Two-Unicast Wireless Networks: Characterizing the Degrees-of-Freedom
Ilan Shomorony and A. Salman Avestimehr

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 chosen independently according to continuous distributions, then, with probability 1, two-unicast layered Gaussian networks can only have 1, 3/2 or 2 sum degrees-of-freedom (unless both source-destination pairs are disconnected, in which case no degrees-of-freedom can be achieved). We provide sufficient and necessary conditions for each case based on network connectivity and a new notion of source-destination paths with manageable interference. Our achievability scheme is based on forwarding the received signals at all nodes, except for a small fraction of them in at most two key layers. Hence, we effectively create a “condensed network” that has at most four layers (including the sources layer and the destinations layer). We design the transmission strategies based on the structure of this condensed network. The converse results are obtained by developing information-theoretic inequalities that capture the structures of the network connectivity. Finally, we extend this result and characterize the full degrees-of-freedom region of two-unicast layered wireless networks.

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 aims to send the same message to one or more destinations, using multiple relay nodes. Since, in this scenario, all destination nodes are interested in the same message, there is effectively only one information stream in the network. Starting from the max-flow-min-cut theorem of Ford-Fulkerson [2], there has been significant progress on this problem. For wireline networks, the maximum multicast flow was characterized in [3]. In [4, 5], it was further shown that this maximum flow can be achieved using linear network codes.

In [6], 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 [6], hence providing a constant-gap approximation of the capacity. Tighter capacity approximations were later derived in [7, 8].

The second research direction focuses on multi-flow wireless networks with only one-hop between the sources and the destinations, i.e., the interference channel. While the capacity of the interference channel remains unknown (except for special cases, such as [9–15]), there has been a variety of capacity approximations derived, such as constant-gap capacity approximations [16–18] and degrees-of-freedom characterizations [19–24].

However, once we go beyond single-hop, there is much less known about the capacity of multi-flow networks. Even in the simplest case with two sources and two destinations there are very few general results, such as [25], where the maximum flow in two-unicast undirected wireline networks is characterized. For two-unicast directed wireline networks, [26–28] have provided graph-theoretic and cut-set based conditions under which rate (1, 1) can be achieved. In the wireless realm, constant-gap approximations of the capacity of specific two-hop networks (the ZZ and ZS networks) were obtained in [29]. Furthermore, it was recently shown that the network resulting from the concatenation of two or more fully connected interference channels (the XX structure) admits the maximum of two degrees-of-freedom [30]. The achievability scheme relies on the notion of real interference alignment, which was introduced in [22].

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 (for each existing link) are independently drawn from a continuous distribution and remain fixed during the course of communication. Moreover, we assume that all channel gains are fully known at all nodes. Under these assumptions, we will 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 degrees-of-freedom (unless the source-destination pairs are disconnected, in which case we have 0 degree-of-freedom). Furthermore, we will extend this result and show that there are only five possible degrees-of-freedom regions for two-unicast layered networks, and we will provide necessary and sufficient conditions for each case that are based only on properties of the network graph.

Paper outline and description of main contributions:
In Section II, we provide some basic definitions and state our main result. Then, in Section III, we give a high-level
description of the proof techniques and the intuition behind some of the arguments. We proceed to describing the networks in which only one degree-of-freedom can be achieved in Section IV. More specifically, if we let \((s_1, d_1)\) and \((s_2, d_2)\) be the pairs of corresponding source and destination, we will show that the maximum achievable degrees-of-freedom is one if and only if we are in one of the following two cases: (i) the network contains a node \(v\) whose removal disconnects pairs \((s_1, d_1)\), \((s_2, d_2)\) and at least one of \((s_1, d_2)\) and \((s_2, d_1)\); or (ii) the network contains an edge \((v_2, v_1)\) such that the removal of \(v_1\) disconnects a destination from both sources and the removal of \(v_2\) disconnects the non-corresponding source from both destinations. The conditions we present can be seen as a generalization of the graph-theoretic conditions given in [27] which characterize when a two-unicast wireline network does not support rate \((1, 1)\).

Then, in Section V, we consider the cases in which two degrees-of-freedom can be achieved. We will show that if our network graph contains a Butterfly or a Grail subgraph, then two degrees-of-freedom can be achieved. In order to describe the third class of networks which admit two degrees-of-freedom, we introduce the notion of manageable interference. We will say that two disjoint source-destination paths have manageable interference if, intuitively, all the interference between them can be either avoided or neutralized. Once again, it is interesting to compare the description of the networks with two degrees-of-freedom to the graph-theoretic description of the wireline networks which support rate \((1, 1)\). While in the wireline case it is possible to achieve rate \((1, 1)\) in networks which contain a Butterfly, a Grail or two edge-disjoint paths from each source to its corresponding destination (see [27, 28]), in the wireless case, it is possible to achieve two degrees-of-freedom in networks which contain a Butterfly, a Grail or two vertex-disjoint paths from each source to its destination that have manageable interference.

In order to describe general achievability schemes that work for an arbitrary number of layers, we propose a new method which involves building a condensed network, by identifying specific key layers which will perform non-trivial relaying operations. All the nodes which do not belong to the key layers will be assumed to simply forward their received signals at all times. Therefore, an effective transfer matrix between any pair of consecutive key layers can be obtained and it can be used to define the edges and the channel gains of our condensed network. To achieve two degrees-of-freedom, we will consider two distinct relaying schemes for the nodes in the key layers. If our condensed network is a \(2 \times 2 \times 2\) interference channel, then we will resort to the real interference alignment schemes provided in [30]. Otherwise, we will show that a linear coding scheme will suffice to achieve the sum degrees-of-freedom. Notice that, since we assume single antennas at all nodes, the cut-set bound tells us that we cannot hope to achieve more than two degrees-of-freedom, and this case requires no converse proof.

In Section VI, we address all the networks which do not fall into the cases considered in Sections IV and V. We will show that they all have \(3/2\) degrees-of-freedom. Our achievability scheme is based on defining two distinct modes of operation for the network. During the first mode, specific nodes act as buffers, storing all the received signals in order to use them during the second mode of operation. Then, in the second mode, these stored signals can be either forwarded towards the destinations or used to neutralize the interference. This way, it is possible to achieve \(3/2\) degrees-of-freedom by evenly dividing the amount of time the network operates in each mode. The converse result is obtained by finding information-theoretic inequalities which capture the fact that the interference, in this case, is not completely manageable.

In Section VII, we describe how the results regarding the sum degrees-of-freedom of two-unicast layered Gaussian networks can be extended to obtain the full degrees-of-freedom region. We show that there are only five possible degrees-of-freedom regions (assuming each source is connected to its destination) and we provide necessary and sufficient conditions for a network to have each of these regions. We do this by using the outer bound provided by the sum degrees-of-freedom and by describing achievability schemes for some specific extreme points in the degrees-of-freedom region, using real interference alignment.

Finally, in Section VIII, we provide some concluding remarks.

II. DEFINITIONS AND MAIN RESULTS

A multiple-unicast Gaussian network \(\mathcal{N} = (G, L)\) consists of a directed graph \(G = (V, E)\), where \(V\) is the vertex (or 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 will focus on two-unicast (two-source two-destination) Gaussian networks, which means that \(L = \{(s_1, d_1), (s_2, d_2)\}\), for distinct vertices \(s_1, s_2, d_1, d_2\). Moreover, we will assume that the network is layered, meaning that the vertex set \(V\) can be partitioned into \(r\) subsets \(V_1, V_2, ..., V_r\) (called 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 a vertex \(v \in V_j\), we will let \(\mathcal{I}(v) \triangleq \{u \in V_{j-1} : (u, v) \in E\}\) (the input nodes) and \(\mathcal{O}(v) \triangleq \{u \in V_{j+1} : (u, v) \in E\}\) (the output nodes). Furthermore, we will let \(\ell(v)\) be the index corresponding to the layer containing \(v\), i.e., \(v \in V_{\ell(v)}\). Notice that the layers induce a natural ordering of the nodes. Thus, we may say, for example, that \(v_a\) occurs before \(v_b\) if \(\ell(v_a) < \ell(v_b)\).

A real-valued channel gain \(h_e\) is associated with each edge \(e \in E\). Since we will often be referring to vertices by \(v_i\), for \(i \in \mathbb{N}\), we will also use \(h_{i,j}\) to represent the channel gain associated with edge \((v_i, v_j)\). We will assume that the channel coefficients \(h_e\) are independently drawn from continuous distributions and are fixed during the course of communication. We also assume that all channel gains are fully known at all nodes. At time \(m\), each node \(v_i\) (with the exception of \(d_1\) and \(d_2\)) transmits a real-valued signal \(X_{vi}[m]\) (or simply \(X_{i}[m]\), when there is no ambiguity), which must satisfy an average power constraint \(\frac{1}{n} \sum_{m=1}^{n} E \left[ X_{i}[m]^2 \right] \leq P\), \(\forall v_i \in V\), for a communication session of duration \(n\), where the expectation is taken with respect to any possible randomization involved.
The signal received by node $v_j$ at time $m$ is given by

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

where $N_j[m]$ is the zero mean unit variance Gaussian discrete-time white noise process associated with node $v_j$. The transmitted signal from node $v_j$ (with the exception of $s_1$ and $s_2$) at time $m$ must be a (possibly randomized) function of its past received signals $Y_j[k]$, for $k = 1, \ldots, m - 1$. Source $s_i$ picks a message $W_i$ that it wishes to communicate to $d_i$, and transmits signals $X_i[m]$, $m = 1, \ldots, n$, which are a function of $W_i$, for $i = 1, 2$. Each destination uses a decoder, which is a mapping $g_i : \mathbb{R}^n \rightarrow \{1, \ldots, |W_i|\}$ from the $n$ received signals to the source message indices (of $W_i$) which are the messages that can be chosen. We say that rates $R_i = \log |W_i|$ for $i = 1, 2$ are achievable if the probability of error in the decoding of both messages by their corresponding destinations can be made arbitrarily close to 0 by choosing a sufficiently large $n$. The sum-capacity $C_S(G)$ is the supremum of the achievable sum-rates for power constraint $P$.

**Definition 1.** The sum degrees-of-freedom $D_S$ of a two-unicast Gaussian network is defined as

$$D_S \equiv \lim_{P \to \infty} \frac{C_S(P)}{\frac{1}{2} \log P}.$$

**Remark:** $D_S$ will in general depend on $H = \{h_e : e \in E\}$. However, we will show that with probability 1, $D_S$ only depends on the network graph $G$, and not on the values of $H$.

We now consider several definitions which will be used throughout the paper.

**Definition 2.** A (directed) path between $v_1 \in V$ and $v_k \in V$ is an ordered set of nodes $\{v_1, v_2, \ldots, v_k\}$ such that $(v_i, v_{i+1}) \in E$ for $i = 1, \ldots, k - 1$. We will commonly refer to a path between $v_1$ and $v_k$ by $P_{v_1,v_k}$. We write $v_1 \leadsto v_k$ if there is a path between $v_1$ and $v_k$. Notice that for any node $v \in V$, $v \leadsto v$.

For simplicity, we will assume that any $v \in V$ belongs to at least one path $P_{s_i,d_j}$ for $i \in \{1, 2\}$ and $j \in \{1, 2\}$. This is reasonable since a node that does not belong to any source-destination path does not alter the achievable rates in the network and can be removed. Moreover, we will always assume that $s_i \leadsto d_j$ for $i = 1, 2$, since $s_i \not\leadsto d_1$ implies that $R_1 = 0$. In order to be able to “cut and paste” path segments we will also consider the following path operations. For a path $P_{v_a,v_b} = \{v_a, v_{a+1}, \ldots, v_b\}$, we will let $P_{v_a,v_b}[a,c] = \{v_c, v_{c+1}, \ldots, v_d\}$ if $a \leq c \leq d \leq b$. Moreover, if we have paths $P_{v_a,v_j}$ and $P_{v_j,v_b}$, we will let $P_{v_a,v_b} = P_{v_a,v_j} \cup P_{v_j,v_b}$ be the path which results from concatenating $P_{v_a,v_j}$ and $P_{v_j,v_b}$.

**Definition 3.** Paths $P_{v_a,v_b}$ and $P_{v_c,v_d}$ are said to be disjoint if $P_{v_a,v_b} \cap P_{v_c,v_d} = \emptyset$. (Notice that such paths are usually called vertex-disjoint, but here we will refer to them as simply disjoint)

**Definition 4.** For a subset of the vertices $S \subset V$, 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.** We say that $N' = (G', L')$ is a subnetwork of $N = (G, L)$, if $G' = G[S]$, for some $S \subset V$ such that $L \subset S \times S$, and $L' = L$.

For the next definitions, we assume we have two disjoint paths $P_{s_1,d_1}$ and $P_{s_2,d_2}$. Since we will often make statements which work for both $P_{s_1,d_1}$ and $P_{s_2,d_2}$, we will let $i = 2$ if $i = 1$ and $i = 1$ if $i = 2$.

**Definition 6.** We will say that a node $v_a \not\in P_{s_i,d_i}$ causes interference on $P_{s_i,d_i}$ and write $v_a \leadsto P_{s_i,d_i}$ if we can find a node $v_b \in P_{s_i,d_i}$ such that $(v_a, v_b) \in E$ and a path $P_{s_i,v_a}$ between $s_i$ and $v_a$ such that $P_{s_i,v_a} \cap P_{s_i,d_i} = \emptyset$, for $i = 1, 2$. Moreover, we will say the interference is direct, and write $v_a \leadsto P_{s_i,d_i}$ if, in addition, $v_a \in P_{s_i,d_i}$. Otherwise, we call the interference indirect.

Consider a subnetwork $(G[S], \{(s_1, d_1), (s_2, d_2)\})$ for some $S \supset (P_{s_1,d_1} \cup P_{s_2,d_2})$. We will define $n_i(G[S], P_{s_i,d_i}) = |\{v \in S : v \leadsto P_{s_i,d_i}\}|$, for $i = 1, 2$. Notice that, in the definition of $n_i$, the path implied by $v \leadsto P_{s_i,d_i}$ must exist in the subnetwork with graph $G[S]$. Moreover, we define $n(D_i(P_{s_1,d_1}, P_{s_2,d_2})) = \{v \in V : v \leadsto P_{s_i,d_i}\}$. When there is no ambiguity in the choice of our two disjoint paths $P_{s_1,d_1}$ and $P_{s_2,d_2}$, we will simplify the notation by using $n_i(G[S])$ and $n_i$.

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

The following example illustrates the definitions above.

**Example 1.** Consider the network depicted in Figure 1. We have two disjoint paths from each source to its corre-

Fig. 1. Two-unicast layered network considered in Example 1.

sponding destination, given by $P_{s_1,d_1} = \{s_1, v_1, v_2, v_3, d_1\}$ and $P_{s_2,d_2} = \{s_2, v_7, v_8, v_9, d_2\}$. If we consider the entire network $N = (G, L)$, then we have that $v_5 \leadsto P_{s_1,d_1}$ (since we have a path $P_{s_2,v_5} = \{s_2, v_4, v_5\}$ disjoint from $P_{s_1,d_1}$) and $v_7 \leadsto P_{s_2,d_2}$ (since we have a path $P_{s_2,v_7} = \{s_2, v_7\}$ disjoint from $P_{s_1,d_1}$ and $v_7 \in P_{s_2,d_2}$). Similarly, we have that $v_6 \leadsto P_{s_2,d_2}$. Thus, we conclude that $n_1(P_{s_2,d_2}, P_{s_1,d_1}) = 1$, $n_2(P_{s_1,d_1}, P_{s_2,d_2}) = 0$, $n_1(G[P_{s_1,d_1}]) = 2$, and $n_2(G, P_{s_2,d_2}) = 1$. If instead we consider the subnetwork $N' = (G[S], L)$, where $S = (P_{s_1,d_1} \cup P_{s_2,d_2})$, we have $n_1(G[S], P_{s_1,d_1}) = n_1(P_{s_2,d_2}, P_{s_1,d_1}) = 1$ and $n_2(G[S], P_{s_2,d_2}) = n_2(P_{s_1,d_1}, P_{s_2,d_2}) = 0$. Finally, we consider the subnetwork $N' = (G[S'], L)$, where $S' = V \setminus \{v_6\}$. 
Then we have \( n_1(G[S'], P_{s_1, d_1}) = 2 \) and \( n_2(G[S'], P_{s_2, d_2}) = 0 \), and we conclude that \( P_{s_1, d_1} \) and \( P_{s_2, d_2} \) have manageable interference.

Now we state our main results.

**Theorem 1.** For a two-unicast layered Gaussian network \( N = (G = (V, E), \{(s_1, d_1), (s_2, d_2)\}) \), where the channel gains are independently drawn from continuous distributions, with probability 1, the sum degrees-of-freedom of \( N, D_\Sigma \), are given by

\[
D_\Sigma = \begin{cases} 
1 & \text{if } N \text{ contains a node } v \text{ whose removal disconnects } d_i \text{ from both sources and } s_i \text{ from both destinations, for } i = 1 \text{ or } i = 2, \\
1 & \text{if } N \text{ contains an edge } (v_2, v_1) \in E \text{ such that the removal of } v_1 \text{ disconnects } d_i \text{ from both sources and the removal of } v_2 \text{ disconnects } s_i \text{ from both destinations, for } i = 1 \text{ or } i = 2, \\
2 & \text{if } N \text{ contains two disjoint paths } P_{s_1, d_1} \text{ and } P_{s_2, d_2} \text{ with manageable interference (see Definition 7),} \\
2 & \text{if } N \text{ or any subnetwork does not contain two disjoint paths } P_{s_1, d_1} \text{ and } P_{s_2, d_2}, \text{ but does not fall into case (A),} \\
\frac{3}{2} & \text{in all other cases.} 
\end{cases}
\]

We also characterize the full degrees-of-freedom region of two-unicast layered Gaussian networks. We first define some basic notions.

**Definition 8.** The capacity region \( C(P) \) of a two-unicast Gaussian wireless network \( N \) with power constraint \( P \) is the closure of the set of all pairs of achievable rates \((R_1, R_2)\).

**Definition 9.** The degrees-of-freedom region of a two-unicast Gaussian network \( N \) is given by

\[
D = \left\{ (D_1, D_2) \in \mathbb{R}_+^2 : \forall w_1, w_2 \in \mathbb{R}_+, w_1 D_1 + w_2 D_2 \leq \lim_{P \to \infty} \left( \sup_{(R_1, R_2) \in C(P)} \frac{w_1 R_1 + w_2 R_2}{\frac{1}{2} \log P} \right) \right\}. 
\]

In order to simplify the characterization of the networks according to their degrees-of-freedom regions, we also consider the following definition.

**Definition 10.** Two disjoint paths \( P_{s_1, d_1} \) and \( P_{s_2, d_2} \) have \((s_1, d_1)\)-manageable interference if we can find \( S \subset V \) such that \( P_{s_1, d_1}, P_{s_2, d_2} \subset S, n_i(G[S]) \neq 1, \) for \( i = 1 \) or \( i = 2 \).

For the degrees-of-freedom region of two-unicast Gaussian networks, we have the following result.

**Theorem 2.** For a two-unicast layered Gaussian network \( N = (G = (V, E), \{(s_1, d_1), (s_2, d_2)\}) \), where the channel gains are chosen according to independent continuous distributions, with probability 1, the degrees-of-freedom region \( D \) is given by

\[
\text{I. } D = \left\{ (D_1, D_2) \in \mathbb{R}_+^2 : D_1 + D_2 \leq 1 \right\} \text{ if } N \text{ falls in cases (A) or (A') in Theorem 1,} \\
\text{II. } D = \left\{ (D_1, D_2) \in \mathbb{R}_+^2 : D_i \leq 1, D_2 \leq 1 \right\} \text{ if } N \text{ falls in cases (B) or (B') in Theorem 1,} \\
\text{III. } D = \left\{ (D_1, D_2) \in \mathbb{R}_+^2 : D_1 \leq 1, D_2 \leq 1, D_1 + D_2 \leq \frac{3}{2} \right\} \text{ if } N \text{ is not in cases I, II and contains disjoint paths } P_{s_1, d_1} \text{ and } P_{s_2, d_2} \text{ whose interference is } (s_1, d_1) \text{- and } (s_2, d_2)\text{-manageable,} \\
\text{IV. } D = \left\{ (D_1, D_2) \in \mathbb{R}_+^2 : D_1 \leq 1, D_1 + 2D_2 \leq 2 \right\} \text{ if } N \text{ is not in cases I, II and III and contains paths } Q_{s_1, d_1}, Z_{s_1, d_1} \text{ and } P_{s_2, d_2} \text{ such that } Q_{s_1, d_1} \text{ and } P_{s_2, d_2} \text{ are disjoint and have } (s_1, d_1)\text{-manageable interference, and } Z_{s_1, d_1} \text{ and } P_{s_2, d_2} \text{ are disjoint and have } (s_2, d_2)\text{-manageable interference,} \\
\text{V. } D = \left\{ (D_1, D_2) \in \mathbb{R}_+^2 : D_2 \leq 1, 2D_1 + D_2 \leq 2 \right\} \text{ if } N \text{ is not in cases I, II and III and contains paths } P_{s_1, d_1}, Q_{s_2, d_2} \text{ and } Z_{s_2, d_2} \text{ such that } Q_{s_2, d_2} \text{ and } P_{s_1, d_1} \text{ are disjoint and have } (s_1, d_1)\text{-manageable interference, and } Z_{s_2, d_2} \text{ and } P_{s_1, d_1} \text{ are disjoint and have } (s_2, d_2)\text{-manageable interference.} 
\]

Moreover, any two-unicast layered Gaussian network \( N \) falls into one of the cases described above. These five regions are depicted in Figure 2.

**III. Proof overview**

Even though Theorem 1 can be seen as a simple consequence of Theorem 2, we will first prove Theorem 1. Theorem 2 will then follow as an extension of it. We will consider cases (A), (A'), (B), (B') and (C) sequentially. The intuition behind (A) is as follows. Let \( W_1 \) be the message from \( s_1 \) and \( W_2 \) be the message from \( s_2 \). If the removal of \( v \) disconnects \( d_1 \) from both sources, then by knowing the received signal at \( v \) we should be able to decode \( W_1 \). Then, since \( v \) also disconnects \( d_2 \) from \( s_2 \), loosely speaking, all the information about \( W_2 \) goes through \( v \). Therefore, \( v \) can use the knowledge about \( W_1 \) to remove any interference due to signals about \( W_1 \), thus being able to decode \( W_2 \) as well. Since a single node can decode both messages, we have that \( D_\Sigma \leq 1 \), and it follows that \( D_\Sigma = 1 \), since 1 degree-of-freedom is trivially achievable from the fact that \( s_1 \sim d_1 \) and \( s_2 \sim d_2 \). The intuition behind (A') is similar. If the removal of \( v_1 \) disconnects \( d_1 \) from both sources, then by knowing the received signal at \( v_1 \) we should be able to decode \( W_1 \). Since the removal of \( v_2 \) disconnects \( s_2 \) from both terminals, all the information regarding \( W_2 \) goes through \( v_2 \). This means that all the information received at \( v_1 \) which does not come from \( v_2 \) is about \( W_1 \) and, thus, by knowing the received signal at \( v_1 \), one can remove the part regarding \( W_1 \) and obtain the part of the transmitted signal at \( v_2 \) regarding \( W_2 \). But this implies that from \( v_1 \) we should be able to decode both \( W_1 \) and \( W_2 \), which implies \( D_\Sigma \leq 1 \).
An example of a network that would fall in (A') is shown in Figure 3.

