Sum Degrees-of-Freedom of Two-Unicast Wireless Networks

Ilan Shomorony and A. Salman Avestimehr
School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA
is256@cornell.edu avestimehr@ece.cornell.edu

***Abstract*—We consider two-source two-destination (i.e., two-unicast) multi-hop wireless networks that have a layered structure with arbitrary connectivity. We show that, if the channel gains are independently drawn from continuous distributions, then, with probability 1, two-unicast layered Gaussian networks can only have 1, $3/2$ or 2 sum degrees-of-freedom[1]. We provide necessary and sufficient conditions for each case based on the network connectivity and a new notion of source-destination paths with manageable interference.**

I. Introduction

Characterizing network capacity is one of the central problems in network information theory. While this problem is in general unsolved, there has been considerable success in two research fronts. The first one focuses on single-flow multi-hop networks, in which one source sends the same message to one or more destinations. In this scenario, all destination nodes require the same message, and there is effectively only one information stream in the network. Starting from the max-flow min-cut theorem of Ford-Fulkerson [1], there has been significant progress on this problem. For wireline networks, the maximum multicast flow was characterized in [2] using random coding and in [3, 4] using linear network coding. In [5], the max-flow min-cut theorem was generalized for a class of linear deterministic networks with broadcast and interference. Inspired by this generalization, the multicast capacity of wireless networks was then characterized to within a gap that does not depend on the channel gains ([5]).

The second research direction focuses on single-hop multi-flow wireless networks, i.e., the interference channel (IFC). While the capacity of the IFC remains unknown (except in special cases, such as [6–9]), several approximations have been derived, such as constant-gap capacity approximations [10, 11] and degrees-of-freedom (DoF) characterizations ([12–14]).

However, once we go beyond single-hop, much less is known about the capacity of multi-flow networks. Even for two-source two-destination networks there are few general results, such as [15], where the maximum flow in two-unicast undirected wireline networks is characterized. For two-unicast directed wireline networks, [16–18] provided graph-theoretic conditions under which rate $(1,1)$ can be achieved. In the wireless realm, constant-gap capacity approximations for certain two-hop networks were obtained in [19]. In [20], it was shown that the network resulting from the concatenation of two or more *fully connected* IFCs admits two DoF.

In this paper, we consider two-unicast multi-hop wireless networks that have a layered structure with *arbitrary connectivity*. We consider an AWGN channel model and assume that the channel gains are independently drawn from continuous distributions and remain fixed during the course of communication. Moreover, we assume that all channel gains are known at all nodes. Under these assumptions, we show that, with probability 1 over the choice of the channel gains, two-unicast layered Gaussian networks can only have 1, $3/2$ or 2 sum DoF. Moreover, we provide necessary and sufficient conditions for each case that are based only on properties of the network graph. We state our main result in Section II and describe its proof in Sections III and IV. Due to space limitations, we omit some proof details and refer to [21] for complete proofs.

II. Definitions and Main Result

A *multiple-unicast Gaussian network* $\mathcal{N} = (G, L)$ consists of a directed graph $G = (V, E)$, where V is the node set and $E \subset V \times V$ is the edge set, and a set of source-destination pairs $L \subset V \times V$. We consider two-unicast Gaussian networks, i.e., $L = \{(s_1, d_1), (s_2, d_2)\}$, for distinct $s_1, s_2, d_1, d_2 \in V$. We will assume that the network is *layered*, i.e., the node set V can be partitioned into r subsets $V_1, V_2, ..., V_r$ (the layers) in such a way that $E \subset \bigcup_{i=1}^{r-1} V_i \times V_{i+1}$, and $V_1 = \{s_1, s_2\}$, $V_r = \{d_1, d_2\}$. For $v \in V_j$, we let $\mathcal{I}(v) \triangleq \{u \in V_{j-1} : (u, v) \in E\}$ and $\mathcal{O}(v) \triangleq \{u \in V_{j+1} : (v, u) \in E\}$. Furthermore, we let $\ell(v)$ be the index of the layer containing v, i.e., $v \in V_{\ell(v)}$. Notice that the layers induce a natural ordering of the nodes.

For each edge $e = (v_i, v_j)$, we associate a real-valued channel gain h_e (or simply $h_{i,j}$). We will assume that the h_e's are independently drawn from continuous distributions and are fixed during the course of communication. We also assume that all h_e's are fully known at all nodes. At time m, each $v_i \in V \setminus \{d_1, d_2\}$ transmits a real signal $X_{v_i}[m]$ (or $X_i[m]$), which must satisfy an average power constraint P. The signal received by $v_j \in V \setminus \{s_1, s_2\}$ at time m is

$$Y_j[m] = \sum_{v_i \in \mathcal{I}(v_j)} h_{i,j} X_i[m] + N_j[m], \text{ for } m = 1, 2, \dots ,$$

where $N_j[m]$ is the zero-mean unit-variance Gaussian discrete-time noise process at v_j. The transmitted signal from $v_j \in V \setminus \{s_1, s_2\}$ at time m must be a function of its past received signals $Y_j[k]$, for $k = 1, ..., m-1$. Source s_i has a message W_i

[1]Unless the source-destination pairs are disconnected, in which case no degrees-of-freedom can be achieved

that it wishes to send to d_i, and encodes it into transmit signals $X_{s_i}[m]$, $m = 1, ..., n$, for $i = 1, 2$, for a communication session of duration n. We say that rates $R_i \triangleq \frac{\log|W_i|}{n}$ for $i = 1, 2$ are achievable if the probability of error in the decoding of both messages by their destinations can be made arbitrarily small by choosing a sufficiently large n. The sum-capacity $C_\Sigma(P)$ is the supremum of the achievable sum-rates.

Definition 1. *The sum degrees-of-freedom d_Σ is defined as*

$$d_\Sigma \triangleq \lim_{P\to\infty} \frac{C_\Sigma(P)}{\frac{1}{2}\log P}.$$

Definition 2. *A path P_{v_1,v_k} from v_1 to v_k is an ordered set $\{v_1, v_2, ..., v_k\} \subset V$ such that $(v_i, v_{i+1}) \in E$ for $i = 1, ..., k-1$. We write $v_1 \rightsquigarrow v_k$, if there is a path P_{v_1,v_k}.*

Definition 3. *Paths P_{v_a,v_b} and P_{v_c,v_d} are said to be disjoint if $P_{v_a,v_b} \cap P_{v_c,v_d} = \emptyset$.*

Definition 4. *For $S \subset V$, we say that $G[S]$ is the graph induced by S on G, if $G[S] = (S, E_s)$, where $E_s = \{(v_i, v_j) \in E : v_i, v_j \in S\}$.*

Definition 5. *$\mathcal{N}' = (G', L)$ is a subnetwork of $\mathcal{N} = (G, L)$, if $G' = G[S]$, for some $S \subset V$ such that $L \subset S \times S$.*

Next, we assume that we have two disjoint paths P_{s_1,d_1} and P_{s_2,d_2}. We let $\bar{i} = 2$ if $i = 1$ and $\bar{i} = 1$ if $i = 2$.

Definition 6. *We say that $v_a \notin P_{s_i,d_i}$ causes interference on P_{s_i,d_i} and write $v_a \overset{I}{\rightsquigarrow} P_{s_i,d_i}$ if we can find $v_b \in P_{s_i,d_i}$ and a path $P_{s_{\bar{i}},v_a}$ such that $(v_a, v_b) \in E$ and $P_{s_{\bar{i}},v_a} \cap P_{s_i,d_i} = \emptyset$, for $i = 1$ or 2. We write $v_a \xrightarrow{I} P_{s_i,d_i}$, if, in addition, $v_a \in P_{s_{\bar{i}},d_{\bar{i}}}$.*

Consider a subnetwork $(G[S], \{(s_1, d_1), (s_2, d_2)\})$ for $S \supset P_{s_1,d_1} \cup P_{s_2,d_2}$. Let $n_i(G[S], P_{s_i,d_i}) \triangleq |\{v \in S : v \overset{I}{\rightsquigarrow} P_{s_i,d_i}\}|$ and $n_i^D(P_{s_{\bar{i}},d_{\bar{i}}}, P_{s_i,d_i}) \triangleq |\{v \in V : v \xrightarrow{I} P_{s_i,d_i}\}|$, for $i = 1, 2$. Notice that the path implied by $v \overset{I}{\rightsquigarrow} P_{s_i,d_i}$ must exist in the subnetwork. If there is no ambiguity in the choice of P_{s_1,d_1} and P_{s_2,d_2}, we will simply use $n_i(G[S])$ and n_i^D.

Definition 7. *Two disjoint paths P_{s_1,d_1} and P_{s_2,d_2} have manageable interference if we can find $S \subset V$ so that $P_{s_1,d_1}, P_{s_2,d_2} \subset S$, $n_1(G[S]) \neq 1$ and $n_2(G[S]) \neq 1$.*

Theorem 1. *For a two-unicast layered Gaussian network $\mathcal{N} = (G = (V, E), \{(s_1, d_1), (s_2, d_2)\})$ where the channel gains are chosen according to independent continuous distributions, with probability 1, d_Σ is given by*

A) $d_\Sigma = 1$ *if $\mathcal{N}$ contains a node v whose removal disconnects d_i from $\{s_1, s_2\}$ and $s_{\bar{i}}$ from $\{d_1, d_2\}$, for $i = 1$ or 2,*

A′) $d_\Sigma = 1$ *if $\mathcal{N}$ contains an edge (v_2, v_1) such that the removal of v_1 disconnects d_i from $\{s_1, s_2\}$ and the removal of v_2 disconnects $s_{\bar{i}}$ from $\{d_1, d_2\}$, for $i = 1$ or 2,*

B) $d_\Sigma = 2$ *if $\mathcal{N}$ contains two disjoint paths P_{s_1,d_1} and P_{s_2,d_2} with manageable interference (see Definition 7),*

B′) $d_\Sigma = 2$ *if $\mathcal{N}$ or any subnetwork does not contain two disjoint paths P_{s_1,d_1} and P_{s_2,d_2}, but is not in case (A),*

C) $d_\Sigma = \frac{3}{2}$ *in all other cases.*

III. Networks with two degrees-of-freedom

In this section, we describe achievability schemes for networks in cases (B) and (B′). First, we will identify *key layers*, whose nodes will perform non-trivial relaying operations. All nodes which do not belong to key layers will forward their received signal. This allows us to build a *condensed network* $\mathcal{N}_c$, which only contains the key layers, V_1 and V_r. The connectivity and channel gains are determined according to the effective transfer matrices between consecutive layers of $\mathcal{N}_c$. An example is shown in Figure 1. We refer to the channel

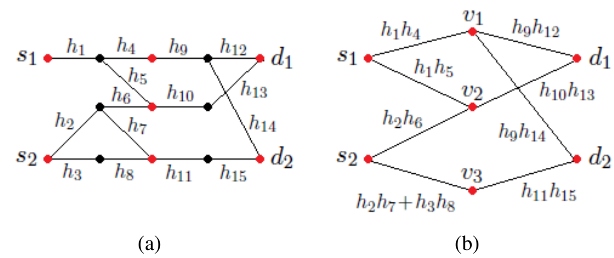


Fig. 1. A 5-layer network (a) and its 3-layer condensed version (b)

gain of edge (v, u) from $\mathcal{N}_c$ by $\hat{h}(v, u)$.

