}^\dagger \bar{\mathbf{V}}\mathbf{u} + n_k \\ &= \mathbf{h}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{k} u_k + \sum_{j \neq k} \mathbf{h}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{j} u_j + n_k \end{aligned} \tag{7}$$

Due to imperfect MMSE estimation at the BS and the choice of ZF beamforming unit vectors, we have

$$\mathbf{h}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{j} = \hat{\mathbf{h}}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{j} + \tilde{\mathbf{h}}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{j} = \tilde{\mathbf{h}}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{j} \tag{8}$$

hence the received signal at k-th user becomes

$$\begin{aligned} y_k &= \mathbf{h}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{k} u_k + \sum_{j \neq k} \tilde{\mathbf{h}}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{j} u_j + n_k \\ &= g_{k,k} u_k + \sum_{j \neq k} g_{k,j} u_j + n_k \end{aligned} \tag{9}$$

$g_{k,k}$ is the effective scalar channel for user k and $g_{k,j}$ are the coefficients which arise due to imperfect ZF beamforming as BS had no access to perfect channel realizations. But until this point, users have no knowledge of their channel.

C. Downlink Training Phase

We assume a very simple downlink training strategy. If the BS had the perfect knowledge of the forward channels to all users, due to ZF beamforming vectors each user would only receive the signal from the beam directed to it and no interference from any other beam would be observed. Here BS estimates the users' channels and therefore channel estimates and the corresponding ZF beamforming vectors are imperfect so each user receives some unwanted signal contribution from the beam directed to any other user. But this interference is of the same order as of the channel noise so for this DL training phase, BS activates all beams simultaneously for T_2 symbols times. So in each symbol interval, every user receives through its effective scalar channel, the Gaussian noise of the channel and the interference due to imperfect channel estimates and ZF beamforming vectors.

$$y_k = g_{k,k} u_k + \sum_{j \neq k} g_{k,j} u_j + n_k \tag{10}$$

Based upon this received signal and the known pilots, k-th user can form the MMSE estimate of the effective scalar channel $g_{k,k}$ which is given by

$$\begin{aligned} \hat{g}_{k,k} &= \frac{\mathbb{E}[g_{k,k} y_k^\dagger]}{\mathbb{E}[y_k y_k^\dagger]} y_k \\ &= \frac{\sqrt{\frac{PT_2}{M}}}{\frac{PT_2}{M} + \frac{PT_2}{M}(M-1)\sigma_1^2 + 1} y_k \end{aligned} \tag{11}$$

(See Appendix 1 for the details of the derivation of this estimator.)

As $\bar{\mathbf{v}}_\mathbf{k}$ is a unit vector independent of $\mathbf{h}_\mathbf{k}$, so effective scalar channel $g_{k,k} = \mathbf{h}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{k}$ is zero mean complex Gaussian with unit variance. As a result, MMSE estimate $\hat{g}_{k,k}$ and the estimation error $\tilde{g}_{k,k}$ both are complex Gaussian

$$g_{k,k} = \hat{g}_{k,k} + \tilde{g}_{k,k} \tag{12}$$

$$\hat{g}_{k,k} \sim \mathcal{CN}\left(0, \frac{\frac{PT_2}{M}}{\frac{PT_2}{M} + \frac{PT_2}{M}(M-1)\sigma_1^2 + 1}\right)$$

$$\tilde{g}_{k,k} \sim \mathcal{CN}\left(0, \frac{\frac{PT_2}{M}(M-1)\sigma_1^2 + 1}{\frac{PT_2}{M} + \frac{PT_2}{M}(M-1)\sigma_1^2 + 1}\right)$$

The estimation error variance in estimating this scalar effective channel is inversely proportional to the downlink power constraint. When this second phase of downlink training ends, both the BS and all of the users have estimates for the channel and coherent transmission with imperfect CSIT and CSIR is possible.

C:5 The length of the downlink training phase T_2 is independent of the number of transmit antennas M at the BS and the number of users K.

