On Degrees of Freedom of Layered Two Unicast Networks with Delayed CSIT

I-Hsiang Wang
EPFL, Lausanne, Switzerland
i-hsiang.wang@epfl.ch

Suhas Diggavi
UCLA, Los Angeles, USA
suhasdiggavi@ucla.edu

***Abstract*—In this paper we study the two unicast information flow problem over layered Gaussian networks with arbitrary number of nodes and connectivity, under the model of delayed channel state information (CSI) at transmitters and instantaneous CSI at receivers. We show that similar to the case with instantaneous CSI at transmitters (CSIT), the degrees of freedom (DoF) region is strictly larger than the time-sharing DoF region if and only if there is no *omniscient node*, definition of which only depends on the topology of the network. Moreover, as in the case with instantaneous CSIT, $2/3$ DoF per user is always achievable when there is no omniscient node in the network.**

I. Introduction

Characterizing the capacity region of multiple unicast sessions over multi-source-multi-destination networks (namely, multiple unicast networks) has been a long-standing open question in network information theory. For a Gaussian network with arbitrary number of nodes and connectivity, except for the single unicast/multicast problem where the approximate capacity is characterized [1], very little is known.

When there are multiple information flows in the network, they cause *(message) interference* to one another, and in general this will reduce the number of *degrees of freedom* (DoF) available for each user. Recently it has been shown that one can make use of the intermediate nodes to facilitate *interference neutralization* [2] [3] [4] [5], thereby reducing the effect of message interference. In particular, [5] fully characterized the DoF region of layered two unicast Gaussian networks with constant channel gains. In some situations, full DoF can be achieved for each user.

Interference neutralization opens various interesting possibilities for managing interference in multiple unicast networks. However it requires *instantaneous* channel state information at transmitters (CSIT) to work. In practice, channel gains are measured at receivers and then fed back to the transmitters with certain *delay*. In situations when the delay is larger than the channel coherence time, the channel state could have changed. In such cases, the transmitters can only have *strictly causal* knowledge about the channel states. Under such constraints, interference neutralization is no longer feasible, and a natural question is whether such delayed CSIT can be helpful. The value of such delayed CSIT was demonstrated for MISO broadcast channels in [6], where it enabled a boost of DoF strictly beyond that when no CSIT is available. For MIMO interference channels with delayed CSIT [7] [8] [9], and the $2 \times 2 \times 2$ network [9] [10], a strict improvement over the case without CSIT was also demonstrated.

However, in a network with arbitrary number of nodes and connectivity, how to make use of delayed transmitter CSI and what is the best-achievable DoF is unsettled even for two unicast Gaussian networks. In this paper, we make some progress, by answering the following question: in layered two unicast Gaussian networks, what is the necessary and sufficient condition when one can make use of delayed CSIT to increase the DoF region strictly beyond the time-sharing DoF region? It turns out that the answer coincides with that under instantaneous CSIT [5], and the condition is that there is no *omniscient node* [4] in the network[1]. Moreover, as in the case under instantaneous CSIT [5], once beyond $1/2$ DoF per user (the time-sharing DoF) is achievable, so is $2/3$ DoF per user. However, when there is no omniscient node in the network, unlike the case under instantaneous CSIT where either $(1/2, 1)$ or $(1, 1/2)$ DoF pair is achievable, there are examples where neither can be achieved while $(2/3, 2/3)$ is achievable (*e.g.,* the $2 \times 2 \times 2$ network, the DoF region of which is completely characterized in [9] [10]).

The converse part of the result follows straightforwardly from that in the case with instantaneous CSIT: if there is an omniscient node in the network, even with instantaneous CSIT, sum DoF of greater than 1 is not achievable. Our main contribution lies in the achievability proof when there is no omniscient node, where we extend the scheme proposed in [6] [10] to exploit delayed CSIT in two unicast networks with arbitrary number of nodes and connectivity. Unlike in specific networks such as the two-user interference channel or the $2 \times 2 \times 2$ network, the key idea in dealing with networks with general connectivity is to identify the special nodes that have to make carefully designed (linear) operations; these nodes are identified in a systematical way. To achieve $2/3$ DoF per user, most nodes in the network can just perform oblivious random linear coding that does not require CSI, except for these special nodes which do the carefully designed linear operations to exploit delayed CSI.

The work of Suhas Diggavi was partially supported by NSF-CPS award 1136174. The work of I-Hsiang Wang was supported by EU project CONECT FP7-ICT-2009-257616.

[1] It is straightforward to verify that this condition is equivalent to the condition given in [5].

The paper is organized as follows. Section II describes the model of the network and the channel state information. Section III summarizes a few definitions and the main theorem. In Section IV, we demonstrate several key ideas through a motivating example. Then Section V is devoted to prove the achievability. Finally we conclude the paper with Section VI.