To prove (B) and (B'), we will provide several achievability schemes for 2 degrees-of-freedom. For networks in (B), i.e., networks which contain two disjoint paths with manageable interference, depending on the network topology, we will either consider simple amplify-and-forward schemes or schemes based on real interference alignment, as described in [30]. If the network is in (B'), we will first restrict ourselves to the subnetwork which satisfies the description in (B'). Then, we will use a result from the double unicast problem for wireline networks to claim that the subnetwork must contain one of the three structures shown in Figure 4. But since we are assuming that the subnetwork has no two disjoint paths, we must have either the structure in Figure 4b or the structure in Figure 4c. We provide an amplify-and-forward achievability scheme in each case.

For case (C), we only need to consider networks which have two disjoint paths \( P_{s_1,d_1} \) and \( P_{s_2,d_2} \), but do not have two disjoint paths with manageable interference. This is because all networks which do not contain two disjoint paths \( P_{s_1,d_1} \) and \( P_{s_2,d_2} \) must fall into (A) or (B'). Moreover, any network that has two disjoint paths with manageable interference will fall into (B). We will identify two main classes of networks in (C), depicted in Figure 5, and for each of these classes we will first provide an achievability scheme, based on two separate modes of operation for the network, which achieves \( \frac{3}{2} \) degrees-of-freedom. Then, we will show that the non-existence of two disjoint paths with manageable interference implies that either the network falls into (B') or \( D_2 \leq \frac{3}{2} \).

We will then build upon the result from Theorem 1 to obtain Theorem 2. For networks in cases (A), (A'), (B) and (B'), we will notice that the degrees-of-freedom region can be readily obtained from the sum degrees-of-freedom. For networks in case (C), we use the fact that the they must contain two disjoint paths that do not have manageable interference to infer properties about the network connectivity. Then, we combine the outer bound provided by the sum degrees-of-freedom with achievability schemes for the extreme points to characterize the degrees-of-freedom region. Some of the extreme points will require the use of real interference alignment schemes.

IV. NETWORKS WITH ONLY ONE DEGREE-OF-FREEDOM

In this section, we will provide converse results for networks that fall in cases (A) and (A'). For the converse proofs, necessary for (A), (A') and (C), we will derive information inequalities which allow us to bound the achievable sum-rates, and thus the degrees-of-freedom. 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 assume that the communication session lasts \( n \) time steps, and for a node \( v_j \in V \), we let \( X_j^n \), \( Y_j^n \) and \( N_j^n \) be length \( n \) vectors whose entries are, respectively, the transmitted signals \( X_j[1],...,X_j[n] \), the received signals \( Y_j[1],...,Y_j[n] \) and the noise terms \( N_j[1],...,N_j[n] \). For a set of nodes \( S \), we will define \( X_S \) to be the set of all \( X_i \)'s, for \( v_i \in S \). Then, if we have \( X^n_2 \), we have a set of length \( n \) vectors. 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

\[
nR_1 = H(W_1) = I(W_1;Y^n_{d_1}) + H(W_1|Y^n_{d_1}) \leq I(W_1;Y^n_{d_1}) + n\epsilon_n \leq I(X^n_{s_1};Y^n_{v_1}) + n\epsilon_n \tag{2}
\]

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 \( W_1 \leftrightarrow X^n_{s_1} \leftrightarrow Y^n_{v_1} \leftrightarrow Y^n_{d_1} \).

Fig. 4. Three categories of networks which are not in case (A). Notice that not all nodes are explicitly shown for the sake of generality. Each line represents a path, not necessarily a link, with any number of nodes.
For $R_2$, we have

$$nR_2 = H(W_2) = I(W_2; Y^n_{d_2}) + H(W_2|Y^n_{d_2})$$

$$(i) \leq I(W_2; Y^n_{d_2}) + n\epsilon_n \leq I(X^n_{s_2}; Y^n_{v}, X^n_{s_1}) + n\epsilon_n$$

$$(ii) \leq I(X^n_{s_2}; Y^n_{v}|X^n_{s_1}) + n\epsilon_n$$

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 $W_2 \leftrightarrow X^n_{s_2} \leftrightarrow (Y^n_{v}, X^n_{s_1}) \leftrightarrow Y^n_{d_2}$; and $(ii)$ follows since $X^n_{s_1}$ is independent of $X^n_{s_2}$. Now, by adding inequalities (2) and (3), we obtain

$$n(R_1 + R_2) = I(X^n_{s_1}; Y^n) + I(X^n_{s_2}; Y^n; X^n_{s_1}) + n\epsilon_n$$

$$= I(X^n_{s_1}; X^n_{s_2}; Y^n) + n\epsilon_n$$

$$= h(Y^n) - h(Y^n|X^n_{s_1}, X^n_{s_2}, X^n_{s_2}(v)) + n\epsilon_n$$

$$\leq h(Y^n) - h(N^n) + n\epsilon_n$$

$$\leq \frac{n}{2} \log \left( 1 + \left( \sum_{u \in \mathcal{I}(v)} |h_{u,v}| \right)^2 P \right)$$

$$- \frac{n}{2} \log(2\pi e) + n\epsilon_n$$

$$= \frac{n}{2} \log \left( 1 + \left( \sum_{u \in \mathcal{I}(v)} |h_{u,v}| \right)^2 P \right)$$

$$= \frac{2}{\pi} e$$

$$+ n\epsilon_n \leq \frac{n}{2} \log(\beta P) + n\epsilon_n,$$

where $\beta$ is a constant which does not depend on $P$, for $P$ sufficiently large. Therefore we conclude that

$$D_S \leq \lim_{P \to \infty} \lim_{n \to \infty} \frac{\log(\beta P) + 2\epsilon_n}{\log P} = 1.$$ 

In order to simplify the converse proofs for $(\Lambda')$ and $(C)$, we will consider a decomposition of the additive Gaussian noise $N_j$ associated with each node $v_j$. More specifically, if $m = |\mathcal{I}(v_j)|$, we break the noise at node $v_j$ into $m$ independent noise components, each with variance $1/m$. Then we associate each of these components with one of the incoming edges, and we can define, for $v_j \in \mathcal{I}(v_j)$,

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

where $N_{i,j}$ is the noise term associated with the edge $(v_i, v_j)$. Clearly, we have $N_j = \sum_{i: i \in \mathcal{I}(v_j)} N_{i,j}$, and $N_j$ has unit variance. Notice that we can now write, for a node $v_j$, $Y_j = \sum_{i: i \in \mathcal{I}(v_j)} \tilde{X}_{i,j}$. Moreover, we will define

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

As before, we let $\tilde{X}_S$ be the set of all $\tilde{X}_i$’s, for $v_i \in S$, and $\tilde{X}_i^n$ be a length $n$ vector with all the $\tilde{X}_i[m]$’s, for $m = 1, \ldots, n$.

In order to find upper bounds to the rates, we will often be interested in showing that certain conditional mutual information terms can be upper bounded by a constant. In particular, if we have a $Z$ structure across two layers in the network, such as the one shown in Figure 6a, we would like to say that $I(X^n;s) \leq I(X^n;s) \leq I(X^n;s)$ can be upper bounded by a constant that does not depend on $P$. Intuitively, the reason is that given $X^n$ and $Y^n$, one can subtract $X^n_{a,b}$ from $Y^n$ and obtain $X^n_{a,b}$. This means that “almost all” information in $X^n$ can be deduced from $(Y^n, X^n)$, and thus the conditional mutual information cannot be very large. This reasoning is formalized in the following lemma, where we generalize the $Z$ structure to one where $|\mathcal{I}(v_b)| \geq 2$ and $|\mathcal{O}(v_c)| \geq 2$, as shown in Figure 6b. Moreover, we generalize this notion to the case where the mutual information may be conditioned on other signals as well, provided that these signals do not contain information about $X^n_{a,b}$, for some $v_d \in D$. The proof can be found in Appendix A.

**Lemma 1.** Suppose we have nodes $v_b$ and $v_c$ such that $(v_c, v_b) \in E$, and let $A = \mathcal{I}(v_b) \setminus \{v_c\}$ and $D = \mathcal{O}(v_c) \setminus \{v_b\}$. Suppose, in addition, that we have a set of nodes $S$ such that, if $u \in \mathcal{O}(v_c)$ and $w \in S$, we have $u \not\sim w$, and a set of nodes $T$ with the property that, if $u \in D$ and $w \in T$, then $u \not\sim w$. Then, we have

$$I(X^n_S; X^n_c, X^n_b) \leq nK,$$

where $K$ is a constant that is only a function of the channel gains and the network graph $G$.

**Remarks:** If, in the statement of Lemma 1, we condition the mutual information on $X^n_T$ instead of $X^n_T$ the same result holds. Also, if instead of conditioning on $X^n_A$ and $X^n_B$ we condition on $X^n_{a,b}$, the same result holds, since, in the proof, we use $X^n_A$ and $X^n_B$ to construct $X^n_{a,b}$. We will consider these cases to be covered by Lemma 1 as well.
We can now proceed to the proof of case (A’) in Theorem 1. 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 \(\mathcal{A} \triangleq \{v \in V : s_2 \not\sim v\}\), and we notice that \(\mathcal{I}(v_1) \setminus \{v_2\} \subset A\), since, otherwise, we would have a node \(v_0 \in \mathcal{I}(v_1) \setminus \{v_2\}\) such that \(s_2 \not\sim v_0\), and this would contradict the fact that the removal of \(v_2\) disconnects \(s_2\) from \(d_1\). Moreover, \(v_2 \not\in A\), because all paths from \(s_2\) to \(d_2\) contain \(v_2\) and we must have at least one such path. Thus we have

\[
\begin{align*}
nR_1 & \leq I(W_1; Y^n_{d_1}) + n\epsilon_n \leq I(\hat{X}_A^n; Y^n_{d_1}) + n\epsilon_n \\
& \leq I(\hat{X}_A^n, X^n_{s_2}; Y^n_{d_1}) - I(X^n_{s_2}; Y^n_{d_1} | \hat{X}_A^n) + n\epsilon_n \\
& \leq \frac{n}{2} \log P + nK_1 - I(X^n_{s_2}; Y^n_{d_1} | \hat{X}_A^n) + n\epsilon_n, \quad (5)
\end{align*}
\]

where (i) follows because \(v_1\) disconnects \(d_1\) from both sources and \(s_1 \in A\), thus we have \(W_1 \leftrightarrow \hat{X}_A^n \leftrightarrow Y^n_{d_1} \leftrightarrow Y^n_{d_2}\); and (ii) follows because \(\mathcal{I}(v_1) \setminus \{v_2\} \subset A\) and \(v_2 \not\in A\), hence we can upper-bound \(I(\hat{X}_A^n, X^n_{s_2}; Y^n_{d_1})\) as

\[
I(\hat{X}_A^n, X^n_{s_2}; Y^n_{d_1}) = h(Y^n_{d_1}) - h(Y^n_{d_1} | \hat{X}_A^n, X^n_{s_2})
\]

\[
= h(Y^n_{d_1}) - h(N^n_{d_2})
\]

\[
\leq \frac{n}{2} \log \left( \left( 1 + \frac{\sum_{u \in \mathcal{I}(v_1)} h_{u,v_1}^i} {2\epsilon / |\mathcal{I}(v_1)|} \right)^2 \right)
\]

\[
\leq \frac{n}{2} \log P + nK_1, \quad (6)
\]

where \(\gamma\) and \(K_1\) are constants which are 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

\[
\begin{align*}
nR_2 & \leq I(W_2; Y^n_{d_2}) + n\epsilon_n \leq I(W_2; \hat{X}_A^n, \hat{X}_A^n) + n\epsilon_n \\
& \leq I(W_2; \hat{X}_A^n, \hat{X}_A^n) + n\epsilon_n \\
& \leq I(X^n_{s_2}; X^n_{s_2} | \hat{X}_A^n) + n\epsilon_n \\
& \leq I(X^n_{s_2}; Y^n_{d_1} | \hat{X}_A^n) + n\epsilon_n \\
& = I(X^n_{s_2}; Y^n_{d_1} | \hat{X}_A^n) + I(X^n_{s_2}; X^n_{s_2} | \hat{X}_A^n, Y^n_{d_1}) + n\epsilon_n \\
& \leq I(X^n_{s_2}; Y^n_{d_1} | \hat{X}_A^n) + nK_2 + n\epsilon_n, \quad (7)
\end{align*}
\]

where (i) follows from the fact that the removal of \(v_2\) and \(A\) disconnects \(d_2\) from both sources, which implies \(W_2 \leftrightarrow (\hat{X}_A^n, \hat{X}_A^n) \leftrightarrow Y^n_{d_2}\); (ii) follows from the fact that \(W_2\) is independent of \(\hat{X}_A^n\); (iii) follows from the fact that, given \(\hat{X}_A^n\), we have \(W_2 \leftrightarrow \hat{X}_A^n \leftrightarrow Y^n_{d_2}\); (iv) follows from the application of Lemma 1 to \(I(X^n_{s_2}; X^n_{s_2} | \hat{X}_A^n, Y^n_{d_1})\), since \(\mathcal{I}(v_1) \setminus \{v_2\} \subset A\).

Finally, by adding (5) and (7) we obtain

\[
n(R_1 + R_2) \leq \frac{n}{2} \log P + n(K_1 + K_2) + n\epsilon_n,
\]

and we conclude that \(D_2 \leq 1\). Since one degree-of-freedom is trivially achievable, we have \(D_2 = 1\) for both (A) and (A’).

V. NETWORKS WITH TWO DEGREES-OF-FREEDOM

In this section, we will provide achievability schemes for the networks which fall into cases (B) and (B’). In order to describe these schemes we will proceed as follows. We will first identify the key layers, whose nodes will be responsible for performing non-trivial relaying operations. All the nodes which do not belong to the key layers will simply forward their received signal. This will allow us to build a condensed version of the network. The condensed network only contains the nodes in the key layers, \(V_1\) and \(V_r\). The edges and respective channel gains are determined according to the effective transfer matrices between two consecutive layers of the condensed network, which are obtained by assuming that all intermediate nodes which are not in the key layers are essentially the same condensed network, and therefore we may use the key layers and the condensed network. If the resulting condensed network is a condensed network. We will refer to the effective channel gains of the edges in the condensed network by \(\hat{h}(\nu, u)\), where \(\nu\) is the starting node and \(u\) is the ending node. For example, in Figure 7, we have \(\hat{h}(s_2, v_3) = h_2h_7 + h_3h_8\) and \(\hat{h}(v_2, d_2) = 0\). Notice that, in the condensed network, the effective additive noises at the nodes are not necessarily independent and identically distributed. However, they are still drawn from continuous distributions, which will be sufficient for us.

The condensed networks will be useful since we will conclude that entire classes of layered networks will possess essentially the same condensed network, and therefore we may describe a single achievability scheme for all the networks in that class. We will describe achievability schemes for \(D_{12} = 2\) in essentially two ways, according to the structure of the condensed network. If the resulting condensed network is a \(2 \times 2\) interference channel, then we will use the scheme described in [30] to achieve \(D_{12} = 2\). Otherwise, we will describe a simple amplify-and-forward scheme that guarantees that the end-to-end transfer matrix for the condensed network (and thus for the original network as well) is of the form

\[
\begin{bmatrix}
\beta_1 & 0 \\
0 & \beta_2
\end{bmatrix}
\]

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 \(d_i\). Since the scaling factors used at the key layers and the noise variances are functions of the channel gains only (and not the power \(P\) of the signals transmitted by the sources), we have essentially two parallel point-to-point AWGN channels.

![Fig. 7. A 5-layer network (a) and its 3-layer condensed version (b)](image-url)
In order to make sure that the output power constraint is satisfied at all nodes, we will restrict the sources to using power $\alpha P$, for some $\alpha \in (0,1)$. It is not difficult to see that, for $P$ sufficiently large, $\alpha$ can be chosen independent of $P$. The effective additive noise at the destination will be linear combinations of the individual Gaussian noises at each node, where the coefficients are functions of the channel gains $h_e$. Therefore, $\sigma^2_e$, the variance of the additive Gaussian noise at destination $d_i$, is a function of $P$, and each source-destination pair $(s_i, d_i)$, for $i = 1, 2$, can use Gaussian random codes to achieve rate

$$R_i = \frac{1}{2} \log \left(1 + \alpha \frac{\beta_i^2 P}{\sigma^2_e} \right),$$

and, therefore, one degree-of-freedom. We conclude that we achieve $D_2 = 2$.

First, we will consider (B), in which case we have two disjoint paths with manageable interference.

A. Two disjoint paths with manageable interference

We let $P_{s_1,d_1}$ and $P_{s_2,d_2}$ be our two disjoint paths such that we have $S \subset V$ containing $P_{s_1,d_1}$ and $P_{s_2,d_2}$ and satisfying $n_1(G[S]) \neq 1$ and $n_2(G[S]) \neq 1$. In general, we will assume that $S$ is chosen to be minimal, and all the nodes in $V \setminus S$ are removed from the network. If we have $n_1(G[S]) = 0$ and $n_2(G[S]) = 0$, then achieving $D_2 = 2$ is trivial: we have two disjoint paths $P_{s_1,d_1}$ and $P_{s_2,d_2}$ with no interference whatsoever. For networks where $n_1(G[S]) \geq 2$, for $i = 1$ or $i = 2$, we will define $v^i_p$ to be the first node on $P_{s_i,d_i}$ whose removal disconnects $d_i$ from $s_i$. Notice that $V_{(v^i_p)}-1$ is the layer containing $I(v^i_p)$. This layer will be used as one of the key layers. Intuitively, this is the last layer where we can choose the scaling used at the nodes so that the interference on $P_{s_i,d_i}$ is canceled. If $n_1(G[S]) \geq 2$ and $n_2(G[S]) = 0$, for $i = 1$ or $i = 2$, our condensed network will be a two-hop network formed by layers $V_1, V_{(v^i_p)}-1$ and $V_e$. If $n_1(G[S]) \geq 2$ and $n_2(G[S]) \geq 2$, our condensed network will be a three-hop network formed by layers $V_1, V_{(v^i_p)}-1, V_{(v^i_p)}-2$ and $V_e$ (unless $\ell(v^i_p) = \ell(v^2_p)$, in which case the condensed network will be a two-hop network). We will need the following technical lemma about $v^i_p$, whose proof can be found in the Appendix.

**Lemma 2.** Assume $n_i(G[S]) \geq 2$, for $i = 1$ or $i = 2$, and let $v^i_p$ be defined as above. Then, there exist two paths $P_{s_1,v^i_p}$ and $P_{s_2,v^i_p}$ such that $P_{s_1,v^i_p} \cap P_{s_2,v^i_p} = \{v^i_p\}$.

The importance of Lemma 2 is that it guarantees that the transfer matrix between $(s_1, s_2)$ and two nodes in $I(v^i_p)$ will be invertible with probability 1. This will be further explained later but, intuitively, it is necessary to give the nodes in $I(v^i_p)$ freedom to cancel the interference from $s_i$ on $P_{s_i,d_i}$. A second useful property about $v^i_p$ is now stated in the form of another Lemma.

**Lemma 3.** Assume $n_i(G[S]) \geq 2$, for $i = 1$ or $i = 2$, and let $v^i_p$ be defined as above. Then, there are (at least) two nodes $v_1, v_2 \in I(v^i_p)$ such that $s_i \sim v_1$ and $s_i \sim v_2$.

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

The importance of the property in Lemma 3 is that it guarantees that, with probability 1, at least two nodes in $I(v^i_p)$ will have in their received signal a component which corresponds to the transmitted signal from $s_i$. Intuitively, this means that, we can cancel the interference from $s_i$ on $P_{s_i,d_i}$, while still allowing the signal from $s_i$ to reach $d_i$. We now consider the case in which we have $n_1(G[S]) \geq 2$ and $n_2(G[S]) = 0$.

1) $n_1(G[S]) \geq 2$, $n_2(G[S]) = 0$: Notice that in this case only $v^1_p$ is defined. Thus, we will consider the condensed network formed by layers $V_1, V_{(v^1_p)}-1$ and $V_e$, with $m = |V_{(v^1_p)}-1|$. Our condensed network should look like the network in Figure 8. The solid lines correspond to edges that must exist in the condensed network, due to the existence of two disjoint paths $P_{s_1,d_1}$ and $P_{s_2,d_2}$. The dashed lines correspond to edges that may or may not exist. To each of the nodes $v_i, i = 1, \ldots, m$ in the intermediate layer, we associate a variable $x_i$ which will be the scaling factor used by node $v_i$.

Our task is to show that the end-to-end transfer matrix, given by

$$T = \begin{pmatrix}
\hat{h}(v_1,d_1) & \hat{h}(v_2,d_1) & \cdots & \hat{h}(v_m,d_1) \\
\hat{h}(v_1,d_2) & \hat{h}(v_2,d_2) & \cdots & \hat{h}(v_m,d_2) \\
0 & x_2 & \cdots & 0 \\
0 & 0 & \cdots & x_m
\end{pmatrix}
\begin{pmatrix}
\hat{h}(s_1,v_1) & \hat{h}(s_2,v_1) \\
\hat{h}(s_1,v_2) & \hat{h}(s_2,v_2) \\
\vdots & \vdots & \ddots & \vdots \\
\hat{h}(s_1,v_m) & \hat{h}(s_2,v_m)
\end{pmatrix},$$

where $T_{jk} = \sum_{i=1}^{m} \hat{h}(s_i,v_i)\hat{h}(v_i,d_k)x_i$, can be made diagonal with non-zero diagonal entries by an appropriate choice of $x_1, \ldots, x_m$. 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 $i = 1, \ldots, m$ and $T_{2,1}$ is always 0. From the use of Lemma 2, we know that for two nodes $v_a, v_b \in I(v^1_p) \subset V_{(v^1_p)}-1$ with associated variables $x_a$ and $x_b$, we must have two disjoint
From Lemma 3, we know that there is a node $v_c \in \mathcal{I}(v_p^1) \subset V_{\ell(v_p^2)}$, such that $s_2 \sim v_c$ and $c \neq m$. We now claim that if the matrices

$$M_1 = \begin{bmatrix} \hat{h}(s_1, v_a) & \hat{h}(v_a, d_1) \\ \hat{h}(s_2, v_a) & \hat{h}(v_a, d_1) \end{bmatrix} \text{ and } M_2 = \begin{bmatrix} \hat{h}(s_2, v_c) & \hat{h}(v_c, d_2) \\ \hat{h}(s_2, v_m) & \hat{h}(v_m, d_2) \end{bmatrix}$$

are both full-rank, then we can choose $x_1, ..., x_m$ so that $T$ is diagonal with non-zero diagonal entries. To see this, we first consider $x' = [x'_1 \ ... \ x'_m]$, 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 guarantees that $T_{1,1} = 1$ and $T_{1,2} = 0$. If $T_{2,2} = 0$ we are done. Otherwise, if $T_{2,2} = 0$, we let $x'' = [x''_1 \ ... \ x''_m]$, where $x''_j = 0$ for $j \neq c, m$ and $[x''_c \ x''_m]^T = M_2^{-1}[0 \ 1]^T$. This choice guarantees that $T_{1,2} = 0$ and $T_{2,2} = 1$. If we have $T_{1,1} \neq 0$, we are done. Otherwise, we set $x''' = x' + x''$. By linearity, this choice will guarantee that $T$ is the identity matrix.