We use two types of transmission strategies, according to the structure of the condensed network. If the condensed network is a $2\times2\times2$ interference channel, we use the scheme described in [20] to achieve $d_\Sigma = 2$. Otherwise, we describe an amplify-and-forward scheme that guarantees that the end-to-end transfer matrix for the condensed network (and also for the original network) is $\begin{bmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{bmatrix}$, for $\beta_1, \beta_2 \neq 0$. Thus we have $Y_{d_i} = \beta_i X_{s_i} + N_{d_i}^{\text{eff}}$, for $i = 1, 2$, where $N_{d_i}^{\text{eff}}$ is the effective additive noise at d_i, and we have two AWGN channels. To satisfy the power constraint at all nodes, we restrict the sources to using power αP, for $\alpha \in (0, 1)$. It can be seen that, for P sufficiently large, α can be chosen independent of P. Since the scaling factors used at the key layers and the variance of $N_{d_i}^{\text{eff}}$, σ_i^2, are functions of the channel gains only (and not P), source-destination pair (s_i, d_i), for $i = 1, 2$, can achieve rate $R_i = \frac{1}{2}\log\left(1 + \frac{\alpha\beta_i^2 P}{\sigma_i^2}\right)$, and, therefore, we achieve $d_\Sigma = 2$.

First, we consider (B). Thus we have disjoint paths P_{s_1,d_1} and P_{s_2,d_2} with manageable interference, i.e., $\exists S \subset V$ such that $P_{s_1,d_1} \cup P_{s_2,d_2} \subset S$, $n_1(G[S]) \neq 1$ and $n_2(G[S]) \neq 1$. We assume S is minimal, and that all nodes in $V \setminus S$ are removed. If $n_1(G[S]) = 0$ and $n_2(G[S]) = 0$, then achieving $d_\Sigma = 2$ is trivial, since there is no interference. If $n_i(G[S]) \geq 2$, for $i = 1$ or 2, we let v_p^i be the first node in P_{s_i,d_i} whose removal disconnects d_i and $s_{\bar{i}}$. If $n_i(G[S]) \geq 2$, $V_{\ell(v_p^i)-1}$ is the last layer where we can choose the scaling factors to cancel the interference on P_{s_i,d_i}, and it will be a key layer. In the following Lemmas, it is assumed that $n_i(G[S]) \geq 2$, for $i = 1$ or 2, and thus v_p^i is defined. The proofs are in Appendix A.

Lemma 1. *There exist two paths P_{s_1,v_p^i} and P_{s_2,v_p^i} such that $P_{s_1,v_p^i} \cap P_{s_2,v_p^i} = \{v_p^i\}$.*

Lemma 2. *There are (at least) two nodes $v_1, v_2 \in \mathcal{I}(v_p^i)$ such that $s_{\bar{i}} \rightsquigarrow v_1$ and $s_{\bar{i}} \rightsquigarrow v_2$.*

We describe the achievability scheme when $n_1(G[S]) \geq 2$ and $n_2(G[S]) = 0$, and thus only v_p^1 is defined. The case where $n_1(G[S]) \geq 2$ and $n_2(G[S]) \geq 2$ is considered in [21]. Our condensed network is formed by layers $V_1, V_{\ell(v_p^1)-1}, V_r$, with $m = |V_{\ell(v_p^1)-1}|$ (see Figure 2(a)). To each $v_i \in V_{\ell(v_p^1)-1}$, $i = 1, ..., m$, we associate a scaling factor x_i. We must show that the end-to-end transfer matrix, given by

$$T = \begin{bmatrix} \sum_{i=1}^m \hat{h}(s_1, v_i)\hat{h}(v_i, d_1)x_i & \sum_{i=1}^m \hat{h}(s_2, v_i)\hat{h}(v_i, d_1)x_i \\ \sum_{i=1}^m \hat{h}(s_1, v_i)\hat{h}(v_i, d_2)x_i & \sum_{i=1}^m \hat{h}(s_2, v_i)\hat{h}(v_i, d_2)x_i \end{bmatrix},$$

can be made diagonal with non-zero diagonal entries by an appropriate choice of $x_1, ..., x_m$. Since, in this case, $n_2(G[S]) = 0$, there is no path from s_1 to d_2, and therefore we must have $\hat{h}(s_1, v_i)\hat{h}(v_i, d_2) = 0$ for $i = 1, ..., m$ and $T_{2,1}$ (the bottom left entry in T) is always 0. From Lemma 1, we can find two nodes $v_a, v_b \in \mathcal{I}(v_p^1) \subset V_{\ell(v_p^1)-1}$ with associated variables x_a and x_b, and two disjoint paths P_{s_1,v_a} and P_{s_2,v_b}. From Lemma 2, we can find $v_c \in \mathcal{I}(v_p^1) \subset V_{\ell(v_p^1)-1}$, such that $s_2 \rightsquigarrow v_c$ and $c \neq m$. We now claim that if the matrices

$$M_1 = \begin{bmatrix} \hat{h}(s_1, v_a)\hat{h}(v_a, d_1) & \hat{h}(s_1, v_b)\hat{h}(v_b, d_1) \\ \hat{h}(s_2, v_a)\hat{h}(v_a, d_1) & \hat{h}(s_2, v_b)\hat{h}(v_b, d_1) \end{bmatrix} \text{ and}$$

$$M_2 = \begin{bmatrix} \hat{h}(s_2, v_c)\hat{h}(v_c, d_1) & \hat{h}(s_2, v_m)\hat{h}(v_m, d_1) \\ \hat{h}(s_2, v_c)\hat{h}(v_c, d_2) & \hat{h}(s_2, v_m)\hat{h}(v_m, d_2) \end{bmatrix}$$

are both full-rank, then we can choose $x_1, ..., x_m$ so that T is as desired. To see this, consider $\mathbf{x}' = [x_1' \ ... \ x_m']$, where $x_j' = 0$ for $j \neq a, b$, and $[x_a' \ x_b']^T = M_1^{-1}[1 \ 0]^T$. This choice of scaling factors results in $T_{1,1} = 1$ and $T_{1,2} = 0$. If $T_{2,2} \neq 0$ we are done. Otherwise, if $T_{2,2} = 0$, we let $\mathbf{x}'' = [x_1'' \ ... \ x_m'']$, where $x_j'' = 0$ for $j \neq c, m$ and $[x_c'' \ x_m'']^T = M_2^{-1}[0 \ 1]^T$. This choice results in $T_{1,2} = 0$ and $T_{2,2} = 1$. If we have $T_{1,1} \neq 0$, we are done. Otherwise, we set $\mathbf{x}''' = \mathbf{x}' + \mathbf{x}''$. By linearity, this choice will guarantee that T is the identity matrix.

Next we show that, with probability 1 over the choice of the h_e's, M_1 and M_2 are full-rank. First we consider the transfer matrix between (s_1, s_2) and (v_a, v_b), given by

$$Z_1 = \begin{bmatrix} \hat{h}(s_1, v_a) & \hat{h}(s_2, v_a) \\ \hat{h}(s_1, v_b) & \hat{h}(s_2, v_b) \end{bmatrix}.$$

The determinant of Z_1 can be seen as a polynomial in the channel gains h_e. If $\det Z_1$ is not identically zero, since the h_e's are drawn independently from continuous distributions, $\det Z_1$ will be non-zero w.p. 1. To see that $\det Z_1$ is not identically zero, notice that the existence of disjoint paths P_{s_1,v_a} and P_{s_2,v_b} guarantees that, if we set $h_e = 1$ if e connects adjacent nodes of P_{s_1,v_a} or P_{s_2,v_b} and $h_e = 0$ otherwise, Z_1 will be the identity matrix. Therefore, Z_1 will be invertible, and $\det Z_1$ cannot be identically zero. Now, we notice that $\det M_1 = \hat{h}(v_a, d_1)\hat{h}(v_b, d_1) \det Z_1$. Since $v_a \rightsquigarrow d_1$ and $v_b \rightsquigarrow d_1$, we have that $\hat{h}(v_a, d_1)\hat{h}(v_b, d_1)$ is also a non-identically zero polynomial in the h_e's, and therefore M_1 is invertible w.p. 1. To show that M_2 is invertible w.p. 1, we follow very similar steps, by noticing that the transfer matrix between $\{v_c, v_m\}$ and $\{d_1, d_2\}$ is full-rank w.p. 1, since we have disjoint P_{v_c,d_1} and P_{v_m,d_2}.

The proof for $n_1(G[S]) \geq 2$, $n_2(G[S]) \geq 2$ follows similar steps, except if our condensed network is a $2 \times 2 \times 2$ interference channel, in which case we apply the real interference alignment scheme described in [20]. If our network $\mathcal{N}$ is in case (B′), we proceed as follows. We use a result provided in [17, 18] to claim that if $\mathcal{N}$ is not in case (A), then it contains two disjoint paths P_{s_1,d_1} and P_{s_2,d_2}, a Butterfly subnetwork or a Grail subnetwork (see Appendix B). Since we have a subnetwork with no two disjoint paths which is not in case (A), we must have a Butterfly or a Grail network. In each case we identify key layers and provide schemes as the one above to achieve $d_\Sigma = 2$. More details are provided in Appendix B.

IV. Networks with 3/2 degrees-of-freedom

In this section, we show that if our network $\mathcal{N}$ is not in cases (A), (A′), (B) and (B′), then $d_\Sigma = \frac{3}{2}$. We start by defining two main categories of networks in (C). If $\mathcal{N}$ is not in (A) nor (B′), then it is easy to see that it must contain two disjoint paths P_{s_1,d_1} and P_{s_2,d_2}. Thus, we assume we have disjoint paths P_{s_1,d_1} and P_{s_2,d_2} that do not have manageable interference (or we would be in (B)). It can then be shown that $\mathcal{N}$ may be assumed to be in one of two cases (see [21], for a proof):

C1. $n_1(G) \geq 2$, $n_1^D = 1$, $n_2(G) = 1$ and $n_2^D = 0$.

