

# Retrospective Interference Alignment over Interference Networks

Hamed Maleki\*, Syed A. Jafar\* and Shlomo Shamai (Shitz)<sup>†</sup>

\*CPCC, EECS Dept., University of California Irvine, Irvine, CA 92697, USA

<sup>†</sup>EE Dept., Technion-Israel Institute of Technology, Technion City, Haifa 32000, Israel

Email: {hmaleki,syed}@uci.edu, sshlomo@ee.technion.ac.il

**Abstract**—Maddah-Ali and Tse recently introduced the idea of retrospective interference alignment, i.e., achieving interference alignment with only outdated (stale) channel state information at the transmitter (CSIT), in the context of the vector broadcast channel. Since the scheme relies on the centralized transmitter’s ability to reconstruct all the interference seen in previous symbols, it is not clear if retrospective interference alignment can be applied in interference networks consisting of distributed transmitters and receivers, where interference is contributed by multiple transmitters, each of whom can reconstruct only the part of the interference caused by themselves. In this work we prove that even in such settings, retrospective interference alignment is feasible. Specifically, we show that it is possible to achieve more than 1 DoF based on only delayed CSIT in the 3 user interference channel and the 2 user X channel consisting of only single antenna nodes. Retrospective interference alignment is also shown to be possible in other settings, such as the 2 user MIMO interference channel and with delayed channel output feedback.

the conventional wisdom that because channel coherence intervals are only relevant in the training and channel estimation at the receivers, they cannot affect capacity when perfect CSIR is assumed (thereby taking training and channel estimation issues at the receiver out of the picture). Indeed in the single user setting channel capacity with no CSIT and perfect CSIR is independent of the channel coherence interval. It is therefore surprising that under no CSIT and perfect CSIR, not only the network capacity but also the network DoF depend on the channel coherence intervals, making a case against the commonly used i.i.d. fading model for studying the capacity of wireless networks. Blind interference alignment shows that heterogenous channel coherence structures, whether present naturally in the form of different channel coherence times and coherence bandwidths experienced by different users [11], or imposed artificially by antenna switching [12], create opportunities for interference alignment. This is mainly because of the insight that because the receivers observe a linear combination of all transmitted symbols with the coefficients determined by the channel realizations, a simple repetition of information symbols from all transmitters will create linearly independent combinations of the transmitted symbols for the receivers that experience changing channel conditions, but will repeat the same linear combination of transmitted symbols for those receivers whose channel conditions remain fixed. Thus, each receiver’s desired symbols should be transmitted across those dimensions where its own channel changes while the other (undesired) receivers’ channels remain fixed. This allows each receiver to resolve its desired symbols which appear in linearly independent combinations each time, while all interference symbols (because they are observed in the same linear combination each time) are aligned along the all-ones vector — thereby achieving interference alignment.

## I. INTRODUCTION

There is much interest in studying the degrees of freedom (DoF) — and thereby exploring the potential for interference alignment (IA) — in wireless networks in the absence of instantaneous CSIT [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13]. On the one hand, there are pessimistic results that include [1], [2], [3], [4], [8], [9], [10] where the DoF are found to collapse due to the inability of the transmitters to resolve spatial dimensions. On the other hand, there are more recent optimistic results [5], [6], [7], [11], [12], [13] where the feasibility of interference alignment is demonstrated under various models of channel uncertainty at the transmitter(s). Closely related to this work are the papers on blind interference alignment [11], [12] and especially the recent work on interference alignment with delayed CSIT [13].

Reference [11] introduces the idea of blind interference alignment, i.e., the ability to achieve interference alignment without any knowledge (not even delayed knowledge) of channel coefficient realizations at the transmitters, by exploiting only statistical knowledge in the form of heterogeneous channel coherence structures associated with different users. The sensitivity of network DoF to channel coherence structures under the assumption of no CSIT and perfect CSIR goes against

More recently, reference [13] has introduced the delayed CSIT model, that will also be the main focus of this paper. The delayed CSIT model assumes i.i.d. fading channel conditions, with no knowledge of current channel state at the transmitter. However, perfect knowledge of channel states is available to the transmitter with some delay. The surprising finding of [13], in the context of the vector broadcast (BC) channel, is that not only is CSIT helpful even when it is outdated, but also that it can have a very significant impact as it is capable of increasing the DoF. The delayed CSIT model studied in [13] is particularly relevant in practice where invariably a delay is involved in any feedback from the receivers to the

The work of Syed Jafar and Hamed Maleki was supported by ONR under grant N00014-12-1-0067 and by NSF CCF under grants 0546860 and 0830809.

The work of S. Shamai was supported by the Israel Science Foundation (ISF).

transmitters. Several recent works point out that channel state information (CSI) can be estimated in principle with sufficient accuracy (estimation error scaling as  $O(SNR^{-1})$ ) to enable the DoF results as SNR becomes large [14], [15]. The main obstacle, from a practical perspective, has been the perceived necessity of delivering this CSI to the transmitter before it becomes outdated. The delayed CSIT model therefore opens a practically meaningful direction to explore the benefits of interference alignment. However, it is one of many possible forms that (delayed) feedback can take in a wireless network.

The terminology for three closely related delayed feedback models is delineated below through a simple example, for ease of reference in the sequel.

*Delayed Feedback Models:* Consider a simple Gaussian channel:

$$Y = HX + N \quad (1)$$

where  $X, Y, N, H$  are the transmitted symbol, the channel output symbol, the additive noise and the i.i.d. time-varying channel, respectively. Perfect channel state information at the receiver (CSIR) is modeled by the assumption that in addition to the channel output symbol  $Y$ , the receiver also receives the instantaneous channel state  $H$  over each channel use. By *delayed* feedback, what is meant is that the information being made available to the transmitter through the feedback channel is based only on past observations at the receivers and, in particular, is independent of the current channel state. Three different settings may be considered.

- 1) **Delayed CSIT:** This is the setting where the feedback provides the transmitters only the values of the past channel states  $H$  but not the output signals.
- 2) **Delayed Output Feedback:** This is the setting where the feedback provides the transmitters only the past received signals  $Y$  seen by the receivers, but not the channel states explicitly.
- 3) **Delayed Shannon Feedback:** This is the setting where the feedback provides the transmitters both the past received signals  $Y$  as well as the past channel states  $H$ .

The definitions extend naturally to multiuser settings, although the amount of feedback, and the possible associations between transmitters and receivers on the feedback channel may give rise to many special cases. Clearly, delayed Shannon feedback is the strongest delayed feedback setting, i.e., it can be weakened to obtain either delayed CSIT or the delayed output feedback model by discarding some of the feedback information. Between delayed CSIT and delayed output feedback, neither is a weakened form of the other because in general the output signals  $Y$  (even while discounting noise) cannot be inferred from the knowledge of channel states  $H$  (e.g., when there is more than one transmitter), and the channel states  $H$  cannot be deduced in general from the channel outputs  $Y$  even if these are noiselessly made available to the transmitter (e.g., when there are more transmit antennas than receive antennas).

In this work, we will be concerned primarily with the delayed CSIT model, and to a lesser extent, with delayed

output feedback. The main question we ask is whether it is possible to achieve more than 1 DoF in interference networks consisting of distributed transmitters and receivers with only delayed CSIT and with only single antenna nodes. For these settings, achieving more than 1 DoF is equivalent to the feasibility of retrospective interference alignment. The main contribution of this paper is to answer this question in the affirmative by presenting a retrospective interference alignment (i.e., interference alignment with delayed CSIT) scheme to achieve more than 1 DoF for the 3 user interference channel and the 2 user X channel. We also note that while the proposed achievable schemes are the first to establish the feasibility of retrospective interference alignment with distributed transmitters, the DoF value remains open in these settings because no tight outer bounds are available, and indeed further improvements in achievable DoF may be possible as shown recently in [18], [19]. An exception is the representative 2 user MIMO interference channel setting solved in this paper where indeed our proposed retrospective IA scheme is recently shown to be DoF optimal [20].

## II. SIMILARITIES BETWEEN BLIND IA AND RETROSPECTIVE IA

While the channel models studied in [11], [12] and [13] are quite different, there are some remarkable *essential* similarities in the achievable schemes used in both works that are further expanded in the present work. We start by pointing out the similar aspects of [11], [12], [13] through a few examples before proceeding to the main contribution of this work.

### A. Vector BC with no instantaneous CSIT

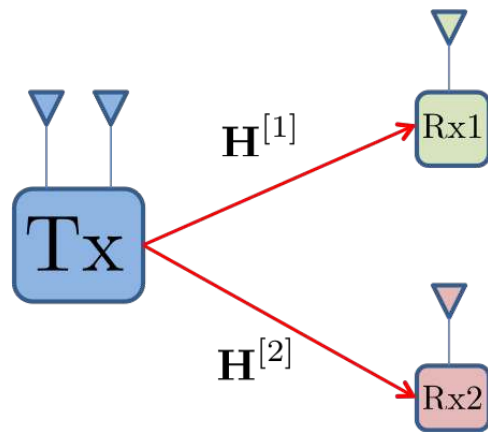


Fig. 1. Vector BC with No Instantaneous CSIT

Consider the broadcast channel (BC) with two single-antenna users, where the transmitter is equipped with two antennas. It is well known that with perfect channel knowledge this channel has 2 DoF, which can be achieved quite simply by zero forcing. However, the DoF are unknown with partial/limited CSIT for most cases and it is believed that no more than 1 DoF may be achievable in general. An outer bound of  $\frac{4}{3}$  DoF has been shown to be applicable to a wide variety of limited CSIT scenarios [4], [5], [13]. The  $\frac{4}{3}$  DoF

outer bound applies to the heterogeneous block fading models studied in [11] where the two users have staggered coherent fading blocks and is shown to be achievable. The same  $\frac{4}{3}$  DoF outer bound applies to the delayed CSIT setting and is achievable in that setting as well [13]. In *both* cases, i.e., both for blind [11] and retrospective [13] IA, the achievable scheme is described as follows:

- Use the channel three times to send two information symbols to each user.
- In one time slot, two symbols for user 1 are sent simultaneously from the two transmit antennas, providing user 1 one linear combination of the two desired symbols. One more linear combination would be needed to resolve the desired symbols at user 1. User 2 simultaneously also obtains a linear combination of these undesired symbols, but does not need to resolve them.
- In another time slot, two symbols for user 2 are sent simultaneously from the two transmit antennas, thus similarly providing user 2 one linear combination of his desired symbols, and providing user 1 a linear combination of these undesired symbols.
- In the final transmission, both users are simultaneously provided another linear combination of their respective desired symbols. The key to alleviate interference in this time slot is that the linear combination of undesired symbols received by a user in the third time slot is *the same linear combination* that he already received earlier, which allows the receiver to cancel the interference and then recover his desired symbols.

The key to alignment in both cases is the final transmission. Because of different channel models, the manner in which the final transmission is accomplished is different in [11] and [13]. In [11] the staggered coherence times ensure that each user receives the final transmission over a channel state that is identical to his channel state when he received interference in a prior time slot, but different from his channel state where he previously received his desired symbols, allowing interference cancellation. In [13] the same effect is accomplished by using delayed CSIT feedback. In the third time slot – because the transmitter now knows the channel states from the previous two time slots – the transmitter is able to send (from only one antenna) a superposition of the linear combinations of the undesired symbols seen previously by the two users. Since these linear combinations were received before, undesired information is easily canceled, leaving only the linear combination of desired symbols that provides the second equation in order to solve for the two desired variables. Since each user is able to resolve his 2 desired symbols over a total of 3 time slots, the DoF of  $\frac{4}{3}$  are achieved.