II. Problem Formulation

A two-source-two-destination layered network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a collection of nodes $\mathcal{V}$ that can be partitioned into $L + 2$ *layers* ($L \geq 0$):

$$\mathcal{V} = \bigcup_{k=0}^{L+1} \mathcal{L}_k,\ \mathcal{L}_k \cap \mathcal{L}_j \neq \emptyset,\ \forall k \neq j,$$

such that for any edge $(\mathsf{u}, \mathsf{v}) \in \mathcal{E}$, $\exists\ k,\ 0 \leq k \leq L$ s.t. $\mathsf{u} \in \mathcal{L}_k, \mathsf{v} \in \mathcal{L}_{k+1}$. The first layer $\mathcal{L}_0 = \{\mathsf{s}_1, \mathsf{s}_2\}$ consists of the two source nodes, and the last layer $\mathcal{L}_{L+1} = \{\mathsf{d}_1, \mathsf{d}_2\}$ consists of the two destination nodes. Without loss of generality we assume each node in the network can be reached by at least one of the source nodes and can reach at least one of the destination nodes. For each node $\mathsf{v} \in \mathcal{V} \setminus \{\mathsf{s}_1, \mathsf{s}_2\}$, we define nodes that can reach v as its *predecessors*. Let $\mathcal{P}(\mathsf{v})$ denote the set of predecessors that can reach v in one step. We will call the nodes in $\mathcal{P}(\mathsf{v})$ as the *parents* of v.

For each edge $(\mathsf{u}, \mathsf{v}) \in \mathcal{E}$, the associated channel gain at time t is given by $h_{\mathsf{vu}}[t] \in \mathbb{R}, t \geq 0$, where the channel gains are i.i.d. drawn from a continuous distribution, over time and $\mathcal{E}$. Let $X_{\mathsf{u}}[t], Y_{\mathsf{u}}[t] \in \mathbb{R}$ denote the transmission and reception of node u respectively. The reception of a node is the superposition of the transmission of its parents along with an additive white Gaussian noise:

$$Y_{\mathsf{v}}[t] = \textstyle\sum_{\mathsf{u} \in \mathcal{P}(\mathsf{v})} h_{\mathsf{vu}}[t] X_{\mathsf{u}}[t] + Z_{\mathsf{u}}[t],$$

where the noise $Z_{\mathsf{u}}[t] \sim \mathcal{N}(0, \sigma^2)$ is i.i.d. over time and $\mathcal{V}$. Each node has the same power constraint P.

At time $t \geq 1$, a node $\mathsf{w} \in \mathcal{V}$ has knowledge of the channel coefficient associated to edge $(\mathsf{u}, \mathsf{v}) \in \mathcal{E}$, h_{vu} up to time $(t - l)$, if the *tail node* of this edge, v, can reach w in l steps in the *undirected* version of $\mathcal{G}$. Note under this CSI model, it remains that a receiver have instantaneous CSI associated to the edges connecting to it, while a transmitter only have unit-delayed CSI associated to the edges emanating from it. For example, in the network depicted in Fig. 1(a), at time t regarding the channel gain associated to the edge $(\mathsf{u}_2, \mathsf{u}_4)$, the source s_1 has its CSI up to time $(t-2)$, simply because the tail node u_4 can reach s_1 in two steps on the undirected version of $\mathcal{G}$.

We are mainly interested in the high-SNR approximate capacity region, namely, the collection of achievable degrees of freedom (DoF). For user $i, i = 1, 2$, its achievable DoF is defined as $d_i := \lim_{P \to \infty} \frac{R_i}{\frac{1}{2}\log(1 + P/\sigma^2)}$, where (R_1, R_2) lies in the capacity region of the two unicast network.

III. Result

To present our main result, we begin with a few definitions, followed by an example network illustrating the definitions.

Definition 1 (Critical Nodes): For each $i = 1, 2$, we define the *critical node* v_i^* as any node with the smallest layer index such that the removal of $\{\mathsf{v}_i^*\}$ separates d_i from $\{\mathsf{s}_1, \mathsf{s}_2\}$. We use $\mathcal{L}_{k_i^*}$ to denote the layer where critical nodes v_i^* lies.

Definition 2 (Cloud): Define the cloud $\mathcal{C}_i$, for $i = 1, 2$, to be the set of nodes that can be reached by v_i^* and that can reach d_i. All nodes in the cloud receive functions of the reception of the critical node and the additive noise terms at their predecessors in the cloud.