Next we show that, with probability 1, $M_1$ and $M_2$ (which are just functions of the channel gains in the original network) 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 where the variables are the channel gains $h_{e}$ from the original network. All we need to show is that this polynomial is not identically zero. Then, since the $h_{e}$’s are drawn independently from continuous distributions, det $Z_1$ will be non-zero with probability 1. To see that this polynomial is not identically zero we notice that the existence of two disjoint paths $P_{s_1, v_a}$ and $P_{s_2, v_b}$ guarantees that, if we set $h_e \equiv 1$ if $e$ connects two consecutive vertices of $P_{s_1, v_a}$ or $P_{s_2, v_b}$ and $h_e \equiv 0$ otherwise, $Z_1$ will be the identity matrix. Therefore, $Z_1$ will be invertible, and thus det $Z_1$ cannot be identically zero. Now, we notice that

$$\text{det } Z_1 = \begin{bmatrix} \hat{h}(s_1, v_a) & \hat{h}(s_2, v_a) \\ \hat{h}(s_1, v_b) & \hat{h}(s_2, v_b) \end{bmatrix} \begin{bmatrix} \hat{h}(v_a, d_1) & \hat{h}(v_a, d_1) \\ \hat{h}(v_b, d_1) & \hat{h}(v_b, d_1) \end{bmatrix} = \hat{h}(v_a, d_1)\hat{h}(v_b, d_1) \text{ det } Z_1.$$

Since $v_a \sim d_1$ and $v_b \sim d_1$, 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 with probability 1. To show that $M_2$ is invertible with probability 1, we will follow very similar steps. We notice that the transfer matrix between $(v_c, v_m)$ and $(d_1, d_2)$ is given by

$$Z_2 = \begin{bmatrix} \hat{h}(v_c, d_1) & \hat{h}(v_m, d_1) \\ \hat{h}(v_c, d_2) & \hat{h}(v_m, d_2) \end{bmatrix}.$$

Since $v_c \in \mathcal{I}(v_p^1)$ and $v_p^1 \in P_{s_1, d_1}$, we clearly have two disjoint paths $P_{v_c, d_1} = (v_c, v_p^1, P_{s_1, d_1}[v_p^1, d_1])$ and $P_{v_m, d_2} = P_{s_2, d_2}[v_m, d_2]$. This implies that det $Z_2$ is non-identically zero, and therefore non-zero with probability 1. Then, we notice that

$$\text{det } M_2 = \begin{bmatrix} \hat{h}(s_2, v_c)\hat{h}(v_c, d_1) & \hat{h}(s_2, v_m)\hat{h}(v_m, d_1) \\ \hat{h}(s_2, v_c)\hat{h}(v_c, d_2) & \hat{h}(s_2, v_m)\hat{h}(v_m, d_2) \end{bmatrix} = \hat{h}(s_2, v_c)\hat{h}(s_2, v_m) \text{ det } Z_2,$$

and since $s_2 \sim v_c$, $s_2 \sim v_m$, we have that $\hat{h}(s_2, v_c)\hat{h}(s_2, v_m)$ is a non-identically zero polynomial in the $h_{e}$’s and therefore so is det $M_2$. This proves that $M_2$ is full-rank with probability 1, and thus we conclude the proof when $n_1(G[S]) \geq 2$, $n_2(G[S]) = 0$. The case where $n_1(G[S]) = 0$, $n_2(G[S]) \geq 2$ follows in the exact same way.

Next, we consider the cases in which $n_1(G[S]) \geq 2$ and $n_2(G[S]) \geq 2$. We will use $V_{l(v_p^2)} - 1$ and $V_{l(v_p^1)} - 1$ as our key layers. We can assume WLOG that $\ell(v_p^2) < \ell(v_p^1)$ and $\ell(v_p^2)^2 < \ell(v_p^1)$ separately.

2) $n_1(G[S]) \geq 2$, $n_2(G[S]) \geq 2$ and $\ell(v_p^2)^2 < \ell(v_p^1)$: We let $m = |V_{l(v_p^2)} - 1|$ and $n = |V_{l(v_p^1)} - 1|$. Our condensed network will be of the form shown in Figure 9a. Once again, the solid lines correspond to edges that must exist in the condensed network, due to the existence of two disjoint paths $P_{s_1, d_1}$ and $P_{s_2, d_2}$, and the dashed lines correspond to edges that may or may not exist. We name the nodes in $V_{l(v_p^2)} - 1$, $v_1, v_2, ..., v_m$ and the nodes in $V_{l(v_p^1)} - 1$, $u_1, u_2, ..., u_n$. Moreover, to each of the nodes $v_i$, $i = 1, ..., m$, we associate a variable $x_i$ which will be the scaling factor used by node $v_i$, and to each of the nodes $u_i$, $i = 1, ..., n$ we associate a variable $y_i$, which will be the scaling factor used by node $u_i$.

We will again show that, with probability 1, there is a choice of $x_1, ..., x_m$ and $y_1, ..., y_n$ such that the effective end-to-end transfer matrix is diagonal with non-zero diagonal entries. This time, however, we will proceed in two steps. First we will show that, with probability 1, we can choose $y_1, ..., y_n$ such that, for some $v_a, v_b \in \mathcal{I}(v_p^1)$, the transfer matrix between $(s_1, s_2)$ and $(v_a, v_b)$ is invertible and the transfer matrix between $(s_1, s_2)$ and $v_m$ is of the form $[0 \ \beta]$ for $\beta \neq 0$. Then, by “suppressing” the key layer $V_{l(v_p^2)} - 1$, we will essentially be in the case we described in V-A1, and thus we can choose $x_1, ..., x_m$ so that the end-to-end transfer matrix is as desired.
In order to describe how we choose \( y_1, ..., y_n \), we must first consider the connectivity between the nodes in \( V_{(v_2^0)} \) and its consecutive layer, \( V_{(v_2^1)} \), in the original network. This layer transition can be depicted as in Figure 9b. We will now show that, with probability 1, it is possible to choose \( y_1, ..., y_n \) all non-zero, such that the transfer matrix \( F \) between \((s_1, s_2)\) and \( v_2^0 \) is of the form \([0 \alpha]\) for \( \alpha \neq 0 \). We first notice that \( F \) is given by

\[
\left( \sum_{i=1}^{n} h(s_1, u_i) h(u_i, v_2) y_i \right) \left( \sum_{i=1}^{n} h(s_2, u_i) h(u_i, v_2) y_i \right).
\]

From Lemma 3, we know that there are at least two nodes \( u_c, u_d \in \mathcal{I}(v_2^0) \) such that \( s_1 \sim u_c \) and \( s_1 \sim u_d \). This implies that \( h(s_1, u_c) h(u_c, v_2) \) and \( h(s_1, u_d) h(u_d, v_2) \), if viewed as polynomials on the channel gains, are not identically zero. Thus, with probability 1, they will be non-zero, and \( \sum_{i=1}^{n} h(s_1, u_i) h(u_i, v_2) y_i \) will have non-zero coefficients in front of \( y_c \) and \( y_d \). This means that we can choose \( y' = (y'_1, ..., y'_n) \), with \( y'_1, ..., y'_n \) all non-zero, so that \( F_1 = \sum_{i=1}^{n} h(s_1, u_i) h(u_i, v_2) y'_i = 0 \). If we have \( F_2 = \sum_{i=1}^{n} h(s_2, u_i) h(u_i, v_2) y'_i \neq 0 \), then we are done. Otherwise, if \( F_2 = 0 \), we proceed as follows. From Lemma 2, we know that we can choose \( u_a, u_b \in \mathcal{I}(v_2^0) \subset V_{(v_2^0)} \) so that we have two disjoint paths \( P_{s_1, u_a} \) and \( P_{s_2, u_b} \). Therefore, the transfer matrix between \((s_1, s_2)\) and \((u_a, u_b)\), given by

\[
K = \begin{bmatrix} h(s_1, u_a) & h(s_1, u_b) \\ h(s_2, u_a) & h(s_2, u_b) \end{bmatrix},
\]

is full-rank with probability 1. This also implies that the matrix

\[
M = \begin{bmatrix} h(s_1, v_2) & h(s_1, v_2) \\ h(s_2, v_2) & h(s_2, v_2) \end{bmatrix}
\]

is full-rank with probability 1, because we have \( \det M = h(s_1, v_2) h(s_2, v_2) \neq 0 \) \( \alpha = 0 \). Moreover, since there exists a path from \( v_2^0 \) to \( v_m \), and there exists no path from \( s_1 \) to \( v_m \) which does not contain \( v_2^0 \), we conclude that, with probability 1, our choice of \( y_1, ..., y_n \) will make the transfer matrix from \((s_1, s_2)\) to \( v_m \) be of the form \([0 \beta]\) for \( \beta \neq 0 \).

Next, we would like to prove that, with this choice of \( y_1, ..., y_n \), there exist nodes \( v_a, v_b \in \mathcal{I}(v_2^0) \), such that the transfer matrix between \((s_1, s_2)\) and \((v_a, v_b)\) is full-rank. First, we notice that, from Lemma 2, there exist two nodes \( v_c, v_f \in \mathcal{I}(v_2^0) \), such that we have two disjoint paths \( P_{s_1, v_c} \) and \( P_{s_2, v_f} \). However, we cannot proceed as before to conclude that the transfer matrix between \((s_1, s_2)\) and \((v_c, v_f)\) is full-rank with probability 1, because our variables \( y_1, ..., y_n \) were not chosen independently from the channel gains. Nonetheless, if we let \( H \) be the set of all \( h(u_i, v_2) \) for \( i = 1, ..., n \) and all the channel gains that appear in \( h(s_1, u_j) \), for \( i = 1, 2 \) and \( j = 1, ..., n \), we notice that our choice of \( y_1, ..., y_n \) only depends on \( H \). Therefore, we assume that all the channel gains in \( H \) are drawn according to their distributions, and are from now on viewed as constants. Then, we can also fix \( y_1, ..., y_n \), following the steps described previously, and view them as constants.

First, we assume that neither \( P_{s_1, v_c} \) nor \( P_{s_2, v_f} \) contains \( v_2^0 \). In this case we will show that we can set \( v_c = v_e \) and \( v_f = v_f \). The determinant of the transfer matrix between \((s_1, s_2)\) and \((v_e, v_f)\) can be seen as a polynomial where the variables are the channel gains which are not in \( H \). Notice that all the channel gains not in \( H \) are still independent (since the choice of \( y_1, ..., y_n \) was made independent of them) and have a continuous distribution. Thus, we will show that, with probability 1 over the choice of the channel gains in \( H \), there exists a choice of the channel gains which are not in \( H \), such that the transfer matrix between \((s_1, s_2)\) and \((v_e, v_f)\) is invertible. Therefore, the determinant of the transfer matrix between \((s_1, s_2)\) and \((v_e, v_f)\) is not identically zero, and will be non-zero with probability 1 over the choice of the channel gains not in \( H \).

Since \( P_{s_1, v_e} \) and \( P_{s_2, v_f} \) are disjoint, there are distinct nodes \( u_e \) and \( u_f \) in \( V_{(v_2^0)} \) such that \( u_e \in P_{s_1, v_e} \) and \( u_f \in P_{s_2, v_f} \). For any \( h_e \notin H \), we will set \( h_e = 1 \) if \( e \) connects two consecutive vertices of \( P_{s_1, v_e} \) or \( P_{s_2, v_f} \) and \( h_e = 0 \) otherwise. Therefore, the transfer matrix between \((u_e, u_f)\) and \((v_e, v_f)\) is the identity matrix. Thus, we have that the transfer matrix between \((s_1, s_2)\) and \((v_e, v_f)\) is given by

\[
\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_e & 0 \\ 0 & y_f \end{bmatrix} \begin{bmatrix} h(s_1, u_e) & h(s_1, u_f) \\ h(s_2, u_e) & h(s_2, u_f) \end{bmatrix}.
\]

The existence of disjoint paths \( P_{s_1, v_e} \) and \( P_{s_2, v_f} \) implies the existence of disjoint paths \( P_{s_1, u_e} = P_{s_1, v_e} \) \((s_1, u_e)\) and \( P_{s_2, u_f} = P_{s_2, v_f} \) \((s_2, u_f)\). Therefore, with probability 1 over the choice of the channel gains in \( H \) (since they were drawn independently first, according to their continuous distributions),

\[
\begin{bmatrix} h(s_1, u_e) & h(s_2, u_e) \\ h(s_1, u_f) & h(s_2, u_f) \end{bmatrix},
\]

is full-rank. Therefore, since we chose \( y_e \) and \( y_f \) to be non-zero, the transfer matrix in \((9)\) must be full-rank, which implies that the transfer matrix between \((s_1, s_2)\) and \((v_e, v_f)\) is full-rank with probability 1 if \( y_1, ..., y_n \) are chosen as described above.

Now, we consider the situations in which either \( P_{s_1, v_c} \) or \( P_{s_2, v_f} \) contains \( v_2^0 \). We will show that, in any case, for some \( v_g, v_h \in \mathcal{I}(v_1^1) \), we can find either

i. two other disjoint paths \( P_{s_1, v_g} \) and \( P_{s_2, v_h} \)

or

ii. two disjoint paths \( P_{s_1, v_g} \) and \( P_{s_2, v_h} \).

If we suppose \( v_2^0 \in P_{s_1, v_e} \), then we are clearly in case ii, by setting \( g = e \) and \( h = f \), and setting \( v_2^0 = P_{s_1, v_e} \). Thus, we suppose that \( v_2^0 \notin P_{s_1, v_e} \). If we let \( v_f \) be the node from \( P_{s_2, v_f} \) in the layer containing \( v_2^0 \), we have two disjoint paths \( P_{s_1, v_2^0} = P_{s_1, v_e} \) \((s_1, v_2^0)\) and \( P_{s_2, v_f} = P_{s_2, v_f} \) \((s_2, v_f)\). We also let \( u_1 \) be the node from \( P_{s_1, d_1} \) in \( V_{(v_2^0)} \). Then we let \( v_1 \) be the last common node between \( P_{s_2, v_f} \) and
Clearly, we have the disjoint paths $P_{v_1,v_2}$ and $P_{v_2,v_3}$. This implies that we have a path $P_{v_1,v_f} = P_{s_1,v_f} \cup P_{s_2,v_f}$ and a path $P_{v_2,v_e} = P_{s_1,v_e} \cup P_{s_2,v_e}$ which are disjoint, and we are in case ii. Note that this case also includes $v_f = v_1$. If, instead, $v_f \in P_{s_2,v_2}$ (Figure 10b), we must have disjoint paths $P_{s_1, v_1}$ and $P_{s_2, v_f}$. We also clearly have two disjoint paths $P_{v_1,v_2}$ and $P_{v_2,v_3}$, and if we choose $y_1, \ldots, y_n$ as described previously, then, for some $v_g, v_h \in \mathcal{I}(v_1)$, the transfer matrix between $(s_1, s_2)$ and $(v_g, v_h)$ will be full-rank with probability 1 over the choice of the channel gains not in $H$. We will look at the determinant of the transfer matrix between $(s_1, s_2)$ and $(v_g, v_h)$ as a polynomial on the channel gains not in $H$, since the channel gains in $H$ and the scaling factors $y_1, \ldots, y_n$ have already been fixed. Then we can show that this determinant is not identically zero by showing that for a specific choice of the channel gains not in $H$, the transfer matrix between $(s_1, s_2)$ and $(v_g, v_h)$ is full-rank. For any $\hat{h}_e$ not in $H$, we will choose $\hat{h}_e = 1$ if $e$ is connecting two consecutive vertices of $P_{s_1,v_g}$ or $P_{s_2,v_h}$, and $\hat{h}_e = 0$ otherwise. This means that the transfer matrix between $(u_g, v_2)$ and $(v_g, v_h)$ is the identity matrix. Then, if we let $u_g$ be the node from $P_{s_1,v_g}$ in layer $V_{l(u_g)-1}$, the transfer matrix between $(s_1, s_2)$ and $(v_g, v_h)$ is given by

\[
\begin{pmatrix}
1 & 0 \\
0 & 1 \\
\end{pmatrix}
\begin{pmatrix}
y_g & 0 \\
0 & 1 \\
\end{pmatrix}
\begin{pmatrix}
\hat{h}(s_1, u_g) & 0 \\
0 & \hat{h}(s_2, u_h) \\
\end{pmatrix}
\begin{pmatrix}
y_1 & y_2 \\
y_2 & y_3 \\
\end{pmatrix}
\begin{pmatrix}
h(s_2, u_g) & h(s_2, u_h) \\
h(s_1, u_g) & h(s_1, u_h) \\
\end{pmatrix}
\begin{pmatrix}
y_g \hat{h}(s_1, u_g) & y_2 \hat{h}(s_2, u_g) \\
0 & 1 \\
\end{pmatrix},
\]

(10)

where we used the fact that our choice of $y_1, \ldots, y_n$ guarantees that the transfer matrix between $(s_1, s_2)$ and $u^2_p$ is $[0 \alpha]$ for some $\alpha \neq 0$. Now, since there exists a path from $s_1$ to $u_g$, $\hat{h}(s_1, u_g)$ is non-zero with probability 1 over the choice of the channel gains in $H$. Therefore, since $y_g$ was chosen to be non-zero, the transfer matrix in (10) is upper-triangular (with non-zero diagonal entries) and thus full-rank.

Therefore we proved that we can find $u_a, v_b \in \mathcal{I}(v_2)$ so that the transfer matrix between $(s_1, s_2)$ and $(u_a, v_b)$ is full-rank with probability 1, after the choice of the scaling factors $y_1, \ldots, y_m$. Next, we consider supressing the layer $V_{l(u_2)-1}$ from the condensed network by incorporating our choice of $y_1, \ldots, y_m$ into the terms $\hat{h}(s_i, v_j)$ for $i = 1, 2$ and $j = 1, \ldots, m$. We will show that the resulting condensed network is equivalent to the one considered in V-A1. As in V-A1, the end-to-end transfer matrix can now be written as

\[
\begin{pmatrix}
\sum_{i=1}^{m} \hat{h}(s_1, v_i) \hat{h}(v_i, d_1) x_i \\
\sum_{i=1}^{m} \hat{h}(s_2, v_i) \hat{h}(v_i, d_2) x_i \\
\end{pmatrix}
\begin{pmatrix}
\sum_{i=1}^{m} \hat{h}(s_1, v_i) \hat{h}(v_i, d_1) x_i \\
\sum_{i=1}^{m} \hat{h}(s_2, v_i) \hat{h}(v_i, d_2) x_i \\
\end{pmatrix},
\]

(11)

As we noted before, the transfer matrix between $(s_1, s_2)$ and $v_m$ is of the form $[0 \beta]$ for some $\beta \neq 0$. This implies that $\hat{h}(s_1, v_m) = 0$ and $\hat{h}(s_2, v_m) = \beta \neq 0$. Moreover, since $v^2_p$ disconnects $d_2$ from $s_1$, we conclude that $\hat{h}(s_1, v_2) \hat{h}(v_2, d_2) = 0$ for $i = 2, \ldots, m$. Otherwise, this would either imply the existence of a path between $s_1$ and $d_2$ not containing $v^2_p$ or contradict the fact that the transfer matrix between $(s_1, s_2)$ and $v^2_p$ is of the form $[0 \alpha]$. Thus, we conclude that $\sum_{i=1}^{m} \hat{h}(s_1, v_i) \hat{h}(v_i, d_2) x_i = 0$. As shown in V-A1, if we can find $v_a, v_b$ and $v_c, c \neq m$, such that the
matrices

\[
M_1 = \begin{bmatrix}
\hat{h}(s_1, v_a)h(v_a, d_1) & \hat{h}(s_1, v_b)h(v_b, d_1) \\
\hat{h}(s_2, v_a)h(v_a, d_1) & \hat{h}(s_2, v_b)h(v_b, d_1)
\end{bmatrix}
\]

and

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

are both full-rank, then it is possible to choose \(x_1, \ldots, x_m\) so that the end-to-end transfer matrix in (11) is diagonal with non-zero diagonal entries. We will choose \(v_a\) and \(v_b\) to be the two nodes in \(\mathcal{I}(v_p^1)\) for which the transfer matrix from \((s_1, s_2)\) to \((v_a, v_b)\)

\[
\begin{bmatrix}
\hat{h}(s_1, v_a) & \hat{h}(s_2, v_a) \\
\hat{h}(s_1, v_b) & \hat{h}(s_2, v_b)
\end{bmatrix}
\]

is full-rank with probability 1. Then, we will notice that

\[
\text{det } M_1 = \begin{vmatrix}
\hat{h}(s_1, v_a)h(v_a, d_1) & \hat{h}(s_1, v_b)h(v_b, d_1) \\
\hat{h}(s_2, v_a)h(v_a, d_1) & \hat{h}(s_2, v_b)h(v_b, d_1)
\end{vmatrix}
\]

\[
= \hat{h}(v_a, d_1)\hat{h}(v_b, d_2)\hat{h}(v_a, d_1)\hat{h}(v_b, d_2)
\]

Since \(v_a, v_b \in \mathcal{I}(v_p^1)\), we have that \(v_a \sim v_1\) and \(v_b \sim v_1\), and \(\hat{h}(v_a, d_1)\hat{h}(v_b, d_1)\) is non-zero with probability 1. Therefore, \(M_1\) is invertible with probability 1.

