How many Users should inform the BS about their Channel Information?

Umer Salim and Dirk Slock
Mobile Communications Department
Eurecom, France

***Abstract*— Simultaneous transmission of multiple data streams from a multiple antenna base station (BS) to multiple single antenna users gives significant gain in spectral efficiency as compared to when a single such user is being served. This simultaneous transmission to multiple users is realizable only if BS knows the forward channels linking its transmitting antennas to these users which requires channel feedback from these users. This feedback overhead could be prohibitively large especially in large user systems, limiting the multi-user transmission gains.**

Exploiting the channel reciprocity in a time-division duplexed (TDD) broadcast channel, we give a simple transmission strategy, where users feedback independent of their channel realizations. We analyze the sum rate of this multi-user system when the channel acquisition load is completely accounted for. We derive a novel lower bound of the sum rate which allows us to optimize over how many users should inform the BS about their channel information, solving the intriguing trade-off of multi-user diversity, interference cancellation and feedback overhead.

I. INTRODUCTION

The very practical scenario of one transmitter transmitting information to multiple users was introduced by Cover in [1], termed as broadcast channel. In multiple-antenna broadcast channels if a 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 [2], [3]. Apart from the multiplexing gain of M, broadcast channels enjoy another gain due to surplus number of users by selecting good users. The term multi-user diversity was coined by Knopp and Humblet in [4]. It has been shown in [5] [6] that the sum capacity of the Gaussian broadcast channel has a scaling factor with the number of users as $M \log(\log(K))$, where K is the total number of users in the system whose channel information is available at the BS.

These promising advantages of broadcast multiple-input multiple-output (MIMO) systems don't come for free. The biggest price to pay to achieve the full multiplexing gain of M is that BS must know the forward channel to at least M users [7]. To further achieve the multi-user diversity gain factor of $M \log(\log(K))$ of broadcast channel sum rate, BS should know the channel state information of all of these K users where normally K could be much larger than M.

Caire et al. studied the achievable rates for multi-user MIMO DL removing all the assumptions of CSIR and CSIT for frequency-division duplexed (FDD) systems in [8]. They compared achievable rates with a wide variety of analog and quantized feedback schemes under the assumption of infinitely large channel coherence lengths and by restricting the number of users (K) equal to M. Later in [9], training and feedback parameters were optimized as a function of channel coherence length and signal to noise ratio, although the number of users was still restricted to M.

In a recent work [10], the authors analyze the trade-off of multi-user diversity and accuracy of channel information at the BS, keeping the total number of feedback bits fixed. They conclude that accurate channel information is more important than having multi-user diversity.

In [6], the authors had given a very innovative scheme coined as Random Beam Forming (RBF) where only a few bits of feedback are required from every user and the sum rate was shown to converge to the sum capacity, obtainable through dirty paper coding (DPC) [11], but for RBF gains to be valid, the number of users in the system should be extremely large.

In [12], it was shown that, with perfect CSIT and with asymptotically large number of users, zero-forcing (ZF) precoding achieves the full multiplexing gain M and the full multi-user diversity gain $M \log(\log(K))$ of the broadcast channel.

A very important aspect, which often gets overlooked in the analysis of multi-user systems, is the consideration of channel coherence time. The channels in practice have finite coherence times which should be carefully partitioned between obtaining feedback and actual data transmission. If all users are bound to feedback their channel information, channel realization might change during the feedback process making the information obtained useless for actual data transmission.

We analyze the cost incurred and the benefit attainable of feedback in a more suitable and meaningful fashion. We make 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 scheduling/precoding/decoding of data. To make the task tractable, we simplify the problem by selecting time-division duplex (TDD) broadcast channel and assuming that perfect reciprocity holds. TDD channel with reciprocity assumption simplifies the acquisition of CSIT as uplink (UL) pilot transmission from users to BS acts as channel feedback [13], [14]. We restrict the CSIT acquisition at the BS only through training and hence we use the terms training and feedback synonymously in the sequel. So we have a fixed bandwidth available, a BS having M transmit antennas and K

single antenna users. Now this fixed bandwidth can be used for UL/DL data transmission or training/feedback. The objective would be to maximize the DL sum rate by optimizing over the number of users who feedback.