Definition 3: A node is s_i*-reachable* if it can be reached by s_i. It is s_i*-only-reachable* if it can be reached by s_i but not s_j, $j \neq i$. It is $\mathsf{s}_1\mathsf{s}_2$*-reachable* if it can be reached by both s_1 and s_2. For each node $\mathsf{v} \in \mathcal{V} \setminus \{\mathsf{s}_1, \mathsf{s}_2\}$, for $i = 1, 2$, let $\mathcal{P}^{\mathsf{s}_i}(\mathsf{v}) \subseteq \mathcal{P}(\mathsf{v})$ denote the set of parents of v that are s_i-*reachable*, and we define for $i = 1, 2$,

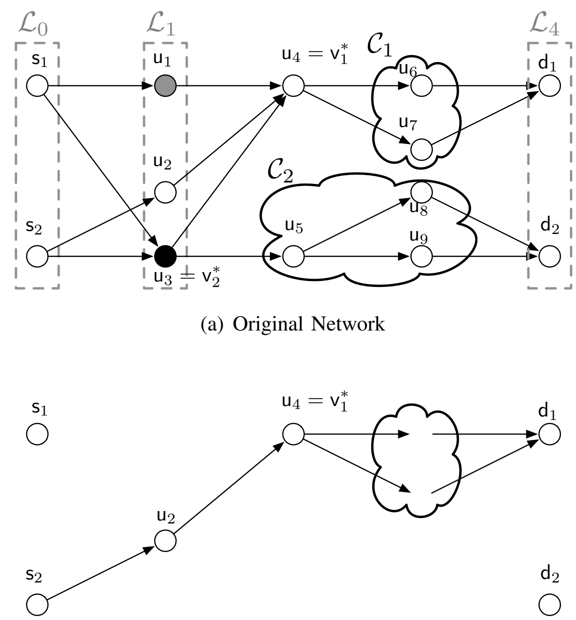
$$\mathcal{K}^{\mathsf{s}_i}(\mathsf{v}) := \begin{cases} \{\mathsf{v}\}, & \text{if } |\mathcal{P}^{\mathsf{s}_i}(\mathsf{v})| \neq 1 \\ \{\mathsf{v}\} \cup \{\mathsf{u} : \mathcal{P}^{\mathsf{s}_i}(\mathsf{u}) = \mathcal{P}^{\mathsf{s}_i}(\mathsf{v})\}, & \text{if } |\mathcal{P}^{\mathsf{s}_i}(\mathsf{v})| = 1 \end{cases} \tag{1}$$

Note that only accounting the contribution from source s_i, the reception of v and a node in $\mathsf{u} \in \mathcal{K}^{\mathsf{s}_i}(\mathsf{v})$ are just a scaled version of each other, *i.e.,* they are s_i-clones.

Definition 4 (Omniscient Node): We say a node $\mathsf{v} \in \mathcal{V}$ is *omniscient* if it satisfies either of (A) or (B) below:

(A) The removal of $\{\mathsf{v}\}$ separates d_1 from $\{\mathsf{s}_1, \mathsf{s}_2\}$ and the removal of $\mathcal{K}^{\mathsf{s}_2}(\mathsf{v})$ separates d_2 from s_2.

(B) The same as above with indice 1 and 2 exchanged.



(a) Original Network

(b) Removal of $\mathcal{K}^{\mathsf{s}_1}(\mathsf{u}_3) = \{\mathsf{u}_1, \mathsf{u}_3\}$

Fig. 1. A Network with Sum DoF = 1

To illustrate the above definitions, let us consider the network illustrated in Fig. 1(a) which consists of five layers: $\mathcal{L}_0, \mathcal{L}_1, \ldots, \mathcal{L}_4$. The second layer $\mathcal{L}_1$ consists of three nodes $\{\mathsf{u}_1, \mathsf{u}_2, \mathsf{u}_3\}$. u_1 is s_1-only-reachable, u_2 is s_2-only-reachable, while u_3 is $\mathsf{s}_1\mathsf{s}_2$-reachable. It is straightforward to see that the removal of u_3 separates d_2 from $\{\mathsf{s}_1, \mathsf{s}_2\}$ and it is the earliest (in terms of layer indices), *i.e.,* u_3 is the critical node for user 2.

Similarly, the node u_4 is the critical node for user 1. The cloud $\mathcal{C}_1 = \{\mathsf{u}_6, \mathsf{u}_7\}$, the reception of which is (roughly) determined by the critical node u_3. Similarly the cloud $\mathcal{C}_2 = \{\mathsf{u}_5, \mathsf{u}_8, \mathsf{u}_9\}$.

The nodes u_1 and u_3 share the same s_1-reachable parent which is s_1. Hence, we see that from (1) the s_1-clone of u_3: $\mathcal{K}^{\mathsf{s}_1}(\mathsf{u}_3) = \{\mathsf{u}_1, \mathsf{u}_3\}$. Now from Fig. 1(b), we see that the removal of $\mathcal{K}^{\mathsf{s}_1}(\mathsf{u}_3)$ separates d_1 from s_1. Hence, u_3 is omniscient since it satisfies Condition (B) of Definition 4.