As we did in V-A1, we use Lemma 3 to guarantee that we can choose \(v_c \in \mathcal{I}(v_p^2)\) such that \(s_2 \sim v_c\) and \(c \neq m\). Then, we notice that the transfer matrix between \((v_c, v_m)\) and \((d_1, d_2)\) is given by

\[
\begin{bmatrix}
\hat{h}(v_c, d_1) & \hat{h}(v_m, d_1) \\
\hat{h}(v_c, d_2) & \hat{h}(v_m, d_2)
\end{bmatrix}
\]

(12)

Since \(v_c \in \mathcal{I}(v_p^2)\) and \(v_p^1 \in \mathcal{P}_{s_1, d_1}\), we clearly have two disjoint paths \(P_{s_1, d_1}\) and \(P_{v_m, d_2}\). This implies that the transfer matrix in (12) is invertible with probability 1. Then, we notice that

\[
\text{det } M_2 = \begin{vmatrix}
\hat{h}(s_2, v_c)h(v_c, d_1) & \hat{h}(s_2, v_m)h(v_m, d_1) \\
\hat{h}(s_2, v_c)h(v_c, d_2) & \hat{h}(s_2, v_m)h(v_m, d_2)
\end{vmatrix}
\]

\[
= \hat{h}(s_2, v_c)\hat{h}(s_2, v_m)\hat{h}(v_c, d_1)\hat{h}(v_m, d_1)
\]

As we noticed before, our choice of \(y_1, \ldots, y_n\) guarantees that \(\hat{h}(s_2, v_m) = \beta \neq 0\). Since \(s_2 \sim v_c\), there must be at least one path \(P_{s_2, v_c}\). If \(P_{s_2, v_c}\) does not contain \(v_p^2\), then the fact that we chose \(y_1, \ldots, y_m\) to be non-zero guarantees that \(\hat{h}(s_2, v_c)\) is non-zero with probability 1. If \(P_{s_2, v_c}\) contains \(v_p^2\), then the fact that the transfer matrix between \((s_1, s_2)\) and \(v_p^2\) is \([0 \alpha]\) for \(\alpha \neq 0\) guarantees that \(\hat{h}(s_2, v_c)\) is non-zero with probability 1. Either way, we conclude that \(M_2\) is invertible with probability 1. This concludes the proof when \(n_1(G[S]) \geq 2\), \(n_2(G[S]) \geq 2\) and \(\ell(v_p^1) > \ell(v_p^2)\).

Next, we consider the situations in which \(\ell(v_p^1) = \ell(v_p^2)\). In this case, our condensed network will only contain three layers, \(V_1, V_{(v_p^1)\rightarrow 1} = V_{(v_p^2)\rightarrow 1}\) and \(V_c\). We will use two different approaches, depending on the size of \(V_{(v_p^1)\rightarrow 1}\).
can be made diagonal with non-zero diagonal entries by an appropriate choice of $x_1, \ldots, x_m$. First, we notice that we can assume that, in the original network, any layer $V_i$ for $i \geq \ell(v_p)$ only contains two nodes. This is because any node in such layer $V_i$ which is not in $P_{s_1,d_1}$ nor $P_{s_2,d_2}$ can be removed since it may not contribute to $v_1(G[S])$ or $v_2(G[S])$ (or that would contradict the fact that $v_1 \hat{\sigma}$ disconnects $d_1$ from $s_2$ and $v_2 \hat{\sigma}$ disconnects $d_2$ from $s_1$). Therefore, the edge configuration between $V_{\ell(v_p) - 1}$ and $\{d_1, d_2\}$ in the condensed network is the same as the edge configuration between $V_{\ell(v_1)} - 1$ and $V_{\ell(v_2)}$ in the original network. It is then easy to see that each $h(v_i, d_j)$, for $i = 1, \ldots, m$ and $j = 1, 2$, when seen as a polynomial in the channel gains, is composed of a single product of variables $h_{c_i}$, one of which is not shared by any other $h$. 

Next we claim that if we can find two sets of nodes $\{v_a, v_b, v_c\} \subset V_{\ell(v_1)} - 1$ and $\{d_a, d_c, v_f\} \subset V_{\ell(v_2)} - 1$, such that the matrices 

$$
M_1 = \begin{bmatrix}
\hat{h}(s_1, v_a)\hat{h}(v_a, d_1) & \hat{h}(s_1, v_b)\hat{h}(v_b, d_1) \\
\hat{h}(s_2, v_a)\hat{h}(v_a, d_1) & \hat{h}(s_2, v_b)\hat{h}(v_b, d_1) \\
\hat{h}(s_1, v_c)\hat{h}(v_c, d_1) & \hat{h}(s_1, v_b)\hat{h}(v_b, d_2) \\
\hat{h}(s_2, v_c)\hat{h}(v_c, d_1) & \hat{h}(s_2, v_b)\hat{h}(v_b, d_2)
\end{bmatrix}
$$

and 

$$
M_2 = \begin{bmatrix}
\hat{h}(s_1, v_a)\hat{h}(v_a, d_2) & \hat{h}(s_1, v_c)\hat{h}(v_c, d_1) \\
\hat{h}(s_2, v_a)\hat{h}(v_a, d_2) & \hat{h}(s_2, v_c)\hat{h}(v_c, d_1) \\
\hat{h}(s_1, v_f)\hat{h}(v_f, d_2) & \hat{h}(s_1, v_c)\hat{h}(v_c, d_1) \\
\hat{h}(s_2, v_f)\hat{h}(v_f, d_2) & \hat{h}(s_2, v_c)\hat{h}(v_c, d_1)
\end{bmatrix}
$$

are full-rank, then we can choose $x_1, \ldots, x_m$ such that the transfer matrix in (13) is diagonal with non-zero diagonal entries. To see this, suppose $M_1$ and $M_2$ are full-rank. Then we can set $x' = [x'_1 \ldots x'_m]$, where $x'_j = 0$ for $j \neq a, b, c$, and 

$$
[x'_a x'_b x'_c]^T = M_1^{-1}[1 0 0]^T.
$$

This guarantees that the transfer matrix in (13) is of the form 

$$
\begin{bmatrix}
1 & 0 \\
0 & \gamma
\end{bmatrix}.
$$

If $\gamma \neq 0$, we achieve our goal with $x'$. If $\gamma = 0$, we set $x'' = [x''_1 \ldots x''_m]$, where $x''_j = 0$ for $j \neq d, e, f$, and $[x''_a x''_b x''_c]^T = M_2^{-1}[1 0 0]^T$. This guarantees that the transfer matrix in (13) is of the form 

$$
\begin{bmatrix}
\delta & 0 \\
0 & 1
\end{bmatrix}.
$$

If $\delta \neq 0$, we achieve our goal with $x''$. If $\delta = 0$, then we let $x''' = x' + x''$, and the transfer matrix in (13) becomes the identity matrix.

Next, we show that we can either find $\{v_a, v_b, v_c\}$ and $\{d_a, d_c, v_f\}$ as described above, or we can remove nodes from $V_{\ell(v_p) - 1}$ so that we have a $2 \times 2$ interference channel (case V-A3). We start by applying Lemma 2 to $\ell(v_p)$. Then we can find $u_a, u_b \in \mathcal{I}(v_p)$ such that there are two disjoint paths $P_{s_1,u_a}$ and $P_{s_2,u_b}$. Then, from Lemma 3 applied to $v_p$, we know that there exist nodes $u_c, u_d \in \mathcal{I}(v_p)$ such that $s_1 \sim u_c$ and $s_1 \sim u_d$. Suppose $\{u_a, u_b\} \neq \{u_c, u_d\}$, then we can assume WLOG that $u_c \neq u_a, u_b$. We choose $\{v_a, v_b, v_c\} = \{u_a, u_b, u_c\}$ and we have 

$$
\det M_1 = \begin{bmatrix}
\hat{h}(s_1, u_a)\hat{h}(u_a, d_1) & \hat{h}(s_1, u_b)\hat{h}(u_b, d_1) & \hat{h}(s_1, u_c)\hat{h}(u_c, d_1) \\
\hat{h}(s_2, u_a)\hat{h}(u_a, d_1) & \hat{h}(s_2, u_b)\hat{h}(u_b, d_1) & \hat{h}(s_2, u_c)\hat{h}(u_c, d_1) \\
\hat{h}(s_1, u_a)\hat{h}(u_a, d_2) & \hat{h}(s_1, u_b)\hat{h}(u_b, d_2) & \hat{h}(s_1, u_c)\hat{h}(u_c, d_2) \\
\hat{h}(s_2, u_a)\hat{h}(u_a, d_2) & \hat{h}(s_2, u_b)\hat{h}(u_b, d_2) & \hat{h}(s_2, u_c)\hat{h}(u_c, d_2)
\end{bmatrix}
$$

The third term in the expansion above can be written as 

$$
\hat{h}(s_1, u_a)\hat{h}(u_a, d_2)\hat{h}(u_a, d_1)\hat{h}(u_b, d_1)\hat{h}(s_1, u_b)\hat{h}(s_2, u_a)\hat{h}(s_2, u_b),
$$

which is a non-identically zero polynomial since $s_1 \sim u_c$, $u_c \in \mathcal{I}(v_p)$, $u_a, u_b \in \mathcal{I}(v_p)$, and there are two disjoint paths $P_{s_1,u_a}$ and $P_{s_2,u_b}$. Moreover, as we noticed before, one of the variables in $h(u_c, d_2)$ is not shared by any other effective channel gain $h$, and therefore, the term above cannot be canceled by the other terms. This allows us to conclude that $M_1$ is full-rank with probability 1. Now, suppose $\{u_a, u_b\} = \{u_c, u_d\}$. This means that the original network must contain the network shown in Figure 14. The curvy lines are used to represent paths. At this point, if $s_2 \sim u_a$, then we can remove the

![Fig. 14. Illustration of a subnetwork contained by the network from V-A4, in the case where \{u_a, u_b\} = \{u_c, u_d\}](image-url)
we have
\[
det M_1 = \begin{vmatrix}
\hat{h}(s_1, u_a)\hat{h}(u_a, d_1) & \hat{h}(s_1, u_b)\hat{h}(u_b, d_1) & \hat{h}(s_1, u'_c)\hat{h}(u'_c, d_1) \\
0 & \hat{h}(s_2, u_a)\hat{h}(u_a, d_1) & \hat{h}(s_2, u'_c)\hat{h}(u'_c, d_1) \\
\hat{h}(s_1, u_a)\hat{h}(u_a, d_2) & \hat{h}(s_1, u_b)\hat{h}(u_b, d_2) & 0
\end{vmatrix}
= -\hat{h}(s_1, u'_c)\hat{h}(u'_c, d_1)\hat{h}(s_2, u_b)\hat{h}(u_b, d_1)\hat{h}(s_1, u_a)\hat{h}(u_a, d_2) \\
+ \hat{h}(s_2, u'_c)\hat{h}(u'_c, d_1)\hat{h}(s_1, u_b)\hat{h}(u_b, d_1)\hat{h}(s_1, u_a)\hat{h}(u_a, d_2) \\
- \hat{h}(s_2, u'_c)\hat{h}(u'_c, d_1)\hat{h}(s_1, u_a)\hat{h}(u_a, d_2)\hat{h}(s_1, u_b)\hat{h}(u_b, d_2).
\]

The last term is non-identically zero since \( s_2 \sim u'_c, u'_c \in \mathcal{I}(v'_d), s_1 \sim u_a, u_a \in \mathcal{I}(v_p^0), s_1 \sim u_b \) and \( u_b \in \mathcal{I}(v_p^0) \). Moreover, as we noticed before, one of the variables in \( \hat{h}(u_a, d_1) \) is not shared by any other effective channel gain \( \hat{h} \), and therefore the last term above cannot be cancelled by the other two terms. Thus, we conclude that \( M_1 \) is invertible with probability 1.

From the symmetry between \( M_1 \) and \( M_2 \) (they simply have \((s_1, d_1)\) and \((s_2, d_2)\) exchanged), the exact same steps can be used to show that either we can find the nodes \( \{v_d, v_c, v_f\} \subset V_{(v_f)} \) such that \( M_2 \) is full-rank with probability 1, or we can remove nodes from \( V_{(v_f)} \) so that we are in the case of \( V-A3 \). This concludes the achievability proof of \( D_2 = 2 \) in the cases where we have two disjoint paths with manageable interference.

Next, we proceed to providing the achievability scheme for \((B')\), in which case we have a subnetwork with no two disjoint paths, and no node \( v \) as described in \((A)\).

**B. The butterfly and the grail**

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 subnetworks in Figure 4. The subnetwork in Figure 4a simply contains two disjoint paths \( P_{s_1, d_1} \) and \( P_{s_2, d_2} \). Next, we formally characterize the other two.

**Definition 11.** 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 the network path consists of a single node, two disjoint paths \( P_{s_1, d_2} \) and \( P_{s_2, d_2} \), 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 15.

![Illustration of a network that contains a Butterfly subnetwork.](image)

**Definition 12.** The network \( \mathcal{N} \) is a Grail network if it contains two disjoint paths \( P_{s_1, d_2} \) and \( P_{s_2, d_2} \), and nodes \( u_0 \in P_{s_1, d_2} \) and \( u_0 \in P_{s_2, d_2} \) such that \( s_2 \sim u_0 \), \( u_0 \sim w_b \), and \( u_b \sim d_2 \). An example is shown in Figure 16.

![Illustration of a network that contains a Grail subnetwork.](image)

Then we can state the following Claim.

**Claim 1.** The absence of a node \( v \) whose removal disconnects \( d_i \) from both sources and \( s_i \) 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.

**Sketch of proof:** We start by building an extended network \( \mathcal{N}' \), by transforming each layer of our original network into two copies of itself, and connecting each node to its copy. Then we notice that the absence of a node \( v \) whose removal disconnects \( d_i \) from both sources and \( s_i \) from both destinations in the original network, for \( i = 1 \) or \( i = 2 \), implies the absence of an edge \( e \) whose removal disconnects \( d_i \) from both sources and \( s_i \) from both destinations, \( i = 1 \) or \( i = 2 \), in the extended network. Therefore, the result obtained in [27, 28] guarantees that \( \mathcal{N}' \) either contains two edge-disjoint paths, a butterfly or a grail. Since any two edge-disjoint paths in \( \mathcal{N}' \) are also vertex-disjoint, we conclude that our original network must contain two vertex-disjoint paths, a butterfly or a grail. A more detailed proof can be found in Appendix C.

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 1 that we may either have a butterfly network or a grail network. We provide achievability schemes for each case separately.

1) **Butterfly network**: We assume we have a subnetwork as described in Definition 11 and that any node which does not belong to \( P_{s_1, d_1} \), \( P_{s_2, d_2} \), \( P_{s_1, d_2} \) or \( P_{s_2, d_1} \) is removed from the network. Moreover, we will assume that, if there are several choices for \( u_0 \) and \( u_1 \), we choose them so that \( u_1 \) is as close as possible to the destinations (i.e., we maximize \( \ell(u_1) \)).

Similar to what we did in the case of two disjoint paths with manageable interference, we will identify a key layer and build a condensed network. Then we will show that by using amplify-and-forward in the nodes in the intermediate key layer, we can make the end-to-end transfer matrix diagonal with non-zero diagonal entries. As our key layer, we will use \( V_{(u)} \). Notice that we are guaranteed to have three nodes in \( V_{(u)} \) (since any extra node would have been removed). The condensed network is shown in Figure 17.

We let the three nodes in \( V_{(u)} \) be called \( v_1, v_2 \) and \( v_3 \) as shown in Figure 17 (notice that \( v_2 = u_1 \)), and associate scaling factors \( x_1, x_2 \) and \( x_3 \) to them. We will follow the same steps that we used in V-A4, except that now our intermediate layer has exactly three nodes. Thus, we will show that either we can remove one of the nodes in \( V_{(u)} \) so that the resulting condensed network falls in case V-A3 (i.e., a \( 2 \times 2 \times 2 \) network could be built).
are full-rank with probability 1. In the latter case, the same steps as in V-A4 guarantee that we can find \( x_1, x_2, x_3 \) such that the end-to-end transfer matrix is diagonal with non-zero diagonal entries. An important property about the Butterfly structure is that for any two nodes \( v_a, v_b \in V_{l(w_1)} \), there exist two disjoint paths between \( \{s_1, s_2\} \) and \( \{v_a, v_b\} \) and two disjoint paths between \( \{v_a, v_b\} \) and \( \{d_1, d_2\} \). Therefore, we see that if \( \hat{h}(s_2, v_1)\hat{h}(v_1, d_1) \) is not a non-identically zero polynomial in the channel gains, we can remove \( v_3 \) and \( v_2 \) and are in V-A3. Similarly if \( \hat{h}(s_1, v_3)\hat{h}(v_3, d_2) \) is non-identically zero, we can remove \( v_1 \) and \( v_2 \) and we are in V-A3. Therefore, we may assume that either \( \hat{h}(s_1, v_3) \) or \( \hat{h}(v_3, d_2) \) is zero, and either \( \hat{h}(s_2, v_1) \) or \( \hat{h}(v_1, d_1) \) is zero. To show that \( M_1 \) is full-rank with probability 1, we first consider the case when \( \hat{h}(v_3, d_2) = 0 \). We notice that the fact that \( \hat{h}(v_3, d_2) = 0 \) and our assumption that \( u_1 \) was chosen as close as possible to the destinations guarantee that there is no path starting on a node in \( P_{v_3,d_2} \) and ending in \( d_2 \). Thus, we see that the first channel gain in the \( P_{v_2,d_1} \) path only appears as a variable in \( \hat{h}(v_2, d_1) \), and no other \( \hat{h} \). Then we notice that

\[
\det M_1 = \\
\begin{bmatrix}
\hat{h}(s_1, v_1)\hat{h}(v_1, d_1) & \hat{h}(s_1, v_2)\hat{h}(v_2, d_1) & \hat{h}(s_1, v_3)\hat{h}(v_3, d_1) \\
0 & \hat{h}(s_2, v_1)\hat{h}(v_1, d_2) & \hat{h}(s_2, v_2)\hat{h}(v_2, d_2) \\
-\hat{h}(s_1, v_1)\hat{h}(v_1, d_1) & \hat{h}(s_1, v_2)\hat{h}(v_2, d_1) & \hat{h}(s_1, v_3)\hat{h}(v_3, d_1) \\
\end{bmatrix}.
\]

The last term is a non-identically zero polynomial, since \( s_1 \sim v_1, v_1 \sim d_2, v_2 \sim d_1, v_3 \sim d_1 \) and there are two disjoint paths \( P_{s_1,v_2} \) and \( P_{s_2,v_3} \). Thus, since \( \hat{h}(v_2, d_1) \) contains a variable which cannot be cancelled by the other term, we conclude that \( \det M_1 \) is non-identically zero, and \( M_1 \) is full-rank with probability 1. If instead we assume that \( \hat{h}(v_3, d_2) \) is not identically zero, then \( \hat{h}(s_1, v_3) = 0 \), and we have that

\[
\det M_1 = \\
\begin{bmatrix}
\hat{h}(s_1, v_1)\hat{h}(v_1, d_1) & \hat{h}(s_1, v_2)\hat{h}(v_2, d_1) & 0 \\
0 & \hat{h}(s_2, v_1)\hat{h}(v_1, d_2) & \hat{h}(s_2, v_2)\hat{h}(v_2, d_2) \\
-\hat{h}(s_1, v_1)\hat{h}(v_1, d_1) & \hat{h}(s_1, v_2)\hat{h}(v_2, d_1) & \hat{h}(s_1, v_3)\hat{h}(v_3, d_1) \\
\end{bmatrix}.
\]

which is not identically zero, since \( s_2 \sim v_3, v_3 \sim d_1, s_1 \sim v_1, s_1 \sim v_2 \) and there are two disjoint paths \( P_{v_1,d_2} \) and \( P_{v_2,d_3} \). Therefore, we conclude that \( M_1 \) is full-rank with probability 1.

2) **Grail network**: We assume that we have a minimal subnetwork which still satisfies Definition 12, i.e., all the unnecessary nodes are removed. As key layers, we will use \( V_{l(w_1)} \) and \( V_{l(w_2)} \). 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 18. We will let the nodes in \( V_{l(w_1)} \) be called \( u_1 \) and \( u_2 \), and the nodes in \( V_{l(w_2)} \) be called \( v_1 \) and \( v_2 \), as shown in Figure 18. Next we will show that either we can suppress one of the two intermediate key layers (by assuming their nodes are just forwarding their received signals) and obtain a network as in V-A3, 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 \( P_{s_2,d_1} \) guarantees that if we suppress \( V_{l(w_1)} \) from the condensed network, we obtain the network in V-A3. Similarly, if \( \hat{h}(v_1, d_1) \) is not identically zero, we can suppress \( V_{l(w_2)} \) from the condensed network, and we are again in the case of V-A3. 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 \( V_{l(w_1)} \), which is given by

\[
F = \\
\begin{bmatrix}
\hat{h}(s_1, u_1)\hat{h}(u_1, v_1)y_1 \\
\hat{h}(s_1, u_1)\hat{h}(u_1, v_1)y_1 \\
\hat{h}(s_2, u_1)\hat{h}(u_1, v_1)y_1 + \hat{h}(s_2, u_2)\hat{h}(u_2, v_1)y_2 \\
\hat{h}(s_2, u_1)\hat{h}(u_1, v_2)y_1 + \hat{h}(s_2, u_2)\hat{h}(u_2, v_2)y_2 \\
\end{bmatrix}.
\]
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
\[
\det M = \begin{bmatrix}
\hat{h}(s_2, u_1)\hat{h}(u_1, v_1) & \hat{h}(s_2, u_2)\hat{h}(u_2, v_1) \\
\hat{h}(s_2, u_1)\hat{h}(u_1, v_2) & \hat{h}(s_2, u_2)\hat{h}(u_2, v_2)
\end{bmatrix} = \hat{h}(s_2, u_1)\hat{h}(s_2, u_2)
\hat{h}(u_1, v_1)\hat{h}(u_2, v_1)
\hat{h}(u_1, v_2)\hat{h}(u_2, v_1),
\]
which is a non-identically zero polynomial on the channel gains, since \(s_2 \sim u_1, s_2 \sim u_2\) and there are two disjoint paths \(P_{s_1, v_1}\) and \(P_{s_2, v_2}\). 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 \(\hat{h}(s_2, u_2)\hat{h}(u_2, v_1) \neq 0\) with probability 1, we are guaranteed that if we choose \(y_1 \neq 0\) and \(y_2 \neq 0\) such that 
P_{2,1} = \hat{h}(s_2, u_1)\hat{h}(u_1, v_1) + \hat{h}(s_2, u_2)\hat{h}(u_2, v_2)\gamma_1 = 0,
then \(F_{1,1} \neq 0, F_{1,2} \neq 0\) and \(F_{2,1} \neq 0\). Notice that, if
\[
F_{1,2} = \hat{h}(s_2, u_1)\hat{h}(u_1, v_1) + \hat{h}(s_2, u_2)\hat{h}(u_2, v_1)\gamma_2 = 0,
\]
we would contradict the fact that the system \(M\gamma = 0\) only has \(\gamma = 0\) as a solution. Therefore, we have that the end-to-end transfer matrix can be expressed as
\[
\begin{bmatrix}
0 & \hat{h}(v_2, d_1) \\
\hat{h}(v_1, d_2) & \hat{h}(v_2, d_2)
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} =
\begin{bmatrix}
\alpha & \beta \\
\gamma & 0
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix},
\]
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. This concludes the achievability proof for the case in which we have a grail subnetwork and thus we conclude all cases in which \(D_\Sigma = 2\) is achievable.

