Fifty-first Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
October 2 - 3, 2013

A Rank Ratio Inequality and the Linear Degrees of Freedom of X-Channel with Delayed CSIT

Sina Lashgari, A. Salman Avestimehr, and Changho Suh

Abstract— **We establish the degrees of freedom of the X-channel with delayed channel knowledge at transmitters (i.e., delayed CSIT), assuming linear coding strategies at the transmitters. We derive a new upper bound and characterize the linear degrees of freedom of this network to be $\frac{6}{5}$. The converse builds upon our development of a general lemma for wireless networks consisting of two transmitters and two receivers, with delayed CSIT. The lemma states that, if the two transmitters employ linear strategies, then the ratio of the rank of received signal subspaces at the two receivers cannot exceed 3/2, due to delayed CSIT. This lemma can also be applied to any arbitrary network, in which a receiver decodes its desired message in the presence of two interferers.**

I. Introduction

Interference management is a key challenge in system design and analysis of wireless networks, and developing intelligent interference management techniques is essential for improving network throughput. The X-channel is a canonical setting for the information-theoretic study of interference management in wireless networks. It consists of two transmitters causing interference at two receivers, and each transmitter aims to communicate an intended message to each receiver. The question is how the transmitters should optimally manage the interference and communicate their messages to the receivers. This problem has been studied extensively in the literature and various interference management techniques have been proposed. In particular, in [1] it is shown that, quite surprisingly, one can significantly improve upon conventional interference management schemes (e.g., orthogonalization) and achieve $4/3$ degrees of freedom (DoF) by using *interference alignment* (IA) [2], [3].

However, in order to perfectly align the interference, the transmitters need to accurately know the *current* state of the channels, which is practically very challenging and may even be impossible (due to, for example, high mobility). Thus, a natural question would be: how can the transmitters optimally manage the interference with only *delayed* knowledge of the channel state information (i.e., delayed CSIT)?

In the context of broadcast channel, Maddah-Ali and Tse in [4] have recently shown that delayed CSIT can still be very useful. In particular, for the multi-antenna broadcast channel with delayed CSIT, they developed an innovative transmission strategy that utilizes the past received signals to create signals of common interest to multiple receivers, hence significantly improving DoF by broadcasting them to the receivers. In a sense, these "signals of common interest" represent aligned interferences in the past receptions.

Subsequently in [5], [6], [7], [8], the impact of delayed CSIT has been explored for a variety of interference networks in which transmit antennas are now distributed at different locations. Unlike multi-antenna broadcast channels, in networks with distributed transmitters, it may not be possible for a transmitter to reconstruct previously received signals, since it may include other transmitters' signals that are not accessible to that transmitter. Hence, although interference alignment has happened in the past receptions, it may not be possible to construct the aligned interference locally at a transmitter and broadcast it to the receivers. Interestingly, even in this setting, delayed CSIT has shown to still provide DoF gains (see e.g., [5], [6], [7], [8], [9]).

In particular, for the X-channel, Jafar and Shamai in [5] developed a scheme, called Retrospective Interference Alignment, that achieves DoF of $\frac{8}{7}$ with delayed CSIT, which is strictly larger than its DoF with no-CSIT (i.e., 1 DoF). Furthermore, Ghasemi-Motahari-Khandani in [8] developed another scheme that achieves DoF of $\frac{6}{5}$ with delayed CSIT. However, given that the only upper bound on the DoF of this network is the one with instantaneous CSIT (i.e., $\frac{4}{3}$ DoF), it remains still open whether $\frac{6}{5}$ is the fundamental limit on the DoF of X-channel with delayed CSIT, or whether there are more efficient interference management techniques.

Our main contribution is to show that the DoF of the Gaussian X- channel with delayed CSIT is indeed $\frac{6}{5}$, under the assumption that only linear encoding schemes are employed at the transmitters. Under that constraint, only a linear combination of information symbols are allowed to be transmitted at each time. In fact, all of the interference management strategies with delayed CSIT that are developed thus far (e.g., [4], [5], [6], [7], [8]) fall into this category.

The key part of the converse is the development of a general lemma that bounds the maximum ratio of the dimensions of received linear-subspaces (at the two receivers) that are created by *distributed* transmitters with delayed CSIT. More specifically, we show that if two distributed transmitters with delayed CSIT employ linear strategies, the ratio of the dimensions of the received signals cannot exceed $\frac{3}{2}$. When instantaneous CSIT is available, this ratio can be as large as 2, and with no CSIT, this ratio is always 1. As a result, this

S. Lashgari and A. S. Avestimehr are with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY (email: sl2232@cornell.edu, avestimehr@ece.cornell.edu); and C. Suh is with the Department of Electrical Engineering, Korea Advanced Institute of Science and Technology, Daejeon, South Korea (email: chsuh@kaist.ac.kr). The research of A. S. Avestimehr and S. Lashgari is supported by NSF Grants CAREER 0953117, CCF-1161720, NETS-1161904, and ONR award N000141310094. The research of C. Suh was funded by the MSIP(Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2013.

978-1-4799-3410-2/13/$31.00 ©2013 IEEE

lemma captures the fundamental impact of delayed CSIT on the dimension of received subspaces. Also, in the case of two centralized transmitters (e.g., multi-antenna BC), this ratio can be as large as 2; therefore, this lemma also captures the impact of *distributed transmitters* on the dimension of received subspaces. The lemma can also be viewed as a generalization of the “entropy leakage Lemma” in [9], which considers a broadcast channel with binary fading, and bounds the maximum ratio of the entropy of received signals at two different receivers.

Other Related Results. For the MIMO broadcast channel with delayed CSIT, a well known upper bound is based on the genie-aided bounding technique. This technique consists of two steps. First, signals of a set of receivers are given to another set of receivers such that the enhanced network becomes a physically degraded broadcast channel. Using the fact that feedback cannot increase capacity for physically degraded broadcast channels [10], we can then take the non-feedback upper bound as that of the original feedback channel. This technique has been used in [4], [11]. A modified technique has also been employed in the MIMO interference channel [6]. Although the technique (and its variant) serves to show the optimality for some special cases, in general it provides an upper bound far from the best known lower bounds. On the other hand, our converse technique provides tighter upper bounds for some scenarios although it is restricted to linear coding strategies.

Notation. We use small letters for scalars, arrowed letters (e.g. $\vec{x}$) for vectors, capital letters (e.g. X) for matrices, and a calligraphic font (e.g. $\mathcal{X}$) for sets. Furthermore, we use bold letters for random entities, and non-bold letters for deterministic values (e.g., realizations of random variables).

II. System Model & Main Results

We consider the Gaussian X-channel depicted in Fig. 1. It consists of two transmitters and two receivers, and each transmitter has a separate message for each of the receivers. Each node is equipped with a single antenna.

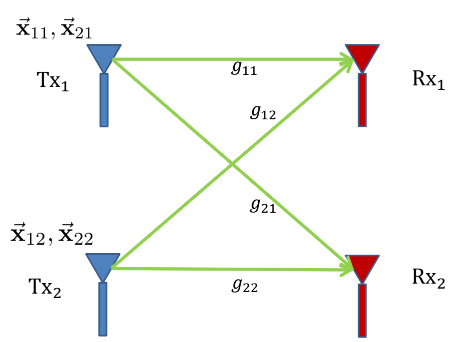

Fig. 1. Network configuration for X-channel. There are two transmitters and two receivers, where each transmitter has a message for each receiver. We assume time-varying channels, with delayed CSIT.

Received signal at Rx_k ($k \in \{1,2\}$) at time t is given by

$$\mathbf{y}_k(t) = \mathbf{g}_{k1}(t)\mathbf{x}_1(t) + \mathbf{g}_{k2}(t)\mathbf{x}_2(t) + \mathbf{z}_k(t), \qquad (1)$$

where $\mathbf{x}_j(t)$ is the transmit signal of Tx_j; $\mathbf{g}_{kj}(t) \in \mathbb{C}$ indicates a channel from Tx_j to Rx_k; and $\mathbf{z}_k(t) \sim \mathcal{CN}(0,1)$. The channel coefficients of $\mathbf{g}_{kj}(t)$'s are i.i.d across time and users, and they are drawn from a continuous distribution. We denote by $\boldsymbol{\mathcal{G}}(t)$ the set of all four channel coefficients at time t. In addition, we denote by $\boldsymbol{\mathcal{G}}^n$ the set of all channel coefficients from time 1 to n, i.e.,

$$\boldsymbol{\mathcal{G}}^n = \{\mathbf{g}_{kj}(t) : k, j \in \{1,2\}, t = 1, \ldots, n\}.$$

Denoting the vector of transmit signals for Tx_j in a block of length n by $\vec{\mathbf{x}}_j^n$, each transmitter Tx_j obeys an average power constraint, $\frac{1}{n}E\{||\vec{\mathbf{x}}_j^n||^2\} \leq P$. We assume delayed channel state information at the transmitters (CSIT). In other words, at time t, only the states of the past $\boldsymbol{\mathcal{G}}^{t-1}$ are known to $\mathrm{Tx}_1, \mathrm{Tx}_2$.