Continuing the comparison, we note that blind IA has an insignificant overhead relative to retrospective IA because blind IA does not require any feedback of channel coefficient values. Blind interference alignment also does not require joint processing of signals by the two transmit antennas, so that the same  $\frac{4}{3}$  DoF are achieved in the 2 user MISO BC as well as the 2 user X channel obtained by separating the transmit antennas. On the other hand the retrospective IA scheme of

[13] does not work with distributed transmit antennas, i.e., in the X channel setting, and while no outer bounds are available, we expect that the loss of DoF in the X channel setting relative to the MISO BC may be fundamental. Further, blind IA can achieve in general higher DoF than retrospective IA even for the MISO BC setting. However, all these benefits of blind IA come at the cost of requiring a suitable heterogeneous coherence structure across users, whereas retrospective IA, in spite of its significant overhead and DoF loss, does not rely on restrictive assumptions on channel fading model.

### B. Vector BC with instantaneous CSIT for User 1

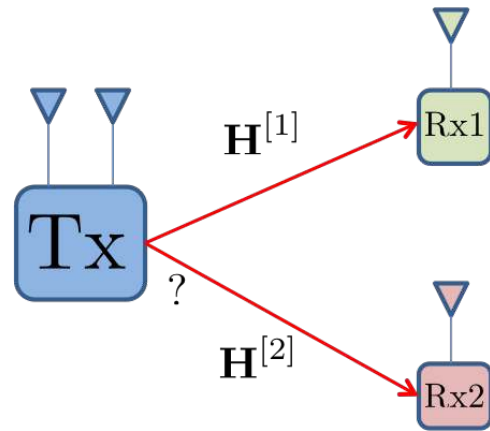


Fig. 2. Vector BC with Instantaneous CSIT for User 1

Consider the same vector BC as before, with the difference that now we assume the channel state of User 1 is instantaneously available to the transmitter. The DoF outer bound in this case is  $\frac{3}{2}$  [5] and is also applicable to a broad class of channel uncertainty models for User 2. In particular, it can be shown to be applicable to the delayed Shannon feedback model, by providing all the information available to Receiver 2 also to Receiver 1, thus making it a physically degraded broadcast channel for which it is known that feedback does not increase capacity [16]. Then, proceeding without feedback, the outer bound arguments for the compound setting in [5] can be extended to this setting, producing the same DoF outer bound of  $\frac{3}{2}$ . Note that an outer bound for delayed Shannon feedback is also an outer bound for delayed CSIT. Similar arguments (without feedback) are applicable to show the  $\frac{3}{2}$  DoF outer bound for blind interference alignment (heterogeneous block fading) model as well. Further, the achievability of  $\frac{3}{2}$  DoF can be shown under both settings of staggered block fading and delayed CSIT in essentially the same manner, as described below.

- Use the channel twice to send two information symbols to User 1 and one information symbol to User 2.
- User 2's information symbol is sent along a beamforming vector orthogonal to User 1's known channel vector, so User 1 sees no interference due to User 2.
- User 1's two symbols are sent twice in a manner that the same linear combination of the two undesired symbols is experienced by User 2 in both timeslots. This allows User

2 to cancel interference from one of the two time slots to recover his one desired symbol. However, User 1 sees two different linear combinations of his desired symbols and no interference, so he is able to recover both symbols.

Thus, 1 DoF is achieved by User 1 and 0.5 DoF by User 2, for a total of 1.5 DoF which is also the outer bound. The key here is to transmit the same linear combination of undesired symbols for User 2 while User 1 sees two different linear combinations. [11] does this by assuming that the channel stays constant for User 2 while it changes for User 1. The delayed CSIT setting can also accomplish the same effect because once the transmitter learns the channel states, it knows the linear combination of undesired symbols already seen by User 2 and re-sends the same linear combination from one antenna. Since this linear combination has not yet been seen by User 1, it gives him the second equation he needs, while for User 2 it is just a repetition of the previously seen interference which can be cancelled. Interestingly, in both cases (staggered block fading or delayed CSIT) the channel knowledge of User 1 (i.e., the user with the known channel) is only needed by the transmitter for one of the two channel uses (because User 2's symbol can be transmitted only once [11]).

### C. The (1,2,3,4) Two User MIMO Interference Channel

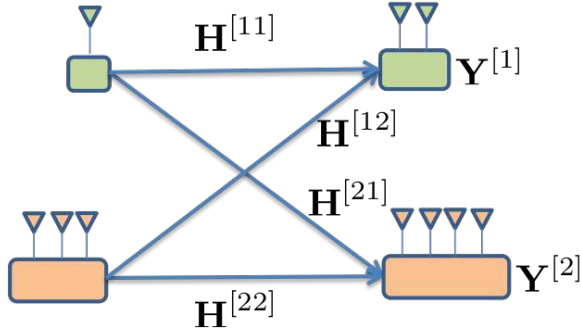


Fig. 3. (1,2,3,4) MIMO IC

Shown in Fig. 3, the (1,2,3,4) MIMO IC is a two user interference channel where User 1 has 1 transmit and 2 receive antennas, while User 2 has 3 transmit and 4 receive antennas. This particular channel configuration was first highlighted in [8] as an example where interference alignment was a possibility. Specifically, the question is, what is the maximum DoF possible for User 2 if User 1 simultaneously achieves his maximum value of 1 DoF? With no interference alignment the result would be only 1 DoF for user 2, but with interference alignment it may be possible to achieve up to 1.5 DoF for User 2. While recent work in [10] showed that under i.i.d. isotropic fading, User 2 cannot achieve more than 1 DoF, thus precluding the possibility of interference alignment, the question remains open for non-iid, non-isotropic models. In particular, [11] shows that under a staggered coherence block fading model, User 2 may indeed achieve 1.5 DoF at the same time that User 1 achieves his maximum DoF =1. Here we briefly summarize how this result is shown in [11] and how the same scheme can be translated to the delayed CSIT setting, once again highlighting the similarity of the two.

- Operate the channel over two time slots, sending a new symbol in each time slot from Transmitter 1 and sending three symbols from Transmitter 2, repeated over two slots.
- The key is to send the three symbols from Transmitter 2 in such a manner that User 1, at his two receive antennas, sees the same two linear combinations of undesired symbols twice. Thus, in each time slot he is able to free up one dimension by cancelling the corresponding interference using the same linear combination received in the other time slot, and recover his desired symbol.

In order to make sure that the same two linear combinations of undesired symbols are seen on the two receive dimensions twice by Receiver 1, reference [11] exploits the assumption that the channel between Transmitter 2 and Receiver 1 stays constant over two time slots. However, in the delayed CSIT setting, after the first time slot Transmitter 2 knows the CSIT from the first time slot and therefore also knows the two different linear combinations of the 3 transmitted symbols observed at the two receive antennas of Receiver 1. In time slot 2 then Transmitter 2 sends the two linear combinations, each from a different transmit antenna. Since only 2 antennas are used in the second time slot, the  $2 \times 2$  channel can be inverted by Receiver 1 to essentially experience the same two linear combinations of undesired symbols on the two resulting receive antennas. Thus, one receive antenna can be cleared of interference in each time slot by using the corresponding linear combination of undesired symbols observed in the other time slot, allowing the desired symbol to be resolved in each time slot.

### III. DIFFERENCES BETWEEN HETEROGENEOUS BLOCK FADING AND DELAYED CSIT SETTINGS

In spite of the strong similarities between the heterogeneous fading model [11] and the delayed CSIT [13] model, the two settings have very marked differences in general. The difference becomes evident as soon as the question of alignment of signals from *different* transmitters comes to light. Note that in all examples described above the signals being aligned are from the same transmitter. However, when we go to more distributed settings, e.g., the  $K$  user interference channels or the  $X$  channels the difference between the two settings is quite stark. With *suitably* staggered coherent times/antenna switching, it is shown in [11], [12] that neither the  $K$  user interference channel, nor the  $X$  network with  $M$  transmitters and  $N$  receivers lose *any* DoF relative to the setting where perfect, global and instantaneous CSIT is available. The  $K$  user interference channel can still achieve  $\frac{K}{2}$  DoF, and the  $M \times N$  user  $X$  network can still achieve  $\frac{MN}{M+N-1}$  DoF, even with no instantaneous (or delayed) CSIT besides the knowledge of the channel coherence structure/antenna switching pattern. However, in the delayed CSIT setting, (albeit only with i.i.d fading) the outer bounds clearly show a loss in DoF. Thus, evidently there is no complete theoretical equivalence between the two settings.

The significance of distributed transmitters brings us to the issue central to the present work. We know that in the delayed

CSIT setting, DoF are lost because the DoF outer bound for the vector BC (say with  $K$  transmit antennas, and  $K$  single antenna users) loses to the  $\frac{K}{2}$  DoF that are achievable in the  $K$  user interference channel both in the compound setting (with finite number of generic states [6]) as well as with suitable staggered coherence block model. However, a very interesting question remains open – Does transmitter cooperation improve DoF even with channel uncertainty at transmitters?

The main difference between distributed transmitters (e.g. the 3 user interference channel and the X channel) and the vector BC is that in the vector BC the transmitters are allowed to cooperate as multiple antennas of one transmit node. With full CSIT it is well known that the vector BC has larger DoF than the 3 user interference channel and the X channel, so the cooperation between transmitters increases DoF. However, with limited CSIT, e.g., in the compound channel setting [5], [6], [7], somewhat surprisingly, several recent results have shown that the DoF are the same whether it is the X channel or the corresponding vector BC. Indeed, with the staggered fading model of [11], the same approach works for the X channel as does for the vector BC. The delayed CSIT setting gives us another framework within which one can hope to gain additional insights into the role of transmitter cooperation in determining the available DoF under channel uncertainty.

The difficulty in dealing with distributed transmitters is underscored by the following observations.

- While the retrospective IA scheme for the vector BC, proposed in [13], requires the transmitter to reconstruct previously seen linear combinations of interference terms based only on delayed CSIT, this cannot be done in the 2 user X channel or the 3 user interference channel because the interference terms are combinations of symbols originating at more than one transmitter and no transmitter has access to the symbols transmitted from other transmitters.
- Aligning interfering symbols from the same transmitter does not produce DoF benefits when all receivers are equipped with at least as many antennas as the transmitters. This is because the channels are invertible, so that if two undesired symbols from the same transmitter are seen aligned at a particular receiver, it must be because they are aligned at the transmitter itself. Now, if these symbols are aligned at the transmitter itself, then they cannot be resolved further downstream at any other receiver, and in particular, at the receiver that wants these two symbols and therefore must not see them aligned together.

The two observations listed above, seem to leave very little hope of achieving interference alignment with distributed transmitters. However, as we show in this paper, this is not the case, and retrospective IA is indeed feasible, even with distributed transmitters. Specifically we show that the 3 user interference channel and the 2 user X channel can each achieve more than 1 DoF. We start with the 3 user interference channel.

#### IV. INTERFERENCE CHANNEL WITH 3 USERS

The interference channel with more than 2 users is one of the earliest settings where interference alignment was first introduced [23], and as such it is natural to ask if interference alignment is possible in this setting with only delayed CSIT?

