Retrospective Interference Alignment for Interference Channels with Delayed Feedback

Lorenzo Maggi, Laura Cottatellucci
EURECOM
Mobile Communications Department
BP193, F-06560 Sophia Antipolis, France
Email: {lorenzo.maggi, laura.cottatellucci}@eurecom.fr

***Abstract*—We deal with interference channels with (*i*) delayed channel state information at the transmitter (CSIT), and (*ii*) delayed output feedback at the receiver and no CSIT. We extend the algorithm by Maleki et al. [8] for $M = 3$ users to any $M \geq 3$ under the constraint that *all* the transmitter/receiver pairs are active simultaneously. We propose a retrospective interference alignment algorithm achieving $M^2/(M^2-1)$ degrees of freedom (DoF) in the case (*i*) and an algorithm attaining $\lceil M/2 \rceil M/(\lceil M/2 \rceil (M-1)+1)$ DoF in the case (*ii*). However, larger DoF - 9/8 in (*i*) and 6/5 in (*ii*) - are easily achievable, for any M, when orthogonal channels (e.g. in time, frequency, etc.) are shared by triplets of transmitter/receiver pairs, and within each subchannel the algorithm by Maleki et al. [8] is applied. Hence, our work suggests that, for all M and both in the case (*i*) and in the case (*ii*), the strategy maximizing the DoF entails that three transmitter/receiver pairs are simultaneously active on the same channel.**

I. Introduction

The capacity of the most recent network architectures, such as femtocells, is severely limited by interference. This has spurred an intense research on the M-user interference channel, in which M transmitter/receiver pairs interfere with each other. A long lasting pessimistic belief about the interference channels (e.g. [1]) assumed that the rate per user in the high signal to noise ratio (SNR) regime scales as M^{-1}, i.e. the total degrees of freedom (DoF) could not be higher than the degrees of freedom of a point-to-point channel. Very striking recent results [2]–[4], though, disproved it by showing that the degrees of freedom per user in the interference channel scale as $1/2$ independently of the number of interferers when a transmission scheme called interference alignment (IN) is applied. These very promising results presuppose a complete knowledge on the channel state information at the transmitters (CSIT). This assumption can be very costly in terms of required feedback frequency band. Additionally, it becomes unrealistic and impractical when the channel fades too rapidly and the feedback becomes rapidly stale. This is fueling intensive works to investigate fundamental questions on the practical applicability of IA such as (*a*) the impact of imperfect [5], partial or no CSI [6], [7], on rate and reliability performance of IA systems, (*b*) the obsolescence of the feedback [8], and (*c*) robustness to imperfect and/or delayed feedbacks. The tradeoff between achievable degrees of freedom and required feedback has been studied when the network nodes are equipped with multiple antennas and IA schemes are applied to spatial dimensions in [9] and [10]. For limited feedback, the full DoF is sustained in constant (non-fading) channels if the number of feedback bits grow logarithmically with SNR. The effective DoF achieved by IA with limited training and feedback is investigated in [11]. The issue of the feedback obsolescence was addressed first for the broadcast channel in [12]. Extensions to the interference channel appeared in [8]. Here, some communication schemes based on (*i*) delayed CSI and (*ii*) delayed output feedback and no CSI at the transmitter (CSIT) are proposed for single-antenna 3-user interference channels and X channels. They are referred to as retrospective IA.

In this paper we generalize the two algorithms proposed in [8] to $M \geq 3$ users, under the constraint that *all* the transmitter/receiver pairs are active at the same time. In Section III, we present an algorithm that achieves $M^2/(M^2–1)$ DoF in the delayed CSIT case. In Section IV, we prove that our second algorithm attains $\lceil M/2 \rceil M/(\lceil M/2 \rceil (M-1)+1)$ DoF in the delayed output feedback setting.

Some notation remarks. Boldface upper ($\mathbf{A}$) and lower case ($\mathbf{a}$) denote matrices and column vectors, respectively. We refer to $(\cdot)^T$ as the matrix transpose. The index vector $\mathbf{i} = a : b$, with $a \leq b$, stands for $[a, a+1, \dots, b-1, b]^T$. The matrix $\mathbf{B} = (\mathbf{A})_{\mathbf{i},\mathbf{j}}$ is obtained by jointly selecting only the rows and the columns of $\mathbf{A}$ indexed by $\mathbf{i}$ and by $\mathbf{j}$, respectively, i.e. $(\mathbf{B})_{p,q} = (\mathbf{A})_{\mathbf{i}_p,\mathbf{j}_q}$. The symbol $*$ indexes either all the rows or all the columns of a matrix. For the sake of conciseness, $(\mathbf{a})_{\mathbf{j}} \equiv (\mathbf{a})_{\mathbf{j},1}$. The matrix $[\mathbf{A}, \mathbf{B}]$ is the juxtaposition of matrices $\mathbf{A}, \mathbf{B}$. The acronym "w.p.1" stands for "with probability one". The indicator function is referred to as $\mathbb{I}$.

II. Interference channel model

We focus on $M \geq 3$ users interference channels. We assume that each user achieves multidimensional signalling by coding only over time, i.e. the transmitters/receivers pairs have only one antenna and one frequency carrier at their disposal. Moreover, no inter-symbol interference is present. The transmission occurs over blocks of N time slots. We call with $\mathbf{H}^{[j,m]}$ the N-by-N complex diagonal channel matrix between transmitter m and receiver j, whose diagonal coefficients are *i.i.d.* and drawn from a continuous distribution. Furthermore, let $\mathbf{x}^{[m]} \in \mathbb{C}^{N\text{x}1}$ be the complex signal sent by transmitter m over N time slots.

