

# NET DEGREES OF FREEDOM OF RECENT SCHEMES FOR THE MISO BC WITH DELAYED CSIT AND FINITE COHERENCE TIME

Yohan Lejosne<sup>\*†</sup>, Dirk Slock<sup>†</sup>, Yi Yuan-Wu<sup>\*</sup>

<sup>\*</sup>Orange Labs, RESA/WASA

38-40 Rue du General Leclerc, 92794 Issy Moulineaux Cedex 9, FRANCE

Email: {yohan.lejosne, yi.yuan}@orange.com

<sup>†</sup>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) which is not a realistic assumption because of feedback delay. A novel approach by Lee and Heath, space-time interference alignment, proves that in the underdetermined (overloaded) MISO BC with  $N_t$  transmit antennas and  $K = N_t + 1$  users  $N_t$  (sum) Degrees of Freedom (DoF) are achievable if the feedback delay is not too big, thus disproving the conjecture that any delay in the feedback necessarily causes a DoF loss. We explain this approach a bit more succinctly and evaluate the net DoF that this scheme can be expected to yield in a realistic system by taking into account the cost of CSIT acquisition (training and feedback). We term the resulting scheme ST-ZF, referring to the use of Space-Time Zero Forcing precoding. The net DoF comparison with TDMA-ZF, MAT-ZF and MAT shows that ST-ZF is also of interest in practice.

## 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]. However, most techniques rely on perfect current CSIT which is not realistic. CSIT is by nature delayed and imperfect. Though interesting results have been found

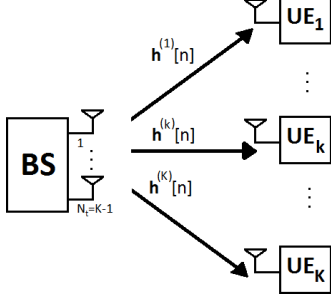
concerning imperfect [5], or quantized [6] CSIT, feedback delay can also be an issue especially if it approaches the coherence time  $T_c$  of the channel. However, a recent study [7] caused a paradigm shift by proposing a scheme yielding more than one degree of freedom (DoF) while relying solely on perfect but outdated CSIT. This technique is referred to hereafter as the Maddah-Ali-Tse (MAT) scheme. MAT allows for some multiplexing gain even if the channel state changes arbitrarily over the feedback delay. The range of coherence time in which the sole use of MAT yields an increased multiplexing gain is determined in [8] and [9] but considering only feedback or only training overheads and not both.

The assumption of totally independent channel variation is overly pessimistic for numerous practical scenarios. Therefore another scheme was proposed in [10] for the time correlated MISO broadcast channel with 2 users. This scheme optimally combines delayed CSIT and current CSIT (both imperfect) but has not been generalized for a larger number of users. Another scheme that simply performs ZF and superposes MAT only during the dead times of ZF has been proposed in [11]. This scheme, hereafter referred to as MAT-ZF recovers the results of optimality of [10] for  $K = 2$  and is valid for any number of users but is based on a block fading model.

It was generally believed that any delay in the feedback necessarily causes a DoF loss. However, Lee and Heath in [12] proposed a scheme that achieves  $N_t$  (sum) DoF in the block fading underdetermined MISO BC with  $N_t$  transmit antennas and  $K = N_t + 1$  users if the feedback delay is small enough ( $\leq \frac{T_c}{K}$ ). We review this ST-ZF scheme and evaluate the net DoF it can be expected to yield in actual systems, accounting for training overhead as well as the DoF consumption due to the feedback on the reverse link. We then compare it to the net multiplexing gains that ZF, MAT, TDMA-ZF and MAT-ZF can be expected to yield in actual systems.

<sup>†</sup> EURECOM's research is partially supported by its industrial members: ORANGE, BMW Group, Swisscom, Cisco, SFR, ST Ericsson, Symantec, SAP, Monaco Telecom.

<sup>\*</sup>Part of this work has been performed in the framework of the FP7 project ICT-317669 METIS, which is partly funded by the European Union. The authors would like to acknowledge the contributions of their colleagues in METIS, although the views expressed are those of the authors and do not necessarily represent the project.