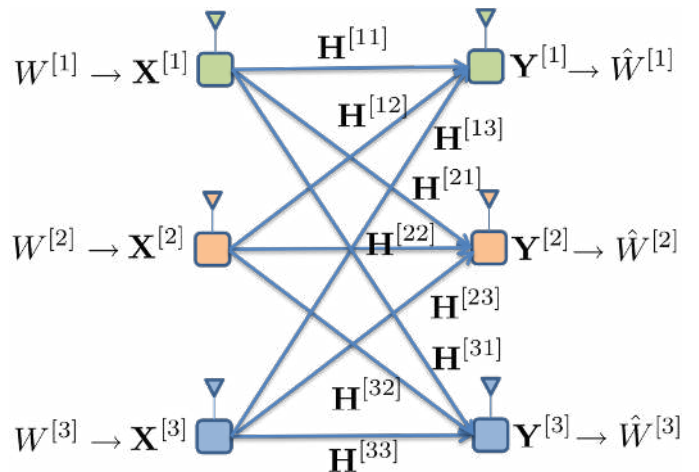


Fig. 4. Interference Channel with 3 Users

The 3 user interference channel, shown in Fig. 4, consists of transmitters 1,2,3, who wish to communicate independent messages  $W^{[1]}, W^{[2]}, W^{[3]}$  to receivers 1,2,3 respectively. We assume that the channels vary in an i.i.d. fashion according to some continuous distribution with support that is bounded away from zero and infinity. The receivers are assumed to have perfect knowledge of all channel states. The transmitters do not know the current channel state, but they do have access to all channel states up to the previous channel use. This model is referred to as the *delayed CSIT* model. The definitions of achievable data rates, capacity, power constraints, DoF are all used in the standard sense as in e.g. [13] and will not be repeated here. We also assume that the reader is familiar with DoF analysis when working with linear beamforming schemes, and in particular the requirements for interference alignment. For these and other standard issues such as – why we ignore noise in this analysis, what are the conditions for desired signals to be recovered in the presence of interference, a literature survey of earlier works on interference alignment and DoF such as [17], [21], [22], [23] is recommended. The following theorem presents an achievability result for the DoF of the 3 user interference channel with delayed CSIT.

*Theorem 1:* The 3 user interference channel with delayed CSIT, can achieve  $\frac{9}{8}$  DoF almost surely.

*Proof:* In order to show the achievability of  $\frac{9}{8}$  DoF, we will consider a 8 symbol extension of the channel. Each user will send 3 information symbols over these 8 channel uses. At each receiver, in addition to the 3 desired symbols, there are 6 interfering symbols. Since the total number of dimensions is only 8, one of the 6 interfering symbols must align within the vector space spanned by the remaining 5, to leave 3 signal dimensions free of interference where the desired signals can be projected.

**Phase I:** As stated earlier, we wish that the 6 interfering symbols should span no more than 5 dimensions. Since we have no instantaneous channel knowledge, let us start by sending random linear combinations of the symbols from each transmitter. Since interference is allowed to fill up 5 dimensions, we can send 5 random linear combinations of the symbols from each transmitter over the first 5 channel

uses without exceeding the quota of 5 dimensions that are allowed to be spanned by interference. This is the end of Phase I. No special effort has been made to align anything so far, and we have exhausted the number of dimensions allowed for interference at each receiver.

At this point, consider the signal seen by Receiver 1 (ignoring noise as usual).

$$\mathbf{Y}^{[1]} = \mathbf{H}^{[11]} \begin{bmatrix} \mathbf{V}_1^{[1]} & \mathbf{V}_2^{[1]} & \mathbf{V}_3^{[1]} \end{bmatrix} \begin{bmatrix} u_1^{[1]} \\ u_2^{[1]} \\ u_3^{[1]} \end{bmatrix} + \mathbf{H}^{[12]} \begin{bmatrix} \mathbf{V}_1^{[2]} & \mathbf{V}_2^{[2]} & \mathbf{V}_3^{[2]} \end{bmatrix} \begin{bmatrix} u_1^{[2]} \\ u_2^{[2]} \\ u_3^{[2]} \end{bmatrix} + \mathbf{H}^{[13]} \begin{bmatrix} \mathbf{V}_1^{[3]} & \mathbf{V}_2^{[3]} & \mathbf{V}_3^{[3]} \end{bmatrix} \begin{bmatrix} u_1^{[3]} \\ u_2^{[3]} \\ u_3^{[3]} \end{bmatrix} \quad (2)$$

Here  $\mathbf{V}_i^{[k]}$  are the  $5 \times 1$  precoding vectors,  $\mathbf{Y}^{[k]}$  is the  $5 \times 1$  vector of received signals so far,  $\mathbf{H}^{[kj]}$  is the  $5 \times 5$  diagonal channel matrix representing the i.i.d. variations of the channel coefficient from Transmitter  $j$  to Receiver  $k$ .

Consider the interference carrying vectors  $\mathbf{V}_i^{[2]}, \mathbf{V}_i^{[3]}, i = 1, 2, 3$ , over the first 5 channel uses. Since these six vectors are only in a five dimensional space, we can identify the  $6 \times 1$  null vector  $\bar{\alpha}^{[1]} = [\alpha_1^{[1]}, \alpha_2^{[1]}, \dots, \alpha_6^{[1]}]$  such that:

$$\begin{bmatrix} \mathbf{H}^{[12]} \mathbf{V}_1^{[2]} & \mathbf{H}^{[12]} \mathbf{V}_2^{[2]} & \mathbf{H}^{[12]} \mathbf{V}_3^{[2]} & \mathbf{H}^{[13]} \mathbf{V}_1^{[3]} \\ & \mathbf{H}^{[13]} \mathbf{V}_2^{[3]} & \mathbf{H}^{[13]} \mathbf{V}_3^{[3]} \end{bmatrix} \bar{\alpha}^{[1]} = \mathbf{0}_{5 \times 1} \quad (3)$$

Similarly, considering Receivers 2, and 3 respectively, we define null vectors  $\bar{\alpha}^{[2]}, \bar{\alpha}^{[3]}$  so that

$$\begin{bmatrix} \mathbf{H}^{[21]} \mathbf{V}_1^{[1]} & \mathbf{H}^{[21]} \mathbf{V}_2^{[1]} & \mathbf{H}^{[21]} \mathbf{V}_3^{[1]} & \mathbf{H}^{[23]} \mathbf{V}_1^{[3]} \\ & \mathbf{H}^{[23]} \mathbf{V}_2^{[3]} & \mathbf{H}^{[23]} \mathbf{V}_3^{[3]} \end{bmatrix} \bar{\alpha}^{[2]} = \mathbf{0}_{5 \times 1} \quad (4)$$

$$\begin{bmatrix} \mathbf{H}^{[31]} \mathbf{V}_1^{[1]} & \mathbf{H}^{[31]} \mathbf{V}_2^{[1]} & \mathbf{H}^{[31]} \mathbf{V}_3^{[1]} & \mathbf{H}^{[32]} \mathbf{V}_1^{[2]} \\ & \mathbf{H}^{[32]} \mathbf{V}_2^{[2]} & \mathbf{H}^{[32]} \mathbf{V}_3^{[2]} \end{bmatrix} \bar{\alpha}^{[3]} = \mathbf{0}_{5 \times 1} \quad (5)$$

**Phase II:** While everything so far is consistent with the approaches of [11], [13], at this point our approach goes into a new direction. As mentioned earlier, because the transmitters are distributed, they do not have access to any of the equations available so far to the receivers – because each equation contains symbols from both transmitters. Thus, unlike the broadcast setting studied in [13] where the transmit antennas are co-located, we cannot construct layer-2 variables out of the equations already available to the receivers. Instead, we will construct layer-2 variables out of the symbols available at each transmitter. Phase 2 consists of the remaining 3 channel uses. It is here that retrospective alignment will be used, based on the knowledge of the channel states from Phase I. No knowledge of channel states, not even delayed CSIT, will be used of the Phase-2 channels. Consider the  $n^{th}$ ,  $n = 6, 7, 8$ , transmission, i.e., any of the three transmissions of Phase 2. Suppose the transmitters choose to send the linear combinations:

$$\text{Transmitter 1: } \rightarrow \mathbf{V}_1^{[1]}(n)u_1^{[1]} + \mathbf{V}_2^{[1]}(n)u_2^{[1]} + \mathbf{V}_3^{[1]}(n)u_3^{[1]} \quad (6)$$

$$\text{Transmitter 2: } \rightarrow \mathbf{V}_1^{[2]}(n)u_1^{[2]} + \mathbf{V}_2^{[2]}(n)u_2^{[2]} + \mathbf{V}_3^{[2]}(n)u_3^{[2]} \quad (7)$$

$$\text{Transmitter 3: } \rightarrow \mathbf{V}_1^{[3]}(n)u_1^{[3]} + \mathbf{V}_2^{[3]}(n)u_2^{[3]} + \mathbf{V}_3^{[3]}(n)u_3^{[3]} \quad (8)$$

where  $\mathbf{V}_i^{[k]}(n)$  are the linear precoding coefficients of symbol  $u_i^{[k]}$  at Transmitter  $k$  and time slot  $n$ ,  $n = 6, 7, 8$ . These linear precoding coefficients can be chosen by the transmitters based on the delayed CSIT, i.e., the knowledge of the channel coefficients from Phase-I.

An important observation here is the following. In order to keep the interference contained in a 5 dimensional space at each receiver, the Phase-2 precoding coefficients must follow the same linear relationship as established in Phase-I. Mathematically, at Receiver 1:

$$\begin{bmatrix} \mathbf{H}^{[12]}(n)\mathbf{V}_1^{[2]}(n) & \mathbf{H}^{[12]}(n)\mathbf{V}_2^{[2]}(n) & \mathbf{H}^{[12]}(n)\mathbf{V}_3^{[2]}(n) \\ \mathbf{H}^{[13]}(n)\mathbf{V}_1^{[3]}(n) & \mathbf{H}^{[13]}(n)\mathbf{V}_2^{[3]}(n) & \mathbf{H}^{[13]}(n)\mathbf{V}_3^{[3]}(n) \end{bmatrix} \bar{\alpha}^{[1]} = \mathbf{0} \quad (9)$$

and similarly at Receivers 2, 3:

$$\begin{bmatrix} \mathbf{H}^{[21]}(n)\mathbf{V}_1^{[1]}(n) & \mathbf{H}^{[21]}(n)\mathbf{V}_2^{[1]}(n) & \mathbf{H}^{[21]}(n)\mathbf{V}_3^{[1]}(n) \\ \mathbf{H}^{[23]}(n)\mathbf{V}_1^{[3]}(n) & \mathbf{H}^{[23]}(n)\mathbf{V}_2^{[3]}(n) & \mathbf{H}^{[23]}(n)\mathbf{V}_3^{[3]}(n) \end{bmatrix} \bar{\alpha}^{[2]} = \mathbf{0} \quad (10)$$

$$\begin{bmatrix} \mathbf{H}^{[31]}(n)\mathbf{V}_1^{[1]}(n) & \mathbf{H}^{[31]}(n)\mathbf{V}_2^{[1]}(n) & \mathbf{H}^{[31]}(n)\mathbf{V}_3^{[1]}(n) \\ \mathbf{H}^{[32]}(n)\mathbf{V}_1^{[2]}(n) & \mathbf{H}^{[32]}(n)\mathbf{V}_2^{[2]}(n) & \mathbf{H}^{[32]}(n)\mathbf{V}_3^{[2]}(n) \end{bmatrix} \bar{\alpha}^{[3]} = \mathbf{0} \quad (11)$$