C2. $n_1(G) = n_1^D = 1$

Next we consider networks in case C1. Case C2 is addressed in Appendix C. Notice that, in case C1, we must have a node $v_1 \notin P_{s_1,d_1} \cup P_{s_2,d_2}$ such that $v_1 \overset{I}{\rightsquigarrow} P_{s_1,d_1}$ and thus we have a path P_{s_2,v_1} disjoint from P_{s_1,d_1}. We let v_m be the last node in $P_{s_2,d_2} \cap P_{s_2,v_1}$, and we have a path P_{v_m,v_1}. If we let $S^* = P_{s_1,d_1} \cup P_{s_2,d_2} \cup P_{v_m,v_1}$, we have $n_1(G[S^*]) \geq 2$. Since P_{s_1,d_1} and P_{s_2,d_2} do not have manageable interference, we must have $n_2(G[S^*]) = 1$. Since $n_2^D = 0$, we conclude that we must have $v_2 \in P_{v_m,v_1} \setminus \{v_m\}$ such that $v_2 \overset{I}{\rightsquigarrow} P_{s_2,d_2}$, and we must have a path $P_{s_1,v_2} \subset S^*$. Thus, we have the subnetwork in Figure 2(b) up to a change in the position of (v_3, v_4).

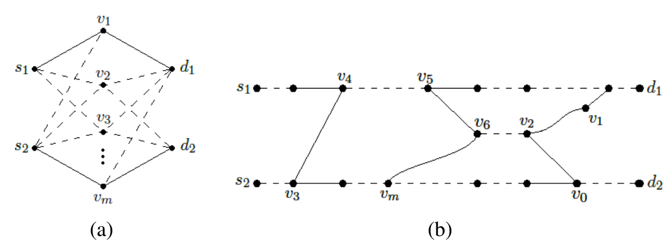


Fig. 2. (a) Illustration of a condensed network with $n_1(G[S])$ and $n_2(G[S]) = 0$. Solid lines represent edges that must exist in the condensed network, while the dashed lines represent edges that may not exist.; (b) Illustration of the network in case C1. The curvy lines and the dashed lines indicate paths (which may be composed by a single edge).

To achieve 3/2 DoF, we use a scheme based on two distinct modes of operation for the network, as illustrated in Figure 3. During Mode 1, we let an intermediate node function as a destination d_2'. Notice that we have disjoint paths P_{s_1,d_1} and $P_{s_2,d_2'}$ with manageable interference. In Mode 2, d_2' becomes a source s_2', and we again have disjoint paths P_{s_1,d_1} and P_{s_2',d_2} with manageable interference. Therefore, if $d_2' = s_2'$ stores the received signals during Mode 1, and forwards them during Mode 2, we can achieve $\frac{3}{2}$ DoF. See [21] for details.

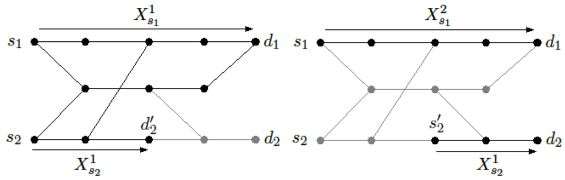


Fig. 3. Depiction of Modes 1 and 2 for the achievability scheme in case C1.

To prove the converse, we name additional nodes as shown in Figure 2(b). We let $v_0 \in P_{s_2,d_2}$ be such that $(v_2, v_0) \in E$. From our previous discussion, there is a path $P_{s_1,v_2} \subset S^*$. We let v_5 be the last node in $P_{s_1,d_1} \cap P_{s_1,v_2}$, and v_6 its consecutive node on P_{s_1,v_2}. To derive the converse inequalities, we consider a decomposition of the unit-variance Gaussian noise N_j at each node v_j into m independent components with variance $1/m$, where $m = |\mathcal{I}(v_j)|$. We associate each component with an incoming edge, and we define, for $v_i \in \mathcal{I}(v_j)$,

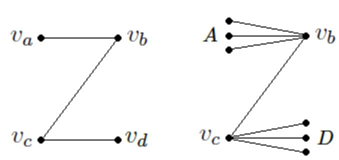
$$\tilde{X}_{i,j} \triangleq h_{i,j} X_i + N_{i,j},$$

where $N_{i,j}$ is the noise term associated with the edge (v_i, v_j). Notice that $N_j = \sum_{i: v_i \in \mathcal{I}(v_j)} N_{i,j}$. We can now write, for a node v_j, $Y_j = \sum_{i: v_i \in \mathcal{I}(v_j)} \tilde{X}_{i,j}$. We also define

$$\tilde{X}_i \triangleq \{\tilde{X}_{i,j} : j \text{ s.t. } v_j \in \mathcal{O}(v_i)\}.$$

We let X_S be the set of all X_i's , for $v_i \in S$, and X_i^n be a length n vector whose entries are the $X_i[m]$'s , for $m = 1, ..., n$. Analogous definitions hold for $\tilde{X}_S$, $\tilde{X}_i^n$, Y_j^n and N_j^n.

Next, we notice that, if we have a Z structure across two layers in the network, as shown in Figure 4(a), then, given $\tilde{X}_a^n$ and Y_b^n, one can subtract $\tilde{X}_{a,b}^n$ from Y_b^n and obtain $\tilde{X}_{c,b}^n$. Therefore, "almost all" information in $\tilde{X}_c^n$ can be deduced from $(Y_b^n, \tilde{X}_a^n)$, and the conditional mutual information $I(X_c^n; \tilde{X}_c^n | Y_b^n, \tilde{X}_a^n)$ cannot be very large. The next Lemma generalizes this notion to the structure in Figure 4(b). The proof is found in Appendix A.



(a) (b)

Fig. 4. The Z structure.

Lemma 3. *Suppose we have nodes v_b and v_c such that $(v_c, v_b) \in E$. Suppose, in addition, that we have sets $A, S \subset V$ such that $\mathcal{I}(v_b) \setminus \{v_c\} \subset A$ and for no $u \in S \cup A$ we have $v_c \rightsquigarrow u$. Then, we have*

$$I(X_S^n; \tilde{X}_c^n | Y_b^n, \tilde{X}_A^n) \le nK,$$

where K is only a function of the h_e's and the network $\mathcal{N}$.

Since there are no two disjoint paths with manageable interference, we can infer the following properties about the network in C1 (Figure 2(b)). The proofs are found in [21].

P1. All paths from s_1 to d_2 contain $\{v_2, v_0\}$
P2. All paths from s_1 to d_2 contain $\{v_5, v_6\}$
P3. All paths from s_2 to d_1 contain $\{v_6, v_2\}$ or $\{v_3, v_4\}$
P4. The removal of v_0 disconnects d_2 from $\{s_1, s_2\}$
P5. The removal of v_5 disconnects s_1 from $\{d_1, d_2\}$
P6. The removal of $\{v_2, v_3\}$ disconnects d_2 from $\{s_1, s_2\}$
P7. The removal of $\{v_2, v_4\}$ disconnects d_1 from $\{s_1, s_2\}$
P8. All paths from s_1 or s_2 to v_2 contain v_6

These properties allow us to derive the inequalities that will build the converse proof. Let $A = \{v \in V : s_2 \not\rightsquigarrow v\}$ and $B = \{v \in V : s_1 \not\rightsquigarrow v\}$. We let W_1 and W_2 be independent random variables corresponding to a uniform choice over the messages from s_1 and s_2 respectively. Then we have

$$\begin{aligned} nR_2 &= H(W_2) = I(W_2; Y_{d_2}^n) + H(W_1 | Y_{d_2}^n) \\ &\overset{(i)}{\le} I(W_2; Y_{d_2}^n) + \epsilon_n \overset{(ii)}{\le} I(\tilde{X}_B^n; Y_0^n) + \epsilon_n \\ &= I(X_2^n, \tilde{X}_B^n; Y_0^n) - I(X_2^n; Y_0^n | \tilde{X}_B^n) + \epsilon_n \\ &\overset{(iii)}{\le} \frac{n}{2} \log P + nK_1 - I(X_2; Y_0 | \tilde{X}_B^n) + \epsilon_n, \end{aligned} \tag{1}$$

where (i) follows from Fano's inequality, and $\epsilon_n \to 0$ as $n \to \infty$; (ii) follows since $W_2 \leftrightarrow \tilde{X}_B^n \leftrightarrow Y_0^n \leftrightarrow Y_{d_2}^n$, which is implied by P4 and the fact that $s_2 \in B$; (iii) follows since $\mathcal{I}(v_0) \setminus \{v_2\} \subset B$ (from P1) and $v_2 \notin B$; thus

$$\begin{aligned} I(X_2^n, \tilde{X}_B^n; Y_0^n) &= h(Y_0^n) - h(Y_0^n | \tilde{X}_B^n, X_2^n) \\ &= h(Y_0^n) - h(N_{2,0}^n) \\ &\le \frac{n}{2} \log \left(\frac{1 + \left(\sum_{u \in \mathcal{I}(v_0)} |h_{u,v_0}| \right)^2 P}{2\pi e / |\mathcal{I}(v_0)|} \right) \\ &\le \frac{n}{2} \log(\gamma P) \le \frac{n}{2} \log P + nK_1, \end{aligned} \tag{2}$$

where γ and K_1 are constants that are independent of P, for sufficiently large P. We also have that

$$\begin{aligned} nR_1 &\le I(W_1; Y_{d_1}^n) + \epsilon_n \overset{(i)}{\le} I(W_1; \tilde{X}_5^n, \tilde{X}_B^n) + \epsilon_n \\ &\overset{(ii)}{=} I(W_1; \tilde{X}_5^n | \tilde{X}_B^n) + \epsilon_n \overset{(iii)}{\le} I(X_5^n; \tilde{X}_5^n | \tilde{X}_B^n) + \epsilon_n \\ &\le I(X_5^n; Y_6^n | \tilde{X}_B^n) + I(X_5^n; \tilde{X}_5^n | \tilde{X}_B^n, Y_6^n) + \epsilon_n \\ &\overset{(iv)}{=} I(X_5^n; Y_6^n | \tilde{X}_B^n) + nK_2 + \epsilon_n, \end{aligned} \tag{3}$$

where (i) follows because, from P5, the removal of v_5 and B disconnects d_1 from $\{s_1, s_2\}$ and thus $W_1 \leftrightarrow (\tilde{X}_5^n, \tilde{X}_B^n) \leftrightarrow Y_{d_1}^n$; (ii) follows since $\tilde{X}_B \perp\!\!\!\perp W_1$; (iii) follows from the fact that, given $\tilde{X}_B^n$, we have $W_1 \leftrightarrow X_5^n \leftrightarrow \tilde{X}_5^n$; (iv) follows from Lemma 3, since P2 implies $\mathcal{I}(v_6) \setminus \{v_5\} \subset B$. For the next inequalities, we assume $\ell(v_4) \le \ell(v_5)$. Similar inequalities are derived in [21] for the case when $\ell(v_4) > \ell(v_5)$. Then we have