D. Coherent Data Phase

The capacity of a channel requires the maximization of the mutual information between the input and the output of that channel over the input distribution under the constraints imposed [9]. The optimization of the mutual information w.r.t. the input density itself is a very vast area of research and very few results are known, hence is certainly out of the scope of this paper. So we adopt the strategy of independent data transmission to all users from the BS with power equally divided among them. So k-th user input signal, u_k is Gaussian i.i.d. i.e, $u_k \sim \mathcal{CN}(0, P/M)$. The intuition is that in case of perfect CSIT and CSIR, Gaussian signals are the optimal ones.

After the two training phases, first in the uplink and second in the downlink direction, both the BS and all users have imperfect channel estimates. So with ZF beamforming employed, the signal y_k received by user k (9) may be expressed as

$$y_k = \hat{g}_{k,k} u_k + \tilde{g}_{k,k} u_k + \sum_{j \neq k} g_{k,j} u_j + n_k \tag{13}$$

The above equation differs a lot from (9) as there user k was unaware of its scalar channel $g_{k,k}$ but (13) effectively represents a point-to-point coherent channel with channel $\hat{g}_{k,k}$ known at user k, although there is Gaussian noise, some interference coming from the ZF beamforming vectors of other users and the noise due to imperfect estimation of the effective channel at user's side.

E. Lower Bound of the Achievable Rate

We are interested in calculating the achievable sum rate of this broadcast channel or its lower bound which could at least point to the number of DOF achievable. If we denote the rate obtained by k-th user as R_k, then it is the mutual information between u_k and y_k with channel $\hat{g}_{k,k}$ known

$$R_k = I(u_k; y_k) \tag{14}$$

In this case, the problem is that we cannot simply use the expression for the mutual information of known scalar channel because of the presence of interference terms whose distributions are unknown. If we combine the noise, the interference and the estimation error contribution in y_k (eq. (13)) in an effective additive noise w_k, then

$$w_k = \tilde{g}_{k,k} u_k + \sum_{j \neq k} g_{k,j} u_j + n_k \quad (15)$$

now the variance of this effective additive noise term conditional upon the effective scalar channel estimate $\hat{g}_{k,k}$ can be calculated to be

$$\mathbb{E}[w_k w_k^\dagger | \hat{g}_{k,k}] = \mathbb{E}[|\tilde{g}_{k,k}|^2]\mathbb{E}[|u_k|^2] + \sum_{j \neq k} \mathbb{E}[|g_{k,j}|^2 | \hat{g}_{k,k}] \mathbb{E}[|u_j|^2] + \mathbb{E}[|n_k|^2]$$

All the expectations in the above equation are already known except $\mathbb{E}[|g_{k,j}|^2 | \hat{g}_{k,k}]$ which is difficult to compute

$$\begin{aligned} \mathbb{E}[w_k w_k^\dagger | \hat{g}_{k,k}] &= \tfrac{P}{M} \frac{\frac{PT_2}{M}(M-1)\sigma_1^2 + 1}{\frac{PT_2}{M} + \frac{PT_2}{M}(M-1)\sigma_1^2 + 1} \\ &+ \tfrac{P}{M} \sum_{j \neq k} \mathbb{E}[|g_{k,j}|^2 | \hat{g}_{k,k}] + 1 \end{aligned}$$

Due to the use of MMSE estimation in the downlink training, we remark that the signal is uncorrelated with the noise and all interfering terms.

$$\mathbb{E}[u_k(\tilde{g}_{k,k} u_k + \sum_{j \neq k} g_{k,j} u_j + n_k)^\dagger] = 0 \quad (16)$$

The above expectation is zero because of the property of uncorrelated MMSE estimation error, the use of independent signals for different users and that the noise is independent of everything else. Now once we have shown that all additive noise terms are uncorrelated with the desired signal, we can invoke Theorem 1 from [15] which states that the worst case uncorrelated noise has the zero mean Gaussian distribution. So we can replace the effective scalar additive noise w_k of unknown distribution with a noise of the same second moment but having Gaussian distribution, it will give a lower bound to the rate R_k of k-th user but we can instantly write the expression for the mutual information as

$$\begin{aligned} R_k &\geq \mathbb{E}_{\hat{g}_{k,k}} \log\left(1 + \frac{|\hat{g}_{k,k}|^2 \mathbb{E}|u_k|^2}{\mathbb{E}[w_k w_k^\dagger | \hat{g}_{k,k}]}\right) \\ &= \mathbb{E}_{\hat{g}_{k,k}} \log\left(1 + \frac{P}{M} \frac{|\hat{g}_{k,k}|^2}{\mathbb{E}[w_k w_k^\dagger | \hat{g}_{k,k}]}\right) \quad (17) \end{aligned}$$