**Fig. 1.** Underdetermined MISO BC, a base station (BS) with  $N_t = K - 1$  antennas and  $K$  single antenna users.

## II. SYSTEM MODEL

We consider a MISO BC with  $K$  single antenna users and a transmitter equipped with  $N_t = K - 1$  antennas. It is typically assumed that  $K = N_t$  since having  $K = N_t$  or  $K > N_t$  single antenna users results in the same maximum sum DoF. However, we will see that having an extra user (underdetermination/overloading) becomes useful when there is some delay in the feedback since space-time precoding (instead of spatial only beamforming) can compensate for the delay in feedback in the MISO BC. An illustration of this BC channel is given in Fig. 1. We categorize this BC as underdetermined because there are more users than transmit antennas, which prevents purely spatial ZF. The signal received by user  $k$  is given by

$$\mathbf{y}^{(k)}[n] = \mathbf{h}^{(k)T}[n]\mathbf{x}[n] + v^{(k)}[n]$$

where  $\mathbf{x}[n] \in \mathbb{C}^{N_t \times 1}$  is the signal sent by the transmitter,  $\mathbf{h}^{(k)}[n] = [h_1^{(k)}[n], h_2^{(k)}[n], \dots, h_{N_t}^{(k)}[n]]^T \in \mathbb{C}^{N_t \times 1}$  and  $v^{(k)}[n] \sim \mathcal{CN}(0, 1)$  respectively denote the channel vector and the independent and identically distributed (i.i.d.) Gaussian noise for user  $k$ . We consider a block fading model: the channel coefficients are constant for the channel coherence time  $T_c$  and change independently between blocks.  $T_{fb}$  is the feedback delay.

The performance metric is (sum) DoF (also called multiplexing gain), it is the prelog of the sum rate. In order to take into account the feedback cost, we define the feedback overhead. Let  $R(P)$  be the ergodic (sum) throughput of a MISO BC with  $N_t$  transmit antennas and  $K$  single antenna receivers and transmit power  $P$  and  $F(P)$  the total feedback rate then

$$\text{DoF} = \lim_{P \rightarrow \infty} \frac{R(P)}{\log_2(P)}; \quad \text{DoF}_{FB} = \lim_{P \rightarrow \infty} \frac{F(P)}{\log_2(P)}.$$

## III. SPACE-TIME ZERO FORCING (ST-ZF) PRECODING FOR THE UNDERDETERMINED MISO BC WITH DELAYED ONLY CSIT

Lee and Heath [12] proposed a (LH) scheme to achieve  $N_t$  DoF in the MISO BC with  $K = N_t + 1$  users when  $\gamma = \frac{T_{fb}}{T_c} \leq \frac{1}{K}$ . We first review the scheme for the general

$K$ -user case when  $\gamma = \frac{1}{K}$  without accounting for the training and feedback costs.

Assuming  $\gamma = \frac{1}{K}$  means that the current CSI is known at the transmitter only after the first  $T_{fb} = \frac{T_c}{K}$  channel uses. The LH scheme allows the transmission of  $N_t$  messages to each of the  $K$  users in  $K$  channel uses scattered over  $K$  coherence blocks. More precisely, the LH scheme uses channel uses  $\{n_1, n_2, \dots, n_K\}$  respectively in blocks  $\{n + 1, n + 2, \dots, n + K\}$ . This results in a transient regime for the first  $K$  blocks after which we have  $KT_{fb}$  instances of the LH scheme in each block assuring the  $N_t$  DoF announced in the steady state. We now focus on one instance of the LH scheme scattered over blocks  $n + 1$  to  $n + K$  for a  $n \geq K$  so that we are in steady state. Only the channel use  $n_1$  in the first block corresponds to the transmitter not having the current CSI.

Messages  $\mathbf{s}^{(k)} = [s_1^{(k)}, \dots, s_{N_t}^{(k)}]^T$ , are intended for user  $k$ ,  $k \in [1, K]$ .  $\mathbf{H}[n] = [\mathbf{h}^{(1)}[n], \dots, \mathbf{h}^{(K)}[n]]^T$  represents the channel matrix during block  $n$  and  $\mathbf{y}[n_j] = [y^{(1)}[n_j], \dots, y^{(K)}[n_j]]^T$  the received signal at the receivers during channel use  $n_j$ . Since we are interested in the DoF provided by the scheme, we hereafter omit the noise variables to be concise. The transmitter always sends a combination of all symbols at each channel use, always the same symbols for an instance of the scheme but with time-varying beamforming matrices  $\mathbf{V}^{(k)}[n_j] \in \mathbb{C}^{N_t \times N_t}$