The references [13] and [15] are related to our work but there are major differences in the scope. They focus on the case when the number of users in the system is less than the number of BS antennas and the users are never trained about their effective channels. Our analysis is for the systems with larger number of users than BS transmit antennas because this setting is certainly much more practical than its opposite counterpart. Moreover, in our scheme, users are explicitly trained about their effective channels after precoding. The other major difference is in the achievable sum rate. Their sum rates are bounded in DL signal to noise ratio (SNR) and hence give zero multiplexing gain, whereas our scheme achieves full multiplexing gain.

We give a transmission strategy where a certain number of users send pilots to the BS, providing CSIT due to TDD reciprocity. These users are chosen independent of their channel realizations (hence termed as **Oblivious Users**). We derive a lower bound of the sum rate for this scheme where CSIT gains (multi-user diversity and interference cancellation) and the cost of CSIT acquisition (power and time overhead) are clearly visible. The maximization of this sum rate gives the optimal number of users who should inform the BS about their channel information, thus solving the CSIT acquisition cost-benefit trade-off.

This paper is organized as follows. Section II describes system model. In section III, the transmission strategy is proposed and the sum rate lower bound is derived. Then results for optimal number of users are obtained in section IV, followed by conclusions in section V.

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 a BS having M transmit antennas and K single-antenna user terminals. In the DL, 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 $\mathbf{H_F^\dagger} = [\mathbf{h_1 h_2} \cdots \mathbf{h_K}]$, so $\mathbf{H_F}$ 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 a transmit power constraint of P i.e., $\mathbb{E}[||\mathbf{x}||^2] \leq P$. P also denotes the DL SNR due to unit variance noise. All users have a peak per symbol power constraint of P_{pk}.

The channel is assumed to be block fading having coherence length of T symbol intervals [16]. The entries of the forward channel matrix $\mathbf{H_F}$ are independent and identically distributed (i.i.d.) complex Gaussian with zero mean and unit variance. Initially all users and BS transmitter are oblivious of the channel realization in each block.

III. Transmission Scheme with Oblivious Users

In this section, first we describe our transmission scheme and then we derive sum rate lower bound. This scheme divides the coherence length of T symbol intervals in three phases, 1) uplink training, 2) downlink training and 3) coherent data transmission.

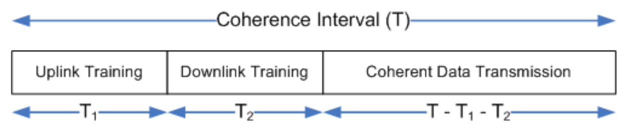

Fig. 1. Transmission Phases with Oblivious Users

The first phase is the UL training phase where K^{obl} of the K users present in the system transmit pilots to the BS. Because of TDD mode, this phase is equivalent to a feedback phase. As K^{obl} single antenna users transmit pilots, the length of this uplink training interval is $T_1 \geq K^{obl}$. Users can use orthogonal codes to be able to transmit pilots simultaneously. For k-th user channel $\mathbf{h_k}$, the BS estimate and corresponding estimation error are denoted by $\hat{\mathbf{h}}_\mathbf{k}$ and $\tilde{\mathbf{h}}_\mathbf{k}$.The mean square error (MSE) in CSIT per channel coefficient, denoted by σ_h^2, is given by

$$\sigma_h^2 = \mathbb{E}[|\mathbf{H}_{ij} - \hat{\mathbf{H}}_{ij}|^2] = \frac{1}{P_{pk}T_1 + 1} \tag{2}$$

See [17] and [13] for details about this phase. This training interval length $T_1 \geq K^{obl}$ is basically the price of obtaining CSIT at the BS through feedback which reduces the effective channel coherence time to $T - T_1$. Thus CSIT acquisition from a large number of users may not be the optimal strategy because feedback (training) from a large number of users will leave almost no time for DL data transmission in each coherence block.

