DEGREES OF FREEDOM IN THE MISO BC WITH DELAYED-CSIT AND FINITE COHERENCE TIME: OPTIMIZATION OF THE NUMBER OF USERS

Yohan Lejosne[*†], *Dirk Slock*[†], *Yi Yuan-Wu*[*]

[*]Orange Labs, RESA/WASA
38-40 Rue du General Leclerc 92794 Issy Moulineaux Cedex 9, FRANCE
Email: {yohan.lejosne, yi.yuan}@orange.com
[†]EURECOM, Mobile Communications Dept.
Campus SophiaTech, 450 route des Chappes, 06410 Biot Sophia Antipolis, FRANCE
Email: dirk.slock@eurecom.fr

ABSTRACT

Most techniques designed for the multi-input single-output (MISO) Broadcast Channel (BC) require accurate current channel state information at the transmitter (CSIT). A different approach has been proposed by Maddah-Ali and Tse (MAT), in which significant increase of the multiplexing gain can be obtained by solely relying on perfect but outdated CSIT. This approach is proven to yield the optimal multiplexing gain when the curent channel state is completely independent of the fed back channel state. Recent work focused on an intermediate case: channels exhibiting some temporal correlation, or equivalently in the block fading model, feedback delays being smaller than the coherence time. We propose modified versions of different schemes for the MISO BC with delayed CSIT in which the number of users to serve is optimized in order to maximize the multiplexing gain when accounting for (CSI acquisition) overhead. Indeed, as the coherence time decreases, CSI acquisition overhead starts to dominate and a reduced number of active users (and hence active Base Station antennas) leads to a better compromise. We consider the MAT scheme, the traditional Zero Forcing (ZF) beamforming scheme, and combinations such as MAT-ZF and TDMA-ZF.

I. INTRODUCTION

Interference is a major limitation in wireless networks and the search for efficient ways of transmitting in this context has been productive and diversified [1]–[3]. Numerous techniques allow the increase of the multiplexing gain. For instance, in a multi-user single-cell context, dirty paper coding allows the transmitter to send information to multiple users while simultaneously pre-canceling interferences [4]. In the interference channel (IC), channel state information at the transmitter (CSIT) can be used to align the interferences from multiple receivers thereby reducing or even eliminating their impact. However, these techniques rely on perfect current CSIT which is not practical. CSIT is by nature delayed and imperfect. Though interesting results have been found concerning imperfect CSIT [5], feedback delay can also be an issue, especially if it approaches the coherence time T_c of the channel. However, a recent study [6] caused a paradigm shift by proposing a scheme yielding a larger (sum) Degree of Freedom (DoF) than TDMA (which has a DoF of one) while relying solely on perfect but outdated CSIT. This allows for some multiplexing gain even if the channel state changes arbitrarily over the feedback delay. This technique is referred to hereafter as the Maddah-Ali-Tse (MAT) scheme. The range of coherence time in which the sole use of the MAT scheme yields an increased multiplexing gain is determined in [7], [8]. In [7] the multiplexing gains provided by MAT, ZF and TDMA are compared in order to determine which strategy is the best as a function of the feedback delay and of the coherence time of the channel. However, even though the feedback delay is accounted for, the training overhead is not and the issue of CSI at the receiver (CSIR) is not addressed either. In [8], MAT and a prediction-based variant of ZF are compared but the CSI acquisition overhead is not accounted for either.