$$\mathbf{x}[n_j] = \sum_{k=1}^K \mathbf{V}^{(k)}[n_j] \mathbf{s}^{(k)}.$$

During the first channel use  $n_1$ , the transmitter does not have any information on the current channel state, so for  $k \in [1, K]$ ,  $\mathbf{V}^{(k)}[n_1] = \mathbf{I}_{N_t}$ , the  $N_t$  by  $N_t$  identity matrix, is as good as any other matrix of full rank. The transmission scheme is summarized as follows

$$\begin{bmatrix} \mathbf{y}[n_1] \\ \mathbf{y}[n_2] \\ \vdots \\ \mathbf{y}[n_K] \end{bmatrix} = \text{diag}(\mathbf{H}[n+1], \mathbf{H}[n+2], \dots, \mathbf{H}[n+K]) * \begin{bmatrix} \mathbf{I}_{N_t} & \dots & \mathbf{I}_{N_t} \\ \mathbf{V}^{(1)}[n_2] & \dots & \mathbf{V}^{(K)}[n_2] \\ \vdots & & \vdots \\ \mathbf{V}^{(1)}[n_K] & \dots & \mathbf{V}^{(K)}[n_K] \end{bmatrix} \begin{bmatrix} \mathbf{s}^{(1)} \\ \mathbf{s}^{(2)} \\ \vdots \\ \mathbf{s}^{(K)} \end{bmatrix} = \begin{bmatrix} \mathbf{H}[n+1] & \dots & \mathbf{H}[n+1] \\ \mathbf{H}[n+2]\mathbf{V}^{(1)}[n_2] & \dots & \mathbf{H}[n+2]\mathbf{V}^{(K)}[n_2] \\ \vdots & & \vdots \\ \mathbf{H}[n+K]\mathbf{V}^{(1)}[n_K] & \dots & \mathbf{H}[n+K]\mathbf{V}^{(K)}[n_K] \end{bmatrix} \begin{bmatrix} \mathbf{s}^{(1)} \\ \mathbf{s}^{(2)} \\ \vdots \\ \mathbf{s}^{(K)} \end{bmatrix}$$

The received signal at user  $i$  and time  $n_1$  is

$$y^{(i)}[n_1] = \sum_{k=1}^K \mathbf{h}^{(i)}[n+1] \mathbf{s}^{(k)} = \mathbf{h}^{(i)}[n+1] \sum_{k=1}^K \mathbf{s}^{(k)}.$$

The beamforming matrices are constructed so that the interference alignment is simply done at each receiver by a subtraction:  $y^{(i)}[n_j] - y^{(i)}[n_1], j \in [2, K]$ . For user  $i$ , at time  $n_j, j \in [2, K]$ , we have

$$y^{(i)}[n_j] - y^{(i)}[n_1] = \sum_{k=1}^K \left( \mathbf{h}^{(i)}[n+j] \mathbf{V}^{(k)}[n_j] - \mathbf{h}^{(i)}[n+1] \right) \mathbf{s}^{(k)}$$

so the interferences are aligned if

$$\mathbf{h}^{(i)}[n+j] \mathbf{V}^{(k)}[n_j] - \mathbf{h}^{(i)}[n+1] = \mathbf{0}, \forall i \neq k.$$

In other words the beamforming matrices  $\mathbf{V}^{(k)}[n_j]$  should transform  $\mathbf{h}^{(i)}[n+j]$  into  $\mathbf{h}^{(i)}[n+1]$  for  $i \neq k$  so that the same interferences are received at any time  $n_j, j \in [1, K]$ . This is done by defining the beamforming matrix for user  $i$  and time  $n_j$  as follows

$$\mathbf{V}^{(i)}[n_j] = \begin{bmatrix} \mathbf{h}^{(1)}[n+j] \\ \vdots \\ \mathbf{h}^{(i-1)}[n+j] \\ \mathbf{h}^{(i+1)}[n+j] \\ \vdots \\ \mathbf{h}^{(K)}[n+j] \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{h}^{(1)}[n+1] \\ \vdots \\ \mathbf{h}^{(i-1)}[n+1] \\ \mathbf{h}^{(i+1)}[n+1] \\ \vdots \\ \mathbf{h}^{(K)}[n+1] \end{bmatrix}$$