We restrict ourselves to linear coding strategies as defined in [12], in which DoF simply represents the dimension of the linear subspace of transmitted signals. More specifically, consider a communication scheme with block length n, in which transmitter Tx_j wishes to transmit a vector $\vec{\mathbf{x}}_{kj} \in \mathbb{C}^{m_{kj}(n)}$ of $m_{kj}(n) \in \mathbb{N}$ information symbols to Rx_k ($j,k \in \{1,2\}$). These information symbols are then modulated with precoding vectors $\vec{\mathbf{v}}_{kj}(t) \in \mathbb{C}^{m_{kj}(n)}$ at times $t = 1, 2, \ldots, n$. Note that the precoding vector $\vec{\mathbf{v}}_{kj}(t)$ depends only upon the outcome of $\boldsymbol{\mathcal{G}}^{t-1}$ due to the delayed CSIT constraint:

$$\vec{\mathbf{v}}_{kj}(t) = f_{k,j,t}^{(n)}\left(\mathcal{G}^{t-1}\right). \qquad (2)$$

Based on this linear precoding, Tx_j will then send $\mathbf{x}_j(t) = \vec{\mathbf{v}}_{1j}(t)^\top \vec{\mathbf{x}}_{1j} + \vec{\mathbf{v}}_{2j}(t)^\top \vec{\mathbf{x}}_{2j}$ at time t. We denote by $\mathbf{V}_{kj}^n \in \mathbb{C}^{n \times m_{kj}(n)}$ the overall precoding matrix of Tx_j for Rx_k, such that the t-th row of $\mathbf{V}_{kj}^n$ is $\vec{\mathbf{v}}_{kj}(t)^\top$). We also denote the precoding functions used by Tx_j by $f_j^{(n)} = \{f_{1,j,t}^{(n)}, f_{2,j,t}^{(n)}\}_{t=1}^n$.

Based on the above setting, the received signal at Rx_k ($k \in \{1,2\}$) after the n time steps of the communication will be

$$\vec{\mathbf{y}}_k^n = \mathbf{G}_{k1}^n(\mathbf{V}_{11}^n\vec{\mathbf{x}}_{11} + \mathbf{V}_{21}^n\vec{\mathbf{x}}_{21}) + \mathbf{G}_{k2}^n(\mathbf{V}_{12}^n\vec{\mathbf{x}}_{12} + \mathbf{V}_{22}^n\vec{\mathbf{x}}_{22}) + \vec{\mathbf{z}}_k^n,$$

where $\mathbf{G}_{kj}^n$ is the $n \times n$ diagonal matrix whose t-th element on the diagonal is $\mathbf{g}_{kj}(t)$. Now, consider the decoding of $\vec{\mathbf{x}}_{kj}$ at Rx_k (i.e., the $m_{kj}(n)$ information symbols of Tx_j for Rx_k). The corresponding interference subspace at Rx_k will be

$$\begin{aligned}\boldsymbol{\mathcal{I}}_{kj} =& \text{colspan}\left(\mathbf{G}_{kj}^n \mathbf{V}_{k'j}^n\right) \cup \text{colspan}\left(\mathbf{G}_{kj'}^n \mathbf{V}_{kj'}^n\right) \\ & \cup \text{colspan}\left(\mathbf{G}_{kj'}^n \mathbf{V}_{k'j'}^n\right),\end{aligned}$$

where $j' = 3 - j, k' = 3 - k$, and colspan(.) of a matrix corresponds to the sub-space that is spanned by its columns. For instance, $\boldsymbol{\mathcal{I}}_{11} = \text{colspan}(\mathbf{G}_{11}^n\mathbf{V}_{21}^n) \cup \text{colspan}(\mathbf{G}_{12}^n\mathbf{V}_{12}^n) \cup \text{colspan}(\mathbf{G}_{12}^n\mathbf{V}_{22}^n)$. Let $\boldsymbol{\mathcal{I}}_{kj}^c \subseteq \mathbb{C}^n$ denote the subspace orthogonal to $\boldsymbol{\mathcal{I}}_{kj}$. Then, in the regime of asymptotically high transmit powers (i.e., ignoring the noise), the decodability of information symbols from Tx_j at Rx_k corresponds to the constraints that the image of $\text{colspan}(\mathbf{G}_{kj}^n\mathbf{V}_{kj}^n)$ on $\boldsymbol{\mathcal{I}}_{kj}^c$ has dimension $m_{kj}(n)$:

$$\dim\left(\mathrm{Proj}_{\mathcal{I}_{kj}^c}\mathrm{colspan}\left(\mathbf{G}_{kj}^n\mathbf{V}_{kj}^n\right)\right)=\dim\left(\mathrm{colspan}\left(\mathbf{V}_{kj}^n\right)\right)$$
$$=m_{kj}(n). \quad (3)$$

Based on this setting, we now define the sum linear degrees of freedom of the X-channel.

Definition 1: Four-tuple $(d_{11},d_{12},d_{21},d_{22})$ degrees of freedom are linearly achievable if there exists a sequence $\{f_1^{(n)},f_2^{(n)}\}_{n=1}^{\infty}$ such that $(\mathbf{V}_{11}^n,\mathbf{V}_{12}^n,\mathbf{V}_{21}^n,\mathbf{V}_{22}^n)$ satisfy the decodability condition of (3) with probability 1, and $\forall(j,k)$,

$$d_{kj}=\lim_{n\to\infty}\frac{m_{kj}(n)}{n}. \quad (4)$$

We also define the linear degrees of freedom region $\mathcal{D}$ as the closure of the set of all achievable 4-tuples $(d_{11},d_{12},d_{21},d_{22})$. Furthermore, the sum linear degrees of freedom ($\mathrm{DoF}_{\text{L-sum}}$) is then defined as follows:

$$\mathrm{DoF}_{\text{L-sum}}=\max\sum_{k,j\in\{1,2\}}d_{kj},\quad \text{s.t. } (d_{11},d_{12},d_{21},d_{22})\in\mathcal{D}. \quad (5)$$

In case transmitters have instantaneous CSIT, it was shown in [2], [13] that the sum degrees of freedom is $\frac{4}{3}$. The achievability uses interference alignment that enables us to deliver four symbols over three timeslots. On the other hand, in the non-CSIT case, one can readily see that the received signals at the two receivers are statistically identical and therefore the DoF collapses to 1, which is that of the multiple access channel. For the case of delayed CSIT, Ghasemi-Motahari-Khandani in [8] develops a new scheme that achieves the sum DoF of $\frac{6}{5}$.

Our main result in this paper is the following theorem, proved in Section III, which states that $\frac{6}{5}$ is the maximum DoF that can be achieved using linear encoding schemes.

Theorem 1: For the X-channel with delayed CSIT,

$$\mathrm{DoF}_{\text{L-sum}}=\frac{6}{5}. \quad (6)$$

Our converse proof builds upon the following key lemma, which is proved in Section III-C.

Lemma 1: For any linear coding strategy $\{f_1^{(n)},f_2^{(n)}\}$, with corresponding $\mathbf{V}_{11}^{(n)},\mathbf{V}_{12}^{(n)}$ as defined in (2),

$$\mathrm{rank}\left[\mathbf{G}_{11}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{12}^n\right]\overset{a.s.}{\leq}\frac{3}{2}\mathrm{rank}\left[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n\right]. \quad (7)$$

Remark 1: Note that this lemma holds for any arbitrary networks with two transmitters and two receivers. It does not require any specific decodability assumption at receivers. (7) says that the ratio of the ranks of received beamforming matrices at Rx_1 and Rx_2 is at most $\frac{3}{2}$. For the case of having instantaneous CSIT, the ratio of $\mathrm{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{12}^n]$ to $\mathrm{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n]$ can be up to 2.[1] Hence, Lemma 1 characterizes the impact of delayed CSIT on the maximum ratio of the ranks of received beamforming matrices.