The assumption of totally independent channel variation is overly pessimistic for numerous practical scenarios. Therefore another scheme was proposed in [9] for the time correlated MISO broadcast channel with 2 users. This scheme optimally combines delayed CSIT and current CSIT (both possibly imperfect) when accounting for feedback delay only, but has not been generalized for a larger number of users. In [10] a scheme was proposed which combines zero-forcing (ZF) beamforming (BF) and the MAT scheme. We shall denote this scheme for the MISO BC with K users by MAT-ZF$_K$. It essentially performs ZF and superposes MAT only during the dead times of ZF. It has been shown that

[†] EURECOM's research is partially supported by its industrial members: ORANGE, BMW Group, Swisscom, Cisco, SFR, ST Ericsson, Symantec, SAP, Monaco Telecom, and also by the EU FP7 projects CROWN, SACRA and WHERE2.

the MAT-ZF scheme recovers the results of optimality of [9] for $K = 2$ but MAT-ZF is valid and optimal in terms of DoF for any number of users. The MAT-ZF scheme is based on a block fading model but it was shown in [10] that stationary fading can be modeled exactly as a special block fading model. The TDMA-ZF scheme is designed similarly, by performing ZF when the transmitter has CSI and TDMA when it does not. TDMA means that only one user is served at a time, it only yields 1 DoF but does not require any CSIT.

We will review the multiplexing gains that ZF, MAT, TDMA-ZF and MAT-ZF can be expected to yield in actual systems, accounting for DoF loss due to training overhead as well as the feedback on the reverse link. As opposed to [8], in the net DoF we also subtract the DoF consumed in the reverse link, as in [11]. In general, weighted net DoF could be considered as in [11] since forward and reverse link rates could have different weights. We consider here unweighted net DoF from which weighted net DoF can easily be extrapolated. Note that the ZF scheme considered in [8] is different and does not have any dead time. However, ZF BF in [8] is based on predicted CSIT only, leading to some DoF loss. Also note that the channel model in [8] is somewhat approximate as it considers the channel variation as piecewise constant in transmit blocks, and a stationary variation between the blocks.

For any of the considered schemes the DoF increases with the number of users but the feedback and training overhead also grows with the number of users. Since we observe that the performances are not always maximum while serving the maximum number of users available, we optimize the number of users to be served to maximize the net multiplexing gain.

II. SYSTEM MODEL

Consider a MISO BC with a base station (BS) with M transmit antennas and K single antenna receivers. Below we shall typically assume $K = M$ (often leading to an interchangeable use of K and M) since this leads to maximum DoF, unless the relative CSI overhead becomes too important in which case the optimal number of users K (and corresponding active BS antennas) decreases and $K < M$. The transmission is modeled by

$$y_k[t] = \mathbf{h}_k^*[t]\mathbf{x}[t] + z_k[t]$$

where $y_k[t]$ is the received signal of user k at symbol time t, depends on $\mathbf{h}_k^*[t] \in \mathcal{C}^{1\times M}$ the channel state vector, $\mathbf{x}[t] \in \mathcal{C}^{M\times 1}$ the transmit signal and $z_k[t]$ the additive white Gaussian noise (AWGN) and $(\cdot)^*$ denotes Hermitian transpose. We assume the received signals to be properly scaled so that the noise variance is unity. The channel matrix is defined as $\mathbf{H}[t] = [\mathbf{h}_1[t], \cdots, \mathbf{h}_K[t]]^* \sim \mathcal{CN}(0,1)^{K\times M}$ and remains constant over T_c symbols and changes independently between blocks.

The performance metric is the sum of number of degrees of freedom (sum DoF) (also called multiplexing gain) but we will refer to it as DoF for simplicity. It is the prelog of the sum rate. Let $R(P)$ be the ergodic throughput of a MISO BC with K receivers and transmit power P then

$$\text{DoF}(K) = \lim_{P\to\infty} \frac{R(P)}{\log_2(P)}\ . \tag{1}$$

In order to take into account the feedback cost we define the feedback overhead

$$\text{DoF}_{FB}(K) = \lim_{P\to\infty} \frac{F(P)}{\log_2(P)} \tag{2}$$

where $F(P)$ is the total feedback rate.