for  $j \in [2, K]$  which assures

$$\begin{bmatrix} y^{(i)}[n_2] - y^{(i)}[n_1] \\ y^{(i)}[n_3] - y^{(i)}[n_1] \\ \vdots \\ y^{(i)}[n_K] - y^{(i)}[n_1] \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{h}^{(i)}[n+2] \mathbf{V}^{(i)}[n_2] - \mathbf{h}^{(i)}[n+1] \\ \mathbf{h}^{(i)}[n+3] \mathbf{V}^{(i)}[n_3] - \mathbf{h}^{(i)}[n+1] \\ \vdots \\ \mathbf{h}^{(i)}[n+K] \mathbf{V}^{(i)}[n_K] - \mathbf{h}^{(i)}[n+1] \end{bmatrix}}_{\mathbf{H}_{\text{eff}}^{(i)}} \mathbf{s}^{(i)}$$

and user  $i$  can decode  $\mathbf{s}^{(i)}$  since the rank of the  $N_t \times N_t$   $\mathbf{H}_{\text{eff}}^{(i)}$  is almost surely  $N_t$  because all channel vectors are independent with a continuous distribution. This scheme allows to transmit a total of  $N_t K$  independent data symbols in  $K$  channels uses thus yielding  $N_t$  DoF.

## IV. NET DoF CHARACTERIZATION

In order to compare the multiplexing gains that MAT, ZF, TDMA-ZF, MAT-ZF and ST-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. Note that for MAT, ZF, TDMA-ZF and MAT-ZF we consider the square case  $K = N_t$  since adding one user would only increase the overhead and not the DoF.

### IV-A. CSI Acquisition Overhead

#### IV-A1. Feedback and Training

Since we are interested in the  $\text{DoF}_{FB}$  which is the scaling of the feedback rate with  $\log_2(P)$  as  $P \rightarrow \infty$ , 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.

In each block a common training of length  $T_{ct} \geq N_t$  is needed to estimate the channel as explained in [9]. To maximize the number of DoF we take  $T_{ct} = N_t$ . A dedicated training of 1 pilot is also needed to insure coherent reception whenever ZF is to be done according to [13].

In Fig. 2 the shape of the blocks with feedback and training (common and dedicated) is presented, D. Tr. stands for dedicated training. The time slots available after accounting for training in each block are divided into two parts: a first one during which the transmitter does not have the current CSI a second during which there is CSIT.

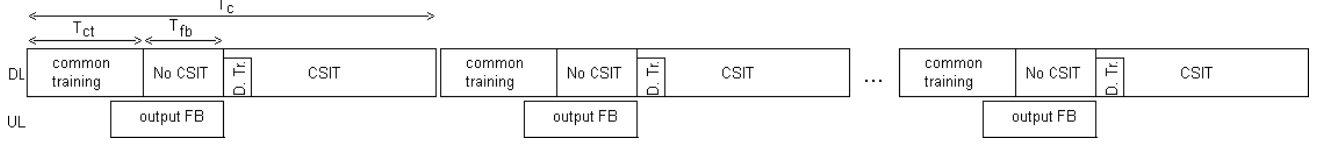
#### IV-A2. MAT CSIR distribution

To perform the MAT scheme each receiver needs the channel of certain other receivers. As a first approximation we consider that after the transmitter received the CSI from all receivers it sends them to all receivers. We refer to this phase as the CSIR distribution. This could be done by broadcasting the channel states however it can be done in a slightly more efficient fashion because each transmitter already knows its own channel.

Let us assume that all the  $K$  receivers need to know  $\{h_1, h_2, \dots, h_K\}$  but receiver  $k, k \in [1, K]$  already knows  $h_k$ . Multicasting all the coefficients  $\{h_1, h_2, \dots, h_K\}$  would take  $K$  channel uses and receiver  $k, k \in [1, K]$  would receive 1 useless message,  $h_k$  that it already has. Instead we broadcast the  $K - 1$  following messages  $\{h_1 + h_2, h_1 + h_3, \dots, h_1 + h_K\}$ . Since receiver 1 already knows  $h_1$  it can subtract it from all the messages it received. Receiver 1 then has  $\{h_2, h_3, \dots, h_K\}$  (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  $i \notin \{1, k\}$ . By doing so the CSIR distribution for  $K$  users can be done in  $N_t(K - 1)$  channel uses instead of  $N_t K$ . The gain is limited but significant for small values of  $K$ . For MAT-ZF another solution is to do the distribution using ZF. This would allow for CSIR distribution in  $N_t(K - 1)$  channel uses too but requires CSIT.