Remark 2: Lemma 1 is inspired by the "entropy leakage Lemma" in [9], which considers a broadcast channel with binary fading, and bounds the maximum ratio of the entropy of received signals at two different receivers. In fact, it can be viewed as an extension of this lemma to the case of two distributed transmitters with linear encoding strategies, in which the entropy is approximated by the rank of the received beamforming matrices.

III. Proof of Theorem 1

A. Achievability

As mentioned in the previous section, the achievability is provided in [8], and utilizes a linear encoding scheme to achieve $\frac{6}{5}$. Here we review the scheme to illustrate how beamforming vectors are chosen. We set $n=5, m_{11}(n)=2, m_{12}(n)=1, m_{21}(n)=1, m_{22}(n)=2$. Let the information symbols of the transmitters be denoted by

$$\vec{\mathbf{x}}_{11}=\begin{bmatrix}a_1\\a_2\end{bmatrix},\quad \vec{\mathbf{x}}_{12}=\begin{bmatrix}b_1\end{bmatrix},$$
$$\vec{\mathbf{x}}_{21}=\begin{bmatrix}c_1\end{bmatrix},\quad \vec{\mathbf{x}}_{22}=\begin{bmatrix}d_1\\d_2\end{bmatrix}. \quad (8)$$

In $t=1$, Tx_1 sends a_1, and Tx_2 sends b_1, which corresponds to choosing the following beamforming vectors at $\mathrm{Tx}_1,\mathrm{Tx}_2$

$$\vec{v}_{11}=\begin{bmatrix}1\\0\end{bmatrix},\vec{v}_{12}=\begin{bmatrix}1\end{bmatrix},\vec{v}_{21}=\begin{bmatrix}0\end{bmatrix},\vec{v}_{22}=\begin{bmatrix}0\\0\end{bmatrix}.$$

In $t=2$, Tx_1 sends a_2, and Tx_2 sends b_1, which corresponds to choosing the following beamforming vectors at $\mathrm{Tx}_1,\mathrm{Tx}_2$

$$\vec{v}_{11}=\begin{bmatrix}0\\1\end{bmatrix},\vec{v}_{12}=\begin{bmatrix}1\end{bmatrix},\vec{v}_{21}=\begin{bmatrix}0\end{bmatrix},\vec{v}_{22}=\begin{bmatrix}0\\0\end{bmatrix}.$$

Therefore, by the end of $t=2$, Rx_2 can cancel b_1 from its received signals to recover an equation only involving a_1 and a_2, denoted by $\vec{\mathbf{m}}_1^\top\vec{\mathbf{x}}_{11}$. It is easy to see that, if this equation is delivered to Rx_1, it can decode all of its desired information symbols (i.e., $\vec{\mathbf{x}}_{11}$ and $\vec{\mathbf{x}}_{12}$). Hence, it is an equation of interest to Rx_1 that is known at Rx_2, and can be created by Tx_1.

A similar schemes is applied in the next two time steps. More specifically, in $t=3$, Tx_1 sends c_1, and Tx_2 sends d_1, which corresponds to choosing the following beamforming vectors at the transmitters

$$\vec{v}_{11}=\begin{bmatrix}0\\0\end{bmatrix},\vec{v}_{12}=\begin{bmatrix}0\end{bmatrix},\vec{v}_{21}=\begin{bmatrix}1\end{bmatrix},\vec{v}_{22}=\begin{bmatrix}1\\0\end{bmatrix}.$$

In $t=4$, Tx_1 sends c_1, and Tx_2 sends d_2, which corresponds to choosing the following beamforming vectors at $\mathrm{Tx}_1,\mathrm{Tx}_2$

$$\vec{v}_{11}=\begin{bmatrix}0\\0\end{bmatrix},\vec{v}_{12}=\begin{bmatrix}0\end{bmatrix},\vec{v}_{21}=\begin{bmatrix}1\end{bmatrix},\vec{v}_{22}=\begin{bmatrix}0\\1\end{bmatrix}.$$

Therefore, by the end of $t=4$, Rx_1 can cancel c_1 from its received signals to recover an equation only involving d_1 and d_2, denoted by $\vec{\mathbf{m}}_2^\top\vec{\mathbf{x}}_{22}$. Again, it is easy to see that, if this equation is delivered to Rx_2, it can decode all of its

[1] To see this, consider the following two-timeslot scheme. In time 1, $\mathrm{Tx}_1,\mathrm{Tx}_2$ send $\mathbf{x}_1,\mathbf{x}_2$ respectively. Rx_2 then gets $\mathbf{g}_{21}(1)\mathbf{x}_1+\mathbf{g}_{22}(1)\mathbf{x}_2$. In time 2, $\mathrm{Tx}_1,\mathrm{Tx}_2$ send $\frac{\mathbf{g}_{21}(1)}{\mathbf{g}_{21}(2)}\mathbf{x}_1$, $\frac{\mathbf{g}_{22}(1)}{\mathbf{g}_{22}(2)}\mathbf{x}_2$ respectively. Rx_2 then gets the same equation as the one received in time 1. On the other hand, Rx_1 gets a new equation almost surely. Therefore, the rank of the received signal at Rx_1 can be twice that of Rx_2. Also one can readily show that the two is the maximum that can be achieved.

desired information symbols (i.e., $\vec{\mathbf{x}}_{21}$ and $\vec{\mathbf{x}}_{22}$). Hence, it is an equation of interest to Rx_2 that is known at Rx_1, and can be created by Tx_2.[2]

Now, in $t = 5$, Tx_1 sends $\vec{\mathbf{m}}_1^\top \vec{\mathbf{x}}_{11}$, and Tx_2 sends $\vec{\mathbf{m}}_2^\top \vec{\mathbf{x}}_{22}$. Since each of these transmit signals is already known at one of the receivers, after this transmission, Rx_1 will recover $\vec{\mathbf{m}}_1^\top \vec{\mathbf{x}}_{11}$ and Rx_2 will recover $\vec{\mathbf{m}}_2^\top \vec{\mathbf{x}}_{22}$. Therefore, all information symbols are delivered to their corresponding receivers, achieving sum DoF of $\frac{6}{5}$.

B. Converse

We will now prove the converse, which is the main contribution of the paper. As mentioned in Section II, the key idea behind the converse is Lemma 1, which we restate below.

Lemma 1. *For any linear coding strategy* $\{f_1^{(n)}, f_2^{(n)}\}$*, with corresponding* $\mathbf{V}_{11}^{(n)}, \mathbf{V}_{12}^{(n)}$ *as defined in (2),*

$$\mathrm{rank}\left[\mathbf{G}_{11}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{12}^n\right] \overset{a.s.}{\leq} \frac{3}{2}\mathrm{rank}\left[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n\right].$$

Before proving Lemma 1, we will first complete the converse proof of Theorem 1. In particular, we prove the following:

$$(d_{11}+d_{12}) + \frac{3}{2}(d_{21}+d_{22}) \leq \frac{3}{2} \tag{9}$$

$$\frac{3}{2}(d_{11}+d_{12}) + (d_{21}+d_{22}) \leq \frac{3}{2}. \tag{10}$$

The desired result follows from summing the above two inequalities. By symmetry, we only need to prove (9). We first state the following lemma, whose proof is based on basic linear algebra and omitted, which is helpful in providing an equivalent condition for decodability of messages in (3).