Next we shall argue intuitively that for any working system on this network, u_3 can decode both users' messages, implying that the sum DoF is at most 1. To see this, we argue first by ignoring the noise. In a working system, since u_3 is a critical node for user 2, it can decode user 2's message; hence transmission from s_2. Since the received signal of u_3 is the sum of the transmitted signals from s_1 and s_2, it therefore knows the transmission from s_1 and the the reception of u_1. Therefore in any working system, d_2-critical node u_3 knows the transmission from source s_2 and $\mathcal{K}^{\mathsf{s}_1}(\mathsf{u}_3) = \{\mathsf{u}_1, \mathsf{u}_3\}$. Finally, the reception of d_1 is determined by the transmissions from $\mathcal{K}^{\mathsf{s}_1}(\mathsf{u}_3)$ and s_2, and hence in a working system, u_3 can decode user 1's message as well.

The following theorem summarizes the main result.

Theorem 1: Under delayed CSIT model, if there is no omniscient node in the network, then the DoF region contains the kite region

$$\mathfrak{K} := \{(d_1, d_2) : d_1 \geq 0, d_2 \geq 0, 2d_1 + d_2 \leq 2, d_1 + 2d_2 \leq 2\}$$

and in particular, $(d_1, d_2) = (2/3, 2/3)$ is achievable.

Conversely, if there is an omniscient node in the network, then the DoF region is the triangular region

$$\mathfrak{T} := \{d_1 \geq 0, d_2 \geq 0, d_1 + d_2 \leq 1\}.$$

Remark: There is an example (the $2 \times 2 \times 2$ interference network) where the DoF region is exactly the kite $\mathfrak{K}$.

The converse part of Theorem 1 follows from the instantaneous CSIT result and hence is omitted here. It can be found in the appendix of the extended version of this paper [11]. In the rest of the paper we focus on the proof of the direct part.

IV. A Motivating Example for Achievability

Before we go into details of the proof, in this section we first present some key intuitions behind the achievability proof via a special case where one of the two users does not experience interference. We illustrate how to identify and code at special nodes of the network. One such special nodes are the *critical nodes* (see Definition 1) which (roughly) determine the reception of their respective destinations, *i.e.,* should be able to decode the desired messages. Moreover, the critical nodes naturally break the overall networks into several effective stages, which play a key role in the scheme.

Consider the special networks where d_1 can only be reached by s_1 and hence only user 2 is interfered. The critical node for user 1 is s_1. The assumption of the non-existence of omniscient node then implies that the critical node for user 2 must lie in a layer with index strictly larger than 1. Therefore we have $k_1^* = 0, k_2^* \geq 2$ (see Definition 1) as illustrated in Fig. 2.

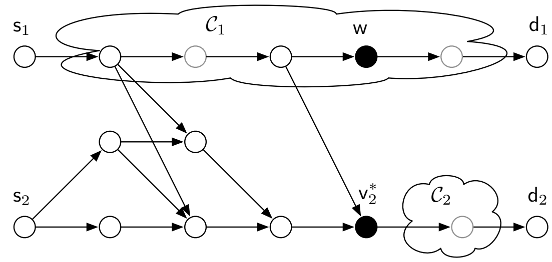


Fig. 2. The case $k_1^* = 0$ and $k_2^* \geq 2$.

First, s_1 can reach the critical node v_2^* since $k_2^* \neq 0$. Second, note that in layer $\mathcal{L}_{k_2^*}$, the layer that user 2's critical node belongs to, only the critical node v_2^* can reach d_2 due to its definition. Therefore other nodes in $\mathcal{L}_{k_2^*}$ belong to cloud $\mathcal{C}_1$. Beside, since v_2^* is not omniscient but its removal will separate d_2 from $\{\mathsf{s}_1, \mathsf{s}_2\}$, the removal of its s_1-clones, $\mathcal{K}^{\mathsf{s}_1}(\mathsf{v}_2^*)$, cannot separate d_1 from s_1. Therefore, there exists a node $\mathsf{w} \in \mathcal{L}_{k_2^*}$ that can reach d_1, such that

i) $\mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*) \neq \mathcal{P}^{\mathsf{s}_1}(\mathsf{w})$, or
ii) $\mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*) = \mathcal{P}^{\mathsf{s}_1}(\mathsf{w})$ and $|\mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*)| \geq 2$.

Now, if instantaneous CSIT is available, the parents of w and v_2^* can choose their transmission according to the channel gains that they are faced with so that interference is completely cancelled "over-the-air" at w and v_2^* [5]. Without the knowledge of these channel coefficients beforehand, however, such cancellation seems impossible.