VI. NETWORKS WITH 3/2 DEGREES-OF-FREEDOM
In this section, we prove that if our network \(N\) does not fall into cases (A), (A'), (B) and (B'), then we have \(D_\Sigma = \frac{3}{2}\). We start by defining two main categories of networks which belong to (C). If \(N\) does not contain a node \(v\) whose removal disconnects \(d_i\) from both sources and \(s_i\) from both terminals, for \(i \in \{1, 2\}\) (i.e., \(N\) is not in (A)), then, from our discussion in V-B, we know that we must have one of the three structures in Figure 4. Moreover, if the network does not contain such a node \(v\) and does not contain two disjoint paths \(P_{s_1, d_1}\) and \(P_{s_2, d_2}\), then we are in (B'). Therefore, all networks in (C) contain two disjoint paths \(P_{s_1, d_1}\) and \(P_{s_2, d_2}\), but do not contain any pair of disjoint paths \(P'_{s_1, d_1}\) and \(P'_{s_2, d_2}\) with manageable interference, or else we would be in case (B).

We will assume that we have two disjoint paths \(P_{s_1, d_1}\) and \(P_{s_2, d_2}\) and we will first show that we can assume that our network \(N\) falls into one of two cases:

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

C2. \(n_1(G) = n_1^P = 1\).

We see this as follows. Since the interference on \(P_{s_1, d_1}\) and \(P_{s_2, d_2}\) is not manageable, we have that either \(n_1(G) = 1\) or \(n_2(G) = 1\). Moreover, we must also have either \(n_1^D = 1\) or \(n_2^D = 1\), because otherwise we can let \(S = P_{s_1, d_1} \cup P_{s_2, d_2}\) and \(n_1(G[S]) = n_1^P\) and \(n_2(G[S]) = n_2^D\). So we assume WLOG that \(n_1^D = 1\). Then, if \(n_1(G) = 1\), we are in case C2. Thus, we assume \(n_1(G) \geq 2\), and we must have \(n_2(G) = 1\). If \(n_2^D = 1\), we are again in case C2 by exchanging the names of \((s_1, d_1)\) and \((s_2, d_2)\). Otherwise, if \(n_2^D = 0\), we are in case C1 (notice that \(n_1^D \leq n_2(G)\)).

We will provide a achievability and a converse for \(D_\Sigma = \frac{3}{2}\) in each case.

A. Achievability for case C1
We will start by considering case C1. Notice that we must have a node \(v_1 \notin P_{s_1, d_1} \cup P_{s_2, d_2}\) such that \(v_1 \sim P_{s_1, d_1}\) and thus we have a path \(P_{s_2, v_1}\) that is disjoint from \(P_{s_1, d_1}\). We let \(v_m\) be the last node in \(P_{s_2, v_1} \cap P_{s_1, v_1}\) and we have the path \(P_{v_m, v_1} = P_{s_2, v_1} \cap v_m, v_1\). Next we consider letting \(S^* = P_{s_1, d_1} \cup P_{s_2, d_2} \cup P_{v_m, v_1}\). This guarantees that \(n_1(G[S^*]) \geq 2\). Since \(P_{s_1, d_1}\) and \(P_{s_2, d_2}\) do not have manageable interference, we must have \(n_2(G[S^*]) = 1\). Moreover, since \(n_2^D = 0\), we conclude that we must have a node \(v_2 \in P_{s_2, v_1} \setminus \{v_m\}\) such that \(v_2 \sim P_{s_2, d_2}\) and we must have a path \(P_{s_1, v_2} \subset S^*\). It can then be seen that our network is as shown in Figure 19 up to a change in the position of the edge \((v_3, v_4)\). The curve lines and the dashed lines indicate paths (which may consist of a single edge or multiple edges). Notice that we may also have \(v_1 = v_2\).

![Illustration of the network in case C1](image)

In order to achieve \(D_\Sigma = \frac{3}{2}\), we will describe a scheme in which we use two different modes of operation for the network. During each mode of operation, only a subset of the nodes will be transmitting, while the others will stay silent. During the first mode of operation, one special node will store its received signals. Then, in the second mode of operation, it will forward the stored signals. We will consider two subcases, according to the position of edge \((v_3, v_4)\) with respect to \(v_2\).

1) \(\ell(v_3) < \ell(v_2)\): In this case, our “special node” will be the node from \(P_{s_2, d_2}\) in \(V_{\ell(v_2)}\). In the first mode of operation, it will function as a virtual destination \(d_2\). Node \(d_2\) and any node \(u \in P_{s_2, d_2}\) such that \(\ell(u) \geq \ell(d_2)\) will stay silent during Mode 1. Then we notice that the two disjoint paths \(P_{s_1, d_1}\) and \(P_{s_2, d_2}\) have manageable interference. This must be the case, since \(n_2(G, P_{s_1, d_1}) = 1\), and this unique interference is caused by \(v_2\) on a node \(u \in P_{s_2, d_2}\) such that \(\ell(u) > \ell(d_2)\), and thus \(n_2(G, P_{s_2, d_2}) = 0\). Moreover, since \(\ell(v_3) < \ell(d_2)\) and \(\ell(v_m) < \ell(d_2)\), we have \(n_1(G, P_{s_1, d_1}) \geq 2\).

Therefore, by using the amplify-and-forward scheme described in V-A1, 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. Notice that, even though \(d_1\) and \(d_2'\) are not in the same layer, one could create a virtual path between \(d_2'\) and a virtual node \(d_2'' \in V_r\) which does not receive nor cause any interference. Then we can use the scheme from V-A1 to guarantee that the transfer matrix between \((s_1, s_2)\) and \((d_1, d_2'')\) is diagonal with non-zero diagonal entries. Then it is easy to see that the same would hold for the transfer matrix between \((s_1, s_2)\) and \((d_1, d_2')\). 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 \(P_{s_1, d_1}\) and \(P_{s_2, d_2}.\) Now we clearly have two disjoint paths with no interference. Therefore, it is clear that 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 \(s_2\) and \(d_2\) and one of them between \(s_1\) and \(d_1.\) Therefore, we achieve \(D_\Sigma = \frac{3}{2}.\) A visual representation of the scheme is shown in Figure 20.

The assumption that there are no two disjoint paths with manageable interference, then \(D_\Sigma \leq \frac{3}{2}.\) We will start by naming some extra nodes which will be important to us, as shown in Figure 22. We will let \(v_0\) be the node on \(P_{s_2, d_2}\) such that \((v_2, v_0) \in E.\) From our discussion in VI-A, we know that we have a path \(P_{s_1, v_2},\) which must be entirely contained in \(S^* = P_{s_1, d_1} \cup P_{s_2, d_2} \cup P_{v_m, v_1}.\) Thus, we let \(v_5\) be the last node in \(P_{s_1, d_1} \cap P_{s_2, d_2},\) and we let \(v_6\) be its consecutive node on \(P_{s_1, v_2}\) (which must be part of \(P_{v_m, v_1}\) as well).

The second mode of operation should last for the same number of time steps as the first one. During Mode 2, \(d_1'\) will become a virtual source \(s_1'.\) Any node \(v \in P_{s_1, d_1}\) such that \(\ell(v) < \ell(s_1')\) will stay silent. Notice that the paths \(P_{s_1, d_1}\) and \(P_{s_2, d_2}\) have manageable interference. Therefore, by assuming the existence of a virtual node \(s_1'' \in V_1\) which is connected to \(s_1'\) through a virtual path that does not receive nor cause any interferences, we can use the linear scheme from V-A1 to guarantee that the transfer matrix from \((s_1', s_2)\) to \((d_1, d_2)\) is diagonal with non-zero diagonal entries. Thus, by letting node \(d_1'' = s_1''\) 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 \(s_2\) and \(d_2\) and one of them between \(s_1\) and \(d_1.\) Therefore, we achieve \(D_\Sigma = \frac{3}{2}.\) A visual representation of the scheme is shown in Figure 21.

The assumption that there are no two disjoint paths with manageable interference allows us to infer some important connectivity properties about networks in case C1, illustrated in Figure 22. Next, we state and prove these properties.

B. Converse for case C1

In this section, we will show that if a network falls in C1, but does not contain two disjoint paths with manageable interference, then \(D_\Sigma \leq \frac{3}{2}.\) We will start by naming some extra nodes which will be important to us, as shown in Figure 22. We will let \(v_0\) be the node on \(P_{s_2, d_2}\) such that \((v_2, v_0) \in E.\) From our discussion in VI-A, we know that we have a path \(P_{s_1, v_2},\) which must be entirely contained in \(S^* = P_{s_1, d_1} \cup P_{s_2, d_2} \cup P_{v_m, v_1}.\) Thus, we let \(v_5\) be the last node in \(P_{s_1, d_1} \cap P_{s_2, d_2},\) and we let \(v_6\) be its consecutive node on \(P_{s_1, v_2}\) (which must be part of \(P_{v_m, v_1}\) as well).
It is easy to see that if we have a path \( P_{s_1,d_2} \) not containing \( \{v_2,v_0\} \), then we must have a node \( v_d \in P_{s_1,d_2} \) such that \( v_a \xleftarrow{} v_d \xrightarrow{} P_{s_2,d_2} \), and thus we would have \( n_2(G,P_{s_2,d_2}) \geq 2 \), which is a contradiction.

P2. All paths from \( s_1 \) to \( d_2 \) contain nodes \( v_5 \) and \( v_6 \). First consider the path \( Q_{s_2,d_2} = P_{n_2,d_2}[s_2,v_m] \oplus P_{v_m,v_1}[v_m,v_2] \oplus (v_2,v_0) \oplus P_{s_2,d_2}[v_0,d_2] \). Clearly, \( Q_{s_2,d_2} \cap P_{s_1,d_1} = \emptyset \) and \( v_5 \xleftarrow{} Q_{s_2,d_2} \). If we have a path \( P_{s_1,d_2} \) not containing \( \{v_5,v_6\} \) we conclude that \( n_2(G,Q_{s_2,d_2}) \geq 2 \). But since \( n_1(G,P_{s_1,d_1}) \geq 2 \) we contradict the fact that there are no two disjoint paths with manageable interference.

P3. All paths from \( s_2 \) to \( d_1 \) contain \( \{v_6,v_2\} \) or \( \{v_3,v_4\} \). Suppose there is a path \( P_{s_2,d_1} \) not containing \( \{v_6,v_2\} \) nor \( \{v_3,v_4\} \). Then we let \( S = P_{s_1,d_1} \cup P_{s_2,d_2} \cup P_{s_2,d_1} \) and we have \( n_1(G[S],P_{s_1,d_1}) \geq 2 \). But since \( P_1 \) and \( P_2 \) imply that any path from \( s_1 \) to \( d_2 \) must contain \( \{v_6,v_2\} \), and \( \{v_6,v_2\} \not\subseteq S \), we must have \( n_2(G[S],P_{s_2,d_2}) = 0 \), contradicting the fact that there are no two disjoint paths with manageable interference.

P4. The removal of \( v_0 \) disconnects \( d_2 \) from both sources. From \( P_1 \), the removal of \( v_0 \) disconnects \( d_2 \) from \( s_1 \). So suppose the removal of \( v_0 \) does not disconnect \( d_2 \) from \( s_2 \) and we have a path \( Q_{s_2,d_2} \) not containing \( v_0 \). We know that \( Q_{s_2,d_2} \) must be disjoint from \( P_{s_1,d_1} \), since otherwise we would contradict the fact that the removal of \( v_0 \) disconnects \( d_2 \) from \( s_1 \) (\( P_1 \)). Moreover, if we let \( S = P_{s_1,d_1} \cup Q_{s_2,d_2} \), since \( v_0 \notin S \), we must have \( n_2(G[S],Q_{s_2,d_2}) = 0 \). If \( n_1(G[S],P_{s_1,d_1}) \neq 1 \), we contradict the assumption of no two disjoint paths with manageable interference. However, if \( n_1(G[S],P_{s_1,d_1}) = 1 \), we must have a direct interference from \( Q_{s_2,d_2} \) on \( P_{s_1,d_1} \), and we will have \( n_2(G[V \setminus \{v_0\}],P_{s_1,d_1}) \geq 2 \) and \( n_2(G[V \setminus \{v_0\}],Q_{s_2,d_2}) = 0 \), and we again reach a contradiction.

P5. The removal of \( v_5 \) disconnects \( s_1 \) from both destinations. From \( P_2 \), the removal of \( v_5 \) disconnects \( s_1 \) from \( d_2 \). So suppose the removal of \( v_5 \) does not disconnect \( s_1 \) from \( d_1 \) and we have a path \( Q_{s_1,d_1} \), not containing \( v_5 \). The path \( Q_{s_1,d_1} \) must be disjoint from \( P_{s_2,d_2} \), or else we would contradict the fact that the removal of \( v_5 \) disconnects \( s_1 \) from \( d_2 \) (\( P_2 \)). So first we let \( S = Q_{s_1,d_1} \cup P_{s_2,d_2} \), and, since \( v_5 \notin S \), we have \( n_2(G[S],P_{s_2,d_2}) = 0 \). If we have \( n_1(G[S],Q_{s_1,d_1}) \neq 1 \), we contradict the assumption of no two disjoint paths with manageable interference. However, if \( n_1(G[S],Q_{s_1,d_1}) = 1 \), we must have a direct interference from \( P_{s_2,d_2} \) on \( Q_{s_1,d_1} \), and we will have \( n_2(G[V \setminus \{v_3\}],Q_{s_1,d_1}) \geq 2 \) and \( n_2(G[V \setminus \{v_3\}],P_{s_2,d_2}) = 0 \), and we again reach a contradiction.

P6. The removal of \( v_2 \) and \( v_3 \) disconnects \( d_2 \) from both sources. From \( P_1 \), the removal of \( v_2 \) disconnects \( d_2 \) from \( s_1 \). So suppose the removal of \( v_2 \) and \( v_3 \) does not disconnect \( d_2 \) from \( s_2 \) and we have a path \( Q_{s_2,d_2} \) not containing \( v_2 \) or \( v_3 \). We know that \( Q_{s_2,d_2} \) is disjoint from \( P_{s_1,d_1} \), or else we would contradict the fact that the removal of \( v_2 \) disconnects \( s_1 \) from \( d_2 \) (\( P_1 \)). Then, we set \( S = P_{s_1,d_1} \cup Q_{s_2,d_2} \). Since \( v_2,v_3 \notin S \), from \( P_1 \), we must have \( n_2(G[S],Q_{s_2,d_2}) = 0 \), and from \( P_3 \), we must have \( n_1(G[S],P_{s_1,d_1}) = 0 \). But this contradicts our assumption of no two disjoint paths with manageable interference.

P7. The removal of \( v_2 \) and \( v_4 \) disconnects \( d_1 \) from both sources. From \( P_3 \), the removal of \( v_2 \) and \( v_4 \) disconnects \( d_1 \) from \( s_2 \). Thus, we assume that we have a path \( P_{s_1,d_1} \), not containing \( v_2 \) nor \( v_4 \). The path \( P_{s_1,d_1} \) must be disjoint of \( P_{s_2,d_2} \), or else we contradict \( P_3 \). Thus we set \( S = Q_{s_1,d_1} \cup P_{s_2,d_2} \). Since \( v_2,v_4 \notin S \), from \( P_1 \), we must have \( n_2(G[S],P_{s_2,d_2}) = 0 \), and from \( P_3 \), we must have \( n_1(G[S],Q_{s_1,d_1}) = 0 \). But this contradicts our assumption of no two disjoint paths with manageable interference.

P8. All paths from \( s_1 \) or \( s_2 \) to \( v_0 \) contain \( v_6 \). This follows easily from \( P_1 \), \( P_2 \) and \( P_3 \), since \( v_2 \sim d_1 \) and \( v_2 \sim d_2 \).

These properties allow us to infer the information inequalities that will build the converse proof. For these derivations, we will let \( A \triangleq \{v \in V : s_2 \not\sim v\} \) and \( B \triangleq \{v \in V : s_1 \not\sim v\} \), and 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. Before we formally derive the inequalities, we will describe some of the intuition that leads to them, for a specific network example, shown in Figure 23.

![Diagram](image)

Fig. 23. Example of a network in case C1.

We will consider, for a given \( n \), the quantities:

\[
\alpha \triangleq \liminf_{P \to \infty} \frac{I(Y^n_4;X^n_B|X^n_B)}{\frac{1}{2} \log P} \quad \text{and} \quad \beta \triangleq \liminf_{P \to \infty} \frac{I(X^n_5;Y^n_6|X^n_B)}{\frac{1}{2} \log P}.
\]

It is easy to see that \( 0 \leq \alpha, \beta \leq 1 \). Intuitively, since all the information from the sources must go through either \( v_4 \) or the nodes in \( B \) to reach \( v_2 \), \( \alpha \) can be thought of as the number of useful degrees-of-freedom (i.e., carrying information about the sources) transmitted by \( v_2 \). Similarly, \( \beta \) can be thought of as the number of degrees-of-freedom transmitted by \( v_5 \), but only counting the degrees-of-freedom with information about message \( W_1 \) (since we condition on \( X^n_B \)). Based on these quantities we will state three inequalities related to the degrees-of-freedom that can be achieved, and for each one we will provide an intuitive explanation. The formal proof is omitted, but it follows from the information inequalities we will derive later based on properties P1-P8. In the sense of Definition 9, we let \( D_i \) be the degrees-of-freedom assigned to \( (s_i,d_i) \), for \( i = 1, 2 \). First, we have

\[
D_1 \leq \beta,
\]

since all information from \( W_1 \) must flow through \( v_5 \). Next, we claim that both \( W_1 \) and \( W_2 \) can be decoded from \( Y^n_4 \) and...
\( \tilde{X}_2^n \), and thus
\[
D_1 + D_2 \leq 1 + \alpha. \tag{16}
\]
To see this, we first notice that, since the removal of \( v_4 \) and \( v_2 \) disconnects \( d_1 \) from \( \{s_1, s_2\} \), from \( Y_1^n \) and \( \tilde{X}_2^n \), \( W_1 \) can be decoded. Then, \( W_1 \) can be used to approximately obtain \( X_A^n \) (since the nodes in \( A \) cannot be influenced by \( W_2 \)), and, by removing its contribution from \( Y_1^n \), we can obtain a noisy version of the transmit signal from \( v_3 \). But since all the information about \( W_2 \) must flow through \( v_3 \), this allows one to use \( Y_2^n \) and \( X_2' \) to decode \( W_2 \) as well. For the third inequality we claim that, from \( Y_0^n \), we can decode \( W_2 \) completely and \((\alpha + \beta - 1)\) degrees-of-freedom of \( W_1 \), and thus
\[
D_2 + (\alpha + \beta - 1) \leq 1. \tag{17}
\]
To see this, we first notice that, since the removal of \( v_0 \) disconnects \( d_2 \) from \( \{s_1, s_2\} \), from \( Y_0^n \), we can decode \( W_2 \), and thus \( X_2' \) approximately. By removing its contribution from \( Y_0^n \), we obtain a noisy version of the transmit signal from node \( v_2 \), which allows us to decode the \( \alpha \) degrees-of-freedom transmitted by it. Now we ask ourselves how many of the \( \alpha \) degrees-of-freedom transmitted by \( v_2 \) must be carrying information about \( W_1 \). To answer this question, we notice that all the degrees-of-freedom transmitted by \( v_2 \) must have come through node \( v_0 \). Since node \( v_0 \) receives \( \beta \) degrees-of-freedom with information about \( W_1 \) from \( v_0 \), at most \( 1 - \beta \) of its degrees-of-freedom can not be about \( W_1 \). Thus, any number of degrees-of-freedom above \( 1 - \beta \) that \( v_2 \) transmits, i.e., \( \alpha - (1 - \beta) = \alpha + \beta - 1 \), must contain information about \( W_1 \). Finally, by adding inequalities (15), (16) and (17), we obtain \( 2(D_1 + D_2) \leq 3 \), and therefore \( D_2 \leq 3/2 \).

Next, we formally derive information inequalities that can be used to show that \( D_2 \leq 3/2 \) for all networks in case C. The intuition is similar to that of inequalities (15), (16) and (17), but the inequalities are somewhat different since they need to hold for any network in case C. First we have
\[
nR_2 \leq I(W_2; Y_0^n) + n\epsilon_n \tag{i}
\leq I(\tilde{X}_2^n; Y_0^n) + n\epsilon_n
\leq I(X_2^n; \tilde{X}_2^n; Y_0^n; Y_0^n) + n\epsilon_n
\leq \frac{1}{2} \log P + nK_3 - I(X_2^n; \tilde{X}_2^n; Y_0^n) + n\epsilon_n, \tag{18}
\]
where \((i)\) follows from the Markov chain \( W_2 \leftrightarrow \tilde{X}_2^n \leftrightarrow Y_0^n \leftrightarrow Y_0^n \), which is implied by P4 and the fact that \( s_2 \in B \); \((ii)\) follows from the fact that \( I(X_2^n; \tilde{X}_2^n; Y_0^n) \) can be upper bounded by \( h(Y_0^n) - h(Y_0^n) \) by following the steps in (6), where \( K_3 \) is a positive constant, independent of \( P \), for \( P \) sufficiently large. We also have that
\[
nR_1 \leq I(W_1; Y_d^n) + n\epsilon_n \tag{i}
\leq I(W_1; \tilde{X}_2^n; \tilde{X}_2^n) + n\epsilon_n
\leq I(W_1; \tilde{X}_2^n; \tilde{X}_2^n) + n\epsilon_n
\leq I(X_2^n; Y_d^n; \tilde{X}_2^n; \tilde{X}_2^n) + I(X_2^n; \tilde{X}_2^n; \tilde{X}_2^n) + n\epsilon_n
\leq I(X_2^n; Y_d^n; \tilde{X}_2^n; \tilde{X}_2^n) + nK_4 + n\epsilon_n, \tag{19}
\]
where \((i)\) follows because P5 and the fact that \( s_2 \in B \) imply that the removal of \( v_3 \) and \( B \) disconnects \( d_1 \) from both sources and thus \( W_1 \leftrightarrow (\tilde{X}_2^n, \tilde{X}_2^n) \leftrightarrow Y_d^n \); \((ii)\) follows from the fact that \( \tilde{X}_B \) is independent of \( W_1 \); \((iii)\) follows from the fact that, given \( X_0^n \), we have \( W_1 \leftrightarrow \tilde{X}_2^n \leftrightarrow \tilde{X}_2^n \); \((iv)\) follows from Lemma 1, since P2 implies that \( \mathcal{I}(v_0) \setminus \{v_2\} \subset B \). To obtain the next inequalities, we consider two cases, according to the position of \( v_4 \) and \( v_5 \).