Lemma 2: Consider two matrices $A_{n\times m}, B_{n\times p}$, with corresponding subspaces $\mathcal{A} = \mathrm{colspan}(A)$, $\mathcal{B} = \mathrm{colspan}(B)$ defined by the span of their columns. Denote by $\mathcal{A}^c$ the orthogonal subspace of $\mathcal{A}$ (in $\mathbb{C}^n$). Then,

$$\begin{aligned}\dim(\mathrm{Proj}_{\mathcal{A}^c}\mathcal{B}) = \dim(\mathcal{B}) \quad &\text{if and only if} \\ \mathrm{rank}[A] + \mathrm{rank}[B] &= \mathrm{rank}[A \quad B]. \end{aligned}\tag{11}$$

Now, note that

$$\dim\left(\mathrm{colspan}(\mathbf{V}_{kj}^n)\right) \overset{a.s.}{=} \dim\left(\mathrm{colspan}(\mathbf{G}_{kj}^n\mathbf{V}_{kj}^n)\right) \tag{12}$$

due to the continuous distribution of $\mathbf{g}_{kj}(t)$ for any t. Therefore, by (12) and Lemma 2, we conclude that (3) occurs with probability 1 if and only if for $j,k \in \{1,2\}$ and $j' = 3-j, k' = 3-k$,

$$\begin{aligned}&\mathrm{rank}[\mathbf{G}_{kj}^n\mathbf{V}_{k'j}^n \quad \mathbf{G}_{kj'}^n\mathbf{V}_{kj'}^n \quad \mathbf{G}_{kj'}^n\mathbf{V}_{k'j'}^n] + \mathrm{rank}[\mathbf{G}_{kj}^n\mathbf{V}_{kj}^n] \\ &\overset{a.s.}{=} \mathrm{rank}[\mathbf{G}_{k1}^n\mathbf{V}_{k1}^n \quad \mathbf{G}_{k2}^n\mathbf{V}_{k2}^n \quad \mathbf{G}_{k1}^n\mathbf{V}_{k'1}^n \quad \mathbf{G}_{k2}^n\mathbf{V}_{k'2}^n].\end{aligned}\tag{13}$$

Thus, we consider (13) as the equivalent decodability condition, which consist of the following four equations:

$$\begin{aligned}&\mathrm{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n] + \mathrm{rank}[\mathbf{G}_{11}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{22}^n] \\ &\overset{a.s.}{=} \mathrm{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{11}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{22}^n] \quad (14)\\ &\mathrm{rank}[\mathbf{G}_{12}^n\mathbf{V}_{12}^n] + \mathrm{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{11}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{22}^n] \\ &\overset{a.s.}{=} \mathrm{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{11}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{22}^n] \quad (15)\\ &\mathrm{rank}[\mathbf{G}_{21}^n\mathbf{V}_{21}^n] + \mathrm{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{22}^n] \\ &\overset{a.s.}{=} \mathrm{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{21}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{22}^n] \quad (16)\\ &\mathrm{rank}[\mathbf{G}_{22}^n\mathbf{V}_{22}^n] + \mathrm{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{21}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n] \\ &\overset{a.s.}{=} \mathrm{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{21}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{22}^n]. \quad (17)\end{aligned}$$

Before continuing with the converse we also state the following claim, whose proof is based on the sub-modularity of the rank function, and will be used later in the proof.

Claim 1: Suppose that for four matrices A, B, C, D with the same number of rows,

$$\begin{aligned}\mathrm{rank}[A] + \mathrm{rank}[B \quad C \quad D] &= \mathrm{rank}[A \quad B \quad C \quad D] \\ \mathrm{rank}[B] + \mathrm{rank}[A \quad C \quad D] &= \mathrm{rank}[A \quad B \quad C \quad D].\end{aligned}\tag{18}$$

Then,

$$\mathrm{rank}[A] + \mathrm{rank}[B] + \mathrm{rank}[C \quad D] = \mathrm{rank}[A \quad B \quad C \quad D].$$[3]

We continue proving the converse. Suppose $(d_{11}, d_{12}, d_{21}, d_{22}) \in \mathcal{D}$, i.e., there exists a sequence $\{f_1^{(n)}, f_2^{(n)}\}_{n=1}^\infty$ resulting in linearly achieving $\{m_{11}(n), m_{12}(n), m_{21}(n), m_{22}(n)\}_{n=1}^\infty$ with probability 1, and $d_{kj} = \lim_{n\to\infty} \frac{m_{kj}(n)}{n}$. Hence, by (14), (15), Claim 1,

$$\begin{aligned}&\mathrm{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n] + \mathrm{rank}[\mathbf{G}_{12}^n\mathbf{V}_{12}^n] \\ &\overset{a.s.}{=} \mathrm{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{11}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{22}^n] \\ &- \mathrm{rank}[\mathbf{G}_{11}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{22}^n].\end{aligned}\tag{19}$$

In addition, by (16), (17), and Claim 1,

$$\begin{aligned}&\mathrm{rank}[\mathbf{G}_{21}^n\mathbf{V}_{21}^n] + \mathrm{rank}[\mathbf{G}_{22}^n\mathbf{V}_{22}^n] \\ &\overset{a.s.}{=} \mathrm{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{21}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{22}^n] \\ &- \mathrm{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n].\end{aligned}\tag{20}$$

[2] One can check that $\vec{\mathbf{m}}_1 = [\mathbf{g}_{22}(2)\mathbf{g}_{21}(1) \quad -\mathbf{g}_{22}(1)\mathbf{g}_{21}(2)]^\top$, and $\vec{\mathbf{m}}_2 = [\mathbf{g}_{12}(3)\mathbf{g}_{11}(4) \quad -\mathbf{g}_{11}(3)\mathbf{g}_{12}(4)]^\top$.

[3] To see the proof, first note that $\mathrm{rank}[A] + \mathrm{rank}[B] + \mathrm{rank}[C \quad D] \geq \mathrm{rank}[A \quad B \quad C \quad D]$. Hence, in order to prove Claim 1, we only need to prove inequality in the other direction. Now, according to the assumptions in claim, and using sub-modularity of the rank, we have

$$\begin{aligned}\mathrm{rank}[A] + \mathrm{rank}[B] &\overset{(18)}{=} \mathrm{rank}[A \quad B \quad C \quad D] - \mathrm{rank}[B \quad C \quad D] \\ &\quad + \mathrm{rank}[A \quad B \quad C \quad D] - \mathrm{rank}[A \quad C \quad D] \\ &\overset{\text{sub-modularity}}{\leq} \mathrm{rank}[A \quad B \quad C \quad D] - \mathrm{rank}[B \quad C \quad D] \\ &\quad + \mathrm{rank}[B \quad C \quad D] - \mathrm{rank}[C \quad D] \\ &= \mathrm{rank}[A \quad B \quad C \quad D] - \mathrm{rank}[C \quad D].\end{aligned}$$

Therefore, we have

$$m_{11}(n)+m_{12}(n)+\frac{3}{2}(m_{21}(n)+m_{22}(n))$$
$$\overset{a.s.}{=} \text{rank}[\mathbf{V}_{11}^n]+\text{rank}[\mathbf{V}_{12}^n]+\frac{3}{2}(\text{rank}[\mathbf{V}_{21}^n]+\text{rank}[\mathbf{V}_{22}^n])$$
$$\overset{a.s.}{=} \text{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n]+\text{rank}[\mathbf{G}_{12}^n\mathbf{V}_{12}^n]+\frac{3}{2}(\text{rank}[\mathbf{G}_{21}^n\mathbf{V}_{21}^n]$$
$$+\text{rank}[\mathbf{G}_{22}^n\mathbf{V}_{22}^n])$$
$$\overset{(19),(20)}{\underset{a.s.}{=}}\text{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{11}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{22}^n]$$
$$-\text{rank}[\mathbf{G}_{11}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{22}^n]$$
$$+\frac{3}{2}(\text{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{21}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{22}^n]$$
$$-\text{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n]) \overset{(a)}{\leq} \text{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{12}^n]$$
$$+\frac{3}{2}\text{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{21}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{22}^n]$$
$$-\frac{3}{2}\text{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n]$$
$$\overset{\text{(Lemma 1)}}{\underset{\leq}{a.s.}}\frac{3}{2}\text{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{21}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{22}^n]$$
$$\leq \frac{3}{2}n, \tag{21}$$

where (a) follows from the fact that $\text{rank}[\mathbf{A} \quad \mathbf{B}] \leq \text{rank}[\mathbf{A}]+\text{rank}[\mathbf{B}]$. Therefore, by dividing both sides of the inequality in (21) by n, and letting $n \to \infty$ we get

$$d_{11}+d_{12}+\frac{3}{2}(d_{21}+d_{22}) \leq \frac{3}{2}. \tag{22}$$

■

We will next prove Lemma 1.

C. Proof of Lemma 1

Consider a linear coding strategy $\{f_1^{(n)}, f_2^{(n)}\}$, with corresponding $\mathbf{V}_{11}^{(n)}, \mathbf{V}_{12}^{(n)}$ as defined in (2). For notational simplicity in the proof, we denote $\mathbf{V}_{11}^n$ by $\mathbf{V}_1^n$, and $\mathbf{V}_{12}^n$ by $\mathbf{V}_2^n$. We first state some definitions.