III. NET DoF CHARACTERIZATION

In order to compare the multiplexing gains that MAT, ZF, TDMA-ZF and MAT-ZF can be expected to obtain in actual systems, we derive their netDoFs, accounting for training overhead as well as the DoF loss due to the feedback on the reverse link. In other words we evaluate how many DoF are available for data on the forward link (we account for delay and training) and subtract the DoF spent on the reverse link for the feedback.

III-A. CSI Acquisition Overhead

III-A1. Feedback and Training

Since we are interested in the DoF_{FB} which is the scaling of the feedback rate with $\log_2(P)$ as $P \to \infty$, the noise in the fed back channel estimate can be ignored in the case of analog feedback or of digital feedback of equivalent rate. The feedback can be considered accurate, suffering only from the delay T_{fb}. We consider analog output feedback, the receivers directly feed back the training signal they receive and the transmitter performs the (downlink) channel estimation. This minimizes the delay in the feedback and takes N_t time slots assuming joint detection at the transmitter.

In each block a common training of length $T_{ct} \geq N_t$ is needed to estimate the channel as explained in [8]. To maximize the number of DoF we take $T_{ct} = N_t$. A dedicated training of 1 time slot is also needed to insure coherent reception whenever ZF is to be done according to [12].

III-A2. MAT CSIR distribution

To perform the MAT scheme each receiver needs to know the channel to all receivers. After the transmitter receives the overall CSI, it needs to be distributed to all receivers. We refer to this phase as the CSIR distribution. Since all the receivers need to have all the CSI, this could be done by broadcasting the channel states. However we can do this CSIR distribution more efficiently because each transmitter already knows its own channel (from the common training).

Let us assume that all K receivers need to know $\{h_1, h_2, \cdots, h_K\}$ but receiver $k, k \in [1, K]$, already knows h_k. Broadcasting (multicasting) all the coefficient vectors $\{h_1, h_2, \cdots, h_K\}$ would take KM channel uses and receiver

$k, k \in [1, K]$ would receive 1 useless message. Instead we broadcast the $K-1$ following messages $\{h_1 + h_2, h_1 + h_3, \cdots, h_1 + h_K\}$. Since receiver 1 already knows h_1 it can subtract it from all the messages it received. Receiver 1 then has $\{h_2, h_3, \cdots, h_K\}$ (and h_1 it already had). Similarly receiver k gets h_1 from the $k-1$st message $h_1 + h_k$ by subtracting h_k and then can extract the other h_i for $i \notin \{1, \ldots, k\}$. By doing so the CSIR distribution for K users can be done in $M(K-1)$ channel uses instead of MK. The gain is significant for small values of K.

Another solution would be to do ZF BF. This would allow to take care of the CSIR distribution in $M(K-1)$ channel uses too.

III-B. ZF

When CSI is available at the transmitter full multiplexing gain can be achieved with ZF [13], in other words it is possible to transmit 1 symbol per channel use per user with this technique. It merely relies on the transmitter using a pseudo inverse of the channel as precoder thereby zero-forcing all inter-user interference. Doing only ZF would allow to transmit 1 symbol per channel use to each user when the transmitter has CSIT and nothing otherwise, thus yielding

$$\text{DoF}(\text{ZF}_K) = K\text{DoF}(\text{ZF}_1) = K\left(1 - \frac{T_{fb}}{T_c}\right). \tag{3}$$

without taking feedback and training costs into account. Once the common training and the output feedback have been done, the transmitter has the CSI and an additional dedicated training of only one channel use, $T_{dt} = 1$, is needed to insure coherent reception according to [12]. This results in $K+1$ channel uses per block dedicated to the training yielding for K users a training overhead of

$$\text{Tr}(ZF_K) = \frac{K(K+1)}{T_c}$$

and M channels uses for the output feedback yielding a feedback overhead of

$$\text{DoF}_{FB}(ZF_K) = \frac{KM}{T_c} = \frac{K^2}{T_c}.$$

The net multiplexing gain is then:

$$\text{netDoF}(\text{ZF}_K) = K\left(1 - \frac{T_{fb}}{T_c} - \frac{2K+1}{T_c}\right). \tag{4}$$