Since the current channel states  $\mathbf{H}^{[**]}(n)$  are not known to the transmitters, the only way to guarantee the above equations for all current channel realizations is to choose  $\mathbf{V}_i^{[k]}$  so that

$$\mathbf{V}_1^{[1]}(n)\alpha_1^{[2]} + \mathbf{V}_2^{[1]}(n)\alpha_2^{[2]} + \mathbf{V}_3^{[1]}(n)\alpha_3^{[2]} = 0 \quad (12)$$

$$\mathbf{V}_1^{[1]}(n)\alpha_1^{[3]} + \mathbf{V}_2^{[1]}(n)\alpha_2^{[3]} + \mathbf{V}_3^{[1]}(n)\alpha_3^{[3]} = 0 \quad (13)$$

$$\mathbf{V}_1^{[2]}(n)\alpha_1^{[1]} + \mathbf{V}_2^{[2]}(n)\alpha_2^{[1]} + \mathbf{V}_3^{[2]}(n)\alpha_3^{[1]} = 0 \quad (14)$$

$$\mathbf{V}_1^{[2]}(n)\alpha_4^{[3]} + \mathbf{V}_2^{[2]}(n)\alpha_5^{[3]} + \mathbf{V}_3^{[2]}(n)\alpha_6^{[3]} = 0 \quad (15)$$

$$\mathbf{V}_1^{[3]}(n)\alpha_4^{[1]} + \mathbf{V}_2^{[3]}(n)\alpha_5^{[1]} + \mathbf{V}_3^{[3]}(n)\alpha_6^{[1]} = 0 \quad (16)$$

$$\mathbf{V}_1^{[3]}(n)\alpha_4^{[2]} + \mathbf{V}_2^{[3]}(n)\alpha_5^{[2]} + \mathbf{V}_3^{[3]}(n)\alpha_6^{[2]} = 0 \quad (17)$$

Now consider the precoding coefficients  $\mathbf{V}_i^{[1]}(n), i = 1, 2, 3$  that are to be chosen by Transmitter 1. Based on equations (12), (13), we can express  $\mathbf{V}_2^{[1]}(n)$  and  $\mathbf{V}_3^{[1]}(n)$  as linear functions of  $\mathbf{V}_1^{[1]}(n)$ , say

$$\mathbf{V}_2^{[1]}(n) = \frac{\begin{bmatrix} -\alpha_1^{[2]} & \alpha_3^{[2]} \\ -\alpha_1^{[3]} & \alpha_3^{[3]} \end{bmatrix}}{\begin{bmatrix} \alpha_2^{[2]} & \alpha_3^{[2]} \\ \alpha_2^{[3]} & \alpha_3^{[3]} \end{bmatrix}} \mathbf{V}_1^{[1]}(n) \quad (18)$$

$$\mathbf{V}_3^{[1]}(n) = \frac{\begin{bmatrix} \alpha_2^{[2]} & -\alpha_1^{[2]} \\ \alpha_2^{[3]} & -\alpha_1^{[3]} \end{bmatrix}}{\begin{bmatrix} \alpha_2^{[2]} & \alpha_3^{[2]} \\ \alpha_2^{[3]} & \alpha_3^{[3]} \end{bmatrix}} \mathbf{V}_1^{[1]}(n) \quad (19)$$

Thus, Transmitter 1 is forced to send:

$$\begin{aligned} \mathbf{X}^{[1]}(n) &= \mathbf{V}_1^{[1]}(n)u_1^{[1]} + \mathbf{V}_2^{[1]}(n)u_2^{[1]} + \mathbf{V}_3^{[1]}(n)u_3^{[1]} \quad (20) \\ &= c((\alpha_2^{[2]}\alpha_3^{[3]} - \alpha_3^{[2]}\alpha_2^{[3]})u_1^{[1]} + (\alpha_3^{[2]}\alpha_1^{[3]} - \alpha_3^{[3]}\alpha_1^{[2]})u_2^{[1]} \\ &\quad + (\alpha_1^{[2]}\alpha_2^{[3]} - \alpha_2^{[2]}\alpha_1^{[3]})u_3^{[1]}) \\ &= cs^{[1]} \end{aligned}$$

where  $c$  is any constant, and without loss of generality we can set it to unity. The new information variable

$$\begin{aligned} s^{[1]} &= (\alpha_2^{[2]}\alpha_3^{[3]} - \alpha_3^{[2]}\alpha_2^{[3]})u_1^{[1]} + (\alpha_3^{[2]}\alpha_1^{[3]} - \alpha_3^{[3]}\alpha_1^{[2]})u_2^{[1]} \\ &\quad + (\alpha_1^{[2]}\alpha_2^{[3]} - \alpha_2^{[2]}\alpha_1^{[3]})u_3^{[1]} \quad (21) \end{aligned}$$

is precisely our Phase-2 variable, available only to Transmitter 1, and composed of variables only available to Transmitter 1. Thus, in Phase-2, even though the transmitter has three information symbols to send, it can only send scaled versions of the same effective scalar symbol  $s^{[1]}$  in order to keep the interference aligned within 5 dimensions at each receiver. Similarly, we can define the effective variables  $s^{[2]}$  and  $s^{[3]}$  to be sent by transmitters 2 and 3 over Phase 2.

$$s^{[2]} = (\alpha_2^{[1]} \alpha_6^{[3]} - \alpha_3^{[1]} \alpha_5^{[3]}) u_1^{[2]} + (\alpha_4^{[3]} \alpha_3^{[1]} - \alpha_6^{[3]} \alpha_1^{[1]}) u_2^{[2]} + (\alpha_1^{[1]} \alpha_5^{[3]} - \alpha_2^{[1]} \alpha_4^{[3]}) u_3^{[2]} \quad (22)$$

$$s^{[3]} = (\alpha_5^{[1]} \alpha_6^{[2]} - \alpha_6^{[1]} \alpha_5^{[2]}) u_1^{[3]} + (\alpha_6^{[1]} \alpha_4^{[2]} - \alpha_6^{[2]} \alpha_4^{[1]}) u_2^{[3]} + (\alpha_4^{[1]} \alpha_5^{[2]} - \alpha_4^{[2]} \alpha_5^{[1]}) u_3^{[3]} \quad (23)$$

Since there are only 3 Phase-2 symbols, and there are 3 channel uses, the operation over Phase-2 can be simply interpreted as each transmitter repeating its own effective information symbol, so that the channel variations provide each receiver with a different linear combination of the 3 effective Phase-2 symbols each time, so that at the end of Phase-2, each receiver is able to decode all three symbols  $s^{[1]}$ ,  $s^{[2]}$ ,  $s^{[3]}$ .

Thus, we have completely determined the resulting precoding vectors sent over the 8 symbols. Putting everything together, the transmitted symbols from e.g., Transmitter 1 are shown below:

*Retrospective Interference Alignment:* As mentioned above, the novelty of *retrospective interference alignment* lies in the construction of layer-2 variables. The first 5 channel uses correspond to Phase 1. All these precoding coefficients  $\mathbf{V}_i^{[1]}(n)$  are chosen independently, randomly, before the beginning of communication and with no knowledge of CSIT. The last three channel uses correspond to Phase 2, and can be easily seen to be repetitions of the Phase-2 variable  $s^{[1]}$ . The transmitted symbols for all other transmitters can be described similarly. Interference alignment is accomplished because this choice of precoding vectors satisfies equations (12)-(17), acting retrospectively.

Keeping the interference aligned within 5 dimensions at each receiver, allows the receiver to null out the 5 interference dimensions and recover the 3 desired symbols from the remaining 3 dimensions from the overall 8 dimensional vector space. While the construction above guarantees that interference is restricted within 5 dimensions, one must also show that the desired signal vectors are not aligned within the interference or aligned among themselves. This is proven, by constructing the  $8 \times 8$  matrix consisting of 3 desired signal vectors and 5 interference vectors that span the interference space received at each receiver (lets call it  $M_i$  at Receiver  $i$ ), and showing that the determinant of these matrices, which is equivalent to a polynomial in Phase 1 variables ( $\mathbf{V}, \mathbf{H}$ ), is not identically a zero polynomial. So  $\det(M_1) \det(M_2) \det(M_3)$  is either identically the zero polynomial, or it is non-zero almost surely for all realizations of  $\mathbf{V}, \mathbf{H}$ . To prove that it is non-zero almost surely, it suffices to show that it is not the zero polynomial, which in turn is established if there exists any non-zero evaluation of this determinant. One of such evaluations is as follows

$$\mathbf{V}^{[1]} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 2 & 1 \\ 2 & 1 & 0 \end{bmatrix} \quad \mathbf{V}^{[2]} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix} \quad \mathbf{V}^{[3]} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 2 & 1 \end{bmatrix}$$

$$\mathbf{H}^{[12]} = \mathbf{H}^{[13]} = \mathbf{H}^{[23]} = \mathbf{H}^{[32]} = \mathbf{I}_{5 \times 5}$$

$$\mathbf{H}^{[21]} = \mathbf{H}^{[31]} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{H}^{[11]} = \mathbf{H}^{[22]} = \mathbf{H}^{[33]} = \begin{bmatrix} 5 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

where  $\mathbf{I}_{5 \times 5}$  is a  $5 \times 5$  identity matrix. Based on these choices for  $\mathbf{V}$  and  $\mathbf{H}$  we have

$$\alpha^{[1]} = [-1 \quad -1 \quad -1 \quad 1 \quad 1 \quad 1]^T$$

$$\alpha^{[2]} = [-1 \quad 1 \quad -3 \quad 1 \quad -1 \quad 3]^T$$

$$\alpha^{[3]} = [-1 \quad 1 \quad -1 \quad 1 \quad -1 \quad 1]^T$$

and the precoding coefficients in the last three channel uses are determined as

$$\begin{bmatrix} \mathbf{V}_1^{[1]}(n) & \mathbf{V}_2^{[1]}(n) & \mathbf{V}_3^{[1]}(n) \end{bmatrix} = [1 \quad 1 \quad 0] \quad n = 6, 7, 8$$

$$\begin{bmatrix} \mathbf{V}_1^{[2]}(n) & \mathbf{V}_2^{[2]}(n) & \mathbf{V}_3^{[2]}(n) \end{bmatrix} = [-1 \quad 0 \quad 1] \quad n = 6, 7, 8$$

$$\begin{bmatrix} \mathbf{V}_1^{[3]}(n) & \mathbf{V}_2^{[3]}(n) & \mathbf{V}_3^{[3]}(n) \end{bmatrix} = [-2 \quad 1 \quad 1] \quad n = 6, 7, 8$$

The channel coefficients in the last 3 channel uses are chosen as follows

$$\mathbf{H}^{[11]} = \mathbf{H}^{[21]} = \mathbf{H}^{[31]} = \mathbf{I}_{3 \times 3}$$

$$\mathbf{H}^{[12]} = \mathbf{H}^{[22]} = \mathbf{H}^{[32]} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{H}^{[13]} = \mathbf{H}^{[23]} = \mathbf{H}^{[33]} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

To be more specific, lets compute the  $8 \times 8$  matrix  $M_i$  by eliminating the last column of interference (i.e. column related to  $u_3^{[3]}$  for Rx1 and Rx2, and column related to  $u_3^{[2]}$  for Rx3), since each column of interference is linearly dependent to the remaining 5 interference columns. We can evaluate  $\det(M_1) \det(M_2) \det(M_3) = 28.24 \cdot (-8) = -5376$  for these chosen values of  $\mathbf{V}$  and  $\mathbf{H}$  which is nonzero and hence  $\det(M_1) \det(M_2) \det(M_3)$  is a nonzero polynomial.

