Interference Channels with Source Cooperation

Vinod Prabhakaran and Pramod Viswanath
Coordinated Science Laboratory
University of Illinois, Urbana-Champaign
Urbana, IL 61801.

May 19, 2009

Abstract

Interference is a fundamental feature of the wireless channel. Cooperation among the radios has been shown to manage interference in a multiple unicast wireless network in a near optimal scaling law sense [4]. In this paper, we study the two user Gaussian interference channel where the source nodes can *both* talk and listen. Focusing on the full-duplex mode of operation, we characterize the sum capacity up to 18 bits. Novel inner and outer bounds are derived in arriving at our main result.

1 Introduction

The standard engineering approach to dealing with interference in wireless systems is to orthogonalize the transmissions and/or treat interference as noise at the receivers. However, these strategies can be far from optimal in several canonical scenarios, including the classical two user Gaussian interference channel [1, 2]. Superposition coding and interference alignment has been shown to perform well in interference channels (where the sources only transmit and destinations only receive) [2, 3].

A different approach to interference management is available when the wireless nodes can cooperate among themselves (this situation is not feasible in the classical interference channel when sources only transmit and destinations only receive). A coarse result (scaling of symmetric capacity as the number of radios grows) derived in [4] shows that distributed cooperation (a so-called *hierarchical MIMO* strategy) can manage interference between the different traffic flows so well as to get near the performance of (co-located) MIMO operation among the nodes. While this is a strong result, it is coarse – in the asymptotic regime of a very large number of radios.

Our goal in this paper is to better understand the role of cooperation in providing interference management gains. We take a fundamental (information theoretic) approach by studying the capacity region of a simple (yet, canonical) wireless network: we start with the classical two user Gaussian interference channel, but endow the sources with the capability to listen as well (apart from the usual capability of transmission). The cooperative links are over the same frequency band as the rest of the links. We treat only the sum-rate under a full-duplex mode of operation in this paper. The main result is a characterization of the sum-rate within 18 bits.

Our main focus while deriving our main result is on constructing novel achievable schemes as well as new converse techniques. Our achievable scheme is generic, in the sense that it can be described in the context of a general discrete memoryless interference channel with source cooperation. This general nature of the scheme (as well as the need to calculate a uniform gap between our inner and outer bounds) results in a large gap (presently 18 bits). In specific instances our characterization is readily sharpened – we provide an example where the gap reduces to 6 bits, for instance. We also provide a few general recipes to improve the inner and outer bounds (though calculating the improvement explicitly is rather involved).

The superposition scheme of Han and Kobayashi [5] for the two user interference channel involves the two destinations partially decoding the interference they receive. In order to facilitate this, the sources encode their messages as a superposition of two partial messages. One of these partial messages, termed the public message, is decoded by the destination where it appears as interference along with the two partial messages which are meant for this destination. The other partial message, called the private message, from the interfering source is treated as noise. Our achievable scheme preserves these two types of messages and employs a further two types of messages:

- a cooperative-public message which is decoded not only by both the destinations, but also by the other source which aids its own destination in decoding it, and
- a cooperative-private message which is decoded by the destination to which it is intended and by the other source which cooperates with the original source in its transmission.

These two messages have similarities to the suggestions of Tuninetti [6], Cao and Chen [7], and Yang and Tuninetti [8], but differ in the details of implementation which have a bearing on the rates achieved.

Other related works include [9] which studies the two-user interference channel under cooperation, [10] which investigates the symmetric two-user interference channel under noiseless feedback, and [11, 12] which study a two-stage, two-source interference network.

2 Problem Statement

We consider the following channel model for source cooperation (see Figure 1). At each discrete-time instance – indexed by $t = 1, 2, \ldots$ – the source nodes 1 and 2 send out, respectively, $X_1(t)$ and $X_2(t) \in \mathbb{C}$. The source nodes 1 and 2, and the destination nodes 3 and 4

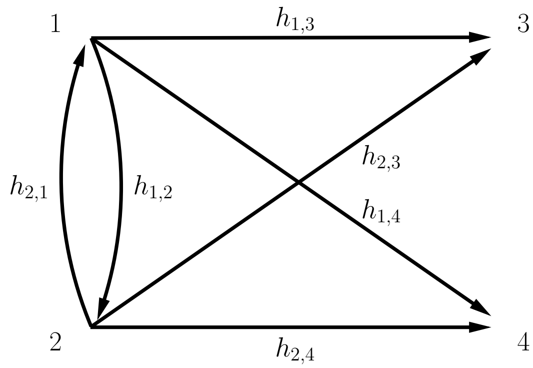


Figure 1: Problem Setup

receive respectively

$$\begin{aligned}
Y_1(t) &= h_{2,1}X_2(t) + N_1(t),\\
Y_2(t) &= h_{1,2}X_1(t) + N_2(t),\\
Y_3(t) &= h_{1,3}X_1(t) + h_{2,3}X_2(t) + N_3(t),\\
Y_4(t) &= h_{2,4}X_2(t) + h_{1,4}X_1(t) + N_4(t),
\end{aligned}$$

where the channel coefficients h's are complex numbers and $N_k(t),\ k = 1, 2, 3, 4,\ t = 1, 2, \ldots$ are independent and identically distributed (i.i.d.) zero-mean Gaussian random variables with unit variance. It is easy to see that, without loss of generality, we may consider a channel where the channel coefficients $h_{1,3}, h_{1,2}, h_{2,4}, h_{2,1}$ are replaced by their magnitudes $|h_{1,3}|, |h_{1,2}|, |h_{2,4}|, |h_{2,1}|$, and the channel coefficient $h_{1,4}$ is replaced by $|h_{1,4}|e^{j\theta/2}$ and $h_{2,3}$ is replaced by $|h_{2,3}|e^{j\theta/2}$, where $\theta \stackrel{\text{def}}{=} \arg(h_{1,4}) + \arg(h_{2,3}) - \arg(h_{1,3}) - \arg(h_{2,4})$. We will consider this channel. We will also assume that $|h_{1,2}| = |h_{2,1}| = h_C$, say, which models the reciprocity of the link between nodes 1 and 2. Further, we consider unit power constraints without loss of generality.