Under delayed CSIT, we shall prove that $(d_1, d_2) = (1, 1/2)$ is always achievable. Our scheme is linear in the sense that every node transmits a linear transformation of its reception. Since the performance measure of interest is DoF, in the rest of the paper we neglect the additive noise terms and focus on the linear combinations that nodes receive and transmit. The goal is to deliver a **block** (length 2) of two symbols $\{a_1, a_2\}$ from s_1 to d_1 and a **block** of one symbol $\{b\}$ from s_2 to d_2 over two time slots. Our scheme breaks the layered network into several stages, each of which consists of several consecutive layers. When we describe the scheme in the following, we focus on how transmitted signals in each stage are formed according to the reception from the previous stage. Moreover, some special nodes that exploit delayed transmitter CSI have to wait until the relevant CSI is gathered. Therefore, the time index t in the following does not refer to "absolute time" but the "order index" of the transmitted signal associated to each block of symbols. Different blocks will be pipelined in a standard way.

Let us describe the scheme for the stage from Layer $\mathcal{L}_0$ to $\mathcal{L}_{(k_2^*-1)}$ and the stage from $\mathcal{L}_{(k_2^*-1)}$ to $\mathcal{L}_{k_2^*}$, and show that DoF $(d_1, d_2) = (1, 1/2)$ is achievable.

Phase 1: Stage $\mathcal{L}_0$ to $\mathcal{L}_{(k_2^*-1)}$, Time $t = 1, 2$: At each of the two time slots, each node in this stage (excluding $\mathcal{L}_{(k_2^*-1)}$) transmits a randomly chosen linear transformation of its reception at that time slot, with s_1 sending $\{a_1, a_2\}$ and s_2 sending $\{b, b\}$ at time $t = 1$ and $t = 2$ respectively.

Phase 2: Stage $\mathcal{L}_{(k_2^*-1)}$ to $\mathcal{L}_{k_2^*}$, Time $t = 1$: Each node in layer $\mathcal{L}_{(k_2^*-1)}$ transmits a randomly chosen linear transformation of its receptions over time slots $t = 1$ and $t = 2$. A node

$u \in \mathcal{L}_{(k_2^*-1)}$ transmits

$$\underline{\beta}_u[1]^T \begin{bmatrix} Y_u[1] \\ Y_u[2] \end{bmatrix} = \begin{bmatrix} \beta_u^{(1)}[1] & \beta_u^{(2)}[1] \end{bmatrix} \begin{bmatrix} g_{us_1}[1]a_1 + g_{us_2}[1]b \\ g_{us_1}[2]a_2 + g_{us_2}[2]b \end{bmatrix}$$

For $i = 1, 2$ and any node $u \in \mathcal{V}$, $g_{us_i}[t]$ denote the end-to-end effective channel coefficient from source s_i to node u at time t, taking the linear transformations at intermediate nodes into account. $\underline{\beta}_u[1]$ denote the linear transformation at time $t = 1$.

Let us denote the received linear combination at v_2^*, $Y_{v_2^*}[1]$, by $A_2 + B_2$, where $A_2 = G^T_{v_2^* s_1} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and $B_2 = g_{v_2^* s_2} b$. Note that $g_{v_2^* s_2} \neq 0$ with high probability since s_2 can reach v_2^*.

Key Observation:

Since v_2^* is not omniscient, in layer $\mathcal{L}_{k_2^*}$ there exists a node $w \in \mathcal{C}_1$ such that in its received signal $Y_w[1] = A_w = G^T_{ws_1} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$, the column vector G_{ws_1} is not aligned to $G_{v_2^* s_1}$ with high probability. This is because either

i) $\mathcal{P}^{s_1}(v_2^*) \neq \mathcal{P}^{s_1}(w)$, or

ii) $\mathcal{P}^{s_1}(v_2^*) = \mathcal{P}^{s_1}(w)$ and $|\mathcal{P}^{s_1}(v_2^*)| \geq 2$.

In both cases, due to the independent random mixing at the parent nodes as well as the i.i.d. fading channel gains, the above claim follows straightforwardly as in the case with instantaneous CSIT [4] [5].

Phase 3: Stage $\mathcal{L}_{(k_2^*-1)}$ to $\mathcal{L}_{k_2^*}$, Time $t = 2$: Each s_2-only-reachable node in layer $\mathcal{L}_{(k_2^*-1)}$ transmits a randomly chosen scaled copy of b. Each s_1-reachable node in $\mathcal{L}_{(k_2^*-1)}$ that cannot reach v_2^* transmits a randomly chosen linear transformation of its receptions over time slots $t = 1$ and $t = 2$, as in Phase 2.

On the other hand, the s_1-reachable parents of v_2^* are those have to take care in forming their transmit linear combinations. Each of them, say, u, will combine its received linear combinations over the time slots $t = 1, 2$ so that in the combination, the contribution from source s_1 is equal to A_2. Mathematically, if $A_2 = G^T_{v_2^* s_1} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = g^{(1)}_{v_2^* s_1} a_1 + g^{(2)}_{v_2^* s_1} a_2$, then u will combine $Y_u[1] = g_{us_1}[1]a_1 + g_{us_2}[1]b$ and $Y_u[2] = g_{us_1}[2]a_2 + g_{us_2}[2]b$ in the following way:

$$X_u[2] = g^{(1)}_{v_2^* s_1} \left(Y_u[1]/g_{us_1}[1] \right) + g^{(2)}_{v_2^* s_1} \left(Y_u[2]/g_{us_1}[2] \right)$$

This operation is feasible under the CSI model, since the effective channel gains $g^{(1)}_{v_2^* s_1}$ and $g^{(2)}_{v_2^* s_1}$ (which determine the linear combination A_2) only contain past channel coefficients, and this knowledge is known to the s_1-reachable parents of v_2^* at this point.