Definition 2: Consider a linear coding strategy $\{f_1^{(n)}, f_2^{(n)}\}$, with corresponding $\mathbf{V}_1^{(n)} \triangleq \mathbf{V}_{11}^{(n)}, \mathbf{V}_2^{(n)} \triangleq \mathbf{V}_{12}^{(n)}$. Define the random set $\boldsymbol{\mathcal{T}}_{\{f_1^{(n)},f_2^{(n)}\}}(\boldsymbol{\mathcal{G}^n})$ with its alphabet being the power set of $\{1,2,\ldots,n\}$ as follows. For any realization of channels $\boldsymbol{\mathcal{G}^n} = \mathcal{G}^n$, resulting in $\mathbf{G}_{21}^n = G_{21}^n, \mathbf{G}_{22}^n = G_{22}^n, \mathbf{G}_{11}^n = G_{11}^n, \mathbf{G}_{12}^n = G_{12}^n$, and $\mathbf{V}_1^n = V_1^n, \mathbf{V}_2^n = V_2^n$, define

$$\mathcal{T}_{\{f_1^{(n)},f_2^{(n)}\}}(\mathcal{G}^n) \triangleq \{t|[\vec{v}_1(t)^\top \quad \vec{0}_{1\times m_2(n)}], [\vec{0}_{1\times m_1(n)} \quad \vec{v}_2(t)^\top] \in \text{rowspan}[G_{21}^{t-1}V_1^{t-1} \quad G_{22}^{t-1}V_2^{t-1}]\}.$$

In words, $\boldsymbol{\mathcal{T}}_{\{f_1^{(n)},f_2^{(n)}\}}(\boldsymbol{\mathcal{G}^n})$ represents the set of random timeslots (random due to the randomness in channels), where the beamforming vectors transmitted by the two transmitters are already individually recoverable by Rx_2 using its received beamforming vectors in the previous timeslots. Since the code $\{f_1^{(n)}, f_2^{(n)}\}$ is fixed in the proof, for notational simplicity, from now on we denote $\boldsymbol{\mathcal{T}}_{\{f_1^{(n)},f_2^{(n)}\}}(\boldsymbol{\mathcal{G}^n})$ by $\boldsymbol{\mathcal{T}}$.

Definition 3: Consider a linear coding strategy $\{f_1^{(n)}, f_2^{(n)}\}$, with corresponding $\mathbf{V}_1^{(n)} \triangleq \mathbf{V}_{11}^{(n)}, \mathbf{V}_2^{(n)} \triangleq \mathbf{V}_{12}^{(n)}$. Define random variables $\mathbf{r}_1(\boldsymbol{\mathcal{G}}^n), \mathbf{r}_2(\boldsymbol{\mathcal{G}}^n)$ in $\{1,\ldots,n\}$ as follows. For any realization of channels $\boldsymbol{\mathcal{G}^n} = \mathcal{G}^n$, which results in $\mathbf{G}_{21}^n = G_{21}^n, \mathbf{G}_{22}^n = G_{22}^n, \mathbf{G}_{11}^n = G_{11}^n, \mathbf{G}_{12}^n = G_{12}^n$, and $\mathbf{V}_1^n = V_1^n, \mathbf{V}_2^n = V_2^n$, define

$$r_i(\mathcal{G}^n) \triangleq \dim(\text{span}(\mathcal{E}_i(\mathcal{G}^n))), \quad i=1,2,$$

where

$$\mathcal{E}_1(\mathcal{G}^n) \triangleq \{\vec{s}_{m_1(n)\times 1}| \quad \exists \vec{l}_{n\times 1} \quad s.t. \quad [\vec{s}^\top \quad \vec{0}_{1\times m_2(n)}] = \vec{l}^{\,\top}[G_{21}^nV_1^n \quad G_{22}^nV_2^n]\},$$
$$\mathcal{E}_2(\mathcal{G}^n) \triangleq \{\vec{s}_{m_2(n)\times 1}| \quad \exists \vec{l}_{n\times 1} \quad s.t. \quad [\vec{0}_{1\times m_1(n)} \quad \vec{s}^\top] = \vec{l}^{\,\top}[G_{21}^nV_1^n \quad G_{22}^nV_2^n]\}.$$

In words, $\mathbf{r}_1(\boldsymbol{\mathcal{G}}^n)$ can be interpreted as the number of linearly independent equations that Rx_2 can recover from its received signal, which only involve symbols of Tx_1. Hereafter, we denote $\mathbf{r}_1(\boldsymbol{\mathcal{G}}^n), \mathbf{r}_2(\boldsymbol{\mathcal{G}}^n)$ simply by $\mathbf{r}_1, \mathbf{r}_2$.

We will now state the following lemma, proved in Appendix I, which is the key to proving Lemma 1.

Lemma 3: For any linear coding strategy $\{f_1^{(n)}, f_2^{(n)}\}$, with corresponding $\mathbf{V}_1^{(n)} \triangleq \mathbf{V}_{11}^{(n)}, \mathbf{V}_2^{(n)} \triangleq \mathbf{V}_{12}^{(n)}$ defined in (2),

- $\text{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n \quad \mathbf{G}_{12}^n\mathbf{V}_2^n] - \text{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{22}^n\mathbf{V}_2^n] \overset{a.s.}{\leq} \text{rank}[\mathbf{G}_{11}^{\boldsymbol{\mathcal{T}}}\mathbf{V}_1^{\boldsymbol{\mathcal{T}}} \quad \mathbf{G}_{12}^{\boldsymbol{\mathcal{T}}}\mathbf{V}_2^{\boldsymbol{\mathcal{T}}}]$
- $\text{rank}[\mathbf{V}_j^{\boldsymbol{\mathcal{T}}}] \leq \mathbf{r}_j, \quad j=1,2$
- $\mathbf{r}_j \overset{a.s.}{\leq} \text{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{22}^n\mathbf{V}_2^n] - \text{rank}[\mathbf{V}_{3-j}^n], \quad j=1,2$

where $\boldsymbol{\mathcal{T}}$ is defined in Definition 2, $\mathbf{V}_i^{\boldsymbol{\mathcal{T}}}$ represents the random submatrix of $\mathbf{V}_i^n$ derived by keeping rows whose indices are in $\boldsymbol{\mathcal{T}}$, and $\mathbf{r}_1, \mathbf{r}_2$ are defined in Definition 3.

We are now ready to prove Lemma 1. We will first use Lemma 3 to find an upper bound on the difference between $\text{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n \quad \mathbf{G}_{12}^n\mathbf{V}_2^n]$ and $\text{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{22}^n\mathbf{V}_2^n]$.

$$\text{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n \quad \mathbf{G}_{12}^n\mathbf{V}_2^n] - \text{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{22}^n\mathbf{V}_2^n]$$
$$\overset{\text{(Lemma 3)}}{\underset{\leq}{a.s.}}\text{rank}[\mathbf{G}_{11}^{\boldsymbol{\mathcal{T}}}\mathbf{V}_1^{\boldsymbol{\mathcal{T}}} \quad \mathbf{G}_{12}^{\boldsymbol{\mathcal{T}}}\mathbf{V}_2^{\boldsymbol{\mathcal{T}}}]$$
$$\leq \text{rank}[\mathbf{G}_{11}^{\boldsymbol{\mathcal{T}}}\mathbf{V}_1^{\boldsymbol{\mathcal{T}}}] + \text{rank}[\mathbf{G}_{12}^{\boldsymbol{\mathcal{T}}}\mathbf{V}_2^{\boldsymbol{\mathcal{T}}}] \overset{a.s.}{=} \text{rank}[\mathbf{V}_1^{\boldsymbol{\mathcal{T}}}] + \text{rank}[\mathbf{V}_2^{\boldsymbol{\mathcal{T}}}]$$
$$\overset{\text{(Lemma 3)}}{\leq} \mathbf{r}_1+\mathbf{r}_2 \overset{\text{(Lemma 3)}}{\underset{\leq}{a.s.}}\text{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{22}^n\mathbf{V}_2^n]$$
$$-\text{rank}[\mathbf{V}_2^n] + \text{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{22}^n\mathbf{V}_2^n] - \text{rank}[\mathbf{V}_1^n].$$

Therefore,

$$\text{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n \quad \mathbf{G}_{12}^n\mathbf{V}_2^n] \overset{a.s.}{\leq} 3\times\text{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{22}^n\mathbf{V}_2^n] - \text{rank}[\mathbf{V}_1^n] - \text{rank}[\mathbf{V}_2^n]. \tag{23}$$

On the other hand, note that

$$\text{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n \quad \mathbf{G}_{12}^n\mathbf{V}_2^n] \leq \text{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n] + \text{rank}[\mathbf{G}_{12}^n\mathbf{V}_2^n] \overset{a.s.}{=} \text{rank}[\mathbf{V}_1^n] + \text{rank}[\mathbf{V}_2^n]. \tag{24}$$

Summing up (23) and (24),

$$\text{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n \quad \mathbf{G}_{12}^n\mathbf{V}_2^n] \overset{a.s.}{\leq} \frac{3}{2}\times\text{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{22}^n\mathbf{V}_2^n],$$

and we complete the proof of Lemma 1.