$$
\begin{aligned}
&nR_2 \leq I(W_2; Y_{d_2}^n) + \epsilon_n \\
&\stackrel{(i)}{\leq} I(X_{s_2}^n; \tilde{X}_2^n, \tilde{X}_3^n) + \epsilon_n \stackrel{(ii)}{\leq} I(X_{s_2}^n; \tilde{X}_2^n, \tilde{X}_3^n | \tilde{X}_A^n) + \epsilon_n \\
&\leq I(X_{s_2}^n; \tilde{X}_3^n, Y_4^n | \tilde{X}_A^n) + I(X_{s_2}^n; \tilde{X}_2^n | \tilde{X}_A^n, \tilde{X}_3^n, Y_4^n) + \epsilon_n \\
&\stackrel{(iii)}{\leq} I(X_{s_2}^n; Y_4^n | \tilde{X}_A^n) + nK_3 + I(X_{s_2}^n, \tilde{X}_3^n; \tilde{X}_2^n | \tilde{X}_A^n, Y_4^n) + \epsilon_n \\
&\stackrel{(iv)}{\leq} I(X_B^n; Y_4^n | \tilde{X}_A^n) + I(X_{s_2}^n, \tilde{X}_3^n; \tilde{X}_2^n | \tilde{X}_A^n, Y_4^n) + nK_3 + \epsilon_n \\
&\leq I(X_B^n; Y_4^n | \tilde{X}_A^n) + I(X_B^n; \tilde{X}_2^n | \tilde{X}_A^n, Y_4^n) \\
&\qquad + I(X_{s_2}^n, \tilde{X}_3^n; \tilde{X}_2^n | \tilde{X}_A^n, Y_4^n, X_B^n) + nK_3 + \epsilon_n \\
&\stackrel{(v)}{\leq} I(X_B^n; Y_4^n | \tilde{X}_A^n) + I(X_B^n; \tilde{X}_2^n | \tilde{X}_A^n, Y_4^n) + nK_3 + \epsilon_n \\
&\leq I(X_B^n; Y_4^n, \tilde{X}_2^n | \tilde{X}_A^n) + nK_3 + \epsilon_n, \qquad (4)
\end{aligned}
$$

where (i) follows because P6 implies $W_2 \leftrightarrow X_{s_2}^n \leftrightarrow (\tilde{X}_2^n, \tilde{X}_3^n) \leftrightarrow Y_{d_2}^n$; (ii) follows since $\tilde{X}_A^n \perp\!\!\!\perp X_{s_2}^n$; (iii) follows by applying Lemma 3 to $I(X_{s_2}^n; \tilde{X}_3^n | \tilde{X}_A^n, Y_4^n)$, since $\ell(v_4) \leq \ell(v_5)$ implies $\mathcal{I}(v_4) \setminus \{v_3\} \subset A$, or else we contradict P3; (iv) follows since $s_2 \in B$; and (v) follows because we have $(X_{s_2}^n, \tilde{X}_3^n) \leftrightarrow (\tilde{X}_A^n, Y_4^n, X_B^n) \leftrightarrow \tilde{X}_2^n$, since the removal of A, v_4 and B disconnects s_2 and $\mathcal{O}(v_3)$ from v_2. This is seen as follows. From P8, all paths from $\{s_2, v_3\}$ to v_2 must contain a node in $\mathcal{I}(v_6)$. From P2, we have $\mathcal{I}(v_6) \setminus \{v_5\} \subset B$. From P3, we know that any path from $\{v_3, s_2\}$ to v_5 must contain v_4. Since $\ell(v_4) < \ell(v_6)$, we have that $v_3 \notin \mathcal{I}(v_6)$; thus, any path from s_2 or $\mathcal{O}(v_3)$ to v_2 must either contain v_4 or a node in B. Notice that we considered $\mathcal{O}(v_3)$ instead of v_3, because we have $\tilde{X}_3^n$, and not X_3^n. Next, we have that

$$
\begin{aligned}
nR_1 &\leq I(W_1; Y_{d_1}^n) + \epsilon_n \stackrel{(i)}{\leq} I(W_1; Y_4^n, \tilde{X}_2^n) + \epsilon_n \\
&\stackrel{(ii)}{\leq} I(\tilde{X}_A^n; Y_4^n, \tilde{X}_2^n) + \epsilon_n \\
&= I(\tilde{X}_A^n, X_B^n; Y_4^n, \tilde{X}_2^n) - I(X_B^n; Y_4^n, \tilde{X}_2^n | \tilde{X}_A^n) + \epsilon_n \\
&\stackrel{(iii)}{\leq} \frac{n}{2} \log P + nK_4 + I(\tilde{X}_A^n, X_B^n, Y_4^n; \tilde{X}_2^n) \\
&\qquad - I(X_B^n; Y_4^n, \tilde{X}_2^n | \tilde{X}_A^n) + \epsilon_n,
\end{aligned}
$$

where (i) follows since P7 implies $W_1 \leftrightarrow (Y_4^n, \tilde{X}_2^n) \leftrightarrow Y_{d_1}^n$; (ii) follows since $s_1 \in A$; (iii) follows from the fact that $I(\tilde{X}_A^n, X_B^n; Y_4^n)$ can be upper bounded as in (2). The second term in the inequality above can be bounded as

$$
\begin{aligned}
&I(\tilde{X}_A^n, X_B^n, Y_4^n; \tilde{X}_2^n) \stackrel{(i)}{\leq} I(\tilde{X}_A^n, \tilde{X}_B^n, Y_4^n; \tilde{X}_2^n) \\
&= I(\tilde{X}_B^n; \tilde{X}_2^n) + I(\tilde{X}_A^n, Y_4^n; \tilde{X}_2^n | \tilde{X}_B^n) \\
&\stackrel{(ii)}{\leq} I(\tilde{X}_B^n; Y_6^n) + I(\tilde{X}_A^n, Y_4^n; \tilde{X}_2^n | \tilde{X}_B^n) \\
&\stackrel{(iii)}{\leq} I(\tilde{X}_B^n; Y_6^n) + I(X_2^n; \tilde{X}_2^n | \tilde{X}_B^n) \\
&\stackrel{(iv)}{\leq} I(X_5^n, \tilde{X}_B^n; Y_6^n) - I(X_5^n; Y_6^n | \tilde{X}_B^n) + I(X_2^n; Y_0^n | \tilde{X}_B^n) + nK_5 \\
&\stackrel{(v)}{\leq} \frac{n}{2} \log P - I(X_5^n; Y_6^n | \tilde{X}_B^n) + I(X_2^n; Y_0^n | \tilde{X}_B^n) + nK_5 + nK_6 \qquad (5)
\end{aligned}
$$

where (i) follows since $(\tilde{X}_A^n, X_B^n, Y_4^n) \leftrightarrow (\tilde{X}_A^n, \tilde{X}_B^n, Y_4^n) \leftrightarrow \tilde{X}_2^n$; (ii) follows since P8 implies $\tilde{X}_B^n \leftrightarrow Y_6^n \leftrightarrow \tilde{X}_2^n$; (iii) follows since, given X_B^n, $(\tilde{X}_A^n, Y_4^n) \leftrightarrow X_2^n \leftrightarrow \tilde{X}_2^n$; (iv) follows by applying Lemma 3 to $I(X_2^n; \tilde{X}_2^n | \tilde{X}_B^n, Y_0^n)$, since $\mathcal{I}(v_0) \setminus \{v_2\} \subset B$, from P1; (v) follows by upper-bounding $I(X_5^n, \tilde{X}_B^n; Y_6^n)$ as in (2). We obtain

$$
\begin{aligned}
nR_1 &\leq n \log P - I(X_5^n; Y_6^n | \tilde{X}_B^n) + I(X_2^n; Y_0^n | \tilde{X}_B^n) \\
&\quad - I(X_B^n; Y_4^n, \tilde{X}_2^n | \tilde{X}_A^n) + n(K_4 + K_5 + K_6) + \epsilon_n. \qquad (6)
\end{aligned}
$$

Now we add (1), (3), (4), (6) and divide by $n \log P$, to obtain

$$
\frac{R_1 + R_2}{\frac{1}{2} \log P} \leq \frac{3}{2} + \frac{\sum_{j=1}^{6} K_j + \frac{1}{n} \epsilon_n}{\log P}.
$$

If we let $n \to \infty$ and then $P \to \infty$, we obtain $d_\Sigma \leq \frac{3}{2}$.

For networks that fall in cases (A) and (A$'$), we derive similar inequalities, and we conclude that $d_\Sigma \leq 1$. This is shown in Appendix D. Since 1 DoF is trivially achievable if $s_i \rightsquigarrow d_i$, for $i = 1$ or 2, we conclude the proof of Theorem 1.

References

[1] L. R. Ford and D. R. Fulkerson, "Maximal flow through a network," *Canadian Journal of Mathematics*, 1956.

[2] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. on Information Theory*, 2000.

[3] S.-Y. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. on Info. Theory*, Feb. 2003.

[4] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Trans. Netw.*, vol. 11, no. 5, pp. 782–795, 2003.

[5] S. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow: a deterministic approach," *to appear in IEEE Transactions on Information Theory*, 2011.

[6] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," *IEEE Trans. on Info. Theory*, Feb. 2009.

[7] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum–rate capacity for Gaussian interference channels," *IEEE Trans. on Info. Theory*, 2009.

[8] V. S. Annapureddy and V. V. Veeravalli, "Gaussian interference networks: Sum capacity in the low-interference regime and new outer bounds on the capacity region," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3032–3050, July 2009.

[9] A. E. Gamal and M. Costa, "The capacity region of a class of deterministic interference channels," *IEEE Transactions on information Theory*, vol. 28, no. 2, pp. 343–346, March 1982.

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

[11] G. Bresler, A. Parekh, and D. Tse, "The approximate capacity of the many-to-one and one-to-many Gaussian interference channels," *Proc. of Allerton Conference*, 2007.

[12] V. Cadambe and S. Jafar, "Interference alignment and spatial degrees of freedom for the k user interference channel," *IEEE International Conference on Communications*, 2008.

[13] A. S. Motahari, S. O. Gharan, and A. K. Khandani, "Real interference alignment with real numbers," *Submitted to IEEE Trans. on Info. Theory*.

[14] M. A. Maddah-Ali, S. A. Motahari, and A. K. Khandani, "Communication over mimo x channels: Interference alignment, decomposition, and performance analysis," *IEEE Transactions on Information Theory*, 2008.

[15] T. C. Hu, "Multi-commodity network flows," *Operations Research*, vol. 11, no. 3, pp. 344–360, 1963.

[16] C. C. Wang and N. B. Shroff, "Beyond the butterfly: A graph-theoretic characterization of the feasibility of network coding with two simple unicast sessions." *ISIT*, 2007.

[17] S. Shenvi and B. K. Dey, "A simple necessary and sufficient condition for the double unicast problem," *in Proceedings of ICC*, 2010.

[18] K. Cai, K. B. Letaief, P. Fam, and R. Feng, "On the solvability of 2-pair unicast networks — a cut-based characterization," *arXiv*, 2010.

[19] S. Mohajer, S. N. Diggavi, C. Fragouli, and D. Tse, "Approximate capacity of gaussian interference-relay networks with weak cross links," *arXiv*, 2010.