1) \( \ell(v_4) \leq \ell(v_5) \): In this case, we have
\[
nR_2 \leq I(W_2; Y_0^n) + n\epsilon_n \tag{i}
\leq I(X_2^n; \tilde{X}_2^n; Y_0^n) + n\epsilon_n
\leq I(W_2; Y_0^n) + n\epsilon_n
\leq I(X_2^n; \tilde{X}_2^n; Y_0^n) + n\epsilon_n
\leq I(X_2^n; \tilde{X}_2^n; Y_0^n) + n\epsilon_n
\leq I(X_2^n; \tilde{X}_2^n; Y_0^n) + n\epsilon_n
\]
where \((i)\) follows because P6 implies the Markov chain \( W_2 \leftrightarrow X_2^n \leftrightarrow (\tilde{X}_2^n, \tilde{X}_2^n) \leftrightarrow Y_0^n \); \((ii)\) follows from the fact that \( X_2^n \) is independent of \( X_0^n \); \((iii)\) follows by applying Lemma 1 to the second term, since \( \ell(v_4) \leq \ell(v_5) \) implies that \( \mathcal{I}(v_4) \setminus \{v_2\} \subset A \), or else we contradict P3; \((iv)\) follows from the fact that \( s_2 \in B \); and \((v)\) follows because we have \( (X_2^n, \tilde{X}_2^n) \leftrightarrow (X_0^n, \tilde{X}_2^n) \leftrightarrow X_2^n \), since the removal of \( A, v_4 \) and \( B \) disconnects \( s_2 \) and \( \mathcal{O}(v_3) \) from \( v_2 \). This can be seen as follows. From P8, all paths from \( s_2 \) or \( v_3 \) to \( v_2 \) must contain a node in \( \mathcal{I}(v_4) \). From P2, we know that \( \mathcal{I}(v_4) \setminus \{v_2\} \subset B \). From P3, we know that any path from \( v_3 \) or \( s_2 \) to \( v_5 \) must contain \( v_4 \). Finally, since \( \ell(v_4) < \ell(v_5) \), we have that \( v_3 \notin \mathcal{I}(v_5) \), and, therefore, 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 had to consider \( \mathcal{O}(v_3) \) instead of simply \( v_3 \), because we have \( X_2^n \), and not \( X_2^n \). Next, we have that
\[
nR_1 \leq I(W_1; Y_d^n) + n\epsilon_n \tag{i}
\leq I(X_2^n; \tilde{X}_2^n) + n\epsilon_n
\leq I(X_2^n; \tilde{X}_2^n) + n\epsilon_n
\leq I(X_2^n; \tilde{X}_2^n) + n\epsilon_n
\leq I(X_2^n; \tilde{X}_2^n) + n\epsilon_n
\]
where \((i)\) follows because P7 implies the Markov chain \( W_1 \leftrightarrow (Y_d^n, \tilde{X}_2^n) \leftrightarrow Y_d^n \); \((ii)\) follows since \( s_1 \in A \); \((iii)\) follows
from the fact that $I(\tilde{X}_n, X_B^n, Y_B^n)$ can be upper bounded by $h(Y_B^n) - h(N_{3,4}^n)$ by following the steps in (6), where $K_6$ is a positive constant, independent of $P$, for $P$ sufficiently large. The second term in the inequality above can be bounded as

$$I(\tilde{X}_n, X_B^n, Y_B^n; \tilde{X}_2^n) \leq I(\tilde{X}_n, \tilde{X}_2^n; \tilde{X}_2^n) = I(\tilde{X}_n, Y_B^n; \tilde{X}_2^n) + I(\tilde{X}_n, Y_B^n, \tilde{X}_2^n | \tilde{X}_2^n)$$

where (i) follows because the Markov chain $(X_A^n, X_B^n, Y_B^n) \leftrightarrow (\tilde{X}_n, \tilde{X}_2^n, Y_B^n) \leftrightarrow \tilde{X}_2^n$; (ii) follows since $X_B^n$, we have $(\tilde{X}_n, Y_B^n) \leftrightarrow \tilde{X}_2^n$; (iii) follows by applying Lemma 1 to $I(X^n, \tilde{X}_2^n, Y_B^n)$. Similarly, we obtain $X^n, \tilde{X}_2^n$ independent of $X_B^n$. This holds true because $P3$ implies that any path from a node in $B$ to a node in $C$ must contain $v_2$, and, thus, the removal of $A$ and $v_2$ disconnects $C$ from both sources. Notice that (vi) is only non-trivial in the cases where $C \not\subset A$ (when $\ell(v_4) > \ell(v_1) + 1$). Next, we have that

$$nR_1 \leq I(W_1; Y_{d_1}^n) + n\epsilon_n \leq I(W_1; \tilde{X}_n^n, Y_{d_4}^n, \tilde{X}_2^n) + n\epsilon_n$$

where (i) follows because $P6$ implies that the removal of $O(v_3)$ and $v_2$ disconnects $d_2$ from both sources. Then, since $P1$ implies that all paths from $v_4$ to $d_2$ contain $v_2$, we know that the removal of $D$ and $v_2$ also disconnects $d_2$ from both sources, and we have the Markov chain $W_2 \leftrightarrow X_2^n \leftrightarrow (X_2^n, X_{d_4}^n) \leftrightarrow Y_{d_4}^n$; (ii) follows from the fact that $X_2^n$ is independent of $X_{d_4}^n$; (iii) follows since $s_2 \in B$; (iv) follows from applying Lemma 1 to $I(X_2^n, X_{d_4}^n)$ since, in case II, if $u \in D \setminus \{v_4\}$, then $u \not\leftrightarrow v_2$, or else we contradict $P8$; (v) follows since $s_2 \in B$; and (vi) follows from the fact that, given $X_2^n$ and $X_A^n, X_C^n$ is independent of $X_{d_4}^n$. This holds true because $P3$ implies that any path from a node in $B$ to a node in $C$ must contain $v_2$, and, thus, the removal of $A$ and $v_2$ disconnects $C$ from both sources. Notice that (vi) is only non-trivial in the cases where $C \not\subset A$ (when $\ell(v_4) > \ell(v_1) + 1$). Next, we have that

$$nR_2 \leq I(W_2; Y_{d_2}^n) + n\epsilon_n \leq I(W_2; Y_{d_2}^n, \tilde{X}_2^n, \tilde{X}_2^n, \tilde{X}_2^n) + n\epsilon_n$$

where (i) follows because $P7$ implies that the removal of $O(v_3)$ and $v_2$ disconnects $d_2$ from both sources. Then, since $P1$ implies that all paths from $v_4$ to $d_2$ contain $v_2$, we know that the removal of $D$ and $v_2$ also disconnects $d_2$ from both sources, and we have the Markov chain $W_1 \leftrightarrow X_2^n \leftrightarrow (X_2^n, X_{d_4}^n) \leftrightarrow Y_{d_4}^n$; (ii) follows from the fact that $X_2^n$ is independent of $X_{d_4}^n$; (iii) follows since $s_2 \in B$; (iv) follows from applying Lemma 1 to $I(X_2^n, X_{d_4}^n)$ since, in case II, if $u \in D \setminus \{v_4\}$, then $u \not\leftrightarrow v_2$, or else we contradict $P8$; (v) follows since $s_2 \in B$; and (vi) follows from the fact that, given $X_2^n$ and $X_A^n, X_C^n$ is independent of $X_{d_4}^n$. This holds true because $P3$ implies that any path from a node in $B$ to a node in $C$ must contain $v_2$, and, thus, the removal of $A$ and $v_2$ disconnects $C$ from both sources. Notice that (vi) is only non-trivial in the cases where $C \not\subset A$ (when $\ell(v_4) > \ell(v_1) + 1$). Next, we have that
obtain
\[
\begin{aligned}
    nR_1 &\leq n \log P - I(X_n^0, Y_n^0; X_B^\infty) + I(X_n^1, Y_n^1; X_B^\infty) \\
        &- I(X_n^0, X_{n+1}^1; X_{B_A}^C) \\
        &- I(X_n^1, X_{n+1}^0; X_{A}^C) + n(K_{10} + K_{11} + K_{12}) + n\epsilon_n.
\end{aligned}
\]

Finally, by adding equations (18), (19), (20) and (22) for case I, and (18), (19), (23) and (24) for case II, we obtain
\[
2n(R_1 + R_2) \leq \frac{3n}{2} \log P + 6nK_{\text{max}} + n\epsilon_n
\]
\[
\Rightarrow R_1 + R_2 \leq \frac{3}{2} + \frac{6K_{\text{max}} + \epsilon_n}{\log P},
\]
where \(K_{\text{max}} = \max_j K_j\). Thus, as we let \(n \to \infty\) and then \(P \to \infty\), we obtain
\[
D_\Sigma \leq \frac{3}{2}.
\]

We now proceed to considering C2. We will show that if our network \(\mathcal{N}'\) does not fall in cases (A), (A'), (B), and (B'), then \(D_\Sigma = \frac{3}{2}\).

C. Achievability for case C2

In this section, we will show that if we are in C2 and no edge as in (A') exists, then we can also achieve \(\frac{3}{2}\) degrees-of-freedom. We start by proving properties about the connectivity of our network, if we are in C2. Notice that, if for some choice of two disjoint paths \(P'_{s_1,d_1}\) and \(P'_{s_2,d_2}\) we are in C1, our previous result shows that \(D_\Sigma = \frac{3}{2}\). Therefore, we may assume that for no choice of two disjoint paths we are in C1. So we suppose we have two disjoint paths \(P_{s_1,d_1}\) and \(P_{s_2,d_2}\), but no two disjoint paths with manageable interference. In addition, we assume that we do not have an edge as in (A'). Since we are in C2, we have that \(n_1(G, P_{s_1,d_1}) = n_1^D(P_{s_2,d_2}, P_{s_1,d_1}) = 1\) and we let \((v_2, v_1) \in E\) be the unique edge such that \(v_2 \in P_{s_2,d_2}\) and \(v_1 \in P_{s_1,d_1}\).

P1. All paths from \(s_2\) to \(d_1\) contain \(v_2\) and \(v_1\).

If we have a path \(P_{s_2,d_1}\) not containing \((v_2, v_1)\), then we must have \(n_1(G, P_{s_1,d_1}) \geq 2\), thus contradicting the fact that we are in C2.

P2. There exists a path \(Q_{s_1,d_1}\) such that \(v_i \notin Q_{s_1,d_1}\), and \(Q_{s_1,d_1} \cap P_{s_2,d_1} = \emptyset\), for \(i = 1\) or \(2\).

Since we have no edge as in (A'), we may assume that the removal of \(v_1\) does not disconnect \(d_1\) from both sources, or the removal of \(v_2\) does not disconnect \(s_2\) from both destinations. However, from P1, the removal of \(v_1\) or \(v_2\) disconnects \(s_2\) from \(d_1\). Therefore, we must have a path \(Q_{s_1,d_1}\) such that \(v_i \notin Q_{s_1,d_1}\), for \(i = 1\) or \(2\). Moreover, if \(Q_{s_1,d_1}\) is not disjoint of \(P_{s_2,d_1}\), we would contradict P1, since there would be a path \(P_{s_2,d_1} \subset Q_{s_1,d_1} \cup P_{s_2,d_1}\) and \(v_i \notin Q_{s_1,d_1} \cup P_{s_2,d_1}\).

P3. If \(i = 1\), we have \(n_1^D(P_{s_2,d_2}, Q_{s_1,d_1}) = 0\) and \(n_1^D(Q_{s_1,d_1}, P_{s_2,d_2}) = 1\), and if \(i = 2\), we have \(n_1^D(Q_{s_2,d_2}, P_{s_1,d_1}) = 0\) and \(n_1^D(P_{s_1,d_1}, Q_{s_2,d_2}) = 1\).

Since \(v_1 \notin Q_{s_1,d_1} \cup P_{s_2,d_1}\), we must have \(n_1^D(P_{s_2,d_2}, Q_{s_1,d_1}) = 0\) (if \(i = 1\)) or \(n_1^D(Q_{s_2,d_2}, P_{s_1,d_1}) = 0\) (if \(i = 2\)), or else we would have a path from \(s_2\) to \(d_1\) not containing \(v_i\), and we would contradict P1. Then, since \(Q_{s_1,d_1}\) and \(P_{s_1,d_1}\) do not have manageable interference, we must have \(n_2^D(Q_{s_1,d_1}, P_{s_2,d_2}) = 1\) (if \(i = 1\)) and \(n_2^D(P_{s_1,d_1}, Q_{s_2,d_2}) = 1\) (if \(i = 2\)).

Since \(n_2^D(Q_{s_1,d_1}, P_{s_2,d_2}) = 1\) (if \(i = 1\)) or \(n_2^D(P_{s_1,d_1}, Q_{s_2,d_2}) = 1\) (if \(i = 2\)), we can assume we have an edge \((v_3, v_4) \in E\) such that \(v_{2+i} \in Q_{s_1,d_i}\) and \(v_{2+i} \in P_{s_i,d_i}\). However, we claim that we must have \(v_0 = v_2\) and \(v_5 = v_1\). If \(v_i \notin Z_{s_i,d_i}\), this is obvious because \(v_i \notin P_{s_i,d_i}\). If \(v_i \notin Z_{s_i,d_i}\), then, if \((v_6, v_5) \neq (v_2, v_1)\), we would have a path from \(s_2\) to \(d_1\) not containing \(v_i\), thus contradicting P1.
Next, we notice that we can assume WLOG that \( i = 1 \). If \( i = 2 \), we can first switch the names of \((s_1, d_1)\) and \((s_2, d_2)\). Then, we also switch the names of \( Z_{s_1, d_1} \) and \( Q_{s_1, d_1} \), and of \((v_2, v_1)\) and \((v_3, v_4)\), and we obtain the case where \( i = 1 \). Thus, from now on we assume \( i = 1 \). We will build our achievability scheme based on the paths \( Z_{s_1, d_1} \), \( Q_{s_1, d_1} \) and \( P_{s_2, d_2} \), an edge \((v_2, v_1)\) such that \( v_2 \in P_{s_2, d_2} \) and \( v_1 \in Z_{s_1, d_1} \), but \( v_1 \notin Q_{s_1, d_1} \), and an edge \((v_3, v_4)\) such that \( v_4 \in P_{s_2, d_2} \) and \( v_3 \in Q_{s_1, d_1} \) but \( v_3 \notin Z_{s_1, d_1} \). Two examples of networks in C2 that satisfy P1-P6 for \( i = 1 \) are shown in Figure 24.

We will now consider two cases and provide a scheme to achieve \( \frac{3}{2} \) degrees-of-freedom in each case. Our schemes will once more 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.

1) \( \ell(v_3) \geq \ell(v_1) \): 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 \( d''_2 \). 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 \( D_N = \frac{3}{2} \). A visual representation of the scheme is shown in Figure 25.

2) \( \ell(v_3) < \ell(v_1) \): In Mode 1, we let \( v_1 \) be a virtual destination \( d'_2 \). Then we consider the path \( P_{s_2, d_2} = P_{s_2, d_2}[s_2, v_2] \oplus (v_2, v_1) \). 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 \( \ell(v_3) < \ell(v_1) \), then \( \ell(v_4) \leq \ell(v_2) \) and the edge \((v_3, v_4)\) will guarantee that \( n_2^2(Q_{s_1, d_1}, P_{s_2, d_2}) \geq 1 \). 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 \( n_2^2(Q_{s_1, d_1}, P_{s_2, d_2}) = 0 \), \( Q_{s_1, d_1} \) and \( P_{s_2, d_2} \) have manageable interference. If \( n_2^2(Q_{s_1, d_1}, P_{s_2, d_2}) = 1 \), the direct interference must be due to an edge \((v_3, v_1)\) so that \( v_3 \to P_{s_2, d_2} \). Otherwise, that would contradict the fact that \( n_2^2(Q_{s_1, d_1}, P_{s_2, d_2}) = 1 \). 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 \( D_\Sigma = \frac{3}{2} \). A visual representation
of the scheme is shown in Figure 26.

Fig. 26. Depiction of Mode 1 and Mode 2 for the achievability scheme in
case C2 when \( \ell(v_3) < \ell(v_1) \).

D. Converse for case C2

In this section, we will show that if our network falls in
C2, and does not fall into (A), (A’), (B), (B’) nor C1, then
\( D_\Sigma \leq \frac{3}{2} \). We will start by deriving additional connectivity
properties, under the assumption that properties P1 to P6 are
satisfied for \( i = 1 \).

P7. The removal of \( v_4 \) disconnects \( d_2 \) from both sources
From P4, we know that the removal of \( v_4 \) disconnects
\( d_2 \) from \( s_1 \). If the removal of \( v_4 \) does not disconnect \( d_2 \)
from \( s_2 \), then we must have a path \( Q_{s_2, d_2} \) not containing
\( v_4 \). We have that \( Q_{s_2, d_2} \) may not intersect \( Q_{s_1, d_1} \), since
that would imply the existence of a path from \( s_1 \) to \( d_2 \) not containing \( \{v_3, v_4\} \) and we would contradict P4. Moreover, we must have \( n_2^p(Q_{s_2, d_2}, Q_{s_1, d_1}) = n_2^p(Q_{s_1, d_1}, Z_{s_2, d_2}) = 0 \). Otherwise, since \( v_1, v_4 \notin Q_{s_1, d_1} \cup Q_{s_2, d_2} \), we would contradict either P1 or P4. But this means that \( Q_{s_1, d_1} \) and \( Q_{s_2, d_2} \) have manageable
interference, which is a contradiction.

P8. The removal of \( v_2 \) disconnects \( s_2 \) from both destinations.
From P1, we know that the removal of \( v_2 \) disconnects
\( s_2 \) from \( d_1 \). If the removal of \( v_2 \) does not disconnect
\( d_2 \) from \( s_2 \), then we must have a path \( Z_{s_2, d_2} \) not containing
\( v_2 \). We have that \( Z_{s_2, d_2} \) may not intersect \( Z_{s_1, d_1} \), or else we would have a path \( P_{s_2, d_1} \) not containing
\( v_2 \), thus contradicting P1. Moreover, we must have \( n_1^p(Z_{s_2, d_2}, Z_{s_1, d_1}) = n_2^p(Z_{s_1, d_1}, Z_{s_2, d_2}) = 0 \). Otherwise, since \( v_2, v_3 \notin Z_{s_1, d_1} \cup Z_{s_2, d_2} \), we would contradict either
P1 or P4. Therefore \( Z_{s_1, d_1} \) and \( Z_{s_2, d_2} \) have manageable
interference, which is a contradiction.

P9. The removal of \( v_1 \) and \( v_3 \) disconnects \( d_1 \) from both sources
From P1, the removal of \( v_1 \) disconnects \( d_1 \) from \( s_2 \).
Thus we assume that the removal of \( v_1 \) and \( v_3 \) does not disconnect \( d_1 \) from \( s_1 \), and we have a path \( M_{s_1, d_1} \), which does not contain \( v_1 \) nor \( v_3 \). Then we have that \( M_{s_1, d_1} \cap P_{s_2, d_2} = \emptyset \), or else we would have a path from \( s_2 \) to \( d_1 \) not containing \( v_1 \), thus contradicting P1. Moreover, P1 and the fact that \( v_1 \notin M_{s_1, d_1} \cup P_{s_2, d_2} \) imply that \( n_2^p(P_{s_2, d_2}, M_{s_1, d_1}) = 0 \). Likewise, P4 and the fact that \( v_3 \notin M_{s_1, d_1} \cup P_{s_2, d_2} \) imply that \( n_2^p(M_{s_1, d_1}, P_{s_2, d_2}) = 0 \), and thus \( M_{s_1, d_1} \) and \( P_{s_2, d_2} \) have manageable interference, which is a contradiction.

P10. There is no path from \( v_1 \) to \( v_3 \).
Suppose \( v_1 \sim v_3 \). Then we must have \( \ell(v_1) < \ell(v_3) \). We will show that our network must contain a
grail subnetwork, and must be in (B’). The network on the left of Figure 24 is an example. We consider paths \( P_{s_2, d_1} = P_{s_2, d_2}[v_2, v_3] \cup (v_2, v_1) \cup Z_{s_1, d_1}[v_1, d_1] \) and \( P_{s_1, d_2} = Q_{s_1, d_1}[v_1, v_3] \cup (v_3, v_4) \cup P_{s_2, d_2}[v_4, d_2] \).
We claim that \( P_{s_1, d_2} \) and \( P_{s_2, d_1} \) must be disjoint. Since \( \ell(v_1) < \ell(v_3) \), we must have \( \ell(v_2) < \ell(v_4) \).
Thus, \( P_{s_2, d_2} \) and \( P_{s_2, d_1} \) and \( P_{s_1, d_2} \) must be disjoint. Therefore, \( P_{s_2, d_1} \) and \( P_{s_1, d_2} \) may only intersect if
\( Z_{s_1, d_1}[v_1, d_1] \) and \( Q_{s_1, d_1}[s_1, v_3] \) intersect. However, since \( v_1 \notin Q_{s_1, d_1}[s_1, v_3] \) and \( v_3 \notin Z_{s_1, d_1}[v_1, d_1] \), it is easy to see that if \( Q_{s_1, d_1}[s_1, v_3] \cup Z_{s_1, d_1}[v_1, d_1] \neq \emptyset \), then there exists a path from \( s_1 \) to \( d_1 \) which does not contain \( v_1 \) nor \( v_3 \), thus contradicting P9. Therefore, if we let \( w_a = v_1 \in P_{s_2, d_1} \) and \( w_b = v_3 \in P_{s_1, d_2} \), we have \( s_1 \sim w_a, w_a \sim w_b \) and \( w_b \sim d_2 \), which satisfies the
description of the grail subnetwork, given in V-B2, by exchanging \( (s_1, d_1) \) and \( (s_2, d_2) \). Therefore, if \( v_1 \sim v_3 \), we are in case (B’), and we may assume \( v_1 \not\sim v_3 \).