■

IV. Discussion

In this paper, we characterized the linear degrees of freedom of the X-channel with delayed CSIT. Our main contribution was the development of a general lemma that shows that, if two distributed transmitters employ linear strategies, the ratio of the dimensions of received linear subspaces at the two receivers cannot exceed $\frac{3}{2}$, due to lack of instantaneous knowledge of the channels.

We conjecture that the total DoF of the X-channel with delayed CSIT (without restriction to linear schemes) is also $\frac{6}{5}$. In fact, we conjecture that the generalization of Lemma 1 for general encoding strategies also holds true. Therefore, a future direction would be to remove the linearity restriction on the encoding schemes, and prove (or disprove) the above conjecture, which (if true) will lead to the DoF characterization of the X-channel with delayed CSIT.

Our lemma (Lemma 1) can be applied to any arbitrary network, in which a receiver decodes its desired message in the presence of two interferers. As an example, by applying Lemma 1 to the three-user interference channel with delayed CSIT, we can derive a new upper bound of $\frac{9}{7}$ on its linear DoF. This is the first upper bound that captures the impact of delayed CSIT on the DoF of this network (see [14] for more details).

We also believe that similar techniques could be applied to other important network configurations to gain insight on how delayed CSIT can be used to improve the DoF, and what the limitations on this DoF improvement are. In particular the K-user interference channel [5], [7] and multi-hop interference networks [15], [16], [17], in which there is a large gap between the state-of-the-art inner and outer bounds with delayed CSIT, can be considered.

References

[1] S. A. Jafar and S. Shamai, "Degrees of Freedom Region of the MIMO X Channel," *IEEE Transactions on Information Theory*, vol. 54, pp. 151–170, Jan. 2008.

[2] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication over MIMO X Channels: Interference Alignment, Decomposition, and Performance Analysis," *IEEE Transactions on Information Theory*, vol. 54, pp. 3457–3470, Aug. 2008.

[3] V. R. Cadambe and S. A. Jafar, "Interference Alignment and the Degree of Freedom for the K User Interference Channel," *IEEE Transactions on Information Theory*, vol. 54, pp. 3425–3441, Aug. 2008.

[4] M. A. Maddah-Ali and D. N. Tse, "Completely Stale Transmitter Channel State Information is Still Very Useful," *IEEE Transactions on Information Theory*, vol. 58, July 2012.

[5] S. A. Jafar and S. Shamai, "Retrospective Interference Alignment over Interference Networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, pp. 228–240, June 2012.

[6] C. S. Vaze and M. K. Varanasi, "The Degrees of Freedom Region and Intererence Alingment for the MIMO Inteference Channel with Delayed CSIT," *IEEE Transactions on Information Theory*, vol. 58, pp. 4396–4417, July 2012.

[7] M. Abdoli, A. Ghasemi, and A. Khandani, "On the Degrees of Freedom of K-User SISO Interference and X Channels with Delayed CSIT," *arXiv:1109.4314*, 2011.

[8] A. Ghasemi, A. Motahari, and A. Khandani, "On the Degrees of Freedom of X Channel with Delayed CSIT," *IEEE International Symposium on Information Theory*, 2011.

[9] A. Vahid, M. Maddah-Ali, and A. S. Avestimehr, "Capacity Results for Binary Fading Interference Channels with Delayed CSIT," *arXiv:1301.5309*, 2013.

[10] A. E. Gamal, "The Feedback Capacity of Degraded Broadcast Channels," *IEEE Transactions on Information Theory*, vol. 24, pp. 379–381, May 1978.

[11] L. Georgiadis and L. Tassiulas, "Broadcast Erasure Channel with Feedback - Capacity and Algorithms," *Worshop on Network Coding, Theory, and Applications*, pp. 54–61, 2009.

[12] G. Bresler, D. Cartwright, and D. N. Tse, "Interference Alignment for the MIMO Interference Channel," *arXiv:1303.5678*, 2013.

[13] V. R. Cadambe and S. A. Jafar, "Degrees of Freedom of Wireless X Networks," *IEEE International Symposium on Information Theory*, 2008.

[14] S. Lashgari, A. S. Avestimehr, and C. Suh, "Linear Degrees of Freedom of the X-Channel with Delayed CSIT," *arXiv:1309.0799*, 2013.

[15] J. Abdoli and A. S. Avestimehr, "On Degrees of Freedom Scaling in Layered Interference Networks with Delayed CSI," *IEEE International Symposium on Information Theory*, 2013.

[16] J. Abdoli and A. S. Avestimehr, "Layered Interference Networks with Delayed CSI: DoF Scaling with Distributed Transmitters," *arXiv:1302.4788*, 2013.

[17] I. Shomorony and A. S. Avestimehr, "Degrees of Freedom of Two-Hop Wireless Networks: "Everyone Gets the Entire Cake"," *arXiv:1210.2143*, 2013.

Appendix I
Proof of Lemma 3

A. Proof of $\text{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n \quad \mathbf{G}_{12}^n\mathbf{V}_2^n] - \text{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{22}^n\mathbf{V}_2^n] \overset{a.s.}{\leq} \text{rank}[\mathbf{G}_{11}^{\mathcal{T}}\mathbf{V}_1^{\mathcal{T}} \quad \mathbf{G}_{12}^{\mathcal{T}}\mathbf{V}_2^{\mathcal{T}}]$:

For a fixed linear coding strategy $\{f_1^{(n)}, f_2^{(n)}\}$, with corresponding $\mathbf{V}_1^n, \mathbf{V}_2^n$, let $\mathcal{A}_i, \mathcal{B}_i, \mathcal{C}_i$, $i = 1, 2, \ldots, n$, denote the following sets:

- $\mathcal{A}_i \triangleq \{\mathcal{G}^n | \ \text{rank}[G_{21}^iV_1^i \quad G_{22}^iV_2^i] = \text{rank}[G_{21}^{i-1}V_1^{i-1} \quad G_{22}^{i-1}V_2^{i-1}]\}$.
- $\mathcal{B}_i \triangleq \{\mathcal{G}^n | \ [\vec{v}_1(i)^\top \quad \vec{0}_{1\times m_2(n)}], [\vec{0}_{1\times m_1(n)} \quad \vec{v}_2(i)^\top] \in \text{rowspan}[G_{21}^{i-1}V_1^{i-1} \quad G_{22}^{i-1}V_2^{i-1}]\}$.
- $\mathcal{C}_i \triangleq \{\mathcal{G}^n | \ \text{rank}[G_{11}^iV_1^i \quad G_{12}^iV_2^i] = \text{rank}[G_{11}^{i-1}V_1^{i-1} \quad G_{12}^{i-1}V_2^{i-1}] + 1\}$,

Note that $\mathcal{B}_i$ is equivalent to $\{\mathcal{G}^n | \quad i \in \mathcal{T}(\mathcal{G}^n)\}$. In order to prove Lemma 3 we first state the following lemma, whose proof is provided in [14].

Lemma 4: $\Pr(\boldsymbol{\mathcal{G}}^n \in \cup_{i=1}^n(\mathcal{A}_i \cap \mathcal{B}_i^c)) = 0$.