V. High SNR DOF of the Sum Rate

The rate for k-th user derived in eq. (17) can further be lower bounded as

$$\begin{aligned} R_k &\geq \mathbb{E}_{\hat{g}_{k,k}} \log\left(\frac{P}{M} \frac{|\hat{g}_{k,k}|^2}{\mathbb{E}[w_k w_k^\dagger | \hat{g}_{k,k}]}\right) \\ &= \mathbb{E}_{\hat{g}_{k,k}} \log\left(\frac{P}{M}|\hat{g}_{k,k}|^2\right) - \mathbb{E}_{\hat{g}_{k,k}} \log\left(\mathbb{E}[w_k w_k^\dagger | \hat{g}_{k,k}]\right) \\ &\geq \mathbb{E}_{\hat{g}_{k,k}} \log\left(\frac{P}{M}|\hat{g}_{k,k}|^2\right) - \log\left(\mathbb{E}[w_k w_k^\dagger]\right) \quad (18) \end{aligned}$$

where the last inequality follows from the Jensen's inequality. With this, we only need to compute the 2nd moment of w_k which is readily shown to be

$$\sigma_w^2 = \mathbb{E}[w_k w_k^\dagger] = \frac{P}{M} \frac{\frac{PT_2}{M}(M-1)\sigma_1^2 + 1}{\frac{PT_2}{M} + \frac{PT_2}{M}(M-1)\sigma_1^2 + 1} + (M-1)\frac{P}{M}\frac{1}{P_u T_1 + 1} + 1 \quad (19)$$

because

$$\mathbb{E}[|g_{k,j}|^2] = \mathbb{E}[|\tilde{\mathbf{h}}_{\mathbf{k}}^\dagger \bar{\mathbf{v}}_{\mathbf{j}}|^2] = \frac{1}{P_u T_1 + 1} \quad (20)$$

As all of the users are symmetrically distributed, so the sum rate of this broadcast channel is given by

$$\begin{aligned} R_{\text{sum}}^{\text{FB}} &= \frac{T - T_1 - T_2}{T} M R_k \quad (21) \\ &\geq \tfrac{T - T_1 - T_2}{T} M \left[\mathbb{E}_{\hat{g}_{k,k}} \log\left(\tfrac{P}{M}|\hat{g}_{k,k}|^2\right) - \log(\sigma_w^2)\right] \end{aligned}$$

where we have incorporated the DOF loss in the sum rate due to two training phases in the uplink and the downlink directions.

If we increase the first training phase duration T_1, it improves the quality of the channel estimates at the BS and interference at each user due to beamforming vectors of other users decreases but it gives only a gain in SNR offset (see (21) and (19)) which is logarithmic in nature but the coefficient $(T - T_1 - T_2)$ reduces the DOF of the sum rate linearly with increase in T_1 so the optimal length of the first training phase should be the minimum possible at high SNR, hence $T_1 = M$. This argumentation assumes that power constraint of user terminals P_u is of the same order as that of the BS power constraint P.

About the second training phase in the downlink direction of length T_2, reasoning is not very different. With the increase in this training interval, users are better able to estimate their effective scalar channels which gives SNR gain, logarithmic in nature but increase in T_2 directly hits DOF due to the coefficient $(T - T_1 - T_2)$ in front of the logarithm. So to exploit the maximum number of DOF at high SNR, the optimal (minimal) value of T_2 comes out to be 1. Hence adopting these values, the sum rate becomes

$$R_{\text{sum}}^{\text{FB}} \geq \frac{T - M - 1}{T} M \left[\mathbb{E}_{\hat{g}_{k,k}} \log\left(\frac{P}{M}|\hat{g}_{k,k}|^2\right) - \log(\sigma_w^2)\right] \quad (22)$$

It's trivial to show that σ_w^2 is bounded by a finite constant for large values of P (the BS power constraint) and if power constraints of users are of the same order as that of P. So for limiting value of P, the lower bound to sum rate becomes

$$\lim_{P \to \infty} R_{\text{sum}}^{\text{FB}} \geq \frac{T - (M+1)}{T} M \log(P) + c_1 \quad (23)$$

where c_1 is a finite constant which does not depend upon P for large values of P.