The received vector $\mathbf{y}^{[m]} \in \mathbb{C}^{N \mathrm{x} 1}$ at receiver m can be written as

$$\mathbf{y}^{[m]} = \sum_{j=1}^{M} \mathbf{H}^{[m,j]} \mathbf{x}^{[j]} + \mathbf{z}^{[m]}, \qquad \forall\, m = 1, \ldots, M, \quad (1)$$

where $\mathbf{z}^{[m]}$ is the vector of complex additive white Gaussian noise at receiver m with zero mean and given variance.

Throughout this work, we assume perfect channel state information at the receiver (CSIR). Furthermore, we expect that the Shannon theoretic definitions of achievable rates, capacity, degrees of freedom etc. are known. The interested reader can refer to [3] for their definition and relation to the system model.

III. Interference Channel with delayed CSIT

In this section we introduce the first of the two classes of interference channel considered in this paper. We assume that transmitter m precodes its n independent complex information symbols $\mathbf{u}_1^{[m]}, \ldots, \mathbf{u}_n^{[m]}$ over N time slots, i.e.

$$\mathbf{x}^{[m]} = \sum_{i=1}^{n} \mathbf{u}_i^{[m]} \mathbf{v}^{[m]}(i), \qquad m = 1, \ldots, M,$$

where $\mathbf{v}^{[m]}(i) \in \mathbb{C}^{N \mathrm{x} 1}$ is the complex precoding vector associated to the i-th symbol of transmitter m. We consider here the following two assumptions.

Assumption 1 (Delayed CSIT)**.** *At time $t \geq 2$, each transmitter has knowledge about all the channel coefficients up to time $t-1$, i.e. $\{(\mathbf{H}^{[j,m]})_{i,i}\}_{j,m}$, for all $j, m \in [1, M]$, and $i \in [1; t-1]$.*

Assumption 2 (Global beamforming knowledge)**.** *Each transmitter knows the precoding vectors of all the other users.*

In the following we will see that Assumption 2 is actually very loose in our specific case. Now, we present our main result on the achievable degrees of freedom under Assumptions 1,2.

Proposition III.1. *In the M-user interference channel, under Assumptions 1 and 2, and with CSIR, DoF_{dC} degrees of freedom are achievable w.p.1 by the algorithm described in Section III-A, where*

$$\mathrm{DoF}_{\mathrm{dC}} = \frac{M^2}{M^2 - 1}. \quad (2)$$

This result is based on the communication scheme described in the following subsection and relies on Conjecture III.1 (see Section III-B).

A. Retrospective IA Algorithm

We now describe the communication scheme apt to achieve DoF_{dC} degrees of freedom under Assumptions 1 and 2. Set the positive integers n, K such that

$$K \leq \lfloor n/M \rfloor. \quad (3)$$

The reader will understand the importance of the inequality (3) in the following. The transmission of the block of symbols $\mathbf{u}_1^{[m]}, \ldots, \mathbf{u}_n^{[m]}$, for every user m, lasts for N time slots, where

$$N = Mn - K. \quad (4)$$

We define the precoding N-by-n matrix for transmitter $m = 1, \ldots, M$ as

$$\mathbf{V}^{[m]} = \left[\mathbf{v}^{[m]}(1), \quad \mathbf{v}^{[m]}(2), \quad \ldots, \quad \mathbf{v}^{[m]}(n)\right].$$

where the column vector $\mathbf{v}^{[m]}(i)$ is the i-th precoding vector for transmitter m. The two phases of our retrospective alignment algorithm are described in the following.

Phase I: The Phase I of the algorithm lasts for N_1 time slots, where

$$N_1 = (M-1)n - K.$$

Let us define the N_1-by-N_1 submatrix $\overline{\mathbf{H}}^{[i,j]} \equiv (\mathbf{H}^{[i,j]})_{1:N_1, 1:N_1}$; let the N_1-by-n submatrix $\overline{\mathbf{V}}^{[m]} \equiv (\mathbf{V}^{[m]})_{1:N_1, *}$ and let the N_1-by-1 subvector $\overline{\mathbf{y}}^{[m]} \equiv (\mathbf{y}^{[m]})_{1:N_1}$, for all possible i, j, m.

Phase I of the algorithm is straightforward, and it consists in each transmitter drawing the components of its precoding vectors from a continuous distribution, such that the entries of the following matrix:

$$\left[\overline{\mathbf{V}}^{[1]}, \quad \overline{\mathbf{V}}^{[2]}, \quad \ldots, \quad \overline{\mathbf{V}}^{[M]}\right]$$

are independent and identically distributed. At the end of Phase I, the signal seen at receiver $m = 1, \ldots, M$ is

$$\overline{\mathbf{y}}^{[j]} = \sum_{m=1}^{M} \overline{\mathbf{H}}^{[j,m]} \overline{\mathbf{V}}^{[m]} [\mathbf{u}_1^{[m]}, \ \mathbf{u}_2^{[m]}, \ \ldots, \ \mathbf{u}_n^{[m]}]^T,$$

and the interference vector space at receiver j is spanned by the columns of the matrix $\overline{\mathbf{G}}^{[j]}$, defined as

$$\left[\overline{\mathbf{H}}^{[j,1]} \overline{\mathbf{V}}^{[1]}, \ldots, \overline{\mathbf{H}}^{[j,j-1]} \overline{\mathbf{V}}^{[j-1]}, \overline{\mathbf{H}}^{[j,j+1]} \overline{\mathbf{V}}^{[j+1]}, \ldots, \overline{\mathbf{H}}^{[j,M]} \overline{\mathbf{V}}^{[M]}\right],$$

having N_1 rows and $N_1 + K = (M-1)n$ columns. From the rank nullity Theorem, we know that the null space of matrix $\overline{\mathbf{G}}^{[j]}$, for all $j = 1, \ldots, M$, spans at least K dimensions. Then, it is possible to find K independent $(M-1)n$-by-1 complex vectors, namely $\boldsymbol{\alpha}_1^{[j]}, \ldots, \boldsymbol{\alpha}_K^{[j]}$, such that

$$\overline{\mathbf{G}}^{[j]} \boldsymbol{\alpha}_i^{[j]} = \mathbf{0}_{N_1}, \quad \forall\, i = 1, \ldots, K, \quad (5)$$

for all $j = 1, \ldots, M$, where $\mathbf{0}_{N_1}$ is the N_1-by-1 null vector.