We may now prove that under properties P1 through P10,
\( D_\Sigma \leq \frac{3}{2} \). We will derive information inequalities, as we did
for C1. Once more, 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 \( A \triangleq \{v \in V : s_2 \not\sim v\} \) and \( B \triangleq \{v \in V : s_1 \not\sim v\} \). First we have

\[
n R_2 \leq I(W_2; Y^n_{d_2}) + n\epsilon_n
\]

\[
(i) \leq I(\hat{X}^n_B; Y^n_4) + n\epsilon_n = I(\hat{X}^n_B, X^n_3; Y^n_4) - I(X^n_3; Y^n_4|\hat{X}^n_B) + n\epsilon_n
\]

\[
(ii) \leq \frac{n}{2} \log P + n K_{13} - I(X^n_3; Y^n_4|\hat{X}^n_B) + n\epsilon_n, \quad (25)
\]

where (i) follows from the Markov chain \( W_2 \leftrightarrow \hat{X}^n_B \leftrightarrow Y^n_4 \leftrightarrow Y^n_4 \), which is implied by P7 and the fact that \( s_2 \in B \); (ii) follows from the fact that \( I(\hat{X}^n_B, X^n_3; Y^n_4) \) can be upper bounded by \( h(Y^n_4) - h(N^n_B) \) by following the steps in (6), where \( K_{13} \) is a constant, independent of \( P \), for \( P \) sufficiently
large. Next, we have
\[ nR_2 \leq I(W_2; Y^n_{d_1}) + n\epsilon_n \]

(i) \[ \leq I(W_2; \hat{X}_A^n) + n\epsilon_n \]

(ii) \[ = I(W_2; \hat{X}_A^n | X^n_{d_1}) + n\epsilon_n \]

(iii) \[ \leq I(X^n_2; \hat{X}_A^n | X^n_{d_1}) + n\epsilon_n \]

(iv) \[ = I(X^n_2; \hat{X}_A^n | X^n_{d_1}) + I(X^n_2; Y^n_1 | \hat{X}_A^n) + n\epsilon_n \]

(v) \[ \leq I(X^n_2; Y^n_1 | \hat{X}_A^n) + n\epsilon_n \]

where (i) follows because from P8, 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 $W_2 \leftrightarrow (\hat{X}_A^n, X_{d_1}^n) \leftrightarrow Y^n_1$; (ii) follows since $X^n_{d_1}$ is independent of $W_2$; (iii) follows because, given $\hat{X}_A^n$, we have $W_2 \leftrightarrow X^n_{d_1} \leftrightarrow X^n_{A}$; (iv) follows by applying Lemma 1 to $I(X^n_2; \hat{X}_A^n | X^n_{d_1}, Y^n_1)$, because $I(v_1) \setminus \{v_2\} \subset A$, or else we would contradict P1. Furthermore, we have

\[ nR_1 \leq I(W_1; Y^n_{d_1}) + n\epsilon_n \]

(i) \[ \leq I(W_1; \hat{X}_A^n, Y^n_1) + n\epsilon_n \]

(ii) \[ = I(W_1; \hat{X}_A^n) + I(W_1; Y^n_1 | \hat{X}_A^n) + n\epsilon_n \]

(iii) \[ \leq I(X^n_3, \hat{X}_A^n | \hat{X}_B^n) + I(W_1; Y^n_1 | \hat{X}_A^n) + n\epsilon_n \]

(iv) \[ = I(X^n_3, \hat{X}_A^n | \hat{X}_B^n) + I(X^n_3, Y^n_1 | \hat{X}_A^n) + n\epsilon_n \]

(v) \[ \leq I(X^n_3, \hat{X}_A^n | \hat{X}_B^n) + n\epsilon_n \]

where (i) follows from P9, which implies $W_1 \leftrightarrow (\hat{X}_A^n, Y^n_1) \leftrightarrow Y^n_1$; (ii) follows from the fact that $X^n_{d_1}$ is independent of $W_1$; (iii) follows from the fact that, given $\hat{X}_B^n$, we have $W_1 \leftrightarrow X^n_3 \leftrightarrow X^n_{A}$; (iv) follows from the fact that the sum degrees-of-freedom guarantees that the degrees-of-freedom region is given by $D = \{ (D_1, D_2) \in \mathbb{R}_+^2 : D_1 \leq 1, D_2 \leq 1 \}$, where $K$ is independent of $P$, for $i = 1, 2$. Hence, we conclude that $\mathcal{D}$ is always a subset of $\{ (D_1, D_2) \in \mathbb{R}_+^2 : D_1 \leq 1, D_2 \leq 1 \}$. It is straightforward to show that $\mathcal{D}$ is convex. Therefore, for networks that belong to cases (A) and (A') from Theorem 1, the fact that $D_\Sigma = 1$ guarantees that the degrees-of-freedom region is given by

\[ D = \{ (D_1, D_2) \in \mathbb{R}_+^2 : D_1 \leq 1, D_2 \leq 1 \}. \]

This region is depicted in Figure 2(a). The degrees-of-freedom region for networks in cases (B) and (B') can also be easily obtained from the result in Theorem 1. Since for all networks, in cases (B) and (B'), we have $1, 1) \in \mathcal{D}$, we conclude that the degrees-of-freedom region in these cases is given by

\[ D = \{ (D_1, D_2) \in \mathbb{R}_+^2 : D_1 \leq 1, D_2 \leq 1 \}. \]

as depicted in Figure 2(b).
For networks in case (C), we will once again consider the division into cases C1 and C2, as described in Section VI. If our network $\mathcal{N}$ falls into case C2, then, under the assumption that properties P2, P3, P5 and P6 in Section VI-C are satisfied for $i = 1$, we have inequalities (25), (26) and (27). By adding these three inequalities, we obtain

$$n(R_1 + 2R_2) \leq n \log P + n \sum_{j=13}^{16} K_j + n\epsilon_n,$$

which implies that, if $(D_1, D_2) \in D$, then $D_1 + 2D_2 \leq 2$. Since the achievability scheme described in Section VI-C shows that $(1, 1/2) \in D$, we conclude that the degrees-of-freedom region for networks which belong to case C2 (but do not belong to cases (A), (A'), (B), (B') and C1) and satisfy properties P2, P3, P5 and P6 in Section VI-C for $i = 1$ is given by

$$D = \{(D_1, D_2) \in \mathbb{R}_+^2 : D_1 \leq 1, D_1 + 2D_2 \leq 2\}.$$

This region is depicted in Figure 2(d). If properties P2, P3, P5 and P6 in Section VI-C are instead satisfied for $i = 2$, then it is easy to see that analogous results will hold, and the degrees-of-freedom region is given by

$$D = \{(D_1, D_2) \in \mathbb{R}_+^2 : D_1 \leq 1, 2D_1 + D_2 \leq 2\},$$

as depicted in Figure 2(e). The only networks that we still need to consider are networks which are in case C1. We will show that in these cases the degrees-of-freedom region will be given by

$$D = \{(D_1, D_2) \in \mathbb{R}_+^2 : D_1 \leq 1, D_1 + 2D_2 \leq 3/2\},$$

as shown in Figure 2(c). In order to do that we will assume that our network $\mathcal{N}$ contains disjoint paths $P_{s_1,d_1}$ and $P_{s_2,d_2}$ such that $n_1(G) \geq 2$, $n_1^P = 1$, $n_2(G) = 1$ and $n_2^P = 0$, and that the nodes from $\mathcal{N}$ are named as described in Section VI-A and depicted in Figure 27(a). We will then consider the condensed network formed by layers $V_1$, $V_{(u_2)}$ and $V_r$, which is shown in Figure 27(b). Notice that $u_2$ is the same node as $v_2$ in the original network. Therefore, edges $\hat{h}(s_1, u_3)$ and $\hat{h}(u_1, d_2)$ may not exist due to property P1 in Section VI-B. Edges $\hat{h}(s_2, u_1)$ and $\hat{h}(u_3, d_1)$ may or may not exist, and that will depend on the position of the edge $(v_3, v_4)$ in the original network. We will show that points $(1, 1/2)$ and $(1/2, 1)$ are included in $D$ and, by convexity and the fact that $D_2 \leq 3/2$, $D$ must be as shown in Figure 2(c).

First note that we may assume that exactly one of the edges $\hat{h}(s_2, u_1)$ and $\hat{h}(u_3, d_1)$ exists. Otherwise, by removing $u_2$, $P_{s_1,d_1}$ and $P_{s_2,d_2}$ have manageable interference, thus $(1,1)$ is in $D$ and we are in case (B). Hence, we restrict ourselves to two cases: (1) $\hat{h}(u_3, d_1) \neq 0$; and (2) $\hat{h}(s_2, u_1) \neq 0$. We consider each one separately.

1) $\hat{h}(u_3, d_1) \neq 0$: This network is depicted in Figure 28(a). We see that it falls in the case described in Section VI-A2. The achievability scheme provided for $D_\Sigma = 3/2$ shows that $(1/2, 1) \in D$. In order to achieve the point $(1,1/2)$, we need to use a scheme based on real interference alignment, similar to the ones described in [22] and [30].

At source $s_1$, the message $W_1$ will be split into two submessages $W_1^{(1)}$ and $W_1^{(2)}$, while $s_2$ will have a single message $W_2$. Each of these messages will be encoded using a single codebook with codewords of length $n$, which is obtained by uniform i.i.d. sampling of the set

$$\mathcal{U} = \mathbb{Z} \cap \left[-\gamma P^{1/\pi + \epsilon}, \gamma P \pi^{1/\pi + \epsilon}\right],$$

for a small $\epsilon > 0$ and a constant $\gamma$. The rate of this code, i.e., the number of codewords, will be determined later. We will let $a_j[1]$, $a_j[2]$, ..., $a_j[n]$ be the $n$ symbols of the codeword associated to message $W_j^{(j)}$, $j = 1, 2$, and $b[1]$, $b[2]$, ..., $b[n]$ be the $n$ symbols of the codeword associated to message $W_2$. At time $t \in \{1, ..., n\}$, source $s_1$ will transmit

$$X_{s_1}[t] = G(a_1[t] + Ta_2[t]),$$

where $T$ is an irrational number, and $G = \beta P^{1/\pi \epsilon} + \gamma$ is chosen to satisfy the power constraint for a constant $\beta$ to be determined. Source $s_2$ will transmit

$$X_{s_2}[t] = \frac{\hat{h}(s_1,u_2)}{\hat{h}(s_2,u_2)} b[t].$$

The maximum power of a transmit signal from $s_1$ is upper-bounded by

$$\beta^2 P^{1/\pi \epsilon} (1 + T^2)^2 \gamma^2 P^{1/\pi \epsilon} = \beta^2 (1 + T^2) \gamma^2 P,$$

and the maximum power of a transmit signal from $s_2$ is upper-bounded by

$$\beta^2 P^{1/\pi \epsilon} \frac{\hat{h}(s_1,u_2)^2}{\hat{h}(s_2,u_2)^2} \gamma^2 P^{1/\pi \epsilon} = \beta^2 \frac{\hat{h}(s_1,u_2)^2}{\hat{h}(s_2,u_2)^2} \gamma^2 P.$$

Thus, for any choice of $T$ and $\gamma$, parameter $\beta$ can be chosen so that the maximum transmit power at the sources is less than $P$. Next we write the received signals at $u_1$, $u_2$ and $u_3$. We will drop the time $t$ from the notation for simplicity.

$$Y_{u_1} = G\hat{h}(s_1,u_1)(a_1 + Ta_2) + N_{u_1},$$

$$Y_{u_2} = G\hat{h}(s_1,u_2)(a_1 + b + Ta_2) + N_{u_2},$$

$$Y_{u_3} = \frac{G\hat{h}(s_2,u_3)\hat{h}(s_1,u_2)}{\hat{h}(s_2,u_2)} b + N_{u_3}.$$
where $\alpha$ is a constant and $K_1$ is a function of the channel gains. We will choose $\alpha$ so that the power constraint at each relay is satisfied. For example, consider node $u_1$. The noise term at its received signal, $N_{u_1}$, is a linear combination of noises in the original networks, whose coefficients are products of channel gains. Therefore, we may assume that $E[N_{u_1}^2]$ is a constant $\sigma_u^2$. Thus, the transmitted power at $u_1$ is
\[
E[\alpha^2 Y_{u_1}^2] = \alpha^2 \beta^2 P \frac{1+\gamma}{\gamma+1} h^2(s_1, u_1) + \alpha^2 \gamma^2 P \frac{1+\gamma}{\gamma+1} + \alpha^2 \sigma_u^2.
\]
It is now easy to see that $\alpha$ can be chosen independently of $P$ to make sure that $E[\alpha^2 Y_{u_1}^2] \leq P$, for $P$ sufficiently large. The received signal at node $u_2$ can be seen as a noisy observation of a point in the set
\[
U_{u_2} = G \hat{h}(s_1, u_2) \{ x_1 + T x_2 : x_1 \in U + U, x_2 \in U \},
\]
for $x_1 = a_1 + b$ and $x_2 = a_2$, where $U + U = \{ u_1 + u_2 : u_1, u_2 \in U \} \subset Z \cap \{-2\gamma P \frac{1+\gamma}{\gamma+1}, 2\gamma P \frac{1+\gamma}{\gamma+1}\}$. As explained in [22], the fact that $T$ is irrational guarantees that there is a one-to-one map from the points in $U_{u_2}$ to the points $(x_1, x_2) \in (U+U) \times U$. Moreover, from Theorem 1 of [22] and subsequent remarks, we conclude that, for almost all choices of $T$, the minimum distance between two points in $U_{u_2}$ satisfies
\[
d_{\min} > \frac{G \hat{h}(s_1, u_2) \kappa}{\max_{x \in U} x(1+\gamma)},
\]
for some constant $\kappa$. Thus we have
\[
d_{\min} > \frac{G \hat{h}(s_1, u_2) \kappa}{\max_{x \in U} x(1+\gamma)} = \frac{G \hat{h}(s_1, u_2) \kappa}{\max_{x \in U} x(1+\gamma)} P^{1/2}.
\]
Node $u_2$ will map its received signal to the nearest point in $U_{u_2}$, and then use the fact that this point uniquely determines $(a_1 + b)$ and $a_2$ to decode these two integers. We will refer to the output of this procedure as $\hat{a}_1 + b$ and $\hat{a}_2$. If the variance of $N_{u_2}$ is given by $\sigma_{u_2}^2$, then the probability of a wrong decoding at $u_2$ is given by
\[
Pr[\hat{a}_1 + b \neq a_1 + b, \hat{a}_2 \neq a_2] \leq 2 Q \left( \frac{d_{\min}}{2\sigma_{u_2}^2} \right),
\]
where $\delta$ is a positive constant, independent of $P$. The transmit signal at $u_2$ will then be
\[
X_{u_2} = \alpha G \frac{\hat{h}(s_1, u_1) \hat{h}(u_1, d_1)}{\hat{h}(u_2, d_1)} (\hat{a}_1 + b).
\]
We choose $\alpha$ independently of $P$, so that the power constraints at $u_1$, $u_2$ and $u_3$ are simultaneously satisfied, for $P$ sufficiently large. The received signal at the destination $d_1$ is given by
\[
Y_{d_1} = \hat{h}(u_1, d_1) X_{u_1} + \hat{h}(u_2, d_1) X_{u_2} + \hat{h}(u_3, d_1) X_{u_3} + N_{d_1},
\]
where $N_{d_1} = \alpha G \hat{h}(s_1, u_1) \hat{h}(u_1, d_1) (\hat{a}_1 + \hat{b}) + N_{d_1}^{eff}$, where $N_{d_1}^{eff} = \frac{\alpha G \hat{h}(s_1, u_1) \hat{h}(u_1, d_1) (\hat{a}_1 + \hat{b})}{\hat{h}(u_3, d_1)} b + N_{d_1}^{eff}$.

The destinations will first perform a hard-decoding, similar to the one performed by $u_2$. If we assume that the decoding at node $u_2$ was correct, the signal received by $d_1$ is a noisy version of a point in the set
\[
U_{d_1} = \alpha G \hat{h}(s_1, u_1) \hat{h}(u_1, d_1) \{ 2x_1 + 2T x_2 : x_1, x_2 \in U \},
\]
for $x_1 = a_1$ and $x_2 = a_2$. Thus, by using the same argument used previously, it can be shown that $d_1$ can decode $a_1$ and $a_2$ with probability of error smaller than $\exp(-\delta P^2)$, for some positive constant $\delta$. Assuming that the decoding at node $u_2$ was correct, the signal received by $d_2$ is a noisy version of a point in the set
\[
U_{d_2} = \alpha G \hat{h}(s_1, u_1) \hat{h}(u_1, d_1) \hat{h}(u_2, d_2) \hat{h}(u_3, d_1) \{ 2x_1 + 2T x_2 : x_1, x_2 \in U \},
\]
for $x_1 = a_1 + b$ and $x_2 = b$, where $T_2 = -\frac{h(u_2, d_1)h(u_3, d_2)}{h(u_2, d_2)h(u_3, d_1)}$.

Next we notice that $\hat{h}(u_2, d_1)$, $\hat{h}(u_3, d_2)$, $\hat{h}(u_2, d_2)$ and $\hat{h}(u_3, d_1)$ are each a polynomial on the channel gains $h_e$ of the original network with only coefficients 1. Since each $h_e$ is independently drawn from a continuous distribution, we see that if the polynomials $\hat{h}(u_2, d_1)\hat{h}(u_3, d_2)$ and $\hat{h}(u_2, d_2)\hat{h}(u_3, d_1)$ are not identical, then, with probability 1, $T_2$ is an irrational number. From the description of the original network, given in Section VI-A, we see that $u_2 = v_2$ is on a path $P_{v_0, v_1}$ such that $P_{v_0, v_1} \cap P_{u_2, d_2} = \{v_m\}$, $v_m \neq v_2$ and $v_1 \rightarrow P_{s_1, d_1}$. Therefore, there must exist two disjoint paths $P_{u_2, d_1}$ and $P_{u_2, d_2}$. This implies that the determinant

$$\begin{vmatrix} \hat{h}(u_2, d_1) & \hat{h}(u_3, d_1) \\ \hat{h}(u_2, d_2) & \hat{h}(u_3, d_2) \end{vmatrix}$$

is non-zero with probability 1. This implies that $\hat{h}(u_2, d_1)\hat{h}(u_3, d_2)$ and $\hat{h}(u_2, d_2)\hat{h}(u_3, d_1)$ are not identical polynomials. Hence, we conclude that, with probability 1, $T_2$ is an irrational number and $d_2$ can decode $b$ (and also $a_1 + b$, and thus $a_1$, even though $d_2$ does not require the message encoded by the $a_1$'s) with probability at least $1 - \exp(-\delta_3 P^e)$, for some $\delta_3 > 0$.

By applying these hard-decoders, destination $d_1$ obtains the estimates $\hat{a}_1[t]$ and $\hat{a}_2[t]$, and destination $d_2$ obtains the estimates $\hat{b}[t]$, for $t = 1, \ldots, n$. Then they can apply typicality-based decoders in order to decode the messages $W_1^{(1)}$, $W_1^{(2)}$ and $W_2$.

We now determine the rate of the codebook which is used to encode $W_1^{(1)}$, $W_1^{(2)}$ and $W_2$. We notice that for each of the messages $W_1^{(1)}$, $W_1^{(2)}$ and $W_2$, we effectively have a point-to-point discrete channel with input and output alphabet $\mathcal{U}$. Even though we do not calculate the actual transition probabilities, we know that the error probability is upper-bounded as

$$P_e \leq 1 - (1 - \exp(-\delta P^e))(1 - \exp(-\delta_3 P^e)) \leq 1 - (1 - \exp(-\delta \min P^e))^3 \leq 4 \exp(-\delta \min P^e),$$

where $\delta \min = \min(\delta, \delta_2, \delta_3)$. This allows us to lower bound the mutual information between the input $U$ and the output $\hat{U}$ of this channel, for a uniform distribution over the input alphabet. Using Fano’s inequality, we have

$$I(U; \hat{U}) \geq H(U) - H(U|\hat{U})$$

$$\geq \log |\mathcal{U}| - (1 + P_e \log |\mathcal{U}|)$$

$$= (1 - P_e) \log |\mathcal{U}| - 1$$

$$\geq (1 - 4 \exp(-\delta \min P^e))(\frac{1 - \epsilon}{2 + \epsilon} \log P + 1) - 1$$

$$\geq (1 - 4 \exp(-\delta \min P^e))\frac{1 - \epsilon}{2 + \epsilon} - 4.$$
Nodes $u_1$ and $u_3$ will simply perform amplify-and-forward. More precisely, their transmit signals will be given by

$$X_{u_1} = \frac{a \hat{h}(u_2, d_1)}{\hat{h}(s_1, u_1) h(u_1, d_1)} Y_{u_1} = \alpha \hat{G}(u_2, d_1) \left( a + b_1 + \frac{\hat{h}(s_1, u_2) \hat{h}(s_2, u_1)}{\hat{h}(s_2, u_2) \hat{h}(s_1, u_1)} b_2 \right) + \alpha K_1 N_{u_1},$$

$$X_{u_3} = \frac{-a \hat{h}(s_2, u_1) h(u_2, d_2)}{\hat{h}(s_1, u_1) \hat{h}(s_2, u_2) \hat{h}(s_2, u_3) \hat{h}(u_3, d_2)} Y_{u_3} = \frac{\alpha \hat{G}(u_2, d_2)}{\hat{h}(u_3, d_2)} \left( -b_1 - \frac{\hat{h}(s_1, u_2) \hat{h}(s_2, u_1)}{\hat{h}(s_2, u_2) \hat{h}(s_1, u_1)} b_2 \right) + \alpha K_2 N_{u_3},$$

for some constant $\alpha$, where $K_1$ and $K_2$ are functions of the channel gains only. The received signal at node $u_2$ can be seen as a noisy observation of a point in the set

$$U_{u_2} = \hat{G}(s_1, u_2) \left\{ T x_1 + x_2 : x_1 \in U, x_2 \in u_2 + U \right\},$$

for $x_1 = b_1$ and $x_2 = a + b_2$, where $T = \frac{\hat{h}(s_2, u_2) \hat{h}(s_1, u_1)}{\hat{h}(s_1, u_1) \hat{h}(s_2, u_2) \hat{h}(s_2, u_3) \hat{h}(u_3, d_2)}$. Next, we notice that $\hat{h}(u_2, d_1)$, $\hat{h}(u_3, d_2)$, $\hat{h}(u_2, d_2)$ and $\hat{h}(u_3, d_1)$ are each a polynomial on the channel gains $h_e$ of the original network with only coefficients 1. From the description of the original network from Section VI-A, we see that $u_2 = u_2$ is on a path $P_{v_m, v_1}$ such that $P_{v_m, v_1} \cap P_{s_1, d_1} = \emptyset$ and $v_m \in P_{s_2, d_2}$. Therefore, there must exist two disjoint paths $P_{s_1, u_1}$ and $P_{s_2, u_2}$, which implies that the determinant

$$\begin{vmatrix} \hat{h}(s_1, u_1) & \hat{h}(s_2, u_1) \\ \hat{h}(s_1, u_2) & \hat{h}(s_2, u_2) \end{vmatrix} = \hat{h}(s_2, u_2) \hat{h}(s_1, u_1) - \hat{h}(s_1, u_2) \hat{h}(s_2, u_1)$$

is non-zero with probability 1. As we argued before, this implies that, with probability 1 on the value of $T$, after mapping the received signal to the nearest point in $U_{u_2}$, $u_2$ can decode $b_1$ and $a + b_2$ with probability at least $1 - \exp(-\delta_4 P')$, for some positive constant $\delta_4$. The transmit signal at $u_2$ will then be

$$X_{u_2} = -\alpha \hat{G} b_1,$$

where $b_1$ is the output of the hard-decoding performed by $u_2$. We again notice that $\alpha$ can be chosen independently of $P$, for $P$ sufficiently large, guaranteeing that the power constraints at $u_1$, $u_2$ and $u_3$ are simultaneously satisfied. The received signal at destination $d_1$ is given by

$$Y_{d_1} = \hat{h}(u_1, d_1) X_{u_1} + \hat{h}(u_2, d_1) X_{u_2} + N_{d_1} = \alpha \hat{G} h(u_2, d_1) \left( a + b_1 + \frac{\hat{h}(s_1, u_2) \hat{h}(s_2, u_1)}{\hat{h}(s_2, u_2) \hat{h}(s_1, u_1)} b_2 \right) + N_{d_1}^\text{eff},$$

where $N_{d_1}^\text{eff} = \alpha \hat{h}(u_1, d_1) K_1 N_{u_1} + N_{d_1}$. The received signal at $d_2$ is given by

$$Y_{d_2} = \hat{h}(u_2, d_2) X_{u_2} + \hat{h}(u_3, d_2) X_{u_3} + N_{d_2} = \alpha \hat{G} h(u_2, d_2) \left( -b_1 - \frac{\hat{h}(s_1, u_2) \hat{h}(s_2, u_1)}{\hat{h}(s_2, u_2) \hat{h}(s_1, u_1)} b_2 \right) + N_{d_2}^\text{eff},$$

where $N_{d_2}^\text{eff} = \alpha \hat{h}(u_1, d_2) K_2 N_{u_1} + N_{d_2}$. Notice that with probability at least $1 - \exp(-\delta_4 P')$, $Y_{d_1}$ and $Y_{d_2}$ are given by

$$Y_{d_1} = \alpha \hat{G} h(u_2, d_1) \left( a + \frac{\hat{h}(s_1, u_2) \hat{h}(s_2, u_1)}{\hat{h}(s_2, u_2) \hat{h}(s_1, u_1)} b_2 \right) + N_{d_1}^\text{eff},$$

$$Y_{d_2} = \alpha \hat{G} h(u_2, d_2) \left( -2b_1 - \frac{\hat{h}(s_1, u_2) \hat{h}(s_2, u_1)}{\hat{h}(s_2, u_2) \hat{h}(s_1, u_1)} b_2 \right) + N_{d_2}^\text{eff}.$$