C:6 For a broadcast channel operating under TDD mode, having M BS antennas, same number of symmetric users, block fading channel of coherence interval T and starting from zero channel state information at both ends, our very simple scheme is able to achieve $M[1 - (M+1)/T]$ DOF.

If we compare this multiplexing gain to the multiplexing gain of the same broadcast channel under the restriction of no feedback to the BS (section III) where DOF is only

$(1-1/T)$, we see that even for very practical values of the block coherence interval T in mobile environments, this lower bound $M[1-(M+1)/T]$ is comparatively much larger and to make the BS learn the channel pays off very well.

A. Upper Bound of the Sum Rate

An upper bound to the sum rate of our scheme can be obtained when one sacrifices minimal lengths for both training intervals but then assumes that BS knows the DL channel perfectly and each user perfectly knows its effective scalar channel. This will remove all the interference terms from the received signal but DOF achieved will still be $M[1-(M+1)/T]$.

A general upper bound of the sum rate of non-coherent broadcast channel can be the sum rate with CSIT and CSIR known giving M DOF but this bound is not tight.

A much better upper bound could be obtained by letting all user terminals co-operate among themselves. So we get a single user point-to-point MIMO square channel of M dimensions. For this non-coherent channel, results are available in the literature [12] and the pre-log is given by $M[1-M/T]$. This shows that our scheme which achieves $M[1-(M+1)/T]$ DOF, is very close to this high SNR asymptote.

B. CSIT Quality Refinement

While switching from eq. (22) to eq. (23) which showed that our scheme is able to achieve $M[1-(M+1)/T]$ DOF for this broadcast channel, the boundedness of effective noise variance σ_w^2 with a finite constant required users' power constraint P_u to be of the same order as that of the BS power constraint P. If this is not the case (i.e. $\lim_{P\to\infty} P_u/P = 0$), the channel quality at the BS will be relatively poor. And the interference power at each user due to beams meant for other users (and hence σ_w^2) will go on increasing with the DL power P and hence all DOF will collapse and the sum rate will be bounded in SNR. This result parallels the result of [3] for digital feedback which showed that feedback rate (quality of CSIT) must increase with SNR (in dBs) to achieve DOF of the broadcast channel, here with analog feedback our result says that the uplink power (which governs the quality of CSIT) must scale with the BS power constraint (and hence the DL SNR). Although the rates unbounded in SNR can be achieved by transmitting to a single user or by time-sharing between users with fixed uplink power or even with no feedback to the BS, but DOF of the broadcast channel (due to multiple antennas at the BS and multiple users at the receiving side) are lost. Again to conclude, in case of imperfect channel estimates the CSIT quality must improve with the DL SNR to have rates unbounded in SNR otherwise the system becomes interference limited.

Remark 1: The channels of concern in this paper are fast fading channels which may arise for fast moving mobile users e.g. for user speeds of 100Km/h, carrier frequency of 2GHz and coherence BW of 100KHz, coherence interval will be about 100 symbol intervals [11]. So even for BSs having 16 antennas, training interval minimization becomes really necessary.

Remark 2: In [6], achievable data rates have been analyzed using first uplink training phase and then transmitting to users without making any attempt of users' learning the channel and those data rates are bounded in SNR. But our scheme shows the scaling of the sum rate versus SNR with a very attractive multiplexing gain.

VI. Concluding Remarks

We studied the capacity of a broadcast channel with no assumptions of channel knowledge under two scenarios. First, when the BS is not allowed any channel information, the capacity region was shown to be bounded by the capacity of MISO point-to-point link, hence the pre-log of the sum rate becomes trivially known. In the second case, when the BS may acquire channel information, we analyzed the sum rate with a very simple scheme, achieving considerable DOF even if one accounts for how that channel knowledge is obtained.