The received signal of v_2^* at time $t = 2$ becomes a linear combination of A_2 and b, and due to the i.i.d. fading channel gains, this linear combination is not aligned to $A_2 + B_2$ with high probability. Hence v_2^* can decode b. Then within the cloud $\mathcal{C}_2$, v_2^* can deliver b to d_2 by routing.

As for destination d_1, it relies on the node w to deliver the two symbols $\{a_1, a_2\}$. Note that w at time $t = 2$ receives a scaled version of A_2 plus some other randomly formed linear combination of $\{a_1, a_2\}$ if w has an s_1-reachable parent that cannot reach v_2^*. Therefore, its received linear combination of $\{a_1, a_2\}$ at time $t = 2$ is not aligned to A_w with high probability, as A_2 is not. Combining this linear combination with its received signal at time $t = 1$, A_w, it can decode $\{a_1, a_2\}$. Hence, d_1 can receive $\{a_1, a_2\}$ from w.

Let us summarize the key elements that we develop throughout this motivating example:

- Critical nodes serve as "anchors" in the scheme. The two sources[2], parents of critical nodes, and critical nodes are the places where delayed transmitter CSI is exploited.
- Other nodes in the network simply perform random linear coding in any phases of the scheme.
- Absence of omniscient nodes guarantees the richness of received linear combinations and is the key to the proof.

V. Proof of Achievability

Without loss of generality, let us assume that $k_1^* \leq k_2^*$, that is, the critical node for user 1 lies in a layer with index smaller than that for user 2. Note that if $k_i^* = 1$, then the critical node v_i^* becomes omniscient, violating the assumption. Hence in the following we consider the case $k_i^* \neq 1$, for $i = 1, 2$. Moreover, if $k_1^* = 0$, it means that d_1 is s_1-only-reachable. If $k_2^* = 0$ too, then d_2 is also s_2-only-reachable, and the two users do not interfere with each other. Hence DoF $(d_1, d_2) = (1, 1)$ is achievable without CSIT by routing. Also, in Section IV we have already shown that $(1, 1/2)$ is achievable when $k_1^* = 0$ and $k_2^* \geq 2$. Hence, $(2/3, 2/3)$ can be achieved by time-sharing between $(1, 1/2)$ and $(0, 1)$.

What remains to be dealt with is the case $2 \leq k_1^* \leq k_2^*$. Our scheme ensures that v_i^* can decode what d_i aims to decode, $i = 1, 2$. We focus on linear schemes. The goal is to deliver a **block** of two symbols $\{a_1, a_2\}$ from s_1 to d_1 and a **block** of two symbols $\{b_1, b_2\}$ from s_2 to d_2 over three time slots. Our scheme breaks the layered network into several stages, each of which consists of several consecutive layers.

First we need to make sure that symbols $\{a_1, a_2\}$ can be retrieved by the critical node v_1^*, which can be reached by both sources. Let us begin by describing the scheme for the stage from Layer $\mathcal{L}_0$ to $\mathcal{L}_{(k_1^*-1)}$ and the stage from $\mathcal{L}_{(k_1^*-1)}$ to $\mathcal{L}_{k_1^*}$.

Phase 1: Stage $\mathcal{L}_0$ to $\mathcal{L}_{(k_1^*-1)}$, Time $t = 1, 2$: At each of the two time slots, each node in this stage (excluding $\mathcal{L}_{(k_1^*-1)}$) transmits a randomly chosen linear transformation of its reception at that time slot, with s_1 sending $\{a_1, a_2\}$ and s_2 sending $\{b_1, b_2\}$ at time $t = 1$ and $t = 2$ respectively.

Phase 2: Stage $\mathcal{L}_{(k_1^*-1)}$ to $\mathcal{L}_{k_1^*}$, Time $t = 1$: Each node in layer $\mathcal{L}_{(k_1^*-1)}$ transmits a randomly chosen linear transformation of its receptions over time slots $t = 1$ and $t = 2$.

Let us denote the received linear combination at v_1^*, $Y_{v_1^*}[1]$, by $A_1 + B_1$, where $A_1 = G^T_{v_1^* s_1} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and $B_1 = G^T_{v_1^* s_2} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$. Since v_1^* is not omniscient, there exists a node $v \in \mathcal{L}_{k_1^*}$ such that in its received signal $Y_v[1] = A_v + B_v = G^T_{vs_1} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} +$

[2]In this special case it is not necessary for the two sources to have delayed transmitter CSI. However, as shown in Section V, it is required in general.