## V. THE X CHANNEL

The X channel consists of two transmitters, two receivers, and four independent messages, one from each transmitter to each receiver. The assumptions regarding the delayed CSIT model are identical to the preceding sections. Our goal is to explore the DoF that can be achieved by the X channel in this setting.

*Theorem 2:* The 2 user X channel with delayed CSIT, can achieve DoF =  $\frac{8}{7}$  almost surely.

*Proof:* We wish to show that  $\frac{8}{7}$  DoF are achievable on the X channel with delayed CSIT.

$$\begin{bmatrix} \mathbf{X}^{[1]}(1) \\ \mathbf{X}^{[1]}(2) \\ \mathbf{X}^{[1]}(3) \\ \mathbf{X}^{[1]}(4) \\ \mathbf{X}^{[1]}(5) \\ \mathbf{X}^{[1]}(6) \\ \mathbf{X}^{[1]}(7) \\ \mathbf{X}^{[1]}(8) \end{bmatrix} = \begin{bmatrix} \mathbf{V}_1^{[1]}(1) & \mathbf{V}_2^{[1]}(1) & \mathbf{V}_3^{[1]}(1) \\ \mathbf{V}_1^{[1]}(2) & \mathbf{V}_2^{[1]}(2) & \mathbf{V}_3^{[1]}(2) \\ \mathbf{V}_1^{[1]}(3) & \mathbf{V}_2^{[1]}(3) & \mathbf{V}_3^{[1]}(3) \\ \mathbf{V}_1^{[1]}(4) & \mathbf{V}_2^{[1]}(4) & \mathbf{V}_3^{[1]}(4) \\ \mathbf{V}_1^{[1]}(5) & \mathbf{V}_2^{[1]}(5) & \mathbf{V}_3^{[1]}(5) \\ \alpha_2^{[2]} \alpha_3^{[3]} - \alpha_3^{[2]} \alpha_2^{[3]} & \alpha_3^{[2]} \alpha_1^{[3]} - \alpha_3^{[3]} \alpha_1^{[2]} & \alpha_1^{[2]} \alpha_2^{[3]} - \alpha_2^{[2]} \alpha_1^{[3]} \\ \alpha_2^{[2]} \alpha_3^{[3]} - \alpha_3^{[2]} \alpha_2^{[3]} & \alpha_3^{[2]} \alpha_1^{[3]} - \alpha_3^{[3]} \alpha_1^{[2]} & \alpha_1^{[2]} \alpha_2^{[3]} - \alpha_2^{[2]} \alpha_1^{[3]} \\ \alpha_2^{[2]} \alpha_3^{[3]} - \alpha_3^{[2]} \alpha_2^{[3]} & \alpha_3^{[2]} \alpha_1^{[3]} - \alpha_3^{[3]} \alpha_1^{[2]} & \alpha_1^{[2]} \alpha_2^{[3]} - \alpha_2^{[2]} \alpha_1^{[3]} \end{bmatrix} \begin{bmatrix} u_1^{[1]} \\ u_2^{[1]} \\ u_3^{[1]} \end{bmatrix} \quad (24)$$

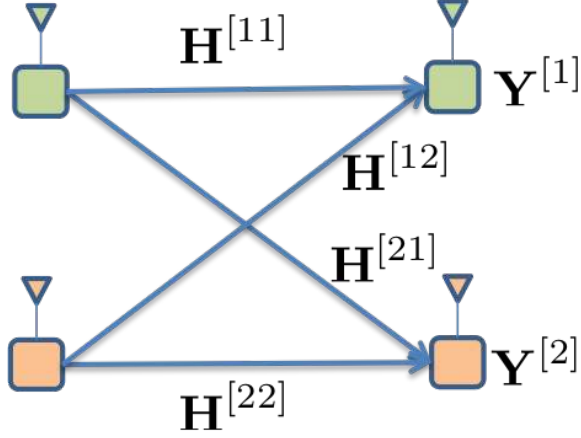


Fig. 5. X Channel

Since we are again dealing with delayed CSIT and distributed transmitters, we will again use the retrospective interference alignment scheme. However, in this section, for the sake of providing a richer understanding of the scheme, we will arrive at it in an alternative fashion.

To this end, consider a 7 symbol extension of the channel, i.e., we will consider 7 channel uses over which the precoding vectors will be designed. Each of the 4 messages  $\bar{W}^{[kj]}$  will be assigned two precoding vectors  $\bar{\mathbf{V}}^{[kj]} = \begin{bmatrix} \bar{\mathbf{V}}_1^{[kj]} & \bar{\mathbf{V}}_2^{[kj]} \end{bmatrix}$ , where  $\bar{\mathbf{V}}_i^{[kj]}$  is the  $i^{\text{th}}$  column of the  $7 \times 2$  matrix  $\bar{\mathbf{V}}^{[kj]}$ ,  $k = 1, 2$ , and will carry the scalar coded information symbol  $u_i^{[kj]}$ . Thus, over the 7 symbol block, the received signal at Receiver  $k$ ,  $k = 1, 2$ , can be expressed as:

$$\begin{aligned} \bar{\mathbf{Y}}^{[k]} &= \bar{\mathbf{H}}^{[k1]} \left( \bar{\mathbf{V}}_1^{[11]} u_1^{[11]} + \bar{\mathbf{V}}_2^{[11]} u_2^{[11]} + \bar{\mathbf{V}}_1^{[21]} u_1^{[21]} + \bar{\mathbf{V}}_2^{[21]} u_2^{[21]} \right) \\ &+ \bar{\mathbf{H}}^{[k2]} \left( \bar{\mathbf{V}}_1^{[12]} u_1^{[12]} + \bar{\mathbf{V}}_2^{[12]} u_2^{[12]} + \bar{\mathbf{V}}_1^{[22]} u_1^{[22]} + \bar{\mathbf{V}}_2^{[22]} u_2^{[22]} \right) + \bar{\mathbf{Z}}^{[k]} \quad (25) \end{aligned}$$

$$= \bar{\mathbf{H}}^{[k1]} \begin{bmatrix} \bar{\mathbf{V}}^{[11]} & \bar{\mathbf{V}}^{[21]} \end{bmatrix} \mathbf{U}^{[*1]} + \bar{\mathbf{H}}^{[k2]} \begin{bmatrix} \bar{\mathbf{V}}^{[12]} & \bar{\mathbf{V}}^{[22]} \end{bmatrix} \mathbf{U}^{[*2]} + \bar{\mathbf{Z}}^{[k]} \quad (26)$$

Here  $\bar{\mathbf{Y}}^{[k]}$ ,  $\bar{\mathbf{Z}}^{[k]}$  are  $7 \times 1$  vectors,  $\bar{\mathbf{H}}^{[kj]}$  are  $7 \times 7$  diagonal matrices and  $u_i^{[kj]}$  are scalar symbols.  $\mathbf{U}^{[*i]} = \begin{bmatrix} u_1^{[i1]}, u_2^{[i1]}, u_1^{[i2]}, u_2^{[i2]} \end{bmatrix}^T$  is the  $4 \times 1$  vector of information symbols originating at transmitter  $i$ .

The key task is to design the precoding vectors  $\bar{\mathbf{V}}_i^{[kj]}$  to achieve interference alignment. The elements of the precoding vector can depend on the channel state only retrospectively, i.e., at time  $n$ ,  $n = 1, 2, \dots, 7$ , the  $n^{\text{th}}$  element of a precoding vector can depend on the past channel states  $\bar{\mathbf{H}}^{[**]}(1), \dots, \bar{\mathbf{H}}^{[**]}(n-1)$  but not the present or future

channel states.

**First 3 Channel Uses:** For the first 3 channel uses each transmitter simply sends random linear combinations (the coefficients are generated randomly offline and shared between all transmitters and receivers before the beginning of communication) of the 4 symbols originating at that transmitter. A different linear combination is sent each time. This gives each receiver 3 equations in 8 variables. Because the combining coefficients are chosen randomly and independently of the channel realizations these equations contain no particular structure and may be regarded as *generic* linear equations. Ignoring noise, the received signals at Receiver 1 at this stage can be expressed as (27):

Note that the channel and precoding vectors without the overbar notation refer to the values over only the first 3 channel uses. The received signal for Receiver 2, is also defined similarly.

$$\begin{aligned} \Rightarrow \mathbf{Y}^{[2]} &= \begin{bmatrix} \mathbf{H}^{[21]} \mathbf{V}_1^{[11]} & \mathbf{H}^{[21]} \mathbf{V}_2^{[11]} & \mathbf{H}^{[22]} \mathbf{V}_1^{[12]} & \mathbf{H}^{[22]} \mathbf{V}_2^{[12]} \end{bmatrix} \\ &\cdot \begin{bmatrix} u_1^{[11]}, u_2^{[11]}, u_1^{[12]}, u_2^{[12]} \end{bmatrix}^T \\ &+ \begin{bmatrix} \mathbf{H}^{[21]} \mathbf{V}_1^{[21]} & \mathbf{H}^{[21]} \mathbf{V}_2^{[21]} & \mathbf{H}^{[22]} \mathbf{V}_1^{[22]} & \mathbf{H}^{[22]} \mathbf{V}_2^{[22]} \end{bmatrix} \\ &\cdot \begin{bmatrix} u_1^{[21]}, u_2^{[21]}, u_1^{[22]}, u_2^{[22]} \end{bmatrix}^T \quad (28) \end{aligned}$$

**Last 4 Channel Uses:** During the last 4 channel uses we will operate over a new effective set of variables. Using the terminology of [13] we can call these the second layer variables. The goal will be to allow each receiver to resolve all layer-2 variables. Since we have only 4 channel uses left, and each channel use will provide only one equation to each receiver, we will choose 4 layer-2 variables.

Let us define our new variables:

$$s_1^{[1]} = u_1^{[11]} - \gamma_1^{[1]} u_2^{[11]} \quad (29)$$

$$s_1^{[2]} = u_1^{[12]} - \gamma_1^{[2]} u_2^{[12]} \quad (30)$$

$$s_2^{[1]} = u_1^{[21]} - \gamma_2^{[1]} u_2^{[21]} \quad (31)$$

$$s_2^{[2]} = u_1^{[22]} - \gamma_2^{[2]} u_2^{[22]} \quad (32)$$

where  $\gamma_i^{[k]}$  are constants whose values will be specified soon after we arrive at the rationale for choosing these values. Consistent with our causality and delayed CSIT constraint, the constants will depend on the channel values only from phase-I, i.e., from the first 3 channel uses. The most important observation here is that the variables  $s_1^{[k]}, s_2^{[k]}$  are available to Transmitter  $k$ .