Appendix 1

We want to estimate $g_{k,k}$ in the equation below when known pilot symbols are transmitted with full power for T_2 symbol intervals

$$y_k = \sqrt{\frac{PT_2}{M}} g_{k,k} + \sqrt{\frac{PT_2}{M}} \sum_{j\neq k} g_{k,j} + n_k \tag{24}$$

$g_{k,k}$ is Gaussian distributed with zero mean and unit variance and $g_{k,j}$ is zero mean Gaussian distributed with variance σ_1^2. Based upon this received signal and the known pilots, k-th user can form the MMSE estimate of the effective scalar channel $g_{k,k}$ which is given by

$$\hat{g}_{k,k} = \frac{\mathbb{E}[g_{k,k} y_k^\dagger]}{\mathbb{E}[y_k y_k^\dagger]} y_k \tag{25}$$

$$\begin{aligned} \mathbb{E}[g_{k,k} y_k^\dagger] &= \sqrt{\frac{PT_2}{M}} \mathbb{E}[|g_{k,k}|^2] + \sqrt{\frac{PT_2}{M}} \sum_{j\neq k} \mathbb{E}[g_{k,k} g_{k,j}^\dagger] \\ &\quad + \mathbb{E}[g_{k,k} n_k^\dagger] \end{aligned} \tag{26}$$

The expectations in the first and the third terms are known and we handle the second term as follows

$$\begin{aligned} \mathbb{E}[g_{k,k} g_{k,j}^\dagger] &\overset{a}{=} \mathbb{E}[\mathbf{h_k^\dagger \bar{v}_k \bar{v}_j^\dagger \tilde{h}_k}] \\ &\overset{b}{=} \mathbb{E}[\mathbf{\tilde{h}_k^\dagger \bar{v}_k \bar{v}_j^\dagger \tilde{h}_k}] + \mathbb{E}[\mathbf{\hat{h}_k^\dagger \bar{v}_k \bar{v}_j^\dagger \tilde{h}_k}] \\ &\overset{c}{=} \mathbb{E}[\mathbf{\tilde{h}_k^\dagger \bar{v}_k \bar{v}_j^\dagger \tilde{h}_k}] + \mathbb{E}[\mathbf{\hat{h}_k^\dagger \bar{v}_k \bar{v}_j^\dagger}]\mathbb{E}[\mathbf{\tilde{h}_k}] \\ &\overset{d}{=} \mathbb{E}[\mathbf{\tilde{h}_k^\dagger \bar{v}_k \bar{v}_j^\dagger \tilde{h}_k}] + \mathbb{E}[\mathbf{\hat{h}_k^\dagger \bar{v}_k \bar{v}_j^\dagger}]\mathbf{0} \\ &\overset{e}{=} \mathbb{E}[\mathbf{\bar{v}_j^\dagger \tilde{h}_k \tilde{h}_k^\dagger \bar{v}_k}] + 0 \\ &\overset{f}{=} \mathbb{E}[\mathbf{\bar{v}_j^\dagger}\mathbb{E}\{\mathbf{\tilde{h}_k \tilde{h}_k^\dagger}\}\mathbf{\bar{v}_k}] \\ &\overset{g}{=} \mathbb{E}[\mathbf{\bar{v}_j^\dagger} \sigma_1^2 \mathbf{I_M \bar{v}_k}] \\ &\overset{h}{=} \sigma_1^2 \mathbb{E}[\mathbf{\bar{v}_j^\dagger \bar{v}_k}] \end{aligned} \tag{27}$$