Phase II: In Phase II we deal with the remaining n channel uses, from time slot $N_1 + 1 = (M-1)n - K + 1$ to time slot N. As in Phase I, we define the n-by-n matrix $\widetilde{\mathbf{H}}^{[i,j]} \equiv (\mathbf{H}^{[i,j]})_{N_1+1:N, N_1+1:N}$ and the n-by-n matrix $\widetilde{\mathbf{V}}^{[m]} \equiv (\mathbf{V}^{[m]})_{N_1+1:N, *}$, for all i, j, m.

Similarly to the definition of $\overline{\mathbf{G}}^{[j]}$, we write $\widetilde{\mathbf{G}}^{[j]}$ as

$$\left[\widetilde{\mathbf{H}}^{[j,1]} \widetilde{\mathbf{V}}^{[1]}, \ \ldots, \ \widetilde{\mathbf{H}}^{[j,j-1]} \widetilde{\mathbf{V}}^{[j-1]}, \widetilde{\mathbf{H}}^{[j,j+1]} \widetilde{\mathbf{V}}^{[j+1]}, \ \ldots, \ \widetilde{\mathbf{H}}^{[j,M]} \widetilde{\mathbf{V}}^{[M]}\right],$$

for all $j = 1, \ldots, M$. Since relation (5) already holds, in order for the interference to span at most $N_1 = (M-1)n - K$ dimensions, the following relations must be satisfied:

$$\mathbf{G}^{[j]}\boldsymbol{\alpha}_i^{[j]} = \mathbf{0}_N, \qquad \forall\, i = 1, \ldots, K,\ \forall\, j = 1, \ldots, M,$$

$$\mathbf{G}^{[j]} = \begin{bmatrix} \overline{\mathbf{G}}^{[j]} \\ \widetilde{\mathbf{G}}^{[j]} \end{bmatrix}, \qquad \forall\, j = 1, \ldots, M.$$

At the beginning of Phase II, each transmitter has at its disposal *all* the channel coefficients up to the N_1-th time slot. For Assumption 2, each transmitter gets to know also the precoding vectors randomly chosen by the other transmitters up to time N_1. This is easily achievable, for example, by sharing among the users the random seeds used to generate the precoding vectors. Hence, all the vectors $\{\boldsymbol{\alpha}_i^{[j]}\}_{i\in[1;K]}$, for all $j = 1, \ldots, M$, are computed by all the transmitters at the beginning of Phase II. In order to align interference, we must set, for all $i = 1, \ldots, K$ and $j = 1, \ldots, M$,

$$\widetilde{\mathbf{G}}^{[j]}\boldsymbol{\alpha}_i^{[j]} = \mathbf{0}_n. \tag{6}$$

We now probe how to ensure the validity of (6). Let us stack the vectors $\boldsymbol{\alpha}$'s in such a way to build the following full rank K-by-$(M-1)n$ matrix, for all $j = 1, \ldots, M$:

$$\mathbf{A}^{[j]} \equiv \begin{bmatrix} \boldsymbol{\alpha}_1^{[j]}, & \boldsymbol{\alpha}_2^{[j]}, & \ldots, & \boldsymbol{\alpha}_K^{[j]} \end{bmatrix}^T.$$

Since the feedback on the channel is delayed at the transmitter and (6) must hold, then each transmitter is compelled to compute its own precoding vectors in such a way that (6) is satisfied for *any* value of the channel coefficients $\widetilde{\mathbf{H}}^{[i,j]}$. To do so, it is sufficient to set up the following system of equations for transmitter m, in the n unknowns stacked in $\widetilde{\mathbf{V}}_{t,*}^{[m]}$:

$$\boldsymbol{\Theta}^{[m]}[\widetilde{\mathbf{V}}_{t,*}^{[m]}]^T = \mathbf{0}, \tag{7}$$

for all $t \in [1; n]$ and for all $m = 1, \ldots, M$, where

$$\boldsymbol{\Theta}^{[m]} \equiv \Big[[\mathbf{A}^{[1]}_{*,(m-2)n+1:(m-1)n}]^T,\ \ldots,\ [\mathbf{A}^{[m-1]}_{*,(m-2)n+1:(m-1)n}]^T,$$
$$[\mathbf{A}^{[m+1]}_{*,(m-1)n+1:mn}]^T,\ \ldots,\ [\mathbf{A}^{[M]}_{*,(m-1)n+1:mn}]^T\Big]^T.$$

Now, we will find a suitable way to choose the precoding vectors. During Phase II, transmitter m has to find the vectors $\widetilde{\mathbf{V}}_{1,*}^{[m]}, \ldots, \widetilde{\mathbf{V}}_{n,*}^{[m]}$ satisfying (7). Since expression (3) holds and the rank of the $(M-1)K$-by-n matrix $\boldsymbol{\Theta}^{[m]}$ is $r_m \le (M-1)K$, then m-th linear system is solved by a null subspace spanning D_m dimensions, where

$$D_m \ge n - (M-1)K \ge n - \frac{M-1}{M}n = \frac{n}{M} \ge 1. \tag{8}$$

We choose a vector $\mathbf{d}_m$ of r_m indices such that the r_m-by-r_m matrix $(\boldsymbol{\Theta}^{[m]})_{*,\mathbf{d}_m}$ is full rank. Let us call with $\overline{\mathbf{d}}_m$ the column indices which are complementary to $\mathbf{d}_m$, i.e. $\overline{\mathbf{d}}_m \cap \mathbf{d}_m = \emptyset$, $\overline{\mathbf{d}}_m \cup \mathbf{d}_m = 1 : n$, for all $m = 1, \ldots, M$. From basic linear algebra, we can express the components of vector $\widetilde{\mathbf{V}}_{t,*}^{[m]}$ indexed by $\mathbf{d}_m$ as a linear function of its components indexed by $\overline{\mathbf{d}}_m$, i.e.