III-C. TDMA-ZF

TDMA-ZF is a direct extension of ZF, it requires the same training and feedback as ZF. The only difference is that while the transmitter is waiting for the CSI, and not sending training symbols (i.e. during the ”dead time”), it performs TDMA since this does not require any CSIT. Thus assuring the transmission of one symbol per channel use during T_{fb} and yielding

$$\begin{aligned} \text{netDoF}(\text{TDMA-ZF}_K) &= \text{netDoF}(\text{ZF}_K) + \frac{T_{fb}}{T_c} \\ &= K\left(1 - \frac{K-1}{K}\frac{T_{fb}}{T_c} - \frac{2K+1}{T_c}\right). \end{aligned} \tag{5}$$

III-D. MAT

The MAT scheme was proposed in [6]. The authors describe an innovative approach that allows to reach a multiplexing gain of

$$\frac{K}{1 + \frac{1}{2} \cdots \frac{1}{K}} = \frac{KD}{Q} \tag{6}$$

with no *current* CSIT at all. Here $\{D, Q\} \in \mathbb{N}^2$ are such that $\frac{1}{1+\frac{1}{2}\cdots\frac{1}{K}} = \frac{D}{Q}$, where D is the least common multiple of $\{1, 2, \cdots, K\}$ and $Q = DH_K$ with $H_K = \sum_{k=1}^{K} \frac{1}{k}$. This scheme allows the transmission of D symbols in Q channel uses for each user as noted in [7].

In the MAT scheme the feedback is needed so that the transmitter can align the interferences for a user in a subspace of dimension $Q - D$. Therefore each receiver needs to feed back its channel vector $Q - D$ times. A total of $N_t(Q - D)$ time slots are needed to do the feedback over Q blocks, yielding a feedback overhead of $(KN_t)(Q - D)/QT_c$ or

$$\text{DoF}_{FB}(\text{MAT}_K) = \frac{K^2(H_K - 1)}{H_K T_c} \tag{7}$$

and the cost of the common training is similar as for ZF, it takes K time slots in each block.

To perform the MAT scheme each receiver needs to know the channels of all receivers resulting in the need for CSIR distribution. The MAT scheme can be decomposed into M phases, phase j is composed of $\frac{D}{j}$ blocks. According to [8] only $M-j+1$ antennas are active during phase j and during each block of phase j only $K-j$ channel states of $M-j+1$ coefficients need to be distributed. With the details given in [6] we understand that each block of phase j is dedicated to a subset of j users and these j users need to get the CSI of the $K - j$ other users. So for a given block the users that need the CSI are not part of the subset of users whose CSI are to be distributed and our CSIR distribution method explained in III-A2 cannot be directly exploited. However for symmetry reasons the total number of CSI to be sent is a multiple of K, for example equal to LK, and among this LK CSI, L CSI are already known at each receiver. Thus by rearranging the CSI in groups of K in which each CSI is already known by a different user we can exploit the method described in III-A2 and reduce the overall number of channel uses needed for the CSIR distribution by a factor $\frac{K-1}{K}$ compared to the one by one broadcasting strategy used in [8].

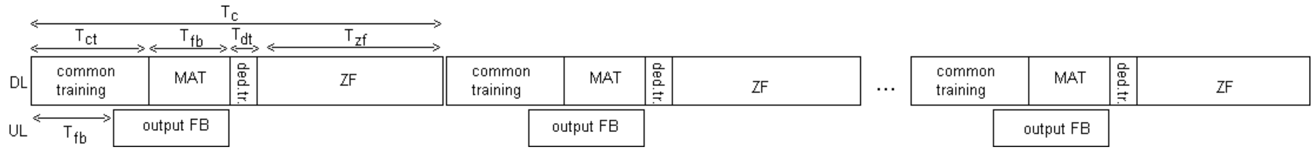

Fig. 1. The MAT-ZF scheme over a few coherence periods.