We adopt ZF precoding at the BS preceded by semi-orthogonal user selection (SUS) algorithm of [12]. In ZF precoding, unit-norm beamforming vector for k-th selected user (denoted as $\bar{\mathbf{v}}_\mathbf{k}$), is selected such that it is orthogonal to the channel vectors of all other selected users. 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 is some residual interference. If we represent ZF beamforming matrix by $\bar{\mathbf{V}} = [\bar{\mathbf{v}}_\mathbf{1} \bar{\mathbf{v}}_\mathbf{2} \cdots \bar{\mathbf{v}}_\mathbf{M}]$, the transmitted signal $\mathbf{x}$ becomes $\mathbf{x} = \bar{\mathbf{V}}\mathbf{u}$ and the signal received by k-th selected user (1) can be expressed as

$$\begin{aligned} y_k &= \mathbf{h_k}^\dagger \bar{\mathbf{V}} \mathbf{u} + n_k \\ &= \mathbf{h_k}^\dagger \bar{\mathbf{v}}_\mathbf{k} u_k + \sum_{j \neq k} \mathbf{h_k}^\dagger \bar{\mathbf{v}}_\mathbf{j} u_j + n_k, \end{aligned} \tag{3}$$

where $\mathbf{u}$ is the data vector with u_k data intended for k-th selected user.

The second phase is the DL training phase where the BS transmits pilots so that the scheduled users estimate their corresponding effective channels. It was shown in [17] that only one symbol interval is sufficient to let the M selected users learn their effective scalar channels ($\mathbf{h_k}^\dagger \bar{\mathbf{v}}_\mathbf{k}$ for user k). Moreover with the fact that BS is able to transmit with sufficient power reducing the estimation error, we assume that selected users are able to estimate their effective scalar channels perfectly and we ignore the overhead of this phase.

When this second phase ends, both sides of the broadcast channel have necessary channel state information, although CSIT is 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. In this third data phase, we adopt independent data transmission with equal power allocation P/M to finally selected M users.

A. Sum Rate Lower Bound

We are interested in getting an expression for the achievable sum rate of this broadcast channel which captures the gain and the cost associated with feedback. The received signal from eq. (3) can be further written as

$$y_k = \hat{\mathbf{h}}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{k} u_k + \tilde{\mathbf{h}}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{k} u_k + \sum_{j \neq k} \tilde{\mathbf{h}}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{j} u_j + n_k. \tag{4}$$

This uses the fact that $\mathbf{h_k}^\dagger \bar{\mathbf{v}}_\mathbf{j} = \tilde{\mathbf{h}}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{j}$ for $k \neq j$ due to ZF beamforming and by splitting the effective channel $\mathbf{h_k}^\dagger \bar{\mathbf{v}}_\mathbf{k}$ in two parts, where $\hat{\mathbf{h}}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{k}$ is perfectly known at the BS. The above equation can be written as

$$y_k = \hat{\mathbf{h}}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{k} u_k + \sum_{j=1}^{M} \tilde{\mathbf{h}}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{j} u_j + n_k. \tag{5}$$

From the above equation, where we have relegated some signal part $\tilde{\mathbf{h}}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{k} u_k$ into interference and by treating all the interference terms which appear due to imperfect CSIT as an additional source of Gaussian noise as in [18] and [19], a lower bound of the SINR of k-th user can be written as

$$\mathrm{SINR}_k^{obl} = \frac{\frac{P}{M} |\hat{\mathbf{h}}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{k}|^2}{1 + \frac{P}{M} \sum_{j=1}^{M} \mathbb{E} |\tilde{\mathbf{h}}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{j}|^2}. \tag{6}$$

The variance of each interference coefficient ($\tilde{\mathbf{h}}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{j}$) can be computed based upon the fact that BS does MMSE estimation which makes estimation error $\tilde{\mathbf{h}}_\mathbf{k}$ independent of any function of channel estimates ($\hat{\mathbf{h}}_\mathbf{k}$) of which beamforming vectors are one particular example.

$$\mathbb{E} |\tilde{\mathbf{h}}_\mathbf{k}^\dagger \bar{\mathbf{v}}_\mathbf{j}|^2 = \mathbb{E} \left[\bar{\mathbf{v}}_\mathbf{j}^\dagger \mathbb{E} \left(\tilde{\mathbf{h}}_\mathbf{k} \tilde{\mathbf{h}}_\mathbf{k}^\dagger \right) \bar{\mathbf{v}}_\mathbf{j} \right] = \sigma_h^2 \mathbb{E} \left[\bar{\mathbf{v}}_\mathbf{j}^\dagger \bar{\mathbf{v}}_\mathbf{j} \right] = \sigma_h^2 \tag{7}$$