$$\widetilde{\mathbf{V}}_{t,\mathbf{d}_m}^{[m]} = -[(\boldsymbol{\Theta}^{[m]})_{*,\mathbf{d}_m}]^{-1}(\boldsymbol{\Theta}^{[m]})_{*,\overline{\mathbf{d}}_m}\widetilde{\mathbf{V}}_{t,\overline{\mathbf{d}}_m}^{[m]}. \tag{9}$$

Finally, we can describe what Phase II consists in, in practice. For all $t = 1, \ldots, n$, the transmitter m, $m = 1, \ldots, M$, draws independently from a continuous distribution the D_m precoding components $\widetilde{\mathbf{V}}_{t,\overline{\mathbf{d}}_m}^{[m]}$, and computes the remaining coefficients $\{\widetilde{\mathbf{V}}_{t,\mathbf{d}_m}^{[m]}\}_{t=1,\ldots,n}$ as in (9).

B. Feasibility of Interference Alignment

By combining (5) and (6), from the rank-nullity Theorem we claim that the total interference seen at each receiver spans at most $N_1 = (M-1)n - K$ dimensions. For example, at receiver 1,

$$\text{rank}\Big[\mathbf{H}^{[1,2]}\mathbf{V}^{[2]}\ ,\ \ldots,\ \mathbf{H}^{[1,M]}\mathbf{V}^{[M]}\Big] \le N_1, \qquad \text{w.p.1.}$$

Hence, at each receiver, the interference subspace shrinks by K dimensions.

We conjecture that, at each receiver, the vector subspace spanned by the useful signal is disjoint (except for the null vector, of course) from the subspace spanned by interference, w.p.1.

Conjecture III.1. *If relation (3) holds, at receiver j, the following $(Mn-K)$-by-Mn matrix:*

$$\mathbf{S}^{[j]} \equiv \begin{bmatrix} \overline{\mathbf{H}}^{[j,1]}\overline{\mathbf{V}}^{[1]}, & \ldots, & \overline{\mathbf{H}}^{[j,M]}\overline{\mathbf{V}}^{[M]} \\ \widetilde{\mathbf{H}}^{[j,1]}\widetilde{\mathbf{V}}^{[1]}, & \ldots, & \widetilde{\mathbf{H}}^{[j,M]}\widetilde{\mathbf{V}}^{[M]} \end{bmatrix} \tag{10}$$

is full rank, w.p.1.

We remark that *all the simulations carried out by the authors confirm conjecture* III.1.

The validity of conjecture III.1 is sufficient to prove Proposition III.1 as the following argument shows.

If $\mathbf{S}^{[j]}$ is full rank, then the interference subspace is disjoint from the useful signal, which can be retrieved by zero-forcing the interference. It follows from (3) that the value of K which optimizes the DoF is $\bar{K} = n/M$, for which the degrees of freedom achieved by our retrospective interference alignment algorithm are

$$\frac{Mn}{N} = \frac{M^2\bar{K}}{(M^2-1)\bar{K}} = \frac{M^2}{M^2-1} = \text{DoF}_{\text{dC}}.$$

We now state that the condition (3) is necessary for conjecture III.1 to hold. Moreover, if (3) holds, then the second horizontal block of matrix (10) is full rank w.p.1.

Lemma III.2. *For all $j = 1, \ldots, M$, the following n-by-Mn matrix:*

$$\widetilde{\mathbf{S}}^{[j]} \equiv \big[\widetilde{\mathbf{H}}^{[j,1]}\widetilde{\mathbf{V}}^{[1]},\ \ldots,\ \widetilde{\mathbf{H}}^{[j,j]}\widetilde{\mathbf{V}}^{[j]},\ \ldots,\ \widetilde{\mathbf{H}}^{[j,M]}\widetilde{\mathbf{V}}^{[M]}\big] \tag{11}$$

is full rank ($= n$) w.p.1 if and only if condition (3) holds.

The proof of Lemma III.2 is in the Appendix. The intuition behind this Lemma is that, even if the matrix $[\widetilde{\mathbf{V}}^{[1]},\ \ldots,\ \widetilde{\mathbf{V}}^{[M]}]$ is not full rank, the introduction of multiplicative independent diagonal matrices $\widetilde{\mathbf{H}}^{[j,i]}$ allows to achieve full rank w.p.1. Nevertheless, if condition (3) fails to holds, i.e. the sum over i of the ranks of the matrices $\widetilde{\mathbf{V}}^{[i]}$ is strictly less than n, then the rank of $\widetilde{\mathbf{S}}^{[j]}$ cannot be equal to n.

IV. Interference Channel with Delayed Output Feedback

In this section we will deal with a second class of interference channel. With respect to Section III, we still assume CSIR, but we drop Assumptions 1 and 2 and we consider the following ones.

Assumption 3 (No CSIT)**.** *The transmitters have neither current nor past channel state information.*

Assumption 4 (Delayed output feedback)**.** *Each transmitter receives a delayed version of the signal received by the respective receiver, i.e. at time $t \geq 2$ user m has at its disposal the signals $\mathbf{y}_p^{[m]}$, for all $p = 1, \ldots, t-1$, $m = 1, \ldots, M$.*

Because of Assumption 3, the precoding vectors technique will not be utilized, since there is no hope for the interfering signals to be aligned. In [8], Maleki, Jafar, and Shamai proposed an algorithm achieving $6/5$ DoF in $M = 3$ user interference channels under Assumptions 3,4. Here we extend their approach to any $M \geq 3$ and we obtain the following result.