Using the CSIR distribution strategy in [8] the CSIR distribution length would be $\mathrm{L}_{CSIR}(MAT_K) = \sum_{j=1}^{K} \frac{D(K-j)(K-j+1)}{j}$, whereas it becomes

$$\mathrm{L}_{CSIR}(\mathrm{MAT}_K) = \frac{K-1}{K} \sum_{j=1}^{K} \frac{D(K-j)(K-j+1)}{j}. \tag{8}$$

using the method we described. The resulting net multiplexing gain is

$$\mathrm{netDoF}(\mathrm{MAT}_K) = K \frac{T_C - K(H_K - 1) - K}{H_K T_c + \frac{K-1}{K} \sum_{j=1}^{K} \frac{(K-j)(K-j+1)}{j}} \tag{9}$$

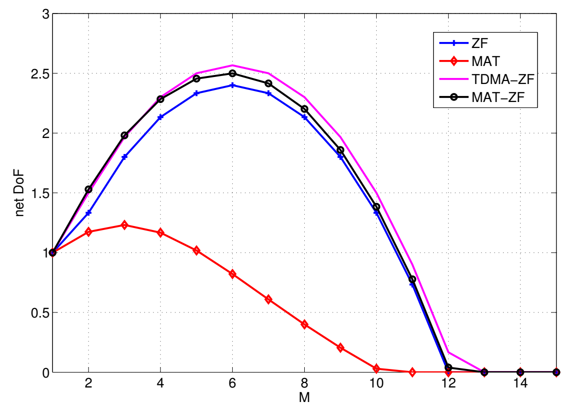

Fig. 2. netDoFs obtained by ZF, MAT, TDMA-ZF and MAT-ZF for $T_c = 30$, $T_{fb} = 5$ as a function of $M = K$.

III-E. MAT-ZF

Let us first ignore the overhead. The idea behind the MAT-ZF scheme is essentially to perform ZF and superpose MAT only during the dead times of ZF. For that purpose we consider Q blocks of T_c symbol periods and split each block into two parts. The first part, the dead times of ZF, spans T_{fb} symbol periods and the second part, the $T_c - T_{fb}$ remaining symbols. We use the first part of each block to perform the MAT scheme T_{fb} times in parallel. During the second part of each block, ZF is performed.

The sum DoF for the MAT-ZF$_K$ scheme is

$$\mathrm{DoF}(\text{MAT-ZF}_K) = K\left(1 - \frac{(Q-D)T_{fb}}{QT_c}\right). \tag{10}$$

Indeed, per user, in QT_c channel uses, the ZF portion transmits $Q(T_c - T_{fb})$ symbols, whereas the MAT scheme transmits DT_{fb} symbols.

Now we need to consider the overhead. In the MAT-ZF scheme we perform ZF and MAT. Since the training for ZF comprises the training needed for MAT, the training cost for MAT-ZF is the same as for ZF, it takes K time slots. In Fig. 1 we illustrate the composition of the blocks of this scheme with feedback and training. In order to perform MAT, the CSIR distribution is required. The scheme was initially meant to be done over Q blocks to perform the MAT scheme but we add more blocks to do the CSIR distribution. We only use the dead times of the additional coherence blocks to do the MAT CSIR distribution while we still perform ZF when the transmitter has CSI in order to avoid any degradation of the ZF DoF. The MAT part then requires $\Delta = \frac{L_{CSIR}(\mathrm{MAT}_K)}{T_{fb}}$ additional blocks. It should actually be the smallest integer no less than this fraction but by repeating the scheme more than once the number of blocks to add per scheme can be reduced to this exact value. Let $\delta = \frac{\Delta}{D}$, then the netDoF of this scheme is $\mathrm{netDoF}(\mathrm{ZF}_K) + \frac{KDT_{fb}}{T_c(Q+\Delta)}$ or