### IV-B. ZF

When CSI is available at the transmitter, the full multiplexing gain can be achieved with ZF [14]. Doing only ZF would allow to transmit 1 symbol per channel use to each user when the transmitter has CSIT and nothing otherwise.



**Fig. 2.** Topology of a few blocks considering the training (common and dedicated) and output feedback.

Without taking feedback and training into account it yields

$$\text{DoF}(\text{ZF}_{N_t}) = N_t \text{DoF}(\text{ZF}_1) = N_t \left(1 - \frac{T_{fb}}{T_c}\right).$$

The needed common and dedicated trainings occupy  $N_t + 1$  time slots and the output feedback of  $N_t$  symbols per user results in a feedback overhead of

$$\text{DoF}_{FB}(\text{ZF}_{N_t}) = \frac{KN_t}{T_c} = \frac{N_t^2}{T_c}.$$

since for ZF the squared case  $K = N_t$  is considered. The net multiplexing gain is then

$$\text{netDoF}(\text{ZF}_{N_t}) = N_t \left(1 - \frac{T_{fb}}{T_c} - \frac{2N_t + 1}{T_c}\right). \quad (1)$$

#### IV-C. TDMA-ZF

TDMA-ZF is a direct extension of ZF. The only difference being that while the transmitter is waiting for the CSI, and not sending training symbols it performs TDMA transmission since this does not require any CSIT, thus yielding

$$\begin{aligned} \text{netDoF}(\text{TDMA-ZF}_{N_t}) &= \text{netDoF}(\text{ZF}_{N_t}) + \frac{T_{fb}}{T_c} \\ &= N_t \left(1 - \frac{N_t - 1}{N_t} \frac{T_{fb}}{T_c} - \frac{2N_t + 1}{T_c}\right) \end{aligned} \quad (2)$$

#### IV-D. MAT

The MAT scheme was proposed in [7]. The authors describe an original approach that yields a multiplexing gain of

$$\frac{N_t}{1 + \frac{1}{2} \cdots \frac{1}{N_t}} = \frac{N_t D}{Q}$$

with no *current* CSIT at all, without accounting for feedback and training overheads. Here  $\{D, Q\} \in \mathbb{N}^2$  are such that  $\frac{1}{1 + \frac{1}{2} \cdots \frac{1}{N_t}} = \frac{D}{Q}$ , where  $D$  is the least common multiple of  $\{1, 2, \dots, N_t\}$  and  $Q = DH_{N_t}$  with  $H_{N_t} = \sum_{m=1}^{N_t} \frac{1}{m}$ . This scheme allows the transmission of  $D$  symbols in  $Q$  time slots for each user as noted in [8].

The MAT scheme is composed of  $N_t$  phases, phase  $j$  is composed of  $\frac{D}{j}$  slots. Multiple instances of the MAT scheme are performed in parallel, the first block is filled with first messages of as many instances of the MAT scheme as possible, then the second block is used for the second message of each instance of the MAT scheme and so on.

Only  $N_t - j + 1$  antennas are active during phase  $j$  therefore the length of the common training needed in each block depends on the phase:  $N_t - j + 1$  for any block of phase  $j$ . This leads to an empty space of  $j - 1$  time slots in each block of phase  $j$ .

Each block of phase  $j$  is dedicated to a subset of  $j$  users and the CSI needed from this block is the CSI of the  $K - j$

other users that will be used to generate the messages for phase  $j + 1$  resulting in a feedback overhead of  $(K - j)(N_t - j + 1)$  per block of phase  $j$

$$\begin{aligned} \text{DoF}_{FB}(\text{MAT}_{N_t}) &= \frac{\sum_{j=1}^K \frac{D}{j} (K - j)(N_t - j + 1)}{QT_c} \\ &= \sum_{j=1}^K \frac{(K - j)(N_t - j + 1)}{jH_{N_t}T_c} \end{aligned}$$