$$\begin{aligned}
\begin{bmatrix} \bar{\mathbf{Y}}^{[1]}(1) \\ \bar{\mathbf{Y}}^{[1]}(2) \\ \bar{\mathbf{Y}}^{[1]}(3) \end{bmatrix} &= \begin{bmatrix} \bar{\mathbf{H}}^{[11]}(1)\bar{\mathbf{V}}_1^{[11]}(1) & \bar{\mathbf{H}}^{[11]}(1)\bar{\mathbf{V}}_2^{[11]}(1) & \bar{\mathbf{H}}^{[12]}(1)\bar{\mathbf{V}}_1^{[12]}(1) & \bar{\mathbf{H}}^{[12]}(1)\bar{\mathbf{V}}_2^{[12]}(1) \\ \bar{\mathbf{H}}^{[11]}(2)\bar{\mathbf{V}}_1^{[11]}(2) & \bar{\mathbf{H}}^{[11]}(2)\bar{\mathbf{V}}_2^{[11]}(2) & \bar{\mathbf{H}}^{[12]}(2)\bar{\mathbf{V}}_1^{[12]}(2) & \bar{\mathbf{H}}^{[12]}(2)\bar{\mathbf{V}}_2^{[12]}(2) \\ \bar{\mathbf{H}}^{[11]}(3)\bar{\mathbf{V}}_1^{[11]}(3) & \bar{\mathbf{H}}^{[11]}(3)\bar{\mathbf{V}}_2^{[11]}(3) & \bar{\mathbf{H}}^{[12]}(3)\bar{\mathbf{V}}_1^{[12]}(3) & \bar{\mathbf{H}}^{[12]}(3)\bar{\mathbf{V}}_2^{[12]}(3) \end{bmatrix} \begin{bmatrix} u_1^{[11]} \\ u_2^{[11]} \\ u_1^{[12]} \\ u_2^{[12]} \end{bmatrix} \\
&+ \begin{bmatrix} \bar{\mathbf{H}}^{[11]}(1)\bar{\mathbf{V}}_1^{[21]}(1) & \bar{\mathbf{H}}^{[11]}(1)\bar{\mathbf{V}}_2^{[21]}(1) & \bar{\mathbf{H}}^{[12]}(1)\bar{\mathbf{V}}_1^{[22]}(1) & \bar{\mathbf{H}}^{[12]}(1)\bar{\mathbf{V}}_2^{[22]}(1) \\ \bar{\mathbf{H}}^{[11]}(2)\bar{\mathbf{V}}_1^{[21]}(2) & \bar{\mathbf{H}}^{[11]}(2)\bar{\mathbf{V}}_2^{[21]}(2) & \bar{\mathbf{H}}^{[12]}(2)\bar{\mathbf{V}}_1^{[22]}(2) & \bar{\mathbf{H}}^{[12]}(2)\bar{\mathbf{V}}_2^{[22]}(2) \\ \bar{\mathbf{H}}^{[11]}(3)\bar{\mathbf{V}}_1^{[21]}(3) & \bar{\mathbf{H}}^{[11]}(3)\bar{\mathbf{V}}_2^{[21]}(3) & \bar{\mathbf{H}}^{[12]}(3)\bar{\mathbf{V}}_1^{[22]}(3) & \bar{\mathbf{H}}^{[12]}(3)\bar{\mathbf{V}}_2^{[22]}(3) \end{bmatrix} \begin{bmatrix} u_1^{[21]} \\ u_2^{[21]} \\ u_1^{[22]} \\ u_2^{[22]} \end{bmatrix} \\
\Rightarrow \mathbf{Y}^{[1]} &= \begin{bmatrix} \mathbf{H}^{[11]}\mathbf{V}_1^{[11]} & \mathbf{H}^{[11]}\mathbf{V}_2^{[11]} & \mathbf{H}^{[12]}\mathbf{V}_1^{[12]} & \mathbf{H}^{[12]}\mathbf{V}_2^{[12]} \end{bmatrix} \begin{bmatrix} u_1^{[11]}, u_2^{[11]}, u_1^{[12]}, u_2^{[12]} \end{bmatrix}^T \\
&+ \begin{bmatrix} \mathbf{H}^{[11]}\mathbf{V}_1^{[21]} & \mathbf{H}^{[11]}\mathbf{V}_2^{[21]} & \mathbf{H}^{[12]}\mathbf{V}_1^{[22]} & \mathbf{H}^{[12]}\mathbf{V}_2^{[22]} \end{bmatrix} \begin{bmatrix} u_1^{[21]}, u_2^{[21]}, u_1^{[22]}, u_2^{[22]} \end{bmatrix}^T \quad (27)
\end{aligned}$$

Once the layer-2 variables are defined, the operation over the last 4 channel uses is very simple. Each transmitter sends a different linear combination of its two layer-2 variables over each channel use. Each receiver sees 4 different linear combinations of the 4 layer-2 variables (2 from each transmitter) over the 4 channel uses, and is therefore able to resolve each variable (again, ignoring noise – for DoF arguments).

From phase-2 we know that both users are able to solve for the layer-2 variables. Now, let us consider the 3 equations accumulated at each receiver over the first 3 channel uses in 8 variables. Let us substitute for four of these variables in terms of the solved values of the layer-2 variables. Specifically, we make the following substitutions:

$$u_1^{[11]} \rightarrow s_1^{[1]} + \gamma_1^{[1]} u_2^{[11]} \quad (33)$$

$$u_1^{[12]} \rightarrow s_1^{[2]} + \gamma_1^{[2]} u_2^{[12]} \quad (34)$$

$$u_1^{[21]} \rightarrow s_2^{[1]} + \gamma_2^{[1]} u_2^{[21]} \quad (35)$$

$$u_1^{[22]} \rightarrow s_2^{[2]} + \gamma_2^{[2]} u_2^{[22]} \quad (36)$$

Note that after these substitutions there are only four unknown variables left (since the  $s_i^{[k]}$  are already known from Phase-2) —  $u_2^{[11]}, u_2^{[12]}, u_2^{[21]}, u_2^{[22]}$ . Since we have four variables and only three equations we will need interference alignment. Out of the 4 remaining unknown variables each receiver only desires 2. As usual on the X channel, the 2 undesired variables will be aligned into one dimension, leaving the remaining two dimensions available to recover the two desired variables. This alignment will be enabled precisely by the choice of the  $\gamma_i^{[k]}$  in the layer-2 variable definitions – thus accomplishing retrospective interference alignment.

Following the substitutions, consider Receiver 1, where we have three equations in the remaining 4 variables (Note that the following quantities – without overbar notation – refer to only the first three channel uses).

$$\begin{aligned}
\mathbf{Y}^{[1]} - \mathbf{H}^{[11]}\mathbf{V}_1^{[11]}s_1^{[1]} - \mathbf{H}^{[12]}\mathbf{V}_1^{[12]}s_1^{[2]} - \mathbf{H}^{[11]}\mathbf{V}_1^{[21]}s_2^{[1]} - \mathbf{H}^{[12]}\mathbf{V}_1^{[22]}s_2^{[2]} \\
= \mathbf{H}^{[11]}\mathbf{V}_1^{[11]}\gamma_1^{[1]}u_2^{[11]} + \mathbf{H}^{[11]}\mathbf{V}_2^{[11]}u_2^{[11]} + \mathbf{H}^{[12]}\mathbf{V}_1^{[12]}\gamma_1^{[2]}u_2^{[12]} \\
+ \mathbf{H}^{[12]}\mathbf{V}_2^{[12]}u_2^{[12]} + \mathbf{H}^{[11]}\mathbf{V}_1^{[21]}\gamma_2^{[1]}u_2^{[21]} + \mathbf{H}^{[11]}\mathbf{V}_2^{[21]}u_2^{[21]} \\
+ \mathbf{H}^{[12]}\mathbf{V}_1^{[22]}\gamma_2^{[2]}u_2^{[22]} + \mathbf{H}^{[12]}\mathbf{V}_2^{[22]}u_2^{[22]} \quad (37)
\end{aligned}$$

The interfering symbols  $u_2^{[21]}, u_2^{[22]}$  arrive along directions:

$$u_2^{[21]} \Rightarrow \mathbf{H}^{[11]}\mathbf{V}_1^{[21]}\gamma_2^{[1]} + \mathbf{H}^{[11]}\mathbf{V}_2^{[21]} \quad (38)$$

$$u_2^{[22]} \Rightarrow \mathbf{H}^{[12]}\mathbf{V}_1^{[22]}\gamma_2^{[2]} + \mathbf{H}^{[12]}\mathbf{V}_2^{[22]} \quad (39)$$

The RHS of the above expressions are the  $3 \times 1$  vectors indicating the direction along which interference is seen by Receiver 1 from the two undesired symbols. For interference alignment we would like these directions to be co-linear.

$$\mathbf{H}^{[11]}\mathbf{V}_1^{[21]}\gamma_2^{[1]} + \mathbf{H}^{[11]}\mathbf{V}_2^{[21]} = \beta \left( \mathbf{H}^{[12]}\mathbf{V}_1^{[22]}\gamma_2^{[2]} + \mathbf{H}^{[12]}\mathbf{V}_2^{[22]} \right) \quad (40)$$

for some constant  $\beta$ . Equivalently

$$\begin{bmatrix} \mathbf{H}^{[11]}\mathbf{V}_1^{[21]} & \mathbf{H}^{[11]}\mathbf{V}_2^{[21]} & \mathbf{H}^{[12]}\mathbf{V}_1^{[22]} & \mathbf{H}^{[12]}\mathbf{V}_2^{[22]} \end{bmatrix}_{3 \times 4} \begin{bmatrix} \gamma_2^{[1]} \\ 1 \\ -\beta\gamma_2^{[2]} \\ -\beta \end{bmatrix} = \mathbf{0}_{3 \times 1} \quad (41)$$

Since the matrix on the left is generic and of size  $3 \times 4$  it has a unique (upto scaling) null vector. The choice of the values of  $\gamma_1^{[2]}, \gamma_2^{[2]}, \beta$  is made precisely to force the vector on the right to be this null vector, thus aligning interference.

Similarly, at Receiver 2, we have three equations in four information symbols:

$$\begin{aligned}
\mathbf{Y}^{[2]} - \mathbf{H}^{[21]}\mathbf{V}_1^{[11]}s_1^{[1]} - \mathbf{H}^{[22]}\mathbf{V}_1^{[12]}s_1^{[2]} - \mathbf{H}^{[21]}\mathbf{V}_1^{[21]}s_2^{[1]} - \mathbf{H}^{[22]}\mathbf{V}_1^{[22]}s_2^{[2]} \\
= \mathbf{H}^{[21]}\mathbf{V}_1^{[11]}\gamma_1^{[1]}u_2^{[11]} + \mathbf{H}^{[21]}\mathbf{V}_2^{[11]}u_2^{[11]} + \mathbf{H}^{[22]}\mathbf{V}_1^{[12]}\gamma_1^{[2]}u_2^{[12]} \\
+ \mathbf{H}^{[22]}\mathbf{V}_2^{[12]}u_2^{[12]} + \mathbf{H}^{[21]}\mathbf{V}_1^{[21]}\gamma_2^{[1]}u_2^{[21]} + \mathbf{H}^{[21]}\mathbf{V}_2^{[21]}u_2^{[21]} \\
+ \mathbf{H}^{[22]}\mathbf{V}_1^{[22]}\gamma_2^{[2]}u_2^{[22]} + \mathbf{H}^{[22]}\mathbf{V}_2^{[22]}u_2^{[22]} \quad (42)
\end{aligned}$$

and the interfering symbols  $u_2^{[11]}, u_2^{[12]}$  arrive along the directions:

$$u_2^{[11]} \Rightarrow \mathbf{H}^{[21]}\mathbf{V}_1^{[11]}\gamma_1^{[1]} + \mathbf{H}^{[21]}\mathbf{V}_2^{[11]} \quad (43)$$

$$u_2^{[12]} \Rightarrow \mathbf{H}^{[22]}\mathbf{V}_1^{[12]}\gamma_1^{[2]} + \mathbf{H}^{[22]}\mathbf{V}_2^{[12]} \quad (44)$$

Thus we would like

$$\begin{bmatrix} \mathbf{H}^{[21]}\mathbf{V}_1^{[11]} & \mathbf{H}^{[21]}\mathbf{V}_2^{[11]} & \mathbf{H}^{[22]}\mathbf{V}_1^{[12]} & \mathbf{H}^{[22]}\mathbf{V}_2^{[12]} \end{bmatrix}_{3 \times 4} \begin{bmatrix} \gamma_1^{[1]} \\ 1 \\ -\delta\gamma_1^{[2]} \\ -\delta \end{bmatrix} = \mathbf{0}_{3 \times 1} \quad (45)$$

As before, we choose the values of  $\gamma_1^{[1]}, \gamma_1^{[2]}, \delta$  to satisfy the equation above, and thereby achieve interference alignment at Receiver 2.