Proposition IV.1. *In the M-user interference channel, under Assumptions 3 and 4,* $\mathrm{DoF}_{\mathrm{dO}}$ *degrees of freedom are achievable w.p.1 by the algorithm described in Section IV-A, where*

$$\mathrm{DoF}_{\mathrm{dO}} = \frac{\lceil M/2 \rceil M}{\lceil M/2 \rceil (M-1) + 1}. \tag{12}$$

The proof of Proposition IV.1 follows from a communication scheme described in the next section. The reader should note that the degree of freedom $\mathrm{DoF}_{\mathrm{dC}}$ obtained in the delayed CSIT case in (2) is strictly lower than $\mathrm{DoF}_{\mathrm{dO}}$, for any value of $M \geq 3$.

A. Algorithm for Delayed Feedback

In this section we will present an algorithm achieving $\mathrm{DoF}_{\mathrm{dO}}$ (12) degrees of freedom w.p.1 in interference channels, under Assumptions 3,4. Each transmitter m sends $n = \lceil M/2 \rceil$ independent information symbols denoted by $\mathbf{u}_1^{[m]}, \ldots, \mathbf{u}_n^{[m]}$. Again, we split the algorithm into two main phases.

Phase I: The Phase I of the algorithm lasts for M time slots. Let us show how to construct the vector of transmitted symbols $\mathbf{x}^{[m]}$, for all transmitters $m = 1, \ldots, M$, in the first M time slots. We stack all the nM information symbols into a nM-by-1 column vector $\mathbf{w}$, such that its k-th element $\mathbf{w}_k$ is the $i = \lceil k/M \rceil$-th information symbol of transmitter $m = [\mathrm{mod}(k-1, M) + 1]$, i.e. $\mathbf{w}_k = \mathbf{u}_i^{[m]}$.

In the k-th slot, $k = 1, \ldots, M$, the n symbols, belonging to n distinct communication flows,

$$\mathbf{w}_{(k-1)n+1},\ \mathbf{w}_{(k-1)n+2},\ \ldots\ ,\ \mathbf{w}_{kn}, \tag{13}$$

are sent by the respective transmitters. We call $\mathcal{T}(k) = \{\mathcal{T}_1(k), \mathcal{T}_2(k), \ldots, \mathcal{T}_n(k)\}$ the set of n transmitters sending information at time slot k, and $\mathcal{T}^c(k)$ is the complement of $\mathcal{T}(k)$, i.e. the set of indices of the users being silent at time step k. The vector $\mathbf{x}$ is built up accordingly to the transmission rule in (13), and $\mathbf{x}_k^{[m]} = 0$ if $m \in \mathcal{T}^c(k)$. Figure 1 illustrates this communication scheme for $M = 5$. In this case, the transmitted vector $\mathbf{x}$ is such that in the first time slot, $\mathbf{x}_1^{[1]} = \mathbf{w}_1$, $\mathbf{x}_1^{[2]} = \mathbf{w}_2$, $\mathbf{x}_1^{[3]} = \mathbf{w}_3$, $\mathbf{x}_1^{[4]} = \mathbf{x}_1^{[5]} = 0$.

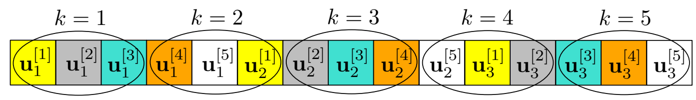

Figure 1. Graphic representation of 15-by-1 vector $\mathbf{w}$ in Phase I with $M = 5$. In the time slot $k = 1, \ldots, 5$, 3 symbols are sent at the same time.

Phase II: Let us first introduce some useful notation. If M is odd, let $\mathcal{F}$ be the set of *all* the received signals of the type $\{\mathbf{y}_k^{[m(k)]}\}$, for all $m(k) \in \mathcal{T}^c(k)$, $k = 1, \ldots, M$. In other words, $\mathcal{F}$ is the set of the signals overheard by the transmitters not communicating over each time slot associated to Phase I. If M is even, then $\mathcal{F}$ is defined as for M odd but it does not include one arbitrary received signal for each receiver. Then, the set $\mathcal{F}$ has cardinality $(n-1)M$ in both cases. Let us stack the elements of $\mathcal{F}$, to be sent during Phase II, into the $(n-1)M$-by-1 vector $\mathbf{f}$, such that the m-th block of $\mathbf{f}$ of dimension $n-1$ collects all the signals received by receiver m and belonging to the set $\mathcal{F}$.

Phase II lasts for $(n-1)M - n + 1$ time slots. By the delayed output feedback Assumption 4, when Phase II starts, every transmitter has at its disposal all the signals received by the respective receiver during Phase I. During the k-th time slot of Phase II, the signals

$$\mathbf{f}_k,\ \mathbf{f}_{k+1},\ \ldots\ ,\ \mathbf{f}_{k+n-1}: \quad k = 1, \ldots, (n-1)M - n + 1, \tag{14}$$

are sent by the concerned transmitters. Transmitters having more than one signal to transmit will send their sum. Figure 2 illustrates the transmission rule for $M = 5$.

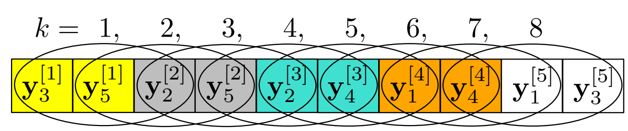

Figure 2. Graphic representation of the 10-by-1 vector $\mathbf{f}$. In each of the 8 time slots assigned to Phase II, 3 output feedback signals are sent at the same time.