[20] T. Gou, S. Jafar, S.-W. Jeon, and S.-Y. Chung, "Aligned interference neutralization and the degrees of freedom of the $2 \times 2 \times 2$ interference channel," *arXiv*, 2010.

[21] I. Shomorony and A. S. Avestimehr, "Two-unicast wireless networks: Characterizing the sum degrees of freedom," *arXiv*:1102.2498v1, 2011.

APPENDIX

A. Proofs of Lemmas

Proof of Lemma 1:

In order to prove Lemma 1, we will first state and prove a claim, which is a simple consequence of the max-flow-min-cut theorem ([1]).

Claim 1. *Suppose we have $A \subset V_{\ell_A}$ and $B \subset V_{\ell_B}$, so that $\ell_A < \ell_B$. If there are no two disjoint paths with starting nodes in A and ending nodes in B, then there exists a node v_d such that $\ell_A \leq \ell(v_d) \leq \ell_B$, whose removal disconnects A from B.*

Proof: We let $G = (V, E)$ be the underlying graph of our original network, and we construct a new graph $G' = (V', E')$ in the following way. We let the layers in V' be V_{ℓ_A},V'_{ℓ_A}, V_{ℓ_A+1},V'_{ℓ_A+1},...,V_{ℓ_B},V'_{ℓ_B}, where V'_i is a copy of V_i. The edges between V'_i and V_{i+1}, for $i = \ell_A, \ell_A + 1, ..., \ell_B - 1$, are the same as the edges between V_i and V_{i+1} in G. To add the edges between V_i and V'_i, for $i = \ell_A, \ell_A+1, ..., \ell_B$, we simply connect each $v \in V_i$ to its copy in V'_i.

It is easy to see that any two edge-disjoint paths between A and B' in G' correspond to two vertex-disjoint paths between A and B in G. Therefore, since we assumed there are no two vertex-disjoint paths between A and B in G, there cannot be two edge-disjoint paths between A and B' in G'. Thus, by the max-flow min-cut Theorem, there exists an edge e_d in G' whose removal disconnects A from B'. It is easy to see that e_d can also be chosen to be an edge between V_i and its copy V'_i, for some i. This is because, if e_d is connecting V'_i and V_{i+1}, for some i, then we can choose the edge in $V_i \times V'_i$ (or $V_{i+1} \times V'_{i+1}$) which is adjacent to e_d, and it will also disconnect A from B', since its removal disconnects e_d from A (or B'). Now this choice of e_d corresponds to a vertex v_d in G whose removal disconnects A from B. ∎

We can now use this claim to prove Lemma 1.

Consider the nodes in $\mathcal{I}(v_i^p)$. Assume, by contradiction, that there are no two paths P_{s_1,v_p^i} and P_{s_2,v_p^i} such that $P_{s_1,v_p^i} \cap P_{s_2,v_p^i} = \{v_p^i\}$. Then, we do not have two vertex-disjoint paths starting in $\{s_1, s_2\}$ and ending in $\mathcal{I}(v_i^p)$. From Claim 1, there exists a node v_d whose removal disconnects $\{s_1, s_2\}$ from $\mathcal{I}(v_i^p)$, and thus from v_p^i. The existence of the path P_{s_i,d_i} containing v_p^i guarantees that $v_d \in P_{s_i,d_i}$. Since the removal of v_p^i disconnects $s_{\bar{i}}$ from d_i, and the removal of v_d disconnects $\{s_1, s_2\}$ from v_p^i, we conclude that the removal of v_d also disconnects $s_{\bar{i}}$ from d_i. But this is a contradiction to the fact that v_p^i was the first such node. ∎

Proof of Lemma 2: Since $n_i(G[S]) \geq 2$, we have that $s_{\bar{i}} \rightsquigarrow d_i$. Thus, since the removal of v_p^i disconnects $s_{\bar{i}}$ from d_i, we must have at least one node $v_1 \in \mathcal{I}(v_p^i)$ such that $s_{\bar{i}} \rightsquigarrow v_1$. If we suppose by contradiction that v_1 is the only such node, then we have that v_1 disconnects $s_{\bar{i}}$ from d_i. If $v_1 \in P_{s_i,d_i}$ we contradict our choice of v_p^i. If $v_1 \notin P_{s_i,d_i}$, then we contradict the fact that $n_i(G[S]) \geq 2$. ∎

Proof of Lemma 3: Let $A' = \mathcal{I}(v_b) \setminus \{v_c\}$ and $D = \mathcal{O}(v_c) \setminus \{v_b\}$. Then we have

$$
\begin{aligned}
& I(X_S^n; \tilde{X}_c^n | Y_b^n, \tilde{X}_A^n) \\
& = I(X_S^n; \{\tilde{X}_{c,j}^n : j \text{ s.t. } v_j \in \mathcal{O}(v_c)\} | Y_b^n, \tilde{X}_A^n) \\
& \overset{(i)}{=} I(X_S^n; \{\tilde{X}_{c,j}^n - \tfrac{h_{c,j}}{h_{c,b}} \tilde{X}_{c,b}^n : j \text{ s.t. } v_j \in \mathcal{O}(v_c)\} | Y_b^n, \tilde{X}_A^n) \\
& \overset{(ii)}{=} I(X_S^n; \{N_{c,j}^n - \tfrac{h_{c,j}}{h_{c,b}} N_{c,b}^n : j \text{ s.t. } v_j \in D\} | Y_b^n, \tilde{X}_A^n) \\
& \leq h(\{N_{c,j}^n - \tfrac{h_{c,j}}{h_{c,b}} N_{c,b}^n : j \text{ s.t. } v_j \in D\}) \\
& \qquad - h(\{N_{c,j}^n - \tfrac{h_{c,j}}{h_{c,b}} N_{c,b}^n : j \text{ s.t. } v_j \in D\} | Y_b^n, \tilde{X}_A^n, X_S^n) \\
& \overset{(iii)}{\leq} \frac{n|D|}{2} \log(2\pi e \kappa) \\
& \qquad - h(\{N_{c,j}^n - \tfrac{h_{c,j}}{h_{c,b}} N_{c,b}^n : j \text{ s.t. } v_j \in D\} | Y_b^n, \tilde{X}_A^n, X_S^n) \\
& \overset{(iv)}{\leq} \frac{n|D|}{2} \log(2\pi e \kappa) \\
& \qquad - h(\{N_{c,j}^n : j \text{ s.t. } v_j \in D\} | N_{c,b}^n, Y_b^n, \tilde{X}_A^n, X_S^n) \\
& \overset{(v)}{=} \frac{n|D|}{2} \log(2\pi e \kappa) - h(\{N_{c,j}^n : j \text{ s.t. } v_j \in D\}) \\
& = n \left(\frac{|D|}{2} \log(2\pi e \kappa) - \sum_{j : v_j \in D} \frac{1}{2} \log \left(\frac{2\pi e}{|\mathcal{I}(v_j)|} \right) \right),
\end{aligned}
$$

where (i) follows since we have $Y_b^n - \sum_{v_a \in A'} \tilde{X}_{a,b}^n = \tilde{X}_{c,b}^n$; (ii) follows since, for $j = b$, $N_{c,j}^n - \frac{h_{c,j}}{h_{c,b}} N_{c,b}^n = 0$; (iii) follows by letting $\kappa \triangleq 1 + (\max_{e,f \in E} h_e/h_f)^2$; (iv) follows because conditioning reduces entropy and thus we can condition on $N_{c,b}^n$; (v) follows from the fact that, since for $w \in D$ and $u \in A \cup S$, $w \not\rightsquigarrow u$, $N_{c,w}^n$ is independent of all the random variables conditioned on. ∎

B. Networks in case (B′)

We start by inferring important properties of the structure of the network, if it does not fall into case (A). We will show that such a network must contain one of the following three structures: (i) two disjoint paths P_{s_1,d_1} and P_{s_2,d_2}; (ii) a Butterfly; or (iii) a Grail. First we formalize the last two.

Definition 8. *The network $\mathcal{N}$ is a Butterfly network if it contains two nodes u_0 and u_1 connected by a path P_{u_0,u_1} (if $u_0 = u_1$, then we assume the path consists of a single node), two disjoint paths P_{s_1,d_2} and P_{s_2,d_1} which do not contain any node from P_{u_0,u_1}, and two paths P_{s_1,d_1} and P_{s_2,d_2} such that $P_{s_1,d_1} \cap P_{s_2,d_2} = P_{u_0,u_1}$. An example is shown in Figure 5.*

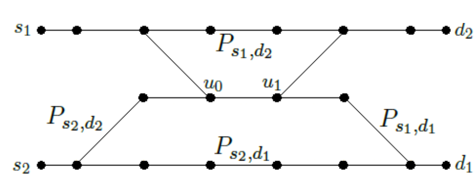


Fig. 5. Illustration of a Butterfly network.

Definition 9. *The network $\mathcal{N}$ is a Grail network if it contains two disjoint paths P_{s_1,d_2} and P_{s_2,d_1} and nodes $w_a \in P_{s_1,d_2}$ and $w_b \in P_{s_2,d_1}$ such that $s_2 \rightsquigarrow w_a$, $w_a \rightsquigarrow w_b$, and $w_b \rightsquigarrow d_2$. An example is shown in Figure 6.*

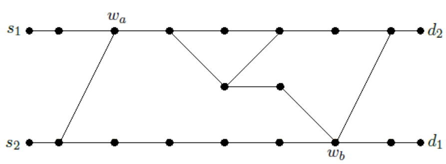


Fig. 6. Illustration of a Grail network.

Then we can state the following Claim.

Claim 2. *The absence of a node v whose removal disconnects d_i from both sources and $s_{\bar{i}}$ from both destinations, for $i = 1$ or $i = 2$, implies that $\mathcal{N}$ must contain (i) two disjoint paths P_{s_1,d_1} and P_{s_2,d_2}, (ii) a Butterfly subnetwork, or (iii) a Grail subnetwork.*

Proof: We let $G = (V, E)$ be the graph of our original network, and we construct an extended network $\mathcal{N}$ with graph $G = (V', E')$ in the following way. We let the layers in V' be $V_1, V_1', V_2, V_2', ..., V_r, V_r'$, where V_j' is a copy of V_j, $j = 1, ..., r$. The edges between V_j' and V_{j+1}, for $j = 1, 2, ..., r-1$, are the same as the edges between V_j and V_{j+1} in G. To add the edges between V_j and V_j', for $j = 1, 2, ..., r$, we simply connect each $v_k \in V_j$ to its copy v_k' in V_j'.