With the details given in [7] we understand that these CSI of the  $K - j$  users also need to be distributed to the  $j$  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 IV-A2 cannot be directly exploited. However for symmetry reasons the total length of CSI to be sent is a multiple of the number of users  $K$ , for example equal to  $LK$ . Among these  $LK$  coefficients,  $L$  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 IV-A2 and reduce the total 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 [9]. Using strategy in [9] the CSIR distribution length would be  $\sum_{j=1}^K \frac{D(K-j)(N_t-j+1)}{j}$  and it becomes

$$L_{CSIR}(\text{MAT}_{N_t}) = \frac{K-1}{K} \sum_{j=1}^K \frac{D(K-j)(N_t-j+1)}{j}$$

using the method we just described. This CSIR distribution can be partially taken care of in the empty space of  $\sum_{j=1}^K \frac{D}{j} (j-1) = D(K - H_K)$  time slots over the  $Q$  blocks due to the decreasing length of the common training in each phase. It leaves  $\frac{K-1}{K} \sum_{j=1}^K \frac{D(K-j)(N_t-j+1)}{j} - D(K - H_K) = D \left( \left( \frac{K-1}{K} \sum_{j=1}^K \frac{(K-j)(N_t-j+1)}{j} \right) + H_k - K \right)$  remaining time slot to be used for the CSIR distribution. This is always greater than 0 for  $K \geq 2$  and we note that there always is more CSIR to be distributed than empty space in any block because there is no empty space in the first phase which takes  $D$  blocks. Since we assumed  $K = N_t$  the resulting net multiplexing gain is

$$\begin{aligned} \text{netDoF}(\text{MAT}_{N_t}) &= \\ &= \frac{N_t(T_c - N_t) - \sum_{j=1}^K \frac{1}{j} (K - j)(N_t - j + 1)}{H_K T_c + \left( \left( \frac{K-1}{K} \sum_{j=1}^K \frac{(K-j)(N_t-j+1)}{j} \right) + H_k - K \right)} \end{aligned} \quad (3)$$

#### IV-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 without accounting for the overhead is

$$\text{DoF}(\text{MAT-ZF}_{N_t}) = N_t \left( 1 - \frac{(Q-D)T_{fb}}{QT_c} \right).$$

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 ZF and MAT are performed. Since the training for ZF comprises the training needed for MAT, the training cost for MAT-ZF is the same as for ZF with a length of  $N_t + 1$  time slots. But in order to perform MAT, the CSIR distribution is also 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 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  $\Delta = \frac{L_{\text{CSIR}}(\text{MAT}_{N_t})}{T_{fb}}$  additional blocks. The case of a non integer  $\Delta$  can be dealt with by repeating the scheme until the total number of blocks to add is an integer. Let  $\delta = \Delta/D$ , then the netDoF of this scheme is  $\text{netDoF}(\text{ZF}_{N_t}) + N_t DT_{fb}/(T_c(Q + \Delta))$  or

$$\text{netDoF}(\text{MAT-ZF}_{N_t}) = \text{netDoF}(\text{ZF}_{N_t}) + \frac{T_{fb}}{T_c} \frac{N_t}{(H_{N_t} + \delta)} \quad (4)$$

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.

#### IV-F. ST-ZF

The LH scheme yields  $N_t$  DoF for  $\gamma = \frac{1}{K}$  with  $K = N_t + 1$  users or for  $\gamma < \frac{1}{K}$  by doing ZF the remaining time. We refer to this scheme as ST-ZF since it is a space-time (ST) precoding, which is combined with ZF for  $\gamma < \frac{1}{K}$ . Simple ZF can only serve  $N_t$  users at a time and as mentioned earlier 1 channel use of dedicated training is needed for synchronization per subset of  $N_t$  users so instead of alternatively ZF to the  $K$  subsets of users  $\{1, \dots, i-1, i+1, \dots, K\}$  in one block, it is less expensive to ZF to only one subset of  $K-1$  out of  $K$  users per block and alternate over different blocks to assure fairness.

To allow the receivers to learn their channels a common training sequence is needed so together with the dedicated training it takes  $(N_t + 1)$  time slots and  $N_t + 1$  time slots for the output feedback (by choosing  $K = N_t + 1$  times a different subset of  $K-1$  users feeding back) resulting in a

part without current CSIT of length  $T_{fb} + 1$  instead of  $T_{fb}$  (therefore reducing the *with CSIT* part of one time slot too). This yields a feedback overhead of

$$\text{DoF}_{FB}(\text{ST-ZF}_{N_t}) = \frac{N_t K}{T_c} = \frac{N_t(N_t + 1)}{T_c}.$$