Lastly, we need to check that the desired symbols are not aligned either with each other or with the interference by this choice of the  $\gamma_*^{[k]}$  constants. Consider Receiver 1, where the desired symbols  $u_2^{[11]}, u_2^{[12]}$  arrive along directions:

$$u_2^{[11]} \Rightarrow \mathbf{H}^{[11]}\mathbf{V}_1^{[11]}\gamma_1^{[1]} + \mathbf{H}^{[11]}\mathbf{V}_2^{[11]} \quad (46)$$

$$u_2^{[12]} \Rightarrow \mathbf{H}^{[12]}\mathbf{V}_1^{[12]}\gamma_1^{[2]} + \mathbf{H}^{[12]}\mathbf{V}_2^{[12]} \quad (47)$$

and the aligned interference arrives along the direction:

$$\mathbf{H}^{[11]}\mathbf{V}_1^{[21]}\gamma_2^{[1]} + \mathbf{H}^{[11]}\mathbf{V}_2^{[21]} \quad (48)$$

Thus, for desired symbols to be resolvable from the interference, we need the following  $3 \times 3$  matrix to have full rank:

$$M_1 = \begin{bmatrix} \mathbf{H}^{[11]}\mathbf{V}_1^{[11]}\gamma_1^{[1]} + \mathbf{H}^{[11]}\mathbf{V}_2^{[11]} & \mathbf{H}^{[12]}\mathbf{V}_1^{[12]}\gamma_1^{[2]} + \mathbf{H}^{[12]}\mathbf{V}_2^{[12]} \\ \mathbf{H}^{[11]}\mathbf{V}_1^{[21]}\gamma_2^{[1]} + \mathbf{H}^{[11]}\mathbf{V}_2^{[21]} & \end{bmatrix} \quad (49)$$

Similarly we can define the  $3 \times 3$  matrix  $M_2$  that also needs to be full rank for Receiver 2 to be able to obtain his desired symbols. Both conditions can be stated together in the following form:

$$\det(M_1)\det(M_2) \neq 0 \quad (50)$$

However, both  $\det(M_1)$  and  $\det(M_2)$  correspond to polynomials in  $\mathbf{V}^{[*]}, \mathbf{H}^{[*]}$  (the  $\gamma$  can be evaluated in terms of  $\mathbf{V}, \mathbf{H}$ ), and therefore, so does the product  $\det(M_1)\det(M_2)$ . Note that  $\mathbf{V}, \mathbf{H}$  are picked independently over complex numbers. Therefore the polynomial  $\det(M_1)\det(M_2)$  is either identically the zero polynomial, or it is non-zero almost surely for all realizations of  $\mathbf{V}, \mathbf{H}$ . To prove that it is non-zero almost surely, it suffices to show that it is not the zero polynomial, which in turn is established if there exists any non-zero evaluation of  $\det(M_1)\det(M_2)$ . Indeed this is easily verified by a numerical example. One such example is as follows

$$\mathbf{V}^{[11]} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \quad \mathbf{V}^{[21]} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\mathbf{V}^{[12]} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \end{bmatrix} \quad \mathbf{V}^{[22]} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 3 \end{bmatrix}$$

$$\mathbf{H}^{[11]} = \mathbf{H}^{[21]} = \mathbf{I}_{3 \times 3}$$

$$\mathbf{H}^{[12]} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{H}^{[22]} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

where  $\mathbf{I}_{3 \times 3}$  is a  $3 \times 3$  identity matrix. Based on these choices for  $\mathbf{V}$  and  $\mathbf{H}$  we have

$$\gamma_1^{[1]} = 4, \quad \gamma_1^{[2]} = 2, \quad \gamma_2^{[1]} = \frac{5}{4}, \quad \gamma_2^{[2]} = \frac{1}{4}$$

The channel coefficients in the last 4 channel uses are chosen randomly because the role of phase two is decoding the 4 Phase-II variables sent by 2 transmitters. We can evaluate  $\det(M_1)\det(M_2) = \frac{11}{4} \cdot \frac{11}{4} = \frac{121}{16}$  for these chosen values of  $\mathbf{V}$  and  $\mathbf{H}$  which is nonzero and hence  $\det(M_1)\det(M_2)$  is a nonzero polynomial.

## VI. DELAYED OUTPUT FEEDBACK

So far we assumed that the only feedback available to the transmitters consists of delayed CSIT. Another commonly studied model for feedback is channel *output* feedback (without explicitly providing the CSI). In this section we study the

3 user interference channel and the X channel with delayed *output* feedback, i.e., the channel output is available to the transmitters only after the channel state associated with the observed output is no longer current. While delayed CSIT created difficulties because of the transmitters' inability to reconstruct the previously received linear combinations of undesired received symbols because of the distributed nature of the information, delayed output feedback automatically provides the transmitters with information that has been previously observed at one of the receivers. Retransmitting this information provides the transmitters an opportunity to provide new observations to the receivers who desire this information, while allowing the receivers who have already observed this interference to cancel it entirely. In this sense, delayed output feedback allows a direct extension of the alignment techniques explored in [13].

### A. 3 User Interference Channel with Delayed Output Feedback

The following theorem presents an achievability result for the 3 user interference channel.

*Theorem 3:* The 3 user interference channel with delayed output feedback available from each receiver to only its corresponding transmitter can achieve  $\frac{6}{5}$  DoF almost surely.

*Proof:*

In order to achieve  $\frac{6}{5}$  DoF we will operate over a 5 channel-use block. Each user will communicate two coded information symbols over these 5 channel uses using linear schemes that can be simply seen as swapping output symbols to help resolve desired signals [13]. A summary of the transmission scheme is described below.

- 1) Over the first time slot, Transmitter 1 sends its first information symbol  $u_1^{[1]}$  and Transmitter 2 simultaneously sends its first information symbol  $u_1^{[2]}$ . Ignoring noise, the received signals are described below:

$$\text{Receiver 1: } Y^{[1]}(1) = H^{[11]}(1)u_1^{[1]} + H^{[12]}(1)u_1^{[2]} \quad (51)$$

$$\text{Receiver 2: } Y^{[2]}(1) = H^{[21]}(1)u_1^{[1]} + H^{[22]}(1)u_1^{[2]} \quad (52)$$

$$\text{Receiver 3: } Y^{[3]}(1) = H^{[31]}(1)u_1^{[1]} + H^{[32]}(1)u_1^{[2]} \quad (53)$$

- 2) Over the second time slot, Transmitter 1 sends its second information symbol  $u_2^{[1]}$  and Transmitter 3 simultaneously sends its first information symbol  $u_1^{[3]}$ . Ignoring noise, the received signals are described below:

$$\text{Receiver 1: } Y^{[1]}(2) = H^{[11]}(2)u_1^{[1]} + H^{[13]}(2)u_1^{[3]} \quad (54)$$

$$\text{Receiver 2: } Y^{[2]}(2) = H^{[21]}(2)u_1^{[1]} + H^{[23]}(2)u_1^{[3]} \quad (55)$$

$$\text{Receiver 3: } Y^{[3]}(2) = H^{[31]}(2)u_1^{[1]} + H^{[33]}(2)u_1^{[3]} \quad (56)$$

- 3) Over the third time slot, Transmitter 2 sends its second information symbol  $u_2^{[2]}$  and Transmitter 3 simultaneously sends its second information symbol  $u_2^{[3]}$ . Ignoring noise, the received signals are described below:

$$\text{Receiver 1: } Y^{[1]}(3) = H^{[11]}(3)u_2^{[2]} + H^{[12]}(3)u_2^{[3]} \quad (57)$$

$$\text{Receiver 2: } Y^{[2]}(3) = H^{[21]}(3)u_2^{[2]} + H^{[22]}(3)u_2^{[3]} \quad (58)$$

$$\text{Receiver 3: } Y^{[3]}(3) = H^{[31]}(3)u_2^{[2]} + H^{[32]}(3)u_2^{[3]} \quad (59)$$

- 4) Over the fourth time slot, Transmitter 3 retransmits  $Y^{[3]}(1)$  and Transmitter 2 retransmits  $Y^{[2]}(2)$ .

- 5) Over the fifth time slot,  $Y^{[3]}(1)$  and  $Y^{[1]}(3)$  are retransmitted from Transmitters 3 and 1 respectively.

Next we explain how every receiver has enough information to recover its two desired symbols.

**Receiver 1:** Consider Receiver 1. From the linear combination of  $Y^{[3]}(1)$  and  $Y^{[1]}(3)$  received over the fifth symbol, this receiver is able to remove  $Y^{[1]}(3)$  which it has previously received, to obtain  $Y^{[3]}(1)$ . Combining  $Y^{[1]}(1)$  and  $Y^{[3]}(1)$  the receiver has enough information to resolve the first received symbol  $u_1^{[1]}$ .

Further, from the linear combination of  $Y^{[3]}(1), Y^{[2]}(2)$  received over the fourth symbol, the receiver removes  $Y^{[3]}(1)$  to obtain a clean  $Y^{[2]}(2)$ . Combining  $Y^{[2]}(2)$  with  $Y^{[1]}(2)$ , the receiver is able to resolve the second desired symbol  $u_2^{[1]}$ .

**Receiver 2:** Consider Receiver 2. From the linear combination of  $Y^{[3]}(1), Y^{[2]}(2)$  received over the fourth symbol, this receiver is able to remove  $Y^{[2]}(2)$  which it has previously received, to obtain  $Y^{[3]}(1)$ . Combining  $Y^{[2]}(1)$  and  $Y^{[3]}(1)$  the receiver has enough information to resolve the first desired symbol  $u_1^{[2]}$ .

Further, from the linear combination of  $Y^{[3]}(1)$  and  $Y^{[1]}(3)$  received over the fifth channel use, the receiver removes  $Y^{[3]}(1)$  to obtain a clean  $Y^{[1]}(3)$ . Combining  $Y^{[1]}(3)$  with  $Y^{[2]}(3)$ , the receiver is able to resolve the second desired symbol  $u_2^{[2]}$ .

**Receiver 3:** Consider Receiver 3. From the linear combination of  $Y^{[3]}(1)$  and  $Y^{[2]}(2)$  received over the fourth symbol, this receiver is able to remove  $Y^{[3]}(1)$  which it has previously received, to obtain  $Y^{[2]}(2)$ . Combining  $Y^{[2]}(2)$  and  $Y^{[3]}(2)$  the receiver has enough information to resolve the first received symbol  $u_1^{[3]}$ .

Further, from the linear combination of  $Y^{[1]}(3), Y^{[3]}(1)$  received over the fifth symbol, the receiver removes  $Y^{[3]}(1)$  to obtain a clean  $Y^{[1]}(3)$ . Combining  $Y^{[1]}(3)$  with  $Y^{[3]}(3)$ , the receiver is able to resolve the second desired symbol  $u_2^{[3]}$ .

Thus, all symbols are resolved and  $\frac{6}{5}$  DoF are achieved on the 3 user interference channel. Note that we are assuming that each receiver has *global* channel knowledge, i.e., it knows not only the channels associated with itself but also the other receivers' channels as well. Further we are once again ignoring noise in this discussion because, as stated earlier, for such linear beamforming schemes, noise does not affect the DoF.