Lemma 4 implies that we need to prove the first inequality in Lemma 3 only for channel realizations $\boldsymbol{\mathcal{G}}^n = \mathcal{G}^n$, such that $\mathcal{G}^n \notin \cup_{i=1}^n(\mathcal{A}_i \cap \mathcal{B}_i^c)$ (since, the rest have probability measure zero). Thus, we only need to show that for any arbitrary channel realization $\boldsymbol{\mathcal{G}}^n = \mathcal{G}^n$ with the corresponding beamforming matrices V_1^n, V_2^n, and $\boldsymbol{\mathcal{T}} = \mathcal{T}$, such that $\mathcal{G}^n \notin \cup_{i=1}^n(\mathcal{A}_i \cap \mathcal{B}_i^c)$, we have

$$\begin{aligned}&\text{rank}[G_{11}^nV_1^n \quad G_{12}^nV_2^n] - \text{rank}[G_{21}^nV_1^n \quad G_{22}^nV_2^n] \\ &\leq \text{rank}[G_{11}^{\mathcal{T}}V_1^{\mathcal{T}} \quad G_{12}^{\mathcal{T}}V_2^{\mathcal{T}}]. \end{aligned} \tag{25}$$

Let $I(.)$ denote the indicator function, we now bound the left hand side of (25) as follows.

$$
\begin{aligned}
&\text{rank}[G_{11}^n V_1^n \quad G_{12}^n V_2^n] - \text{rank}[G_{21}^n V_1^n \quad G_{22}^n V_2^n] \\
=&\sum_{i=1}^{n}(\text{rank}[G_{11}^i V_1^i \quad G_{12}^i V_2^i] - \text{rank}[G_{11}^{i-1} V_1^{i-1} \quad G_{12}^{i-1} V_2^{i-1}]) \\
&- (\text{rank}[G_{21}^i V_1^i \quad G_{22}^i V_2^i] - \text{rank}[G_{21}^{i-1} V_1^{i-1} \quad G_{22}^{i-1} V_2^{i-1}]) \\
\leq&\sum_{i=1}^{n} \max\{(\text{rank}[G_{11}^i V_1^i \quad G_{12}^i V_2^i] \\
&- \text{rank}[G_{11}^{i-1} V_1^{i-1} \quad G_{12}^{i-1} V_2^{i-1}]) - (\text{rank}[G_{21}^i V_1^i \quad G_{22}^i V_2^i] \\
&- \text{rank}[G_{21}^{i-1} V_1^{i-1} \quad G_{22}^{i-1} V_2^{i-1}]), 0\} \\
\overset{(a)}{=}&\sum_{i=1}^{n} I(\text{rank}[G_{11}^i V_1^i \quad G_{12}^i V_2^i] \\
&= \text{rank}[G_{11}^{i-1} V_1^{i-1} \quad G_{12}^{i-1} V_2^{i-1}] + 1) \\
&\times I(\text{rank}[G_{21}^i V_1^i \quad G_{22}^i V_2^i] \\
&= \text{rank}[G_{21}^{i-1} V_1^{i-1} \quad G_{22}^{i-1} V_2^{i-1}]) \\
=&\sum_{i=1}^{n} I(\mathcal{G}^n \in \mathcal{A}_i \cap \mathcal{C}_i) = \sum_{i=1}^{n}(I(\mathcal{G}^n \in \mathcal{A}_i \cap \mathcal{B}_i \cap \mathcal{C}_i) \\
&+ I(\mathcal{G}^n \in \mathcal{A}_i \cap \mathcal{B}_i^c \cap \mathcal{C}_i)) \\
\leq&\sum_{i=1}^{n}(I(\mathcal{G}^n \in \mathcal{B}_i \cap \mathcal{C}_i) + I(\mathcal{G}^n \in \mathcal{A}_i \cap \mathcal{B}_i^c)) \\
\overset{(b)}{=}&\sum_{i=1}^{n} I(\mathcal{G}^n \in \mathcal{B}_i \cap \mathcal{C}_i) \overset{(c)}{=} \sum_{i\in\mathcal{T}} I(\mathcal{G}^n \in \mathcal{C}_i) \\
=&\sum_{i\in\mathcal{T}} I(\text{rank}[G_{11}^i V_1^i \quad G_{12}^i V_2^i] \\
&= \text{rank}[G_{11}^{i-1} V_1^{i-1} \quad G_{12}^{i-1} V_2^{i-1}] + 1),
\end{aligned} \tag{26}
$$

where (a) holds since $\text{rank}[G_{k1}^i V_1^i \quad G_{k2}^i V_2^i] - \text{rank}[G_{k1}^{i-1} V_1^{i-1} \quad G_{k2}^{i-1} V_2^{i-1}] \in \{0,1\}$ for $k = 1,2$; and (b) follows from the assumption that $\mathcal{G}^n \notin (\mathcal{A}_i \cap \mathcal{B}_i^c)$ for $i \in \{1,2,\ldots,n\}$; and (c) follows from the fact that $\mathcal{T} = \{i|\mathcal{G}^n \in \mathcal{B}_i\}$. We now only need to show the following to complete the proof of (25).

$$
\begin{aligned}
&\sum_{i\in\mathcal{T}} I(\text{rank}[G_{11}^i V_1^i \quad G_{12}^i V_2^i] \\
&\quad = \text{rank}[G_{11}^{i-1} V_1^{i-1} \quad G_{12}^{i-1} V_2^{i-1}] + 1) \\
&\leq \text{rank}[G_{11}^{\mathcal{T}} V_1^{\mathcal{T}} \quad G_{12}^{\mathcal{T}} V_2^{\mathcal{T}}].
\end{aligned} \tag{27}
$$

Without loss of generality, let us assume that $\mathcal{T} = \{\tau_1, \tau_2, \ldots, \tau_k\}$ for some k, such that $\tau_1 < \tau_2 < \ldots < \tau_k$. We define $\mathcal{T}_j \triangleq \{\tau_1, \tau_2, \ldots, \tau_j\}$, and use $V_1^{\mathcal{T}_j}$ and $V_2^{\mathcal{T}_j}$ to denote the sub-matrices of V_1^n and V_2^n with rows in $\mathcal{T}_j$. We also use $G_{11}^{\mathcal{T}_j}$ to denote the $|\mathcal{T}_j| \times |\mathcal{T}_j|$ diagonal matrix with channel coefficients of $g_{11}(t)$ at timeslots $t \in \mathcal{T}_j$ on its diagonal (similarly defined for other channel matrices). We now present a claim that will be used to show (27) and complete the proof.

Claim 2: For any $j = 1, 2, \ldots, k$,

$$
\begin{aligned}
&I(\text{rank}[G_{11}^{\tau_j} V_1^{\tau_j} \quad G_{12}^{\tau_j} V_2^{\tau_j}] \\
&\quad = \text{rank}[G_{11}^{\tau_j-1} V_1^{\tau_j-1} \quad G_{12}^{\tau_j-1} V_2^{\tau_j-1}] + 1) \\
&\leq I(\text{rank}[G_{11}^{\mathcal{T}_j} V_1^{\mathcal{T}_j} \quad G_{12}^{\mathcal{T}_j} V_2^{\mathcal{T}_j}] \\
&\quad = \text{rank}[G_{11}^{\mathcal{T}_{j-1}} V_1^{\mathcal{T}_{j-1}} \quad G_{12}^{\mathcal{T}_{j-1}} V_2^{\mathcal{T}_{j-1}}] + 1).
\end{aligned} \tag{28}
$$

Proof: The claim is trivially true when $\text{rank}[G_{11}^{\tau_j} V_1^{\tau_j} \quad G_{12}^{\tau_j} V_2^{\tau_j}] = \text{rank}[G_{11}^{\tau_j-1} V_1^{\tau_j-1} \quad G_{12}^{\tau_j-1} V_2^{\tau_j-1}]$. So, suppose $\text{rank}[G_{11}^{\tau_j} V_1^{\tau_j} \quad G_{12}^{\tau_j} V_2^{\tau_j}] = \text{rank}[G_{11}^{\tau_j-1} V_1^{\tau_j-1} \quad G_{12}^{\tau_j-1} V_2^{\tau_j-1}] + 1$. It means that $[g_{11}(\tau_j)\vec{v}_1(\tau_j)^\top \quad g_{12}(\tau_j)\vec{v}_2(\tau_j)^\top]$ is linearly independent of $\text{rowspan}[G_{11}^{\tau_j-1} V_1^{\tau_j-1} \quad G_{12}^{\tau_j-1} V_2^{\tau_j-1}]$. Since $\mathcal{T}_{j-1} \subseteq \{1, 2, \ldots, \tau_j - 1\}$, then $[g_{11}(\tau_j)\vec{v}_1(\tau_j)^\top \quad g_{12}(\tau_j)\vec{v}_2(\tau_j)^\top]$ is also linearly independent of $\text{rowspan}[G_{11}^{\mathcal{T}_{j-1}} V_1^{\mathcal{T}_{j-1}} \quad G_{12}^{\mathcal{T}_{j-1}} V_2^{\mathcal{T}_{j-1}}]$. Hence,

$$
\begin{aligned}
\text{rank}[G_{11}^{\mathcal{T}_j} V_1^{\mathcal{T}_j} \quad G_{12}^{\mathcal{T}_j} V_2^{\mathcal{T}_j}] =& \text{rank}[G_{11}^{\mathcal{T}_{j-1}} V_1^{\mathcal{T}_{j-1}} \quad G_{12}^{\mathcal{T}_{j-1}} V_2^{\mathcal{T}_{j-1}}] \\
&+ 1.
\end{aligned}
$$