$$\mathrm{netDoF}(\text{MAT-ZF}_K) = \mathrm{netDoF}(\mathrm{ZF}_K) + \frac{T_{fb}}{T_c} \frac{K}{(H_K + \delta)} \tag{11}$$

i.e., the netDoF of ZF plus an additional term, the DoF brought about by MAT but decreased by a factor due to the CSIR distribution. From (11) we can analyze the behavior of the expected gain of MAT-ZF over ZF, it increases with T_{fb}, decreases with T_c and decreases with K since $\delta = O(K^2 \ln K)$, .

IV. OPTIMIZATION OF THE NUMBER OF USERS

In Fig. 2 the netDoFs of ZF, MAT, TDMA-ZF and MAT-ZF are plotted for $T_c = 30$, $T_{fb} = 5$ as a function of $M = K$. We notice that all the 4 schemes reach a maximum netDoF and then decrease. For each scheme there is an optimum number of users and active transmit antennas depending on the system parameters. Indeed first the DoF increases with M until a certain number beyond which the overhead becomes dominating. In this example the maximum netDoF is reached by TDMA-ZF with 6 active antennas (and 6 users).

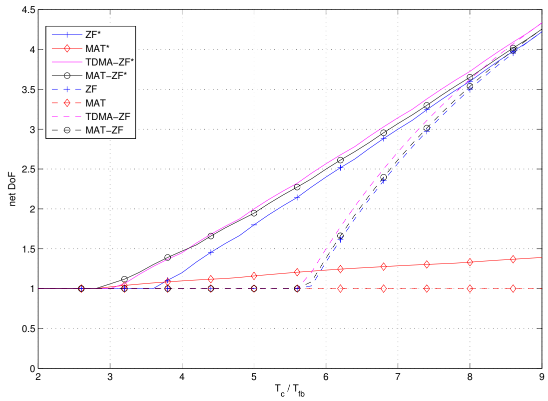

Fig. 3. netDoFs obtained by ZF, MAT, TDMA-ZF and MAT-ZF and their optimized version as a function of $\frac{T_c}{T_{fb}}$ for $M = 10$ and $T_{fb} = 5$.

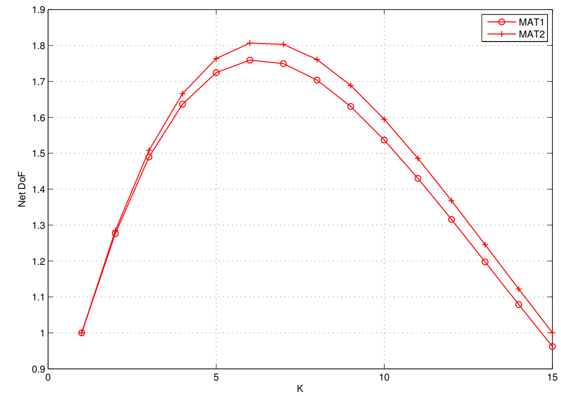

Fig. 4. netDoFs provided by MAT as a function of $K = M$ for $T_{fb} = 5$ and $T_c = 100$.

We define ZF* MAT*, TDMA-ZF* and MAT-ZF*, the respective versions of ZF, MAT, TDMA-ZF and MAT-ZF, in which the number K of served users (equal to the number of active antennas) is optimized to maximize the multiplexing gain. The degrees of freedom reached by the 4 optimized schemes are then

$$\text{netDoF}(\text{ZF}^*_M) = \arg\max_{K \le M} \text{netDoF}(\text{ZF}_K) \quad (12)$$

and similarly for MAT, TDMA-ZF and MAT-ZF

$$\text{netDoF}(\text{MAT}^*_M) = \arg\max_{K \le M} \text{netDoF}(\text{MAT}_K) \quad (13)$$

$$\text{netDoF}(\text{TDMA-ZF}^*_M) = \arg\max_{K \le M} \text{netDoF}(\text{TDMA-ZF}_K) \quad (14)$$

$$\text{netDoF}(\text{MAT-ZF}^*_M) = \arg\max_{K \le M} \text{netDoF}(\text{MAT-ZF}_K) \quad (15)$$