### B. X Channel with Delayed Output Feedback

*Theorem 4:* The X channel with delayed output feedback can achieve  $\frac{4}{3}$  DoF almost surely.

*Proof:* In light of the earlier discussion on the vector BC, it is easily seen that the X channel, with only delayed channel output feedback can easily achieve the  $\frac{4}{3}$  DoF outer bound. Since the achievable scheme is essentially the same as the schemes studied [11] and [13], we only present a brief description.

- Send two symbols to each receiver over three time slots to achieve  $\frac{4}{3}$  DoF.
- In the first time slot, each transmitter sends its own symbol intended for Receiver 1. Receiver 1 observes a linear combination (along with noise) of the two desired

symbols, while Receiver 2 sees a linear combination of undesired symbols (and noise).

- The second time slot is similar to the first time slot, except the information symbols transmitted are for User 2.
- In the third time slot, the transmitters send a superposition of the previous undesired outputs. This is possible due to delayed output feedback as long as each undesired output signal is available to one of the transmitters.

The third time slot provides each receiver with a second linear combination of desired symbols while the interfering undesired information is cancelled because it has been received previously.

*Remark:* While it is remarkable that the achievability results presented in this work for both the 3 user interference channel and the X channel under delayed output feedback correspond to higher DoF than with delayed CSIT, it should be noted that these are only achievability results and in the absence of outer bounds it is not possible to make categorical comparisons between the two settings.

## VII. CONCLUSION

We explored similarities, differences, and the apparent difficulties in achieving interference alignment with channel uncertainty at the transmitters based on recent works that assume two different channel uncertainty models – heterogeneous block fading and delayed CSIT. While there are many shared aspects that allow the schemes to be translated from one setting to another for many cases, overall the two settings are indeed fundamentally different and face different challenges. In particular, the delayed CSIT setting appears to be more sensitive to whether the transmitters are co-located or distributed, unlike previous results where both for compound channels and suitably staggered block fading models the two were found to be equivalent from a DoF perspective.

A retrospective interference alignment scheme is presented for the 3 user interference channel and the 2 user X channel. While the scheme operates in two phases, and with two layers of variables as the scheme proposed in [13], the novelty of retrospective alignment appears in the construction of auxiliary (layer - 2 in the terminology of [13]) variables that aid in the alignment of the previously transmitted information symbols based on only the information symbols available to each transmitter. This scheme was used to prove the achievability of  $\frac{9}{8}$  DoF for the 3 user interference channel with delayed CSIT.

While the 2 user MISO BC with delayed CSIT easily achieves the outer bound of  $\frac{4}{3}$ , it is not known if the same DoF can be achieved on the X channel, i.e., without cooperation between transmitters. Using the same scheme, we were able to show that delayed CSIT is still useful in the X channel from a DoF perspective, as one can achieve  $\frac{8}{7}$  DoF. We also found that the 3 user interference channel and the X channel can achieve  $\frac{6}{5}$  and  $\frac{4}{3}$  DoF respectively when delayed output feedback is available to the transmitters. It is remarkable that with perfect and instantaneous CSIT, output feedback does not increase DoF for interference channels or X networks [24].

## REFERENCES

- [1] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel", *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1691-1706, July 2003.
- [2] S. Jafar and A. Goldsmith, "Isotropic fading vector broadcast channels: the scalar upperbound and loss in degrees of freedom", *IEEE Trans. Inform. Theory*, vol. 51, no. 3, pp. 848-857, March 2005.
- [3] S. Jafar, "Too much mobility limits the capacity of wireless ad-hoc networks," *IEEE Trans. on Information Theory*, vol. 52, no. 7, pp. 3954-3965, November 2005.
- [4] A. Lapidoth, S. Shamai, and M. Wigger, "On the capacity of fading MIMO broadcast channels with imperfect transmitter side-information," *Proceedings of 43rd Annual Allerton Conference on Communications, Control and Computing*, Sep. 28-30, 2005.
- [5] H. Weingarten, S. Shamai, and G. Kramer, "On the compound MIMO broadcast channel", *Proceedings of Annual Information Theory and Applications Workshop UCSD*, Jan. 2007.
- [6] T. Gou, S. Jafar, and C. Wang, "On the Degrees of Freedom of Finite State Compound Wireless Networks," *IEEE Trans. on Information Theory*, Vol. 57, No. 6, Pages: 3286-3308, June 2011.
- [7] M. A. Maddah-Ali, "The degrees of freedom of the compound MIMO broadcast channels with finite states", *Corr*, vol. abs/0909.5006, 2009.
- [8] C. Huang, S. A. Jafar, S. Shamai, and S. Vishwanath, "On degrees of freedom region of MIMO networks without CSIT", *arxiv:0909.4017*, 2009, to appear in *IEEE Trans. on Information Theory*.
- [9] C. S. Vaze and M. K. Varanasi, "The degrees of freedom regions of MIMO broadcast, interference, and cognitive radio channels with no CSIT," *CoRR*, vol. abs/0909.5424, 2009.
- [10] Y. Zhu and D. Guo, "The degrees of freedom of MIMO interference channels without state information at transmitters," *CoRR*, vol. abs/1008.5196, 2010, to appear in *IEEE Trans. on Information Theory*.
- [11] S. Jafar, "Exploiting channel correlations simple interference alignment schemes with no CSIT," *Preprint available on Arxiv arXiv:0910.0555*, Oct. 2009.
- [12] C. Wang, T. Gou, and S. A. Jafar, "Aiming Perfectly in the Dark - Blind Interference Alignment through Staggered Antenna Switching," *IEEE Trans. on Information Theory*, Vol. 59, No. 6, Pages: 2734-2744, June 2011.
- [13] M. A. Maddah-Ali and D. Tse, "On the degrees of freedom of miso broadcast channels with delayed feedback", EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2010-122, Sep 2010. [Online]. Available: <http://www.eecs.berkeley.edu/Pubs/TechRpts/2010/EECS-2010-122.html>.
- [14] N. Jindal, "MIMO broadcast channels with finite rate feedback", *IEEE Trans. on Information Theory*, vol. 52, no. 11, pp. 5045-5049, Nov. 2006.
- [15] G. Caire, N. Jindal, and S. Shamai, "On the required accuracy of transmitter channel state information in multiple antenna broadcast channels," in *Proceedings of the Asilomar Conference on Signals, Systems and Computers*, 2007.
- [16] A. E. Gamal, "The feedback capacity of degraded broadcast channels," *IEEE Trans. Inform. Theory*, vol. 3, pp. 379-381, April 1978.
- [17] L. Zheng and D. N. Tse, "Packing spheres in the Grassmann manifold: A geometric approach to the non-coherent multi-antenna channel," *IEEE Trans. Inform. Theory*, vol. 48, no. 2, pp. 359-383, Feb 2002.
- [18] A. Ghasemi, A. Motahari and A. Khandani, "On the degrees of freedom of X channel with delayed CSIT," in *Proceedings of ISIT*, 2011.
- [19] M. Abdoli, A. Ghasemi and A. Khandani, "On the Degrees of Freedom of K-User SISO Interference and X Channels with Delayed CSIT," *CoRR*, vol. abs/1109.4314, 2011.
- [20] C. Vaze and M. Varanasi, "The degrees of freedom region of the MIMO interference channel with delayed CSIT," in *Proceedings of ISIT*, 2011.
- [21] S. Jafar and S. Shamai, "Degrees of freedom region for the MIMO X channel," *IEEE Trans. on Information Theory*, vol. 54, no. 1, pp. 151-170, Jan. 2008.
- [22] M. Maddah-Ali, A. Motahari, and A. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Trans. on Information Theory*, August 2008, pp. 34573470.
- [23] V. Cadambe and S. Jafar, "Interference alignment and the degrees of freedom of the K user interference channel," *IEEE Trans. on Information Theory*, vol. 54, no. 8, pp. 3425-3441, Aug. 2008.
- [24] V. R. Cadambe and S. A. Jafar, "Degrees of freedom of wireless networks with relays, feedback, cooperation and full duplex operation," *IEEE Transactions on Information Theory*, vol. 55, pp. 2334-2344, May 2009.



**Hamed Maleki** received the B.Sc. degree in Electrical Engineering from Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran and the M.Sc. degree in Electrical Engineering from University of Tehran, Tehran, Iran, in 2006 and 2008, respectively. He is currently working toward the Ph.D. degree at the University of California, Irvine (UC Irvine). His current research interests include multiuser information theory and wireless communications.

Mr. Maleki was a recipient of UC Irvine Graduate Fellowship for the year 2008-09. He was also a recipient of the University of California, Irvine CPMC graduate fellowship for the year 2009-10.



**Syed Ali Jafar** (S'99-M'04-SM'09) received the B. Tech. degree in Electrical Engineering from the Indian Institute of Technology (IIT), Delhi, India in 1997, the M.S. degree in Electrical Engineering from California Institute of Technology (Caltech), Pasadena USA in 1999, and the Ph.D. degree in Electrical Engineering from Stanford University, Stanford, CA USA in 2003. His industry experience includes positions at Lucent Bell Labs, Qualcomm Inc. and Hughes Software Systems. He is currently an Associate Professor in the Department of Electrical Engineering and Computer Science at the University of California Irvine, Irvine, CA USA. His research interests include multiuser information theory and wireless communications.

Dr. Jafar received the NSF CAREER award in 2006, the ONR Young Investigator Award in 2008, the IEEE Information Theory Society paper award in 2009 and the Engineering School Fariborz Maseeh Outstanding Research Award in 2010. He received the annual UC Irvine EECS Professor of the Year award three times, in 2006, 2009 and 2011, for excellence in teaching. Dr. Jafar was the inaugural instructor for the First Canadian School of Information Theory in 2011, a plenary speaker for the IEEE Communication Theory Workshop 2010 and SPCOM 2010, and a Visiting Erskine Fellow to the University of Canterbury New Zealand. He served as Associate Editor for IEEE Transactions on Communications 2004-2009, for IEEE Communications Letters 2008-2009 and is currently serving as Associate Editor for IEEE Transactions on Information Theory.



**Shlomo Shamai (Shitz)** (S'80-M'82-SM'89-F'94) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion—Israel Institute of Technology, in 1975, 1981 and 1986 respectively.

During 1975-1985 he was with the Communications Research Labs, in the capacity of a Senior Research Engineer. Since 1986 he is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, where he is now the William Fondiller Professor of Telecommunications. His research interests encompasses a wide spectrum of topics in information theory and statistical communications.

Dr. Shamai (Shitz) is an IEEE Fellow, and the recipient of the 2011 Claude E. Shannon Award. He is the recipient of the 1999 van der Pol Gold Medal of the Union Radio Scientifique Internationale (URSI), and a co-recipient of the 2000 IEEE Donald G. Fink Prize Paper Award, the 2003, and the 2004 joint IT/COM societies paper award, the 2007 IEEE Information Theory Society Paper Award, the 2009 European Commission FP7, Network of Excellence in Wireless COMMunications (NEWCOM++) Best Paper Award, and the 2010 Thomson Reuters Award for International Excellence in Scientific Research. He is also the recipient of 1985 Alon Grant for distinguished young scientists and the 2000 Technion Henry Taub Prize for Excellence in Research. He has served as Associate Editor for the Shannon Theory of the IEEE Transactions on Information Theory, and has also served on the Board of Governors of the Information Theory Society. He is a member of the Executive Editorial Board of the IEEE Transactions on Information Theory.