Furthermore by using $\hat{\mathbf{h}}_\mathbf{k} = \sqrt{1 - \sigma_h^2} \mathbf{g_k}$ where $\mathbf{g_k} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I_M})$ in the numerator, the SINR becomes

$$\mathrm{SINR}_k^{obl} = \frac{1 - \sigma_h^2}{1 + P\sigma_h^2} \frac{P}{M} |\mathbf{g_k}^\dagger \bar{\mathbf{v}}_\mathbf{k}|^2, \tag{8}$$

where the first term represents the SINR loss factor w.r.t. a system with perfect CSIT and CSIR as $\mathbf{g_k}$'s (and hence $\bar{\mathbf{v}}_\mathbf{k}$'s) are perfectly known at the BS. During the data phase, the lower bound (LB) of the per symbol sum rate can be written as

$$\mathrm{LB} = \sum_{k=1}^{M} \mathbb{E}_{\mathbf{g_k}} \log \left(1 + \frac{1 - \sigma_h^2}{1 + P\sigma_h^2} \frac{P}{M} |\mathbf{g_k}^\dagger \bar{\mathbf{v}}_\mathbf{k}|^2 \right), \tag{9}$$

where the users being transmitted have been selected using SUS algorithm.

If one deals with the same system (K users and M BS antennas) with perfect CSI assumption ($\sigma_h^2 = 0, T_1 = 0$), the sum rate obtained through SUS and ZF beamforming would be

$$R_{\mathrm{ZF}}(K, M, P) = \sum_{k=1}^{M} \mathbb{E}_{\mathbf{g_k}} \log \left(1 + \frac{P}{M} |\mathbf{g_k}^\dagger \bar{\mathbf{v}}_\mathbf{k}|^2 \right). \tag{10}$$

And for large user regime, it was shown in [12] to be well-approximated by

$$R_{\mathrm{ZF}}(K, M, P) \approx M \log \left(1 + \frac{P}{M} \log(K) \right). \tag{11}$$

So the lower bound of the sum rate (during the data phase) can be written in terms of the sum rate of a perfect CSI system as

$$\mathrm{LB} = R_{\mathrm{ZF}}(K^{obl}, M, P_m) \tag{12}$$

where P_m is the reduced power given by

$$P_m = \frac{1 - \sigma_h^2}{1 + P\sigma_h^2} P. \tag{13}$$

By taking into account the loss of coherence interval T due to feedback (training) interval of length T_1 (we restrict $T_1 = K^{obl}$ for simplicity), sum rate lower bound becomes

$$\mathrm{LB}^{obl} = \frac{T - K^{obl}}{T} \sum_{k=1}^{M} \mathbb{E}_{\mathbf{g_k}} \log \left(1 + \frac{1 - \sigma_h^2}{1 + P\sigma_h^2} \frac{P}{M} |\mathbf{g_k}^\dagger \bar{\mathbf{v}}_\mathbf{k}|^2 \right). \tag{14}$$

The biggest virtue of this lower bound is that it gives the achievable sum rate in terms of the sum rate of a perfect CSI system (employing SUS and ZF precoding) with loss appearing as an SNR reduction factor and as reduced multiplexing gain due to feedback interval.

We can make the approximation of large user regime as in eq. (11), following the footsteps of [12] and putting the value of σ_h^2 from eq. (2), the above sum rate will become

$$\mathrm{SR}^{obl} = \frac{T - K^{obl}}{T} M \log \left(1 + \frac{\frac{P}{M} \frac{P_{pk} K^{obl}}{P_{pk} K^{obl} + 1} \log(K^{obl})}{1 + P \frac{1}{P_{pk} K^{obl} + 1}} \right). \tag{15}$$

Due to the approximation made at this final step, this sum rate expression is not necessarily a lower bound. We have verified through extensive simulations that it closely follows the lower bound and the true sum rate of the system even for moderate number of users, although we don't plot these simulation results due to space limitations.

IV. Optimal Users to Feedback

We can maximize the final sum rate expression to evaluate the optimal number of users who should inform the BS about their channel information. Thus our objective function is

$$K^{obl*} = \arg\max_{K^{obl}} \mathrm{SR}^{obl}, \tag{16}$$

where SR^{obl} is given in eq. (15). The analytical solution for the above equation does not bear closed form result but numerical optimization trivially gives the optimal value of K^{obl}.

A. Optimal Feeding back Users vs. DL SNR

We plot the graph of the optimal number of users with SNR in Fig. 2 and also plot corresponding sum rate achieved by using that optimal number of users for each value of SNR in Fig. 3. We take T=1000 symbol intervals, there are 200 users in the system with per user peak power constraint of 5 dB and the BS is equipped with $M = 4$ antennas.