To decode its signal the receiver  $i$  needs to know  $\mathbf{H}_{\text{eff}}^{(i)}$ . The ST-ZF schemes spans over  $K$  (blocks), but the different instances of the scheme overlap: the  $n$ th instance spans over blocks  $n+1$  to  $n+K$ , the  $n+1$ th over blocks  $n+2$  to  $n+K+1$ . So only the last line of  $\mathbf{H}_{\text{eff}}$  is new in each instance. Therefore each receiver actually only needs to get  $N_t$  coefficients of its  $\mathbf{H}_{\text{eff}}^{(i)}$  (the last line). Using the CSIT part to transmit these coefficients takes  $K = N_t + 1$  time slots by sending each time slot 1 message to each user of a different subset of  $N_t = K-1$  users. The net multiplexing gain is then

$$\begin{aligned} \text{netDoF}(\text{ST-ZF}_{N_t}) &= N_t \frac{T_c - 2(N_t + 1)}{T_c} - \frac{N_t(N_t + 1)}{T_c} \\ &= N_t \left( 1 - \frac{3(N_t + 1)}{T_c} \right) \end{aligned} \quad (5)$$

where the subtraction  $N_t \frac{2(N_t+1)}{T_c}$  is due to training and transmission of  $\mathbf{H}_{\text{eff}}$  and the subtraction  $\frac{N_t(N_t+1)}{T_c}$  is due to feedback, as long as  $\frac{T_{fb}+1}{T_c-2(N_t+1)} \leq \frac{1}{K} \Leftrightarrow T_c \geq K(T_{fb}+3)$  since ST-ZF needs a with CSIT part  $K-1 = N_t$  times longer than the no current CSIT part.

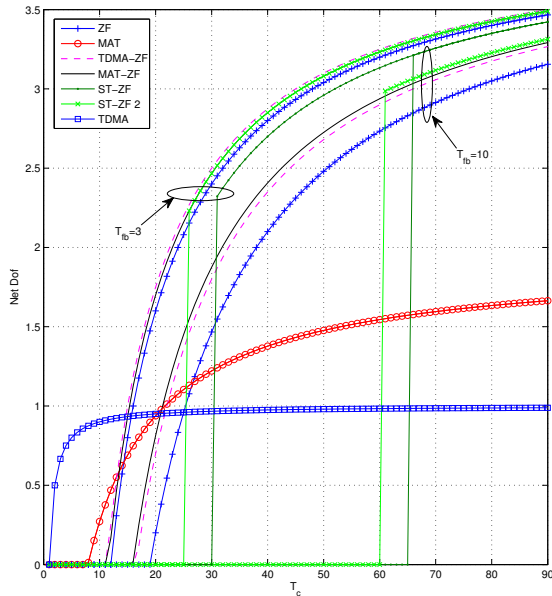
Another way of transmitting  $\mathbf{H}_{\text{eff}}^{(i)}$  to the receivers is to do it in the following blocks in the no current CSIT part as presented for MAT-ZF. It assures a multiplexing gain always greater than that of ZF since it leaves the ZF part of each block untouched. It also enlarges the range of validity of the scheme as it does not reduce the part with CSIT. It takes  $KN_t$  time slots, therefore  $\frac{KN_t}{T_{fb}+1}$  blocks since the dead time in ST-ZF is  $T_{fb} + 1$  time slots long. We refer to this variant as ST-ZF 2. With this strategy the multiplexing gain is

$$\begin{aligned} \text{netDoF}(\text{ST-ZF2}_{N_t}) &= \\ &= \frac{N_t \left( 1 - \frac{2(N_t+1)}{T_c} \right) + \text{netDoF}(\text{ZF}_{N_t}) \frac{KN_t}{T_{fb}+1}}{1 + \frac{KN_t}{T_{fb}+1}} \end{aligned} \quad (6)$$

as long as  $\frac{T_{fb}+1}{T_c-(N_t+1)} \leq \frac{1}{K}$ .

#### IV-G. Numerical results

In Fig. 3 we plot the netDoF provided by ZF, MAT, TDMA-ZF, MAT-ZF, TDMA and ST-ZF for  $N_t = 4$ ,  $T_{fb} \in \{3, 10\}$  as a function of  $T_c$  using (1) for ZF, (3) for MAT, (2) for TDMA-ZF, (4) for MAT-ZF, (5) for ST-ZF and (6) for ST-ZF 2. For TDMA we use  $\frac{T_c-1}{T_c}$ , one pilot per coherence period being needed to insure coherent reception, by keeping the overhead to a minimum TDMA outperforms the other schemes for small  $T_c$ . We notice that even for  $T_{fb} = 3$  and  $T_c = 90$  the net DoF loss of the different