In (b), we use $\mathbf{h_k} = \hat{\mathbf{h}}_\mathbf{k} + \tilde{\mathbf{h}}_\mathbf{k}$, (c) follows as $\tilde{\mathbf{h}}_\mathbf{k}$ is independent of the estimate $\hat{\mathbf{h}}_\mathbf{k}$ and beamforming vectors, (d) follows as estimation error is of zero mean, (f) follows as estimation error is independent of the beamforming vectors and (g) follows because elements of $\tilde{\mathbf{h}}_\mathbf{k}$ are i.i.d. So now we have to compute the expectation of the inner product of two ZF beamforming vectors which needs to be calculated over all the channel vectors. Without loss of generality, we can assume that $k = 1$ and $j = 2$ hence we want to compute $\mathbb{E}[\bar{\mathbf{v}}_\mathbf{2}^\dagger \bar{\mathbf{v}}_\mathbf{1}]$. Conditional upon estimates of the channel vectors $\hat{\mathbf{h}}_\mathbf{3}, \hat{\mathbf{h}}_\mathbf{4} \cdots \hat{\mathbf{h}}_\mathbf{M}$, both of these vectors lie in a 2-D null space of these channel vector estimates. $\hat{\mathbf{h}}_\mathbf{1}$ and $\hat{\mathbf{h}}_\mathbf{2}$ can also be projected in this null space of other channel vectors. Now $\bar{\mathbf{v}}_\mathbf{1}$ will be orthogonal to the projection of $\hat{\mathbf{h}}_\mathbf{2}$ and $\bar{\mathbf{v}}_\mathbf{2}$ will be orthogonal to the projection of $\hat{\mathbf{h}}_\mathbf{1}$. As $\hat{\mathbf{h}}_\mathbf{1}$, $\hat{\mathbf{h}}_\mathbf{2}$ and hence their projections in this 2-D null space are distributed in an independent and isotropic manner, so the same is true for $\bar{\mathbf{v}}_\mathbf{1}$ and $\bar{\mathbf{v}}_\mathbf{2}$. Hence conditional upon $\hat{\mathbf{h}}_\mathbf{3}, \hat{\mathbf{h}}_\mathbf{4} \cdots \hat{\mathbf{h}}_\mathbf{M}$, they are independent and isotropically distributed. But the mean of an isotropically distributed vector is zero.

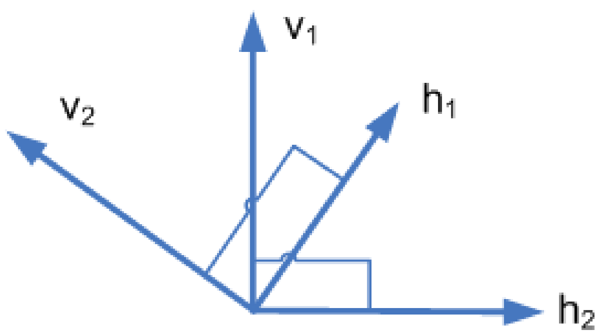

Fig. 2. Channel Projections and corresponding ZF vectors for users 1 and 2 in 2-D null space of all other users' channel vectors

$$\begin{aligned}
\mathbb{E}\left[\bar{\mathbf{v}}_\mathbf{2}^\dagger \bar{\mathbf{v}}_\mathbf{1}\right] &= \mathbb{E}_{\hat{\mathbf{h}}_{\mathbf{3},\mathbf{4},\cdots,\mathbf{M}}}\left[\mathbb{E}_{\hat{\mathbf{h}}_\mathbf{1},\hat{\mathbf{h}}_\mathbf{2}|\hat{\mathbf{h}}_{\mathbf{3},\mathbf{4},\cdots,\mathbf{M}}}\{\bar{\mathbf{v}}_\mathbf{2}^\dagger \bar{\mathbf{v}}_\mathbf{1}\}\right] \\
&= \mathbb{E}_{\hat{\mathbf{h}}_{\mathbf{3},\mathbf{4},\cdots,\mathbf{M}}}\left[\mathbb{E}_{\hat{\mathbf{h}}_{\mathbf{1},\mathbf{2}}|\hat{\mathbf{h}}_{\mathbf{3},\mathbf{4},\cdots,\mathbf{M}}}\{\bar{\mathbf{v}}_\mathbf{2}^\dagger\}\right. \\
&\quad \left.\mathbb{E}_{\hat{\mathbf{h}}_{\mathbf{1},\mathbf{2}}|\hat{\mathbf{h}}_{\mathbf{3},\mathbf{4},\cdots,\mathbf{M}}}\{\bar{\mathbf{v}}_\mathbf{1}\}\right] \\
&= \mathbb{E}_{\hat{\mathbf{h}}_{\mathbf{3},\mathbf{4},\cdots,\mathbf{M}}}\left[\mathbf{0}^\dagger \mathbf{0}\right] \\
&= 0 \qquad (28)
\end{aligned}$$