$G^T_{\mathsf{v}\mathsf{s}_2}\begin{bmatrix}b_1\\b_2\end{bmatrix}$, the column vector $G_{\mathsf{v}\mathsf{s}_2}$ is not aligned to $G_{\mathsf{v}_1^*\mathsf{s}_2}$ with high probability.

Phase 3**: Stage** $\mathcal{L}_0$ **to** $\mathcal{L}_{(k_1^*-1)}$**, Time** $t=3$: Each node in this stage (excluding $\mathcal{L}_{(k_1^*-1)}$) transmits a randomly chosen linear transformation of its reception at time slot $t=3$, with s_1 sending some linear combination of $\{a_1,a_2\}$, A (which is not aligned to A_1), and s_2 sending B_1 respectively.

Phase 4**: Stage** $\mathcal{L}_{(k_1^*-1)}$ **to** $\mathcal{L}_{k_1^*}$**, Time** $t=2,3$: At each of the two time slots, each node in layer $\mathcal{L}_{(k_1^*-1)}$ transmits a randomly chosen linear transformation of its reception at time slot $t=3$. The received signals of v_1^* at time $t=2$ and $t=3$ are

$$Y_{\mathsf{v}_1^*}[2]=\sum_{u\in\mathcal{P}(\mathsf{v}_1^*)}h_{\mathsf{v}_1^*\mathsf{u}}[2]\beta_{\mathsf{u}}[2]\left(g_{\mathsf{u}\mathsf{s}_1}[3]A+g_{\mathsf{u}\mathsf{s}_2}[3]B_1\right)$$
$$Y_{\mathsf{v}_1^*}[3]=\sum_{u\in\mathcal{P}(\mathsf{v}_1^*)}h_{\mathsf{v}_1^*\mathsf{u}}[3]\beta_{\mathsf{u}}[3]\left(g_{\mathsf{u}\mathsf{s}_1}[3]A+g_{\mathsf{u}\mathsf{s}_2}[3]B_1\right).$$

The two linear combinations of $\{A,B_1\}$ are linearly independent with high probability. Hence v_1^* can decode $\{A,B_1\}$. Combined with $Y_{\mathsf{v}_1^*}[1]=A_1+B_1$, v_1^* can decode A_1 and hence $\{a_1,a_2\}$.

So far we have designed the scheme so that v_1^* is able to decode $\{a_1,a_2\}$ over three time slots. Next, the scheme in the later stages has to guarantee that

i) Symbols $\{a_1,a_2\}$ can be delivered to d_1 from v_1^*, and
ii) v_2^* can decode $\{b_1,b_2\}$

We shall distinguish into two cases: $k_1^*=k_2^*$ and $k_1^*<k_2^*$.

A. $2\le k_1^*=k_2^*=k^*$

In this case, the two critical nodes v_1^* and v_2^* are in the same layer $\mathcal{L}_{k^*}$. Due to the definition of critical nodes, we can conclude that $\mathcal{L}_{k^*}=\{\mathsf{v}_1^*,\mathsf{v}_2^*\}$, that is, only these two nodes are in layer $\mathcal{L}_{k^*}$.

Therefore, the node v described in the above scheme must be v_2^*. Denote the received signal at time $t=1$, $Y_{\mathsf{v}_2^*}[1]$, by A_2+B_2, where $A_2=G^T_{\mathsf{v}_2^*\mathsf{s}_1}\begin{bmatrix}a_1\\a_2\end{bmatrix}$ and $B_2=G^T_{\mathsf{v}_2^*\mathsf{s}_2}\begin{bmatrix}b_1\\b_2\end{bmatrix}$.

Since v_1^* is not omniscient, the two column vectors $G_{\mathsf{v}_1^*\mathsf{s}_2}$ and $G_{\mathsf{v}_2^*\mathsf{s}_2}$ are not aligned with high probability. Similarly, since v_2^* is not omniscient, the two column vectors $G_{\mathsf{v}_1^*\mathsf{s}_1}$ and $G_{\mathsf{v}_2^*\mathsf{s}_1}$ are not aligned either with high probability.

We choose the transmitted linear combination of s_1 at time $t=3$, A, to be A_2. Similar to the reasoning why v_1^* can decode $\{a_1,a_2\}$, v_2^* can decode $\{b_1,b_2\}$, by first decoding $\{A_2,B_1\}$ and then combining with $Y_{\mathsf{v}_2^*}[1]=A_2+B_2$ to obtain B_2.