The vector $\mathbf{x}^{[m]}$ transmitted by each transmitter m during Phase II is built up accordingly with the transmission rule in (14). Helped by Figure 2, one can see that in the first time slot of Phase II, $\mathbf{x}_6^{[1]} = \mathbf{f}_1 + \mathbf{f}_2$, $\mathbf{x}_6^{[2]} = \mathbf{f}_3$, $\mathbf{x}_6^{[3]} = \mathbf{x}_6^{[4]} = \mathbf{x}_6^{[5]} = 0$. The transmission ends in the 13-th time slot, i.e. the 8-th time slot of Phase II, with $\mathbf{x}_{13}^{[1]} = \mathbf{x}_{13}^{[2]} = \mathbf{x}_{13}^{[3]} = 0$, $\mathbf{x}_{13}^{[4]} = \mathbf{f}_8$, $\mathbf{x}_{13}^{[5]} = \mathbf{f}_9 + \mathbf{f}_{10}$.

For the assumption of perfect CSIR, by a straightforward sequential detection, each receiver can identify the whole output feedback vector $\mathbf{f}$, w.p.1. Then the generic receiver k, for each symbol $\mathbf{u}_i^{[k]}$, $i = 1, \dots, n$, can now build a n-by-n linear system of equations, where the unknowns are $\mathbf{u}_i^{[k]}$ itself and the $n-1$ symbols that were sent by the other transmitters simultaneously with $\mathbf{u}_i^{[k]}$. The matrix of the (channel) coefficients is full rank w.p.1. Note that, for each symbol, one equation out of n comes from Phase I.

Returning to our example for $M = 5$, the three linear systems of equations for receiver 1 to retrieve $\mathbf{u}_1^{[1]}, \mathbf{u}_2^{[1]}, \mathbf{u}_3^{[1]}$, respectively, are

$$\begin{bmatrix} \mathbf{y}_1^{[1]} \\ \mathbf{y}_1^{[4]} \\ \mathbf{y}_1^{[5]} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{1,1}^{[1,1]} & \mathbf{H}_{1,1}^{[1,2]} & \mathbf{H}_{1,1}^{[1,3]} \\ \mathbf{H}_{1,1}^{[4,1]} & \mathbf{H}_{1,1}^{[4,2]} & \mathbf{H}_{1,1}^{[4,3]} \\ \mathbf{H}_{1,1}^{[5,1]} & \mathbf{H}_{1,1}^{[5,2]} & \mathbf{H}_{1,1}^{[5,3]} \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^{[1]} \\ \mathbf{u}_1^{[2]} \\ \mathbf{u}_1^{[3]} \end{bmatrix},$$

$$\begin{bmatrix} \mathbf{y}_2^{[1]} \\ \mathbf{y}_2^{[2]} \\ \mathbf{y}_2^{[3]} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{2,2}^{[1,1]} & \mathbf{H}_{2,2}^{[1,4]} & \mathbf{H}_{2,2}^{[1,5]} \\ \mathbf{H}_{2,2}^{[2,1]} & \mathbf{H}_{2,2}^{[2,4]} & \mathbf{H}_{2,2}^{[2,5]} \\ \mathbf{H}_{2,2}^{[3,1]} & \mathbf{H}_{2,2}^{[3,4]} & \mathbf{H}_{2,2}^{[3,5]} \end{bmatrix} \begin{bmatrix} \mathbf{u}_2^{[1]} \\ \mathbf{u}_1^{[4]} \\ \mathbf{u}_1^{[5]} \end{bmatrix},$$

$$\begin{bmatrix} \mathbf{y}_4^{[1]} \\ \mathbf{y}_4^{[3]} \\ \mathbf{y}_4^{[4]} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{4,4}^{[1,1]} & \mathbf{H}_{3,4}^{[1,5]} & \mathbf{H}_{4,4}^{[1,2]} \\ \mathbf{H}_{4,4}^{[3,1]} & \mathbf{H}_{3,4}^{[3,5]} & \mathbf{H}_{4,4}^{[3,2]} \\ \mathbf{H}_{4,4}^{[4,1]} & \mathbf{H}_{3,4}^{[4,5]} & \mathbf{H}_{4,4}^{[4,2]} \end{bmatrix} \begin{bmatrix} \mathbf{u}_3^{[1]} \\ \mathbf{u}_2^{[5]} \\ \mathbf{u}_3^{[2]} \end{bmatrix}.$$

Hence, each receiver can decode all the symbols sent by the respective transmitter w.p.1. It is straightforward to check that $\mathrm{DoF}_{\mathrm{dO}}$ (12) degrees of freedom are achieved w.p.1. Hence, Proposition IV.1 is proved.

V. Remarks

The DoF achievable by our two algorithms, in which all the transmitter/receiver pairs are active at the same time, attain their maximum at $M = 3$, which is the case probed by Maleki et al. in [8]. Since the DoF of our algorithms find their maximum at $M = 3$, it is straightforward to see that larger DoF - 9/8 in (*i*) and 6/5 in (*ii*) - are easily achievable, for any number M of users, by letting triplets of transmitter/receiver pairs orthogonally sharing the channel and subsequently applying the algorithms in [8], which coincide with ours for $M = 3$, within each orthogonal subchannel.
Hence, our work suggests that, for all M and in the two cases (*i*) and (*ii*), the strategy maximizing the DoF entails that exactly three transmitter/receiver pairs are simultaneously active.
Nonetheless, there seems to be room for improving our results. At first glance, the case of delayed CSIT seems to be more promising: Phase I might be shortened in order to start exploiting in advance the information about the past channel state. Hence, an additional middle phase should be devised.