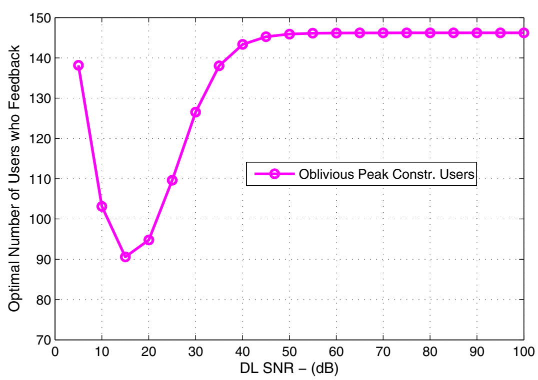

Fig. 2. Optimal Users versus SNR

The behavior of the curve of optimal number of users feeding back versus SNR is quite intriguing. At high SNR (interference limited regime), the scheme requires very good quality CSIT and due to peak power constrained users, it translates to obtaining feedback from each user for longer intervals which comes out to be a lot of users transmitting feedback (users have orthogonal codes and hence can be separated).

At low SNR, system is basically noise limited and multi-user diversity factor is very important hence the users with very strong channels should be scheduled. In our scheme (where users feedback independent of their channel realizations) requires feedback from a large number of users initially to enjoy multi-user diversity but that consumes a lot of coherence time in feedback so number of users who feedback

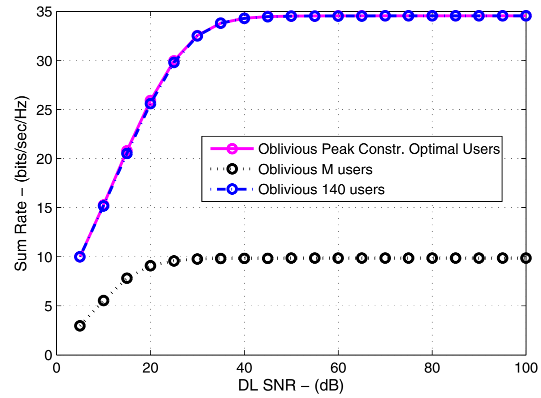

Fig. 3. Sum Rate with Optimal Users versus SNR

decrease further and later they start increasing again to provide high quality CSIT.

We have also plotted the sum rate when 140 users (this is the number of users at high DL SNR) feedback in each coherence interval. This curve overlaps fully the sum rate of the schemes with optimal feedback load (Fig. 3). It indicates that for a fixed channel coherence length, a fixed reasonable value of feeding back users (normally much larger than M) can achieve the cost-benefit trade-off of feedback significantly. In other words, the sum rate as a function of SNR is not very sensitive to the number of users who feedback.

B. Optimal Feeding back Users vs. Channel Coherence Time

We now analyze how the optimal number of users behaves with the change in channel coherence time. So we plot two graphs, one showing the optimal number of users versus coherence interval in Fig. 4 and the other showing the sum rate corresponding to the optimal number of users versus coherence interval in Fig. 5. Here BS has $M = 4$ antennas, its power constraint is 20 dB and there are 500 users in the system with each user restricted to a peak power constraint of 5 dB.

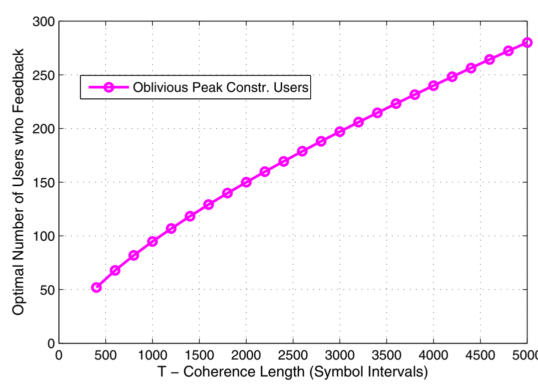

Fig. 4. Optimal Users versus Coherence Length

The curve of optimal number of users versus channel coherence time shows almost linear increase. For smaller