Hence we conclude that

$$\mathbb{E}[g_{k,k} g_{k,j}^\dagger] = 0 \qquad (29)$$

With this, $\mathbb{E}[g_{k,k} y_k^\dagger]$ becomes

$$\mathbb{E}[g_{k,k} y_k^\dagger] = \sqrt{\frac{PT_2}{M}} \qquad (30)$$

The other expectation $\mathbb{E}[y_k y_k^\dagger]$ becomes easy to compute because now we know that $\mathbb{E}[g_{k,k} g_{k,j}^\dagger] = 0$.

$$\begin{aligned}
\mathbb{E}[y_k y_k^\dagger] &= \frac{PT_2}{M}\mathbb{E}[|g_{k,k}|^2] + \frac{PT_2}{M}\sum_{j\neq k}\sum_{l\neq k}\mathbb{E}[g_{k,j} g_{k,l}^\dagger] + 1 \\
&= \frac{PT_2}{M} + \frac{PT_2}{M}\sum_{j\neq k}\mathbb{E}[|g_{k,j}|^2] + 1 \\
&= \frac{PT_2}{M} + \frac{PT_2}{M}(M-1)\sigma_1^2 + 1 \qquad (31)
\end{aligned}$$

Putting the values from eq.(30) and eq.(31) into eq.(25) gives the desired result.

Acknowledgments

The authors would like to acknowledge helpful comments by Randa Zakhour.

Eurecom's research is partially supported by its industrial members: BMW Group Research & Technology, Bouygues Telecom, Cisco, Hitachi, ORANGE, SFR, Sharp, STMicroelectronics, Swisscom, Thales. The research work leading to this paper has also been partially supported by the European Commission under the ICT research network of excellence NewCom++ of the 7th Framework programme and by the French ANR project APOGEE.

References

[1] G. Caire and S. Shamai (Shitz), "On the achievable throughput of a multiantenna gaussian broadcast channel," *IEEE Trans. on Information Theory*, vol. 49, pp. 1691–1706, July 2003.
[2] N. Jindal and A. Goldsmith, "Dirty paper coding versus tdma for mimo broadcast channels," *IEEE Trans. on Information Theory*, vol. 51, pp. 1783–1794, May 2005.
[3] N. Jindal, "Mimo broadcast channels with finite rate feedback," *IEEE Trans. on Information Theory*, vol. 52, pp. 5045–5060, November 2006.
[4] D. Gesbert, M. Kountouris, J. R. W. Heath, C. B. Chae, and T. Salzer, "From single user to multiuser communications: Shifting the mimo paradigm," *IEEE Sig. Proc. Magazine*, 2007.
[5] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, pp. 585–595, November 1999.
[6] T. L. Marzetta, "How much training is required for multiuser mimo?," in *Proc. Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA*, November 2006, pp. 359–363.
[7] T. Marzetta and B. Hochwald, "Capacity of a mobile multiple-antenna communications link in rayleigh flat fading," *IEEE Trans. on Information Theory*, vol. 45, pp. 139–157, January 1999.
[8] N. Jindal, "A high snr analysis of mimo broadcast channels," in *Proc. IEEE Int. Symp. Information Theory, Adelaide, Australia*, 2005, pp. 2310–2314.
[9] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, New York: John Wiley and Sons, 1991.
[10] T. Cover, "Broadcast channels," *IEEE Trans. on Information Theory*, vol. 18, pp. 2–14, January 1972.
[11] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*, Cambridge, U.K. Cambridge Univ. Press, 2005.
[12] L. Zheng and D. N. C. Tse, "Communication on the grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. on Information Theory*, vol. 48, pp. 359–383, February 2002.
[13] T. Marzetta and B. Hochwald, "Fast transfer of channel state information in wireless systems," *IEEE Trans. on Signal Processing*, vol. 54, pp. 1268–1278, April 2006.
[14] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. on Information Theory*, vol. 52, pp. 3936–3964, September 2006.
[15] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Transactions on Information Theory*, vol. 49, pp. 2058–2080, April 2003.