Next we claim that if we have an edge $e \in E'$ whose removal from $\mathcal{N}'$ disconnects d_i' from $\{s_1, s_2\}$ and $s_{\bar{i}}$ from $\{d_1', d_2'\}$, $i \in \{1, 2\}$, then our original network falls in (A). Suppose we have such an edge $e \in E'$. If $e \in V_j \times V_j'$ for some j, then it is easy to see that in the original network, this edge corresponds to a single node in V_j whose removal disconnects d_i from both sources and $s_{\bar{i}}$ from both destinations, and we must be in (A). Otherwise, if $e \in V_j' \times V_{j+1}$ for some j, then the removal of the edge $\tilde{e}$ in $V_j \times V_j'$ (or $V_{j+1} \times V_{j+1}'$) which is adjacent to e must also disconnect d_i' from $\{s_1, s_2\}$ and $s_{\bar{i}}$ from $\{d_1', d_2'\}$. This is because all paths from $\{s_1, s_2\}$ to $\{d_1', d_2'\}$ which contain the nodes in e must also contain the nodes in $\tilde{e}$. Then we notice that $\tilde{e}$ can be translated to a node v in $\mathcal{N}$ whose removal disconnects d_i from both sources and s_i from both destinations, and $\mathcal{N}$ falls into case (A).

Therefore, the absence of a node v as described in (A) in our network $\mathcal{N}$ implies that $\mathcal{N}'$ does not contain an edge whose removal disconnects d_i from both sources and s_i from both destinations for $i = 1$ or 2. Thus, we employ a result for two-unicast networks, shown in both [17][2] and [18], which guarantees that the extended network $\mathcal{N}'$ must contain one of three structures: two edge-disjoint paths $P_{s_1,d_1'}$ and $P_{s_2,d_2'}$, a Butterfly, or a Grail. Moreover, we notice that, in $\mathcal{N}'$, any pair of edge-disjoint paths is also vertex-disjoint, and corresponds to a pair of vertex-disjoint paths in $\mathcal{N}$. Thus, we conclude that if our network $\mathcal{N}$ is not in (A), then it must contain two vertex-disjoint paths P_{s_1,d_1} and P_{s_2,d_2}, a Grail subnetwork or a Butterfly subnetwork. ■

[2]In [17], a fourth kind of network, the augmented half-butterfly, was included among the networks which do not contain such an edge. However, it can be verified that the augmented half-butterfly contains a Grail subnetwork.

Next, we assume that all nodes that do not belong to the subnetwork satisfying the conditions in (B') are removed. Since the resulting network does not contain two disjoint paths, but does not fall in case (A), we conclude from Claim 2 that we may either have a Butterfly network or a Grail network. Next, we provide the achievability scheme for the Grail network. The achievability scheme for the Butterfly network can be found in [21].

Achievability for the Grail network:

We assume that we have a minimal subnetwork which still satisfies Definition 9, i.e., all the unnecessary nodes are removed. Our condensed network will be formed by V_1, $V_{\ell(w_a)}$, $V_{\ell(w_b)}$ and V_r. Notice that if we assume that the subnetwork is chosen to be minimal, each of these layers must contain exactly two nodes. Therefore, our condensed network will be as shown in Figure 7. We will let the nodes in $V_{\ell(w_a)}$

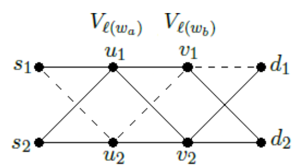


Fig. 7. Illustration of the condensed network of a Grail network. Solid lines represent edges that must exist in the condensed network, while the dashed lines represent edges that may not exist.

be called u_1 and u_2, and the nodes in $V_{\ell(w_b)}$ be called v_1 and v_2, as shown in Figure 7. Next we will show that either we can suppress one of the two intermediate key layers (by assuming their nodes are also just forwarding their received signals) and obtain a $2\times2\times2$ interference channel and use the result from [20], or we can choose scaling factors y_1, y_2, x_1 and x_2 (respectively for u_1, u_2, v_1 and v_2) so that the end-to-end transfer matrix is diagonal with non-zero diagonal entries. We notice that if $\hat{h}(s_1, u_2)$ is not identically zero, then the existence of two disjoint paths P_{s_1,d_2} and P_{s_2,d_1} containing u_1 and u_2 respectively guarantees that if we suppress $V_{\ell(w_b)}$, our condensed network becomes a $2 \times 2 \times 2$ interference channel. Similarly, if $\hat{h}(v_1, d_1)$ is not identically zero, we can suppress $V_{\ell(w_a)}$ from the condensed network, and we again obtain a $2\times2\times2$ interference channel. Therefore, we will assume that $\hat{h}(s_1, u_2) = \hat{h}(v_1, d_1) = 0$, and we will show that there is a choice of y_1, y_2, x_1 and x_2 so that the end-to-end transfer matrix is diagonal with non-zero diagonal entries. In order to do that we first consider the transfer matrix between V_1 and $V_{\ell(w_b)}$, which is given by

$$F = \begin{bmatrix} \hat{h}(s_1, u_1)\hat{h}(u_1, v_1)y_1 & \sum_{j=1}^{2} \hat{h}(s_2, u_j)\hat{h}(u_j, v_1)y_j \\ \hat{h}(s_1, u_1)\hat{h}(u_1, v_2)y_1 & \sum_{j=1}^{2} \hat{h}(s_2, u_j)\hat{h}(u_j, v_2)y_j \end{bmatrix}.$$

Then we notice that if we let

$$M = \begin{bmatrix} \hat{h}(s_2, u_1)\hat{h}(u_1, v_1) & \hat{h}(s_2, u_2)\hat{h}(u_2, v_1) \\ \hat{h}(s_2, u_1)\hat{h}(u_1, v_2) & \hat{h}(s_2, u_2)\hat{h}(u_2, v_2) \end{bmatrix},$$

we have

$$\det M = \begin{vmatrix} \hat{h}(s_2,u_1)\hat{h}(u_1,v_1) & \hat{h}(s_2,u_2)\hat{h}(u_2,v_1) \\ \hat{h}(s_2,u_1)\hat{h}(u_1,v_2) & \hat{h}(s_2,u_2)\hat{h}(u_2,v_2) \end{vmatrix}$$
$$= \hat{h}(s_2,u_1)\hat{h}(s_2,u_2) \begin{vmatrix} \hat{h}(u_1,v_1) & \hat{h}(u_2,v_1) \\ \hat{h}(u_1,v_2) & \hat{h}(u_2,v_2) \end{vmatrix},$$

which is a non-identically zero polynomial on the channel gains, since $s_2 \rightsquigarrow u_1$, $s_2 \rightsquigarrow u_2$ and there are two disjoint paths P_{u_1,v_1} and P_{u_2,v_2}. Thus M is invertible with probability 1. Since we also have that $\hat{h}(s_2,u_1)\hat{h}(u_1,v_2) \neq 0$ and $\hat{h}(s_2,u_2)\hat{h}(u_2,v_2) \neq 0$ w.p. 1, we are guaranteed that if we choose $y_1 \neq 0$ and $y_2 \neq 0$ such that $F_{2,2} = \hat{h}(s_2,u_1)\hat{h}(u_1,v_2)y_1 + \hat{h}(s_2,u_2)\hat{h}(u_2,v_2)y_2 = 0$, then $F_{1,1} \neq 0$, $F_{1,2} \neq 0$ and $F_{2,1} \neq 0$. Notice that, if $F_{1,2} = \hat{h}(s_2,u_1)\hat{h}(u_1,v_1)y_1 + \hat{h}(s_2,u_2)\hat{h}(u_2,v_1)y_2$ were zero, we would contradict the fact that the system $M\mathbf{y} = \mathbf{0}$ only has $\mathbf{y} = \mathbf{0}$ as a solution. Therefore, we have that the end-to-end transfer matrix can be expressed as

$$\begin{bmatrix} 0 & \hat{h}(v_2,d_1) \\ \hat{h}(v_1,d_2) & \hat{h}(v_2,d_2) \end{bmatrix} \begin{bmatrix} x_1 & 0 \\ 0 & x_2 \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ \gamma & 0 \end{bmatrix}$$
$$= \begin{bmatrix} \hat{h}(v_2,d_1)\gamma x_2 & 0 \\ \hat{h}(v_1,d_2)\alpha x_1 + \hat{h}(v_2,d_2)\gamma x_2 & \hat{h}(v_1,d_2)\beta x_1 \end{bmatrix},$$

where $\alpha \neq 0$, $\beta \neq 0$ and $\gamma \neq 0$. Therefore, since $\hat{h}(v_2,d_1)$, $\hat{h}(v_1,d_2)$ and $\hat{h}(v_2,d_2)$ are all non-zero with probability 1, we can choose x_1 and x_2 non-zero to make the end-to-end transfer matrix diagonal with non-zero diagonal entries.

C. Networks in case C2

If we are in case C2, then we have, for two disjoint paths P_{s_1,d_1} and P_{s_2,d_2}, $n_1(G) = n_1^D = 1$. Since we must not have an edge (v_2, v_1) as in case (A′), it is possible to infer some properties about the network graph. In [21], it is shown that, if a network falls in case C2 and for no other choice of disjoint paths P'_{s_1,d_1} and P'_{s_2,d_2} we are in case C1, then we can assume WLOG that we have two other paths Q_{s_1,d_1} and Z_{s_1,d_1}, both disjoint from P_{s_2,d_2}, such that

- $n_1^D(P_{s_2,d_2}, Q_{s_1,d_1}) = 0$ and $n_2^D(Q_{s_1,d_1}, P_{s_2,d_2}) = 1$,
- $n_1^D(P_{s_2,d_2}, Z_{s_1,d_1}) = 1$ and $n_2^D(Z_{s_1,d_1}, P_{s_2,d_2}) = 0$.

Thus, we will let (v_2, v_1) be an edge such that $v_2 \in P_{s_2,d_2}$ and $v_1 \in Z_{s_1,d_1}$ and (v_3, v_4) be an edge such that $v_3 \in Z_{s_1,d_1}$ and $v_4 \in P_{s_2,d_2}$. Moreover, it is shown in [21] that since all pairs of disjoint paths P'_{s_1,d_1} and P'_{s_2,d_2} must be in case C2, we must have the following two properties:

P1. All paths from s_2 to d_1 contain v_2 and v_1.
P2. All paths from s_1 to d_2 contain v_3 and v_4

An example of a network with paths Z_{s_1,d_1}, Q_{s_1,d_1} and P_{s_2,d_2} satisfying all the above properties is shown in Figure 8. We will now consider two cases and provide a scheme to achieve 3/2 degrees-of-freedom in each case. Our schemes will once again be based on using two modes of operation and having nodes store the received signals during the first mode of operation and use them during the second mode of operation.

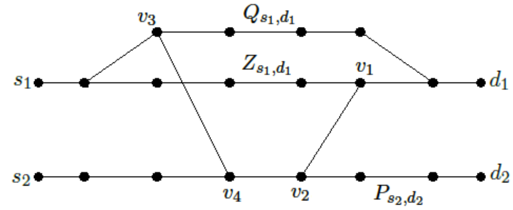


Fig. 8. An example of a network in case C2.