**Fig. 3.** NetDoF of ZF, MAT, TDMA-ZF, MAT-ZF, ST-ZF and TDMA for  $N_t = 4$ ,  $T_{fb} \in \{3, 10\}$  as a function of  $T_c$ .

schemes compared to the optimum  $N_t = 4$  is significant. The DoF yielded by MAT and ST-ZF do not depend on  $T_{fb}$  except that ST-ZF is valid only for  $T_c$  greater or equal to a threshold which grows with  $T_{fb}$ . We observe that the ST-ZF scheme performs better for larger values of  $T_{fb}$ , because the cost of the distribution of  $\mathbf{H}_{\text{eff}}^{(i)}$  in the ST-ZF scheme can be compensated by not losing any DoF on the no CSIT part of each block only if this part is long enough. If we compare analytically TDMA-ZF and ST-ZF we see that for  $T_{fb} = N_t$  both schemes yield about the same multiplexing gain and the multiplexing gain of TDMA-ZF decreases below that of ST-ZF for  $T_{fb}$  sufficiently larger than  $N_t$ . The gain of ST-ZF over the other schemes becomes significant with larger values of  $T_{fb}$ . ST-ZF 2 is better than ST-ZF for small values of  $T_{fb}$ , so the best way to do the distribution of  $\mathbf{H}_{\text{eff}}^{(i)}$  (in the CSIT part or in the DCSIT part) depends on  $T_{fb}$ . We notice that for small  $T_{fb}$  TDMA-ZF performs better than both variants of ST-ZF. However by considering also a certain coherence in the frequency domain the relative cost of the transmission of the  $\mathbf{H}_{\text{eff}}^{(i)}$  in the CSIT part should become negligible even for small values of  $T_{fb}$ . To summarize, for small  $T_c$  TDMA is the best, for large  $T_c$  TDMA-ZF or ST-ZF is better depending on  $T_{fb}$  and for intermediate  $T_c$  TDMA-ZF or MAT-ZF is better depending on  $T_{fb}$ .

## V. CONCLUSION

The ST-ZF scheme is very interesting theoretically as it proved that up to a certain delay in the feedback the full DoF of the MISO BC is still attainable, without accounting for feedback or training costs. By evaluating the multiplexing gain that can be reached by the ST-ZF scheme in actual

systems, i.e. accounting for training overhead and feedback cost we demonstrate that the ST-ZF scheme is also of interest in practice since it outperforms TDMA-ZF and MAT-ZF for feedback delays larger than the number of transmit antennas. Furthermore the NetDoF curves suggest that also for ST-ZF the number of active users needs to be optimized (reduced) as  $T_c$  decreases, as was done in [15] for MAT-ZF.

## VI. 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 W. 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. Cadambe and S. Jafar, "Interference alignment and degrees of freedom of the  $n$ -user interference channel," *Information Theory, IEEE Transactions on*, vol. 54, no. 8, pp. 3425–3441, aug. 2008.
- [4] P. Viswanath and D. 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] B. Makki and T. Eriksson, "Csi feedback in correlated slow-fading channels," *Communications Letters, IEEE*, vol. 15, no. 12, pp. 1294–1297, Dec. 2011.
- [7] M. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," in *Proc. Allerton*, Monticello, IL, USA, Oct. 2010.
- [8] J. Xu, J. Andrews, and S. Jafar, "Miso broadcast channels with delayed finite-rate feedback: Predict or observe?" *Wireless Communications, IEEE Transactions on*, vol. 11, no. 4, pp. 1456–1467, 2012.
- [9] 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.
- [10] 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.
- [11] 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*, Hong Kong, China, Aug. 2012.
- [12] N. Lee and R. W. Heath Jr., "Not too delayed csit achieves the optimal degrees of freedom," in *Proc. Allerton*, Monticello, IL, USA, Nov. 2012.
- [13] U. Salim and D. T. M. Slock, "How much feedback is required for TDD multi-antenna broadcast channels with user selection?" *EURASIP Journal on Advances in Signal Processing*, May 2010.
- [14] N. Jindal, "Mimo broadcast channels with finite-rate feedback," *Information Theory, IEEE Transactions on*, vol. 52, no. 11, pp. 5045–5060, nov. 2006.
- [15] Y. Lejosne, D. Slock, and Y. Yuan-Wu, "Degrees of Freedom in the MISO BC with delayed-CSIT and Finite Coherence Time: Optimization of the number of users," in *Proc. NetG-Coop*, Avignon, France, Nov. 2012.