B. $2\le k_1^*<k_2^*$

In layer $\mathcal{L}_{k_1^*}$, v_1^* is able to decode $\{a_1,a_2,B_1\}$ due to the chosen scheme. If v_1^* cannot reach v_2^*, then inside the cloud $\mathcal{C}_1$, we can deliver $\{a_1,a_2\}$ to d_1 without affecting the reception of v_2^*. If v_1^* can reach v_2^*, then the operation has to be more careful. Hence, we shall distinguish further into two cases: a) v_1^* can reach v_2^*, and b) v_1^* cannot reach v_2^*. In the latter case, a more detailed discussion is required. Due to space constraint, we leave it in the extended version of this paper [11]. Below we focus on the case when v_1^* can reach v_2^*.

In this case all nodes in $\mathcal{L}_{k_1^*}$ keep silent except v_1^* and v. Note that the node v can decode B_1. The important thing is, B_1 and B_{v} (the contribution from s_2 at time $t=1$) are not aligned. We shall make use of v_1^* as the source of providing $\{a_1,a_2\}$ and v as the source of providing $\{b_1,b_2\}$. During the first time slot, we shut user 1 off, that is, keep v_1^* silent, and let v deliver B_1 to d_2 by routing. During the second and the third time slots, we mimic the $(1,1/2)$-achievability scheme in the case $k_1^*=0,k_2^*\ge 2$ in Section IV to deliver symbols $\{a_1,a_2\}$ to d_1 and B_{v} to d_2. v_1^* takes the role of s_1, and v takes the role of s_2. The only difference is that, v may not be able to decode $\{b_1,b_2\}$. As an alternative, during Phase 1 of the scheme (described in Section IV) it sends out $A_{\mathsf{v}}+B_{\mathsf{v}}$. Note that now in Phase 2, the linear combination A_2 will contain the contribution from A_{v}, but it is still a linear combination of $\{a_1,a_2\}$. The rest of the scheme remains the same, and at the end of the three phases, d_2 can decode B_{v}. In summary, at the end of the three time slots, d_1 obtains $\{a_1,a_2\}$ and d_2 obtains $\{B_1,B_{\mathsf{v}}\}$ and hence $\{b_1,b_2\}$.

VI. Concluding Remarks

Though absence of an omniscient node enables performance better than time-sharing, we still do not have a conclusive DoF characterization of arbitrary two-unicast networks with delayed CSIT. This is part of an ongoing study, where we are exploring better outer bounds as well as new coding ideas.

References

[1] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Transactions on Information Theory*, vol. 57, pp. 1872–1905, April 2011.

[2] S. Mohajer, S. N. Diggavi, C. Fragouli, and D. N. C. Tse, "Approximate capacity of a class of Gaussian interference-relay networks," *IEEE Transactions on Information Theory*, vol. 57, pp. 2837–2864, May 2011.

[3] T. Gou, S. A. Jafar, S.-W. Jeon, and S.-Y. Chung, "Aligned interference neutralization and the degrees of freedom of the 2 × 2 × 2 interference channel," *Arxiv Preprint*, December 2010, http://arxiv.org/abs/1012.2350.

[4] I.-H. Wang, S. U. Kamath, and D. N. C. Tse, "Two unicast information flows over linear deterministic networks," *Proceedings of IEEE International Symposium on Information Theory*, pp. 2464–2468, August 2011. Extended version: http://arxiv.org/abs/1105.6326.

[5] I. Shomorony and A. S. Avestimehr, "Two-unicast wireless networks: Characterizing the degrees-of-freedom," *Submitted to IEEE Transactions on Information Theory*, March 2011, http://arxiv.org/abs/1102.2498.

[6] M. A. Maddah-Ali and D. N. C. Tse, "Completely stale transmitter channel state information is still very useful," *Submitted to IEEE Transactions on Information Theory*, September 2011, http://arxiv.org/abs/1010.1499.

[7] H. Maleki, S. A. Jafar, and S. Shamai, "Retrospective interference alignment," *Arxiv Preprint*, September 2010, http://arxiv.org/abs/1009.3593.

[8] C. S. Vaze and M. K. Varanasi, "The degrees of freedom region and interference alignment for the MIMO interference channel with delayed CSI," *Submitted to IEEE Transactions on Information Theory*, March 2011, http://arxiv.org/abs/1101.5809.

[9] R. Tandon, S. Mohajer, H. V. Poor, and S. Shamai, "On interference networks with feedback and delayed CSI," *Submitted to IEEE Transactions on Information Theory*, September 2011, http://arxiv.org/abs/1109.5373.

[10] C. S. Vaze and M. K. Varanasi, "The degrees of freedom of the 2×2×2 interference network with delayed csit and with limited shannon feedback," *Proceedings of Allerton Conference on Communication, Control, and Computing*, pp. 824–831, September 2011.

[11] I.-H. Wang and S. N. Diggavi, "On degrees of freedom of layered two unicast networks with delayed CSIT," *Preprint*, February 2012. Link available at http://sites.google.com/site/ihsiangw/publication.