In Fig. 3 we evaluate the performances of ZF, MAT, TDMA-ZF and MAT-ZF and of their optimized versions ZF*, MAT*, TDMA-ZF* and MAT-ZF* in terms of multiplexing gains for $T_{fb} = 5$ and $M = 10$. Since 1 degree of freedom is always reachable, by simply doing TDMA we replace all the values that are less than 1 by 1. We observe that for a large number of users, $K = 10$, without optimization of the number of users to be served, the MAT scheme is outperformed by a simple SISO transmission (TDMA) at least until $\frac{T_c}{T_{fb}} > 9$. For ZF, TDMA-ZF and MAT-ZF for large value of the ratio $\frac{T_c}{T_{fb}}$ there is no difference between the original schemes and their optimized versions in other words it is best to use the maximum number $M = 10$ of transmit antennas and serve $K = 10$ users. On the other hand for small values of the ratio significant gains are observed by reducing the number of active antennas and served users. Indeed when the ratio $\frac{T_c}{T_{fb}}$ is small the overhead becomes too large, for instance for $T_c = 6T_{fb}$ TDMA-ZF achieves 1.5 DoF whereas its optimized version reaches more than 2.5 DoF by finding the right compromise, serving enough users to increase the DoF but not too many to limit the overhead.

It is also remarkable that TDMA-ZF generally outperforms MAT-ZF. This can be explained by the cost of the CSIR distribution needed for the MAT scheme. TDMA-ZF brings an increase of $\frac{T_{fb}}{T_c}$ DoF over ZF while MAT-ZF brings $\frac{T_{fb}}{T_c} \frac{K}{(H_K+\delta)}$ DoF over ZF. Since $\frac{K}{(H_K+\delta)}$ can only be larger than 1 for very small values of K or large T_{fb}, it is only in those case that MAT can beat TDMA.

V. MAT

We proposed an improvement of the CSIR distribution but MAT still needs a very long coherence time for the CSIR distribution to be less penalizing. In Fig. 4 we plot the net DoF yielded by MAT with the CSIR distribution as in [8] (MAT1) and with our new CSIR distribution according to (8) (MAT2) for $T_c = 100$, $T_{fb} = 5$ as a function of $M = K$. We observe that especially for intermediate values of K the optimization of the CSIR distribution brings a limited but non negligible gain.

In Fig. 2 and Fig. 3 we notice that MAT barely outperforms simple TDMA transmission (1 DoF) for the parameters considered. In Fig. 5 we plot the net DoF obtained by MAT for $T_{fb} = 5$ and $T_c \in \{16, 64, 256, 1024, 4096, 16384\}$ as a function of the number of users K as well as the the asymptotic DoF of MAT, $\frac{K}{H_K}$. We observe that for example for $K = 5$, the asymptotic DoF of 2.19 is almost reached if $T_c \ge 1024$ but decreases by 8% for $T_c = 256$. For $K = 10$, $T_c = 1024$ only allows to reach 93% of the asymptotic DoF. The MAT scheme is impaired by the CSIR distribution overhead that grows quickly with the number of users, even with our optimized broadcast.

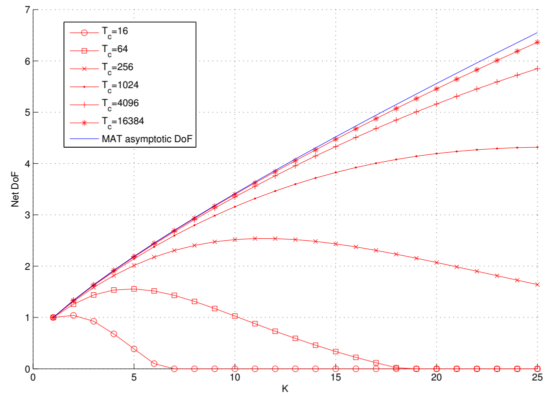

Fig. 5. netDoFs provided by MAT as a function of $K = M$ for $T_{fb} = 5$ and $T_c \in \{16, 64, 256, 1024, 4096, 16384\}$.