1) Achievability scheme if $\ell(v_3) \geq \ell(v_1)$: In Mode 1, we let the node from P_{s_2,d_2} in $V_{\ell(v_1)}$ be a virtual destination d'_2. Any node $v \in P_{s_2,d_2}$ such that $\ell(v) \geq \ell(d'_2)$ will stay silent during Mode 1. Then we notice that the two disjoint paths Q_{s_1,d_1} and P_{s_2,d'_2} have no direct edge between them and thus have manageable interference. Therefore, it is possible to guarantee that the transfer matrix between (s_1, s_2) and (d_1, d'_2) is diagonal with non-zero diagonal entries. During Mode 1, d'_2 will store its received signals.

The second mode of operation should last for the same number of time steps as the first one. In Mode 2, d'_2 will become a virtual source s'_2. Then, we remove all the nodes from the network except those in the paths Z_{s_1,d_1} and $P_{s'_2,d_2}$. We again have two disjoint paths with no direct interference. Therefore, we can have the transfer matrix between (s_1, s'_2) and (d_1, d_2) be diagonal with non-zero diagonal entries. Thus, by letting node $d'_2 = s'_2$ forward each of the signals received during Mode 1 in Mode 2, it is clear that, over the two modes, we create three parallel AWGN channels, two of them between s_1 and d_1 and one of them between s_2 and d_2. Therefore, we achieve 3/2 DoF. A visual representation of the scheme is shown in Figure 9.

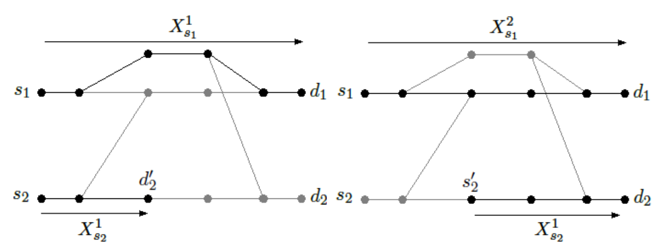


Fig. 9. Depiction of Mode 1 and Mode 2 for the achievability scheme in case C2 if $\ell(v_3) \geq \ell(v_1)$.

2) Achievability scheme if $\ell(v_3) < \ell(v_1)$: In Mode 1, we let v_1 be a virtual destination d'_2. Then we consider the path P_{s_2,d'_2}, formed by concatenating the segment of P_{s_2,d_2} from s_2 to v_2 and the edge (v_2, v_1). Then we notice that Q_{s_1,d_1} and P_{s_2,d'_2} are disjoint paths. Moreover, we claim that if $v_1 = d'_2$ stays silent, Q_{s_1,d_1} and P_{s_2,d'_2} have manageable interference. We must have $n_1(G, Q_{s_1,d_1}) = 0$, since otherwise we would have a path from s_2 to d_1 not containing v_1, and we would contradict P1. If $\ell(v_4) < \ell(v_1)$, then $\ell(v_4) \leq \ell(v_2)$ and the edge (v_3, v_4) will guarantee that $n_2^D(Q_{s_1,d_1}, P_{s_2,d'_2}) \geq 1$. Moreover, since we have a path $Z_{s_1,d'_2} = Z_{s_1,d_1}[s_1, v_1]$ not containing v_3, we must have $n_2(G, P_{s_2,d'_2}) \geq 2$. If $\ell(v_4) = \ell(v_1)$, then

(v_3, v_4) will not cause a direct interference from Q_{s_1,d_1} to $P_{s_2,d_2'}$. Then, if we have $n_2^D(Q_{s_1,d_1}, P_{s_2,d_2'}) = 0$, Q_{s_1,d_1} and P_{s_2,d_2} have manageable interference. If $n_2^D(Q_{s_1,d_1}, P_{s_2,d_2'}) = 1$, the direct interference must be due to an edge (v_3, v_1) so that $v_3 \xrightarrow{I} P_{s_2,d_2'}$. Otherwise, that would contradict the fact that $n_2^D(Q_{s_1,d_1}, P_{s_2,d_2}) = 1$. Therefore, the fact that we have a path $Z_{s_1,d_2'}$ not containing v_3 guarantees that $n_2(G, P_{s_2,d_2'}) \geq 2$. We conclude that, in any case, Q_{s_1,d_1} and $P_{s_2,d_2'}$ have manageable interference. Therefore, during Mode 1, it is possible to use an amplify-and-forward scheme which guarantees that the transfer matrix between (s_1, s_2) and (d_1, d_2') is diagonal with non-zero diagonal entries. During Mode 1, d_2' will store its received signals.

The second mode of operation should last for the same number of time steps as the first one. We will remove all nodes except those in Z_{s_1,d_1} and P_{s_2,d_2}. In Mode 2, s_2 will transmit the same signals it transmitted during Mode 1, while s_1 will transmit new signals. The only interference between the two paths happens through the edge (v_2, v_1). However, node v_1 received, during Mode 1, scaled versions of the transmitted signals at s_2. Therefore, by using the signals received during Mode 1, v_1 is able to remove the interference due to s_2 from its received signal during Mode 2. Hence we can guarantee that the transfer matrix between (s_1, s_2) and (d_1, d_2) during Mode 2 is diagonal with non-zero diagonal entries. Over the two modes, we again create three parallel AWGN channels, two of them between s_1 and d_1 and one of them between s_2 and d_2. Therefore, we achieve $3/2$ DoF. A visual representation of the scheme is shown in Figure 10.

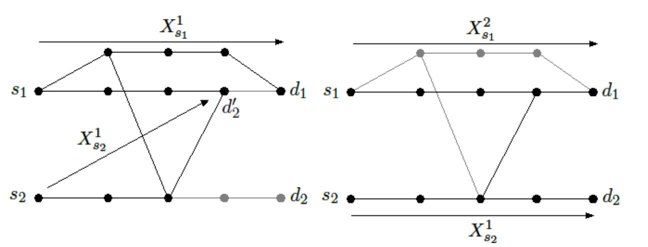


Fig. 10. Depiction of Mode 1 and Mode 2 for the achievability scheme in case C2 when $\ell(v_3) < \ell(v_1)$.

Next, we show that if our network falls in case C2, and does not fall into cases (A), (A′), (B), (B′) nor C1, then $d_\Sigma \leq \frac{3}{2}$. Similar to what we did for C1, we will use the fact that our network does not fall into cases (A), (A′), (B), (B′) nor C1 to infer connectivity properties about the network. We refer to [21] for the proofs of the following properties.

P3. The removal of v_4 disconnects d_2 from $\{s_1, s_2\}$
P4. The removal of v_2 disconnects s_2 from $\{d_1, d_2\}$
P5. The removal of v_1 and v_3 disconnects d_1 from $\{s_1, s_2\}$
P6. There is no path from v_1 to v_3

We now prove that under properties P1 through P6, $d_\Sigma \leq \frac{3}{2}$. We will derive information inequalities, as we did for C1. We let W_1 and W_2 be independent random variables corresponding to a uniform choice over the messages on sources s_1 and s_2 respectively, and we let $A = \{v \in V : s_2 \not\rightsquigarrow v\}$ and $B = \{v \in V : s_1 \not\rightsquigarrow v\}$. First we have

$$\begin{aligned} nR_2 &= H(W_2) = I(W_2; Y_{d_2}^n) + H(W_2|Y_{d_2}^n) \\ &\leq I(W_2; Y_{d_2}^n) + \epsilon_n \overset{(i)}{\leq} I(\tilde{X}_B^n; Y_4^n) + \epsilon_n \\ &= I(\tilde{X}_B^n, X_3^n; Y_4^n) - I(X_3^n; Y_4^n|\tilde{X}_B^n) + \epsilon_n \\ &\overset{(ii)}{\leq} \frac{n}{2}\log P + nK_7 - I(X_3^n; Y_4^n|\tilde{X}_B^n) + \epsilon_n, \end{aligned} \quad (7)$$

where (i) follows since $W_2 \leftrightarrow \tilde{X}_B^n \leftrightarrow Y_4^n \leftrightarrow Y_{d_2}^n$, which is implied by P3 and the fact that $s_2 \in B$; (ii) follows from the fact that $I(\tilde{X}_B^n, X_3^n; Y_4^n)$ can be upper bounded by $h(Y_4^n) - h(N_{3,4}^n)$ by following the steps in (2), where K_7 is a constant, independent of P, for P sufficiently large. Next, we have

$$\begin{aligned} nR_2 &= I(W_2; Y_{d_2}^n) + H(W_2|Y_{d_2}^n) \leq I(W_2; Y_{d_2}^n) + \epsilon_n \\ &\overset{(i)}{\leq} I(W_2; \tilde{X}_2^n, \tilde{X}_A^n) + \epsilon_n \overset{(ii)}{=} I(W_2; \tilde{X}_2^n|\tilde{X}_A^n) + \epsilon_n \\ &\overset{(iii)}{\leq} I(X_2^n; \tilde{X}_2^n|\tilde{X}_A^n) + \epsilon_n \leq I(X_2^n; \tilde{X}_2^n, Y_1^n|\tilde{X}_A^n) + \epsilon_n \\ &= I(X_2^n; Y_1^n|\tilde{X}_A^n) + I(X_2^n; \tilde{X}_2^n|\tilde{X}_A^n, Y_1^n) + \epsilon_n \\ &\overset{(iv)}{\leq} I(X_2^n; Y_1^n|\tilde{X}_A^n) + nK_8 + \epsilon_n, \end{aligned} \quad (8)$$

where (i) follows because from P4, the removal of v_2 disconnects d_2 from s_2, and therefore, the removal of v_2 and A disconnects d_2 from both sources, and we have $W_2 \leftrightarrow (\tilde{X}_2^n, \tilde{X}_A^n) \leftrightarrow Y_{d_2}^n$; (ii) follows since $\tilde{X}_A^n$ is independent of W_2; (iii) follows because, given $\tilde{X}_A^n$, we have $W_2 \leftrightarrow X_2^n \leftrightarrow \tilde{X}_2^n$; (iv) follows by applying Lemma 3 to $I(X_2^n; \tilde{X}_2^n|\tilde{X}_A^n, Y_1^n)$, because $\mathcal{I}(v_1) \setminus \{v_2\} \subset A$, or else we would contradict P1. Furthermore, we have