VI. Acknowledgments

This reasearch was supported by "Agence Nationale de la Recherche" with reference ANR-09-VERS-001, and by the European research project SAPHYRE, which is partly funded by the European Union under its FP7 ICT Objective 1.1 - The Network of the Future. The authors thank Professor Konstantin Avrachenkov, for the fruitful discussions and the constant support.

Appendix

Proof of Lemma III.2

Proof: ($\Rightarrow$) Take $D = \min_m D_m \geq 1$. Suppose that condition (3) holds. We want to prove that the rank of matrix $\widetilde{\mathbf{S}}^{[j]}$ is n, w.p.1. Let us set $j \in [1; M]$. By construction, for all $m = 1, \dots, M$, w.p.1

a) the rank of the n-by-n matrix $\widetilde{\mathbf{V}}^{[m]}$ is at least D;
b) *any* subset of D rows of $\widetilde{\mathbf{V}}^{[m]}$ forms a submatrix with rank D.

Let us call with Ω the event that conditions *a)* and *b)* are jointly satisfied. We call with $\{\widetilde{\mathbf{H}}\}$ and $\{\widetilde{\mathbf{V}}\}$ the collection of all the elements of matrices $\widetilde{\mathbf{H}}^{[j,i]}$ and $\widetilde{\mathbf{V}}^{[i]}$ respectively, with $i = 1, \dots, M$. Let $\Psi(\{\widetilde{\mathbf{H}}\}, \{\widetilde{\mathbf{V}}\})$ be the event that the matrix $\widetilde{\mathbf{S}}^{[j]}$ is full rank when $\{\widetilde{\mathbf{H}}\}$ and $\{\widetilde{\mathbf{V}}\}$ are fixed. Let $f_{\mathbf{H},\mathbf{V}}(.)$ be the joint probability density function associated to $\{\widetilde{\mathbf{H}}, \widetilde{\mathbf{V}}\}$. Then we can write $\mathrm{prob}[\Psi]$ as

$$\int_{\{\widetilde{\mathbf{V}}\}} \int_{\{\widetilde{\mathbf{H}}\}} \mathbb{I}\left[\Psi(\{\widetilde{\mathbf{H}}, \widetilde{\mathbf{V}}\})\right] f_{\mathbf{H},\mathbf{V}}(\{\widetilde{\mathbf{H}}, \widetilde{\mathbf{V}}\}) d\widetilde{\mathbf{H}} d\widetilde{\mathbf{V}} =$$
$$\int_{\{\widetilde{\mathbf{V}}\}:\Omega} \int_{\{\widetilde{\mathbf{H}}\}} \mathbb{I}\left[\Psi(\{\widetilde{\mathbf{H}}, \widetilde{\mathbf{V}}\})\right] f_{\mathbf{H}}(\{\widetilde{\mathbf{H}}\}) f_{\mathbf{V}}(\{\widetilde{\mathbf{V}}\}) d\widetilde{\mathbf{H}} d\widetilde{\mathbf{V}}, \quad (15)$$

since $\{\widetilde{\mathbf{H}}\}$ and $\{\widetilde{\mathbf{V}}\}$ are independent and Ω is verified almost surely. So we suppose that conditions *a)* and *b)* are valid for sure, and we want to prove that

$$\int_{\{\widetilde{\mathbf{H}}\}} \mathbb{I}\left[\Psi(\{\widetilde{\mathbf{H}}, \widetilde{\mathbf{V}}\})\right] f_{\mathbf{H}}(\{\widetilde{\mathbf{H}}\})\, d\widetilde{\mathbf{H}} = 1, \quad (16)$$

for all the coefficients $\{\widetilde{\mathbf{V}}\}$ such that Ω is verified. Indeed, if (16) is verified, then evidently $\mathrm{prob}[\Psi] = 1$ from (15). Therefore, in the following we will fix $\{\widetilde{\mathbf{V}}\}$ and we will find a minor of $\widetilde{\mathbf{S}}^{[j]}$ which is nonnull almost surely. Consider the following set of indices:

$$\mathbf{i}_m = (m-1)D + 1 : \min(n, mD), \qquad m = 1, \dots, \lceil n/D \rceil.$$

For condition *b)*, the rows of the submatrix $\widetilde{\mathbf{V}}^{[m]}_{\mathbf{i}_m,*}$ are linearly independent, hence thanks to relation (3) there exist a collection of indices $\mathbf{c}_m$ with the same cardinality as $\mathbf{i}_m$ such that

$$\det\left(\widetilde{\mathbf{V}}^{[m]}_{\mathbf{i}_m,\mathbf{c}_m}\right) \neq 0, \qquad \forall\, m = 1, \dots, M. \quad (17)$$

Let us build the following n-by-n submatrix of $\widetilde{\mathbf{S}}^{[j]}$:

$$\mathbf{L}^{[j]} \equiv \left[\widetilde{\mathbf{H}}^{[j,1]}\widetilde{\mathbf{V}}^{[1]}_{*,\mathbf{c}_1},\ \widetilde{\mathbf{H}}^{[j,2]}\widetilde{\mathbf{V}}^{[2]}_{*,\mathbf{c}_2},\ \dots,\ \widetilde{\mathbf{H}}^{[j,\lceil n/D\rceil]}\widetilde{\mathbf{V}}^{[\lceil n/D\rceil]}_{*,\mathbf{c}_{\lceil n/D\rceil}}\right].$$

We prove now that the matrix $\mathbf{L}^{[j]}$ is non singular w.p.1. The determinant of $\mathbf{L}^{[j]}$ can be expressed as a polynomial whose (independent random) variables are the diagonal channel coefficients of $\widetilde{\mathbf{H}}^{[j,m]}$, with $m = 1, \dots, \lceil n/D \rceil$, and whose (deterministic) multiplicative coefficients are $\{\widetilde{\mathbf{V}}\}$. We will show that such a polynomial is not identically zero, i.e. there is at least one coefficient which is nonnull.