VI. CONCLUDING REMARKS

We proposed a variant of the MAT-ZF and TDMA-ZF schemes in which the number of users to be served and corresponding number of active antennas is optimized. For the sake of comparison we also optimized the number of users to be served and corresponding number of active antennas in the ZF and MAT algorithms. The optimization brings significants gains when the coherence time gets really short because it assures that the relative overhead will not become too large.

This optimization leads for extremely short coherence times to TDMA with $K = 1$ being optimal, which does not require any CSIT (apart from rate information).

Even with the optimized versions of the schemes we observe that the relative benefit of adding the MAT component over just simply considering the ZF scheme is debatable. While leading to a theoretically optimal combination, the practical complication is often unjustified, given the time span over which the MAT scheme needs to be implemented, especially for a larger number of users K. Moreover MAT-ZF is often outperformed by the simpler association of TDMA and ZF.

VII. REFERENCES

[1] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna gaussian broadcast channel," *Information Theory, IEEE Transactions on*, vol. 49, no. 7, pp. 1691–1706, July 2003.

[2] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and Wei Yu, "Multi-cell mimo cooperative networks: A new look at interference," *Selected Areas in Communications, IEEE Journal on*, vol. 28, no. 9, pp. 1380 –1408, december 2010.

[3] V.R. Cadambe and S.A. Jafar, "Interference alignment and degrees of freedom of the -user interference channel," *Information Theory, IEEE Transactions on*, vol. 54, no. 8, pp. 3425 –3441, aug. 2008.

[4] P. Viswanath and D.N.C. Tse, "Sum capacity of the vector gaussian broadcast channel and uplink-downlink duality," *Information Theory, IEEE Transactions on*, vol. 49, no. 8, pp. 1912 – 1921, aug. 2003.

[5] A. Lapidoth, S. Shamai, and M. A. Wigger, "On the capacity of fading mimo broadcast channels with imperfect transmitter side-information," *Arxiv preprint arxiv:0605079*, 2006.

[6] M.A. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," in *Proc. Allerton*, Monticello, IL, USA, Oct. 2010.

[7] J. Xu, J.G. Andrews, and S.A. Jafar, "Broadcast channels with delayed finite-rate feedback: Predict or observe?," *Arxiv preprint arxiv:1105.3686*, 2011.

[8] M. Kobayashi and G. Caire, " On the net DoF comparison between ZF and MAT over time-varying MISO broadcast channels," in *Proc. ISIT*, Boston, MA, USA, Jun. 2012.

[9] S. Yang, M. Kobayashi, D. Gesbert, and X. Yi, "Degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," in *Proc. ISIT*, Boston, MA, USA, Jun. 2012.

[10] Y. Lejosne, D. Slock, and Y. Yuan-Wu, " Degrees of Freedom in the MISO BC with Delayed-CSIT and Finite Coherence Time: a Simple Optimal Scheme," in *Proc. ICSPCC*, Aug. 2012.

[11] C. Suh, I.H. Wang, and D. Tse, "Two-way Interference Channels," in *Proc. ISIT*, Boston, MA, USA, Jun. 2012.

[12] U. Salim and D. Slock, "How much feedback is required for TDD multi-antenna broadcast channels with user selection ?," *EURASIP Journal on Advances in Signal Processing*, 05 2010.

[13] N. Jindal, "Mimo broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory,*, vol. 52, no. 11, pp. 5045 –5060, nov. 2006.