The destinations will first perform a hard-decoding, similar to the one performed by $u_2$. If we assume that the decoding at node $u_2$ was correct, the signal received by $d_1$ is a noisy version of a point in the set

$$U_{d_1} = \alpha \hat{G} h(u_2, d_1) \left\{ x_1 + T^{-1} x_2 : x_1, x_2 \in U \right\},$$

for $x_1 = a$ and $x_2 = b_2$. Thus, it can be shown that, with probability 1 over the value of $T^{-1}$, $d_1$ can decode $a$ (and also $b_2$) with probability of error smaller than $\exp(-\delta_4 P')$, for some positive constant $\delta_4$. Again assuming that the decoding at node $u_2$ was correct, the signal received by $d_2$ is a noisy version of a point in the set

$$U_{d_2} = \alpha \hat{G} h(u_2, d_2) \left\{ x_1 + T x_2 : x_1 \in 2U, x_2 \in U \right\},$$

for $x_1 = -2b_1$ and $x_2 = b_2$. With probability 1 over the value of $T$, $d_2$ can decode $b_1$ and $b_2$ with probability at least $1 - \exp(-\delta_6 P')$, for some positive constant $\delta_6$ (if the decoding at $u_2$ was also correct). Therefore, destination $d_1$ obtains $a[t]$ and destination $d_2$ obtains both $b_1[t]$ and $b_2[t]$, for $t = 1, \ldots, n$, and, by applying typicality-based decoders, the messages $W_1$, $W_1^{(1)}$ and $W_2^{(2)}$ can be decoded by their intended destinations. By following the same steps as in case (1), our codebook can have rate

$$R = \left( 1 - 4 \exp(-\delta_{\text{min}} P') \right) \frac{1 - \epsilon \log P}{2 \epsilon + 2} - 4,$$

where $\delta_{\text{min}} = \min(\delta_4, \delta_5, \delta_6)$. Thus, each of the messages carries $\frac{1-\epsilon}{2\epsilon+2}$ degrees-of-freedom. Since $\epsilon$ can be chosen arbitrarily small, we conclude that $(1/2, 1) \in D$, and $D$ is as given in (31). This concludes the derivation of the degrees-of-freedom region of all two-unicast layered Gaussian networks.

In order to state the result in a concise way, we will use the notion of disjoint paths with $(s_1, d_1)$-manageable interference (see Definition 10). Notice that if $P_{s_1, d_1}$ and $P_{s_2, d_2}$ have interference that is both $(s_1, d_1)$-manageable and $(s_2, d_2)$-manageable, they do not necessarily have manageable interference, since the latter requires a single set $S$ for which $n_1(G[S]) \neq 1$ and $n_2(G[S]) \neq 1$. We will describe case C1 in terms of $(s_1, d_1)$-manageable interference through the following claim.

Claim 2. A network $N$ is in case C1 and not in cases (A), (A'), (B) and (B') if and only if it has disjoint paths $P_{s_1, d_1}$ and $P_{s_2, d_2}$ with interference that is not manageable, but is both $(s_1, d_1)$-manageable and $(s_2, d_2)$-manageable.

Proof: By definition, if $N$ is in case C1, we have WLOG $n_1(G) \geq 2$, $n_1^D = 1$, $n_2(G) = 1$ and $n_2^D = 0$, which implies that $P_{s_1, d_1}$ and $P_{s_2, d_2}$ have $(s_1, d_1)$-manageable interference. Moreover, we see from Section VI-B (and Figure 27(a)) that...
We are in case $C_1$. Since $P_{s_1,d_1}$ and $P_{s_2,d_2}$ do not have manageable interference but have interference that is both $(s_1,d_1)$-manageable and $(s_2,d_2)$-manageable, then we must be in case $C_1$. Since $P_{s_1,d_1}$ and $P_{s_2,d_2}$ do not have manageable interference, we must have either $n_i^P = 1$ or $n_i^D = 1$. If $n_i^P = 1$, for $i = 1$ or 2, then since $P_{s_1,d_1}$ and $P_{s_2,d_2}$ have $(s_1,d_1)$-manageable interference, we must have $n_i(G) \geq 2$. Therefore, we cannot have $n_i^P = n_i^D = 1$, or else we would have $n_1(G) \geq 2$ and $n_2(G) \geq 2$. We conclude that the only possible case is $n_i(G) \geq 2$, $n_i^P = 1$, $n_i(G) = 1$ and $n_i^D = 0$, and we are in case $C_1$.

For networks in case $C_2$, we can describe whether properties $P_2$, $P_3$, $P_5$ and $P_6$ in Section VI-C are satisfied for $i = 1$ or $i = 2$ in terms of $(s_i,d_i)$-manageable interference with the following claim.

**Claim 3.** A network $\mathcal{N}$ that is in case $C_2$ and not in cases (A), (A'), (B'), and $C_1$ satisfies properties $P_2$, $P_3$, $P_5$ and $P_6$ in Section VI-C for $i = k$ if and only if it contains paths $Z_{s_1,d_1}$, $Q_{s_1,d_1}$, and $P_{s_2,d_2}$ such that $Z_{s_1,d_1}$ and $P_{s_2,d_2}$ are disjoint and have $(s_1,d_1)$-manageable interference, $Z_{s_1,d_1}$ and $P_{s_2,d_2}$ are disjoint and have $(s_2,d_2)$-manageable interference.

**Proof:** If $\mathcal{N}$ satisfies properties $P_2$, $P_3$, $P_5$ and $P_6$ in Section VI-C for $i = k$, then it is easy to see that it must contain paths $Z_{s_1,d_1}$, $Q_{s_1,d_1}$, and $P_{s_2,d_2}$ as in the statement of the claim. Conversely, we want to argue (WLOG) that if we have paths $Z_{s_1,d_1}$ and $Z_{s_2,d_2}$, each disjoint from $P_{s_2,d_2}$, such that $Z_{s_1,d_1}$ and $P_{s_2,d_2}$ have $(s_2,d_2)$-manageable (but not $(s_1,d_1)$-manageable) interference and $Q_{s_1,d_1}$ and $P_{s_2,d_2}$ have $(s_1,d_1)$-manageable (but not $(s_2,d_2)$-manageable) interference, then the network $\mathcal{N}$ must satisfy properties $P_2$, $P_3$, $P_5$ and $P_6$ for $i = 1$. Since $Z_{s_1,d_1}$ and $P_{s_2,d_2}$ have $(s_2,d_2)$-manageable (but not $(s_1,d_1)$-manageable) interference, we must have $n_1(G,Z_{s_1,d_1}) = n_1^P(P_{s_2,d_2},Z_{s_1,d_1}) = 1$. Similarly, we must have $n_2(G,P_{s_2,d_2}) = n_2^D(Q_{s_1,d_1},P_{s_2,d_2}) = 1$. We let $(v_2,v_1)$ be the unique edge such that $v_2 \in P_{s_2,d_2}$ and $v_1 \in Z_{s_1,d_1}$. Now we notice that we must have $n_2^D(P_{s_2,d_2},Q_{s_1,d_1}) = 0$, since otherwise we would have a path $P_{s_2,d_2}$ which does not contain $\{v_2,v_1\}$, and this would imply that $n_1(G,Z_{s_1,d_1}) \geq 2$, which contradicts the fact that $Z_{s_1,d_1}$ and $P_{s_2,d_2}$ do not have $(s_1,d_1)$-manageable interference. Similarly, we must have $n_2^D(Z_{s_1,d_1},P_{s_2,d_2}) = 0$.

The results obtained in this section regarding the complete degrees-of-freedom region are summarized in Theorem 2.

**VIII. Conclusion**

We explored the degrees-of-freedom of two-unicast layered Gaussian networks. Our result shows that, in terms of degrees-of-freedom, there are essentially three categories of such networks. In the first one, the network connectivity creates a bottleneck for the information flow, forcing all the messages to be decodable at a single node. Therefore, we only have one degree-of-freedom. In the second, the interference can be completely avoided or neutralized and we achieve two degrees-of-freedom. Networks which contain a grail or a butterfly achieve their two degrees-of-freedom by “borrowing” their achievability schemes from linear network coding. When no such structures exist, we must find two disjoint paths and verify whether it is possible to completely neutralize their interference. The notion of manageable interference arises quite naturally in these cases: we simply want to make sure that no path receives a single interference from the other path. This way, a path may either receive no interference at all or receive at least two interferences from the other path, which allows for interference cancellation techniques. The remaining networks have exactly 3/2 sum degrees-of-freedom. The intuitive reason is that these networks do not have a single node through which all information must pass restricting the degrees-of-freedom to one, but their interference is not manageable. One way to achieve the 3/2 sum degrees-of-freedom is to perform a sort of “scheduling” of the transmissions, in order to be able to avoid or neutralize the interference. The buffering involved in this scheduling costs us half a degree-of-freedom. This third class of networks can be further subdivided into three subclasses, according to which of the points (1,1/2) and (1/2,1) are included in the degrees-of-freedom region.

The achievability schemes considered show that only a small fraction of the nodes must in fact be careful in their relaying operations. In fact, we achieve the sum degrees-of-freedom by converting any multi-layered network into a condensed network with at most four layers. Thus, the nodes in at most four layers (the source layer, the destination layer, and two intermediate layers) perform relaying operations that require channel state information. All the other nodes are simply forwarding their received signals. Another interesting aspect of the achievability schemes is that, in most cases, a linear scheme suffices to achieve two degrees-of-freedom. In some cases, however, we must resort to a more sophisticated scheme such as real interference alignment. For example, this is the case of the $2 \times 2 \times 2$ interference channel studied in [30], where a simple linear scheme cannot achieve two degrees-of-freedom (unless the channel gains vary with time).

Our assumption of a layered network topology serves mainly to simplify the problem. If non-layered networks are considered, the main issue is that interference may occur not only between signals originated at different sources, but also between signals originated at different times. Therefore, in order to perform interference cancellation, for example, one needs to make sure that the two cancelling signals correspond to the same time-version of the source signal. For non-layered networks, other techniques such as signal delaying and backward decoding must be used to achieve the degrees-of-freedom, and the problem becomes significantly more difficult. However, the layered network assumption is not artificial. Since the layered topology simplifies the analysis and the implementation of coding schemes, it is desirable in practice, and it can actually be simulated in practical contexts by having the transmitters on each layer transmit on a different frequency band, which allows us to assume that the links only exist between consecutive layers. Moreover, a layered structure can also arise from the scheduling of the transmitting nodes in a wireless network with half-duplex nodes. In this context, each hop would capture which nodes are transmitting and which
nodes are receiving at a given time slot, and the same node could appear in multiple layers, since they may be transmitting at multiple time-slots.

The natural extension of this work would be to consider more than two information flows in the network. Recently, in [31], networks with two source-destination pairs where each source has a message to each destination (for a total of four messages) were considered. Interestingly, it was shown that the sum degrees-of-freedom can also take values 4/3 and 5/3, in addition to the values 1, 3/2 and 2 that are possible in the setup considered in this paper. The next step would thus be to consider more than two source-destination pairs. The main issue lies in the combinatorial complexity of a larger number of source-destination pairs. For example, if one were to extend a notion such as manageable interference to more than two source-destination pairs, not only would the number of interfering pairs increase, but also which subset of the other sources contributes to each interference. Moreover, it is not clear whether the non-linear schemes that were necessary for networks such as the 2 x 2 x 2 interference channel can be easily extended to networks with more than two source-destination pairs. In particular, the characterization of degrees-of-freedom of the 3 x 3 x 3 interference channel remains an open problem.

A. Proof of Lemma 1

\[ I(X^n_b;\hat{X}^n_c, A^n_A, Y^n_T) \]
\[ = I(X^n_b;\{X^n_{c,j} : j \text{ s.t. } v_j \in \mathcal{O}(v_c)\}|Y^n_b, \hat{X}^n_A, Y^n_T) \]
\[ \leq I(X^n_b;\{N^n_{c,j} = h_{c,b} N^n_{c,b} : j \text{ s.t. } v_j \in \mathcal{O}(v_c)\}|Y^n_b, \hat{X}^n_A, Y^n_T) \]
\[ \leq h(|\{N^n_{c,j} = h_{c,b} N^n_{c,b} : j \text{ s.t. } v_j \in \mathcal{O}(v_c)\}|) \]
\[ \leq \frac{n|D|}{2} \log(2\pi e) \]

where (i) follows from the fact that \( Y^n_b = \sum_{v_j \in A} \hat{X}^n_{c,b} = \hat{X}^n_{c,b} \); (ii) follows since, for \( j = b \), \( N^n_{c,j} = h_{c,b} N^n_{c,b} = 0 \); (iii) follows by letting \( \kappa = 1 + (\max_{c,f \in E} h_{c,b})^2 \); (iv) follows because conditioning reduces entropy and thus we can condition on \( N^n_{c,b} \); (iv) follows from the fact that, since for \( u \in D \) and \( w \in T \), \( u \nleftrightarrow w \), \( N^n_{c,u} \) is independent of all the random variables conditioned on.

B. Proof of Lemma 2

Consider the nodes in \( I(v^n_p) \). Assume, by contradiction, that there are no two paths \( P_{s_1,v^n_p} \) and \( P_{s_2,v^n_p} \) such that \( P_{s_1,v^n_p} \cap P_{s_2,v^n_p} = \{v^n_p\} \). Then, we do not have two vertex-disjoint paths starting in \( \{s_1, s_2\} \) and ending in \( I(v^n_p) \). From Menger’s Theorem, there exists a node \( v_d \) whose removal disconnects \( \{s_1, s_2\} \) from \( I(v^n_p) \), and thus from \( v^n_p \). The existence of the path \( P_{s_1,d} \) containing \( v^n_p \) guarantees that \( v_d \in P_{s_1,d} \). Since the removal of \( v^n_p \) disconnects \( s_1 \) from \( d_i \), and the removal of \( v_d \) disconnects \( \{s_1, s_2\} \) from \( v^n_p \), we conclude that the removal of \( v_d \) also disconnects \( s_1 \) from \( d_i \).

But this is a contradiction to the fact that \( v^n_p \) was the first such node.

C. Proof of Claim 1

We let \( G = (V, E) \) be the graph of our original network, and we construct an extended network \( N' \) with graph \( G = (V', E') \) in the following way. We let the layers in \( V' \) be \( V_1, V'_2, V'_3, \ldots, V'_r \), where \( V'_j \) is a copy of \( V_j, j = 1, \ldots, r \). The edges between \( V'_j \) and \( V'_{j+1} \), for \( j = 1, 2, \ldots, r-1 \), are the same as the edges between \( V_j \) and \( V_{j+1} \) in \( G \). To add the edges between \( V_j \) and \( V'_{j+1} \), for \( j = 1, 2, \ldots, r \), we simply connect each \( v_j \in V_j \) to its copy in \( V'_{j+1} \). The source-destination pairs of \( N' \) are the same as of \( N' \).

Next we claim that if we have an edge \( e \in E' \) whose removal from \( N' \) disconnects \( d_i \) from both sources \( s_j \) and \( s_i \) from both destinations, \( i \in \{1, 2\} \), then our original network falls in (A). Suppose we have such an edge \( e \in E' \). If \( e \in V_j \times V'_{j+1} \) 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 \( s_j \) and \( s_i \) 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 \( e \) in \( V_j \times V'_{j+1} \) which is adjacent to \( e \) must also disconnect \( d_i \) from both sources \( s_j \) and \( s_i \) from both terminals. This is because all paths from any source to any destination which contain the nodes in \( e \) must also contain the nodes in \( e \). Therefore, \( e \) can be translated to a node \( v \) in \( N' \) whose removal disconnects \( d_i \) from both sources \( s_j \) and \( s_i \) from both destinations, and \( N' \) falls into case (A).

Therefore, the absence of a node \( v \) as described in (A) in our network \( N' \) implies that \( N' \) does not contain an edge whose removal disconnects \( d_i \) from both sources \( s_j \) and \( s_i \) from both destinations for some \( i \in \{1, 2\} \). Thus, we employ a result for double unicast networks, shown in both [27] and [28], which guarantees that the extended network \( N' \) must contain one of the three structures shown in Figure 4: 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 \( N' \), any pair of edge-disjoint paths is also vertex-disjoint, and corresponds to a pair of vertex-disjoint paths in \( N' \). Thus, we conclude that if our network \( N' \) 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 structure or a butterfly structure.
REFERENCES

[1] I. Shomorony and A. S. Avestimehr. Sum degrees-of-freedom of two-unicast wireless networks. Proceedings of IEEE International Symposium on Information Theory, St. Petersburg, Russia, 2011.
[2] L. R. Ford and D. R. Fulkerson. Maximal flow through a network. Canadian Journal of Mathematics, 8:399–404, 1956.
[3] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung. Network information flow. IEEE Transactions on Information Theory, 46(4):1204–1216, July 2000.
[4] S.-Y.R. Li, R. W. Yeung, and N. Cai. Linear network coding. IEEE Transactions on Information Theory, 43(2):371–81, February 2003.
[5] R. Koetter and M. Médard. An algebraic approach to network coding. IEEE/ACM Trans. Netw., 11(5):782–795, 2003.
[6] A. S. Avestimehr, S. Diggavi, and D. Tse. Wireless network information flow: a deterministic approach. IEEE Transactions on Information Theory, 57(4), April 2011.
[7] Ayfer Ozgur and Suhas Diggavi. Approximately achieving gaussian relay network capacity with lattice codes. ISIT Proceedings, 2010.
[8] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung. Noisy network coding. IEEE Information Theory Workshop (ITW), 2010.
[9] A. S. Motahari and A. K. Khandani. Capacity bounds for the Gaussian interference channel. IEEE Transactions on Information Theory, 55(2):620–643, February 2009.
[10] X. Shang, G. Kramer, and B. Chen. A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels. IEEE Transactions on Information Theory, 55(2):689–699, February 2009.
[11] 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, 55(7):3032–3050, July 2009.
[12] S. Sridharan, A. Jafarian, S. Vishwanath, and S. A. Jafar. Capacity of symmetric k-user gaussian very strong interference channels. IEEE Global Telecommunications Conference, 2008.
[13] A. El Gamal and M. Costa. The capacity region of a class of deterministic interference channels. IEEE Transactions on Information Theory, 28(2):343–346, March 1982.
[14] H. Sato. The capacity of the Gaussian interference channel under strong interference. IEEE Trans. on Info. Theory, 27:786–788, November 1981.
[15] A. B. Carleial. Interference channels. IEEE Trans. on Info. Theory, 24:60–70, January 1978.
[16] R. Etkin, D. Tse, and H. Wang. Gaussian interference channel capacity to within one bit. IEEE Transactions on Information Theory, 54(12):5534–5562, December 2008.
[17] 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, Urbana, IL, September 26-28 2007.
[18] Guy Bresler and David Tse. The two-user gaussian interference channel: a deterministic view. European Transactions on Telecommunications, 19(4):333–354, June 2008.
[19] V.R. Cadambe and S.A. Jafar. Interference alignment and spatial degrees of freedom for the k user interference channel. IEEE International Conference on Communications (ICC), pages 971–975, 2008.
[20] R. Etkin and D. Tse. Degrees of freedom in underspread mimo fading channels. IEEE Transactions on Info. Theory, 52:1576–1608, April 2006.
[21] S.A. Jafar and S. Shamai (Shitz). Degrees of freedom region for the mimo x channel. IEEE Transactions on Info. Theory, 54(1):151–170, January 2008.
[22] A. S. Motahari, S. O. Gharan, and A. K. Khandani. Real interference alignment with real numbers. Submitted to IEEE Transactions on Information Theory, 2009.
[23] M. A. Maddah-Ali, S. A. Motahari, and Amir K. Khandani. Communication over mimo x channels: Interference alignment, decomposition, and performance analysis. IEEE Transactions on Information Theory, 54(8):3457–3470, August 2008.
[24] S. A. Jafar and M. Fakhereddin. Degrees of freedom for the mimo interference channel. IEEE Transactions on Info. Theory, 53(7):2637–2642, July 2007.
[25] T. C. Hu. Multi-commodity network flows. Operations Research, 11(3):344–360, 1963.
[26] 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. In Proc. IEEE International Symposium on Information Theory, 2007.
[27] S. Shenvi and B. K. Dey. A simple necessary and sufficient condition for the double unicast problem. in Proceedings of ICC, 2010.
[28] K. Cai, K. B. Letaief, P. Fam, and R. Feng. On the solvability of 2-pair unicast networks — a cut-based characterization. ArXiv, 2010.
[29] S. Mohajer, S. N. Diggavi, C. Fragouli, and D.N.C. Tse. Approximate capacity of gaussian interference-relay networks with weak cross links. ArXiv, 2010.
[30] T. Gou, S.A. 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.
[31] C. Wang, T. Gou, and S. A. Jafar. Multiple unicast capacity of 2-source 2-sink networks. IEEE Global Telecommunications Conference, 2011.