■

Based on this claim, the proof of (27) is as follows.

$$
\begin{aligned}
&\sum_{i\in\mathcal{T}} I(\text{rank}[G_{11}^i V_1^i \quad G_{12}^i V_2^i] \\
&\quad = \text{rank}[G_{11}^{i-1} V_1^{i-1} \quad G_{12}^{i-1} V_2^{i-1}] + 1) \\
=&\sum_{j=1}^{k} I(\text{rank}[G_{11}^{\tau_j} V_1^{\tau_j} \quad G_{12}^{\tau_j} V_2^{\tau_j}] \\
&\quad = \text{rank}[G_{11}^{\tau_j-1} V_1^{\tau_j-1} \quad G_{12}^{\tau_j-1} V_2^{\tau_j-1}] + 1) \\
\overset{\text{Claim 2}}{\leq}&\sum_{j=1}^{k} I(\text{rank}[G_{11}^{\mathcal{T}_j} V_1^{\mathcal{T}_j} \quad G_{12}^{\mathcal{T}_j} V_2^{\mathcal{T}_j}] \\
&\quad = \text{rank}[G_{11}^{\mathcal{T}_{j-1}} V_1^{\mathcal{T}_{j-1}} \quad G_{12}^{\mathcal{T}_{j-1}} V_2^{\mathcal{T}_{j-1}}] + 1) \\
=&\text{rank}[G_{11}^{\mathcal{T}_k} V_1^{\mathcal{T}_k} \quad G_{12}^{\mathcal{T}_k} V_2^{\mathcal{T}_k}] = \text{rank}[G_{11}^{\mathcal{T}} V_1^{\mathcal{T}} \quad G_{12}^{\mathcal{T}} V_2^{\mathcal{T}}].
\end{aligned}
$$

B. *Proof of* $\text{rank}[\mathbf{V}_j^{\mathcal{T}}] \leq \mathbf{r}_j, \quad (j = 1, 2):$

It is sufficient to prove that $\text{rank}[\mathbf{V}_1^{\mathcal{T}}] \leq \mathbf{r}_1$, since the other inequality (i.e. $\text{rank}[\mathbf{V}_2^{\mathcal{T}}] \leq \mathbf{r}_2$) can be proven similarly. We show that for any realization $\mathcal{G}^n = \{G_{kj}^n\}_{k,j\in\{1,2\}}$ with the corresponding values $\mathcal{T}$, r_1, and matrices V_1^n, V_2^n, we have $\text{rank}[V_1^{\mathcal{T}}] \leq r_1$. But according to definition of r_1, it is sufficient to prove

$$
\begin{aligned}
&\text{rowspan}[V_1^{\mathcal{T}}] \subseteq \text{span}(\vec{s}_{m_1(n)\times 1} | \quad \exists \vec{l}_{n\times 1} \\
&\quad s.t. \quad [\vec{s}^\top \quad \vec{0}_{1\times m_2(n)}] = \vec{l}^\top [G_{21}^n V_1^n \quad G_{22}^n V_2^n]).
\end{aligned} \tag{29}
$$

The following proves (29), thereby completing the proof for $\text{rank}[V_1^{\mathcal{T}}] \leq r_1$:

$$\begin{aligned}
&\text{rowspan}[V_1^{\mathcal{T}}]\\
=&\text{span}(\vec{v}_1(i)|1 \leq i \leq n, [\vec{v}_1(i)^\top \quad \vec{0}_{1\times m_2(n)}]\\
&, [\vec{0}_{1\times m_1(n)} \quad \vec{v}_2(i)^\top] \in \text{rowspan}[G_{21}^{i-1}V_1^{i-1} \quad G_{22}^{i-1}V_2^{i-1}])\\
\subseteq&\text{span}(\vec{v}_1(i)|1 \leq i \leq n, [\vec{v}_1(i)^\top \quad \vec{0}_{1\times m_2(n)}]\\
&, [\vec{0}_{1\times m_1(n)} \quad \vec{v}_2(i)^\top] \in \text{rowspan}[G_{21}^{n}V_1^{n} \quad G_{22}^{n}V_2^{n}])\\
\subseteq&\text{span}(\vec{v}_1(i)|1 \leq i \leq n, [\vec{v}_1(i)^\top \quad \vec{0}_{1\times m_2(n)}]\\
&\in \text{rowspan}[G_{21}^{n}V_1^{n} \quad G_{22}^{n}V_2^{n}])\\
\subseteq&\text{span}(\vec{s}_{m_1(n)\times 1}| \quad \exists\vec{l}_{n\times 1} \quad s.t.\\
&[\vec{s}^\top \quad \vec{0}_{1\times m_2(n)}] = \vec{l}^{\,\top}[G_{21}^{n}V_1^{n} \quad G_{22}^{n}V_2^{n}]).
\end{aligned}$$

C. Proof of $\mathbf{r}_j \overset{a.s.}{\leq} \text{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{22}^n\mathbf{V}_2^n] - \text{rank}[\mathbf{V}_{3-j}^n], \quad (j = 1, 2):$

We will show this for $j = 1$; the proof for $j = 2$ will be similar. Since $\text{rank}[\mathbf{G}_{22}^n\mathbf{V}_2^n] \overset{a.s.}{=} \text{rank}[\mathbf{V}_2^n]$, it is sufficient to show that $\mathbf{r}_1 \leq \text{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{22}^n\mathbf{V}_2^n] - \text{rank}[\mathbf{G}_{22}^n\mathbf{V}_2^n]$. To do so, we show that for any realization $\mathcal{G}^n = \{G_{kj}^n\}_{k,j\in\{1,2\}}$ with the corresponding value r_1, and matrices V_1^n, V_2^n, we have $r_1 \leq \text{rank}[G_{21}^nV_1^n \quad G_{22}^nV_2^n] - \text{rank}[G_{22}^nV_2^n]$.

Since $r_1 = \dim(\text{span}(\vec{s}_{m_1(n)\times 1}| \quad \exists\vec{l}_{n\times 1}$ $s.t. \quad [\vec{s}^\top \quad \vec{0}_{1\times m_2(n)}] = \vec{l}^{\,\top}[G_{21}^nV_1^n \quad G_{22}^nV_2^n]))$,

$$\exists L_{r_1\times n} \quad s.t. \quad [S \quad 0_{r_1\times m_2(n)}] = L[G_{21}^nV_1^n \quad G_{22}^nV_2^n], \tag{30}$$

for some $S_{r_1\times m_1(n)}$, such that $\text{rank}[S] = r_1$. This means

$$LG_{22}^nV_2^n = 0_{\mathbf{r}_1\times m_2(n)}, LG_{21}^nV_1^n = S, \text{rank}[LG_{21}^nV_1^n] = r_1. \tag{31}$$

We now state a claim useful in completing the proof.

Claim 3: For three matrices A, B, C where the number of columns in A is equal to the number of rows in B, C,

$$\text{rank}[AB \quad AC] - \text{rank}[AC] \leq \text{rank}[B \quad C] - \text{rank}[C]. \tag{32}$$

Proof: By Frobenius's inequality, for any three matrices X, Y, Z where XY, YZ, and XYZ are defined,

$$\text{rank}[XY] + \text{rank}[YZ] \leq \text{rank}[XYZ] + \text{rank}[Y]. \tag{33}$$

By setting $X = A, Y = [B \quad C], Z = [0 \quad I]^t$, where I is the identity matrix, the desired result follows. ■

Therefore, by setting $A = L, B = [G_{21}^nV_1^n \quad G_{22}^nV_2^n], C = G_{22}^nV_2^n$ in Claim 3, and using (31), we get

$$r_1 - 0 \leq \text{rank}[G_{21}^nV_1^n \quad G_{22}^nV_2^n] - \text{rank}[G_{22}^nV_2^n].$$

■