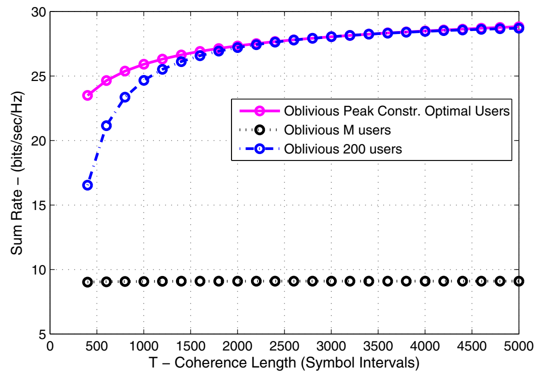

Fig. 5. Sum Rate with Optimal Users versus Coherence Length

values of the coherence interval, small number of users is optimal so that not a lot of coherence interval gets wasted in feedback. For very large values of the coherence interval, feedback from a large number of users is optimal so as to select the best users among them. Thus the number of users, who feedback, must scale up with the increase in channel coherence time.

Sum rate curves for optimal users and when only M users feedback have been plotted at $\text{SNR} = 20$ dB. Sum rate curve has also been plotted for a fixed number of users (200) feeding back. We saw that in sum rate corresponding to the optimal number of users versus SNR, a single suitable number of users feeding back (fixed feedback load) can capture the gain of optimal feedback, here it is not possible to find one such number of users (feeding back) capturing the sum rate gains for a large range of channel coherence length. So the sum rate as a function of T is relatively sensitive to the number of users who feedback.

Remark: When transmission is switched from UL to DL or vice versa, a guard interval (consisting of a few symbol intervals) must be inserted between the two due to hardware constraints. We don't take this guard interval into account as it does not affect our optimization.

V. Conclusions

We studied the sum rate of a broadcast channel under no assumption of CSI. We proposed a transmission strategy where users train the BS about their channels, independent of their channel realizations. We derived a novel lower bound of the sum rate of this scheme where the gains and the cost associated to CSIT feedback are clearly visible. Owing to its simplicity, it can be optimized to evaluate the number of users who should inform the BS about their channel information. In our system setting, the results show that sum rate is not very sensitive to the number of feeding back users as a function of SNR whereas it is sensitive as a function of channel coherence time.

Acknowledgments

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] T. Cover, "Broadcast channels," *IEEE Trans. on Information Theory*, vol. 18, pp. 2–14, January 1972.

[2] P. Viswanath and D. Tse, "Sum capacity of the multiple antenna gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. on Information Theory*, vol. 49, pp. 1912–1921, August 2003.

[3] W. Yu and J. M. Cioffi, "Sum capacity of gaussian vector broadcast channels," *IEEE Trans. on Information Theory*, vol. 50, pp. 1875–1892, September 2004.

[4] R. Knopp and P.A. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. International Conference on Communications, Seattle, USA*, June 1995, pp. 331–335.

[5] M. Sharif and B. Hassibi, "A comparison of time-sharing, DPC and beamforming for MIMO broadcast channels with many users," *IEEE Transactions on Communications*, vol. 55, pp. 11–15, January 2007.

[6] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channels with partial side information," *IEEE Transactions on Information Theory*, vol. 51, pp. 506–522, February 2005.

[7] 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.

[8] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO downlink made practical : Achievable rates with simple channel state estimation and feedback schemes," *Arxiv preprint cs.IT/0710.2642*.

[9] M. Kobayashi, G. Caire, N. Jindal, and N. Ravindran, "How much training and feedback are needed in MIMO broadcast channels?," in *Proc. IEEE Int. Symp. Information Theory*, 2008.

[10] N. Ravindran and N. Jindal, "Multi-user diversity vs. accurate channel feedback for MIMO broadcast channels," in *Proc. IEEE International Conference on Communications, Beijing, China*, 2008, pp. 3684–3688.

[11] 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.

[12] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE Journal on Selected Areas in Communications*, vol. 24, March 2006.

[13] 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.

[14] 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.

[15] J. Jose, A. Ashikhmin, P. Whiting, and S. Vishwanath, "Scheduling and pre-conditioning in multi-user MIMO TDD systems," in *Proc. IEEE International Conference on Communications, Beijing, China*, 2008, pp. 4100–4105.

[16] 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.

[17] U. Salim and D. Slock, "Broadcast channel: Degrees of freedom with no CSIR," in *Proc. Allerton Conf. on Communication, Control, and Computing*, 2008.

[18] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Transactions on Information Theory*, vol. 46, pp. 933–946, May 2000.

[19] 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.