$$\begin{aligned} nR_1 &= H(W_1) = I(W_1; Y_{d_1}^n) + H(W_1|Y_{d_1}^n) \leq I(W_1; Y_{d_1}^n) + \epsilon_n \\ &\overset{(i)}{\leq} I(W_1; \tilde{X}_3^n, Y_1^n) + \epsilon_n = I(W_1; \tilde{X}_3^n) + I(W_1; Y_1^n|\tilde{X}_3^n) + \epsilon_n \\ &\overset{(ii)}{\leq} I(W_1; \tilde{X}_3^n|\tilde{X}_B^n) + I(W_1; Y_1^n|\tilde{X}_3^n) + \epsilon_n \\ &\overset{(iii)}{\leq} I(X_3^n; \tilde{X}_3^n|\tilde{X}_B^n) + I(W_1; Y_1^n|\tilde{X}_3^n) + \epsilon_n \\ &\overset{(iv)}{\leq} I(X_3^n; \tilde{X}_3^n|\tilde{X}_B^n) + I(\tilde{X}_A^n; Y_1^n|\tilde{X}_3^n) + \epsilon_n \\ &= I(X_3^n; \tilde{X}_3^n|\tilde{X}_B^n) + I(\tilde{X}_A^n, X_2^n; Y_1^n|\tilde{X}_3^n) \\ &\qquad - I(X_2^n; Y_1^n|\tilde{X}_3^n, \tilde{X}_A^n) + \epsilon_n \\ &\leq I(X_3^n; \tilde{X}_3^n|\tilde{X}_B^n) + I(\tilde{X}_A^n, X_2^n, \tilde{X}_3^n; Y_1^n) \\ &\qquad - I(X_2^n; Y_1^n|\tilde{X}_3^n, \tilde{X}_A^n) + \epsilon_n \\ &\overset{(v)}{=} I(X_3^n; \tilde{X}_3^n|\tilde{X}_B^n) + I(\tilde{X}_A^n, X_2^n; Y_1^n) - I(X_2^n; Y_1^n|\tilde{X}_A^n) + \epsilon_n \\ &\overset{(vi)}{\leq} I(X_3^n; \tilde{X}_3^n|\tilde{X}_B^n) + \frac{n}{2}\log P + nK_9 - I(X_2^n; Y_1^n|\tilde{X}_A^n) + \epsilon_n \\ &\leq I(X_3^n; Y_4^n|\tilde{X}_B^n) + I(X_3^n; \tilde{X}_3^n|\tilde{X}_B^n, Y_4^n) + \frac{n}{2}\log P + nK_9 \\ &\qquad - I(X_2^n; Y_1^n|\tilde{X}_A^n) + \epsilon_n \\ &\overset{(vii)}{\leq} I(X_3^n; Y_4^n|\tilde{X}_B^n) + \frac{n}{2}\log P + n(K_9 + K_{10}) \\ &\qquad - I(X_2^n; Y_1^n|\tilde{X}_A^n) + \epsilon_n, \end{aligned} \quad (9)$$

where (i) follows from P5, which implies $W_1 \leftrightarrow (\tilde{X}_3^n, Y_1^n) \leftrightarrow Y_{d_1}^n$; (ii) follows from the fact that $\tilde{X}_B^n$ is independent of W_1; (iii) follows from the fact that, given $\tilde{X}_B^n$, we have $W_1 \leftrightarrow X_3^n \leftrightarrow \tilde{X}_3^n$; (iv) follows from the fact that $s_1 \in A$; (v) follows because P1 and P6 imply that $s_2 \not\rightsquigarrow v_3$ and, therefore, $v_3 \in A$; (vi) follows from the fact that $I(\tilde{X}_A^n, X_2^n; Y_1^n)$ can be upper bounded by $h(Y_1^n) - h(N_{2,1}^n)$ by following the steps in (2), where K_{15} is a constant, independent of P, for P sufficiently large; and (vii) follows by applying Lemma 3 to $I(X_3^n; \tilde{X}_3^n | \tilde{X}_B^n, Y_4^n)$, since $\mathcal{I}(v_4) \setminus \{v_3\} \subset B$, or else we contradict P2. In order to bound the sum degrees-of-freedom, we can use the fact that

$$\begin{aligned} nR_1 &= H(W_1) = I(W_1; Y_{d_1}^n) + H(W_1 | Y_{d_1}^n) \\ &\leq I(W_1; Y_{d_1}^n) + \epsilon_n = h(Y_{d_1}^n) - h(Y_{d_1}^n | W_1) \\ &\leq h(Y_{d_1}^n) - h(Y_{d_1}^n | W_1, X_{\mathcal{I}(d_1)}^n) \\ &= h(Y_{d_1}^n) - h(N_{d_1}^n) \leq \frac{n}{2} \log P + nK_{11}, \end{aligned} \tag{10}$$

where the last inequality follows in the same way as (2). Therefore, we can add inequalities (7), (8), (9) and (10), and divide by $n \log P$ to obtain

$$\frac{R_1 + R_2}{\frac{1}{2}\log P} \leq \frac{3}{2} + \frac{\sum_{j=7}^{11} K_j + \frac{1}{n}\epsilon_n}{\log P}.$$

Thus, if we let $n \to \infty$ and then $P \to \infty$, we obtain $d_\Sigma \leq \frac{3}{2}$.

D. Networks in cases (A) and (A′):

The intuition behind the converse results is that there is a single node (v in case (A) and v_1 in case (A′)) which can approximately decode the messages from both sources. We start by considering (A), and we assume WLOG that we have a node v whose removal disconnects d_1 from both sources and s_2 from both destinations. We let W_1 and W_2 be independent random variables corresponding to uniform choices over the messages on sources s_1 and s_2 respectively. Then we have

$$\begin{aligned} nR_1 &= H(W_1) = I(W_1; Y_{d_1}^n) + H(W_1 | Y_{d_1}^n) \\ &\overset{(i)}{\leq} I(W_1; Y_{d_1}^n) + \epsilon_n \overset{(ii)}{\leq} I(X_{s_1}^n; Y_v^n) + \epsilon_n \end{aligned} \tag{11}$$

where (i) follows from Fano's inequality, where $\epsilon_n \to 0$ as $n \to \infty$; and (ii) follows because the removal of v disconnects d_1 from both sources; thus we have $W_1 \leftrightarrow X_{s_1}^n \leftrightarrow Y_v^n \leftrightarrow Y_{d_1}^n$. For R_2, we have

$$\begin{aligned} nR_2 &\leq I(W_2; Y_{d_2}^n) + \epsilon_n \overset{(i)}{\leq} I(X_{s_2}^n; Y_v^n, X_{s_1}^n) + \epsilon_n \\ &\overset{(ii)}{\leq} I(X_{s_2}^n; Y_v^n | X_{s_1}^n) + \epsilon_n \end{aligned} \tag{12}$$

where (i) follows because the removal of v disconnects d_2 from s_2, and, as a consequence, the removal of v and s_1 disconnects d_2 from both sources, and we have $W_2 \leftrightarrow X_{s_2}^n \leftrightarrow (Y_v^n, X_{s_1}^n) \leftrightarrow Y_{d_2}^n$; and (ii) follows since $X_{s_1}^n$ is independent of $X_{s_2}^n$. Now, by adding inequalities (11) and (12), we obtain

$$\begin{aligned} n(R_1 + R_2) &= I(X_{s_1}^n; Y_v^n) + I(X_{s_2}^n; Y_v^n | X_{s_1}^n) + \epsilon_n \\ &= I(X_{s_1}^n, X_{s_2}^n; Y_v^n) + \epsilon_n \\ &\leq I(X_{s_1}^n, X_{s_2}^n, X_{\mathcal{I}(v)}^n; Y_v^n) + \epsilon_n \\ &\leq \frac{n}{2} \log P + nK_{12} + \epsilon_n, \end{aligned} \tag{13}$$

where the last inequality follows as in (2) and K_{12} is a constant which does not depend on P, for P sufficiently large. Therefore we conclude that

$$d_\Sigma \leq \lim_{P \to \infty} \lim_{n \to \infty} \frac{\log P + K_{12} + \frac{2}{n}\epsilon_n}{\log P} = 1.$$

We can now proceed to the proof of case (A′). We assume WLOG that we have an edge $(v_2, v_1) \in E$ such that the removal of v_1 disconnects d_1 from both sources and the removal of v_2 disconnects s_2 from both destinations. We let $A \triangleq \{v \in V : s_2 \not\rightsquigarrow v\}$, and we notice that $\mathcal{I}(v_1) \setminus \{v_2\} \subset A$, since, otherwise, we would have a node $v_a \in \mathcal{I}(v_1) \setminus \{v_2\}$ such that $s_2 \rightsquigarrow v_a$, and this would contradict the fact that the removal of v_2 disconnects s_2 from d_1. Moreover, $v_2 \notin A$, because all paths from s_2 to d_2 contain v_2 and we must have at least one such path. Thus we have

$$\begin{aligned} nR_1 &\leq I(\tilde{X}_A^n, X_2^n; Y_1^n) - I(X_2^n; Y_1^n | \tilde{X}_A^n) + \epsilon_n \\ &\overset{(i)}{\leq} \frac{n}{2} \log P + nK_{13} - I(X_2^n; Y_1^n | \tilde{X}_A^n) + \epsilon_n, \end{aligned} \tag{14}$$

where (i) follows because v_1 disconnects d_1 from both sources and $s_1 \in A$, thus we have $W_1 \leftrightarrow \tilde{X}_A^n \leftrightarrow Y_1^n \leftrightarrow Y_{d_1}^n$; and (ii) follows as in step (iii) of (1), where K_{13} is a constant that is independent of P, for sufficiently large P. Next we notice that, since the removal of v_2 disconnects d_2 from s_2 and the removal of A disconnects d_2 from s_1, the removal of v_2 and A disconnects d_2 from both sources. Thus we have

$$\begin{aligned} nR_2 &\leq I(W_2; Y_{d_2}^n) + \epsilon_n \overset{(i)}{\leq} I(W_2; \tilde{X}_2^n, \tilde{X}_A^n) + \epsilon_n \\ &\overset{(ii)}{=} I(W_2; \tilde{X}_2^n | \tilde{X}_A^n) + \epsilon_n \overset{(iii)}{\leq} I(X_2^n; \tilde{X}_2^n | \tilde{X}_A^n) + \epsilon_n \\ &\leq I(X_2^n; Y_1^n | \tilde{X}_A^n) + I(X_2^n; \tilde{X}_2^n | \tilde{X}_A^n, Y_1^n) + \epsilon_n \\ &\overset{(iv)}{\leq} I(X_2^n; Y_1^n | \tilde{X}_A^n) + nK_{14} + \epsilon_n, \end{aligned} \tag{15}$$

where (i) follows from the fact that the removal of v_2 and A disconnects d_2 from both sources, which implies $W_2 \leftrightarrow (\tilde{X}_2^n, \tilde{X}_A^n) \leftrightarrow Y_{d_2}^n$; (ii) follows from the fact that W_2 is independent of $\tilde{X}_A^n$; (iii) follows from the fact that, given $\tilde{X}_A^n$, we have $W_2 \leftrightarrow X_2^n \leftrightarrow \tilde{X}_2^n$; (iv) follows from the application of Lemma 3 to $I(X_2^n; \tilde{X}_2^n | \tilde{X}_A^n, Y_1^n)$, since $\mathcal{I}(v_1) \setminus \{v_2\} \subset A$. Finally, by adding (14) and (15) we obtain

$$n(R_1 + R_2) \leq \frac{n}{2} \log P + n(K_{13} + K_{14}) + \epsilon_n,$$

and we conclude that $d_\Sigma \leq 1$. Since one degree-of-freedom is trivially achievable, we have that $d_\Sigma = 1$ for (A) and (A′).