By the Laplace Expansion Theorem [13], it is easy to see that one of the terms of the polynomial determinant of $\mathbf{L}^{[j]}$ is

$$\prod_{m=1}^{\lceil n/D \rceil} \det\left(\widetilde{\mathbf{H}}^{[j,m]}_{\mathbf{i}_m,\mathbf{i}_m} \widetilde{\mathbf{V}}^{[m]}_{\mathbf{i}_m,\mathbf{c}_m}\right)$$

$$= \prod_{m=1}^{\lceil n/D \rceil} \det\left(\widetilde{\mathbf{H}}^{[j,m]}_{\mathbf{i}_m,\mathbf{i}_m}\right) \prod_{m=1}^{\lceil n/D \rceil} \det\left(\widetilde{\mathbf{V}}^{[m]}_{\mathbf{i}_m,\mathbf{c}_m}\right). \qquad (18)$$

Hence, the first productory in (18) becomes the product of polynomial variables, while the second productory is the deterministic multiplicative coefficients, which is nonnull by (17). Therefore, expression (16) holds and the rank of matrix $\widetilde{\mathbf{S}}^{[j]}$ is n.

($\Leftarrow$) Conversely, let us suppose that $K > n/M$. In order to ensure that Phase II is still feasible, we also need to impose $D \geq 1$, i.e. $n/M < K \leq n/(M-1)$. We want to prove that the matrix $\widetilde{\mathbf{S}}^{[j]}$ is rank deficient with nonnull probability. We will show that, for any $m = 1, \ldots, M$,

$$\operatorname{rank}\left(\widetilde{\mathbf{S}}^{[m]}\right) \leq MD = M(n-(M-1)K) < n, \qquad (19)$$

with nonnull probability. It is sufficient to prove that, with nonnull probability,

$$\operatorname{rank}\left(\widetilde{\mathbf{H}}^{[i,j]}\widetilde{\mathbf{V}}^{[j]}\right) \leq D, \quad \forall\, i,j. \qquad (20)$$

With probability $\epsilon > 0$, the rank of matrix $\widetilde{\mathbf{V}}^{[j]}$ is D and then there exist $n-D$ independent and nonnull n-by-1 vectors $\beta_1^{[j]}, \ldots, \beta_{n-D}^{[j]}$ such that

$$\beta_k^{[j]\,T}\widetilde{\mathbf{V}}^{[j]} = \mathbf{0}_n^T, \qquad k = 1, \ldots, n-D,\ j = 1, \ldots, M.$$

Since $\widetilde{\mathbf{H}}^{[i,j]}$ is non singular w.p.1, then we can write equivalently

$$\beta_k^{[j]\,T}[\widetilde{\mathbf{H}}^{[i,j]}]^{-1}\widetilde{\mathbf{H}}^{[i,j]}\widetilde{\mathbf{V}}^{[j]} = \mathbf{0}_n^T, \qquad k = 1, \ldots, n-D.$$

The vectors $\{\beta_k^{[j]\,T}[\widetilde{\mathbf{H}}^{[i,j]}]^{-1}\}_k$ are independent and nonnull w.p.1, therefore relation (20) holds and the rank of $\widetilde{\mathbf{S}}^{[j]}$ is strictly less than n with nonnull probability, *q.e.d.*.

■

References

[1] A. Host-Madsen and A. Nosratinia, "The multiplexing gain of wireless networks," in *Proc. of IEEE International Symposium on Information Theory (ISIT)*, Adelaide, Australia, Sep. 2005.

[2] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Transactions on Information Theory*, vol. 12, no. 54, pp. 5534–5562, Dec. 2008.

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

[4] S. A. Jafar and S. Shamai, "Degrees of freedom region of the MIMO x channel," *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 151–170, Jan. 2008.

[5] B. Nosrat-Makouei, J. G. Andrews, and R. W. Heath, "A simple SINR characterization for linear interference alignment over uncertain MIMO channels," in *Proc. of IEEE International Symposium on Information Theory (ISIT)*, 2010, p. 2288.

[6] Y. Zhu and D. Guo, "Isotropic MIMO interference channels without CSIT: The loss of degrees of freedom," in *Proc. Allerton Conf. on Communication, Control and Computing*, Allerton House, UIUC, Illinois, USA, Sep./Oct. 2009, pp. 1338–1343.

[7] S. Jafar, "Exploiting channel correlations - simple interference alignment schemes with no CSIT," in *Proc. of IEEE Global Telecommunications Conference (Globecom)*, Atlanta, GA, Dec. 2010, pp. 671–675.

[8] H. Maleki, S. A. Jafar, and S. Shamai, "Retrospective interference alignment," *arXiv:1009.3593v1 [cs.IT]*, Sep. 2010.

[9] H. Bölcskei and I. Thukral, "Interference alignment with limited feedback," in *Proc. of IEEE International Symposium on Information Theory (ISIT)*, Seoul, Korea, Jun. 2009, pp. 1759–1763.

[10] R. Krishnamachari and M. Varanasi, "Interference alignment under limited feedback for MIMO interference channels," in *Proc. of IEEE International Symposium on Information Theory (ISIT)*, Austin, Texas, Jul. 2010, pp. 619–623.

[11] B. Xie, Y. Li, H. Minn, and A. Nosratinia, "Interference alignment under training and feedback constraints," in *Proc. of IEEE Global Telecommunications Conference (Globecom)*, Apr. 2011.

[12] M. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," in *Proc. Allerton Conf. on Communication, Control and Computing*, Allerton, USA, Sep. 2010, p. 1188.

[13] www.proofwiki.org/wiki/Laplace's_Expansion_Theorem