There is a causality restriction on what the sources are allowed to transmit: it can only depend on the message it sends and everything it has heard up to the previous time instance, *i.e.*,

$$X_k(t) = f_{k,t}(M_k, Y_k^{t-1}),\ k = 1, 2,$$

where M_k is the message to be conveyed by source k and f is a (deterministic) encoding function. A blocklength-T codebook of rate (R_1, R_2) is (for each $k = 1, 2$) a sequence of encoding functions, $f_{k,t},\ t = 1, 2, \ldots, T$ such that

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}|X_k(t)|^2\right] \le 1,$$

with message alphabets $\mathcal{M}_k = \{1, 2, \ldots, 2^{TR_k}\}$ over which the messages M_k are uniformly distributed, and decoding functions $g_{k+2} : \mathcal{C}^T \to \mathcal{M}_k$. We say that a rate (R_1, R_2) is achievable if there is sequence of rate (R_1, R_2) codebooks such that as $T \to \infty$,

$$\mathbb{P}\left(g_{k+2}(Y_{k+2}^T) \ne M_k\right) \to 0,\ k = 1, 2.$$

We would also like to consider a linear deterministic model (introduced in [13]) for the above channel. In order to treat both models together, we will adopt the following:

$$\begin{aligned}
Y_1(t) &= h^*_{2,1}(X_2(t)),\\
Y_2(t) &= h^*_{1,2}(X_1(t)),\\
Y_3(t) &= h_{1,3}(X_1(t)) + h^*_{2,3}(X_2(t)),\\
Y_4(t) &= h_{2,4}(X_2(t)) + h^*_{1,4}(X_1(t)).
\end{aligned}$$

Here the functions with a $*$ are potentially random functions and the others are deterministic functions.

Gaussian case: In the Gaussian case, we specialize to (with some abuse of notation[1]):

$$\begin{aligned}
h^*_{2,1}(X_2) &= h_{2,1}X_2 + N_1,\\
h^*_{1,2}(X_1) &= h_{1,2}X_1 + N_2,\\
h_{1,3}(X_1) &= h_{1,3}X_1,\\
h_{2,4}(X_2) &= h_{2,4}X_2,\\
h^*_{2,3}(X_2) &= h_{2,3}X_2 + N_3,\\
h^*_{1,4}(X_1) &= h_{1,4}X_1 + N_4.
\end{aligned}$$

We will further assume that $h_{1,2} = h_{2,1}$ which is justified by the reciprocity of the links between the sources.

Linear deterministic case: Let $n_{1,2}, n_{1,3}, n_{1,4}, n_{2,1}, n_{2,3}, n_{2,4}$ be non-negative integers and $n \stackrel{\text{def}}{=} \max(n_{1,2}, n_{1,3}, n_{1,4}, n_{2,1}, n_{2,3}, n_{2,4})$. The inputs to the channel X_1 and X_2 are n-length vectors over a finite field $\mathbb{F}$. We define

$$\begin{aligned}
h^*_{2,1}(X_2) &= \mathbf{S}^{n-n_{2,1}}X_2,\\
h^*_{1,2}(X_1) &= \mathbf{S}^{n-n_{1,2}}X_1,\\
h_{1,3}(X_1) &= \mathbf{S}^{n-n_{1,3}}X_1,\\
h_{2,4}(X_2) &= \mathbf{S}^{n-n_{2,4}}X_2,\\
h^*_{2,3}(X_2) &= \mathbf{S}^{n-n_{2,3}}X_2,\\
h^*_{1,4}(X_1) &= \mathbf{S}^{n-n_{1,4}}X_1,
\end{aligned}$$

where

$$\mathbf{S} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0\\ 1 & 0 & 0 & \dots & 0\\ 0 & 1 & 0 & \dots & 0\\ \vdots & \ddots & \ddots & \ddots & \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}_{n\times n}$$

[1]The correct notation would be $Y_1(t) = h^{*(t)}_{2,1}(X_2(t)) = h_{2,1}X_2(t) + N_1(t)$. This t-index in the notation for random functions like $h^{*(t)}_{2,1}$ will be suppressed. We will tacitly assume that application of $*$-ed functions for different values of t result in independent realizations of N's.

is the $n \times n$ shift matrix. Further, to model the reciprocity of the links between the sources, we set $n_{1,2} = n_{2,1} = n_C$, say.

3 Main Results

3.1 Sum-rate Characterization

The following theorems characterize the sum-rates of the interference channels with source cooperation introduced in the previous section.

Theorem 1 Linear deterministic case. *The sum-capacity of the linear deterministic channel with source cooperation is the minimum of the following:*

$$u_1 = \max(n_{1,3} - n_{1,4} + n_C, n_{2,3}, n_C) + \max(n_{2,4} - n_{2,3} + n_C, n_{1,4}, n_C), \tag{1}$$

$$u_2 = \max(n_{1,3}, n_{2,3}) + (\max(n_{2,4}, n_{2,3}, n_C) - n_{2,3}), \tag{2}$$

$$u_3 = \max(n_{2,4}, n_{1,4}) + (\max(n_{1,3}, n_{1,4}, n_C) - n_{1,4}), \tag{3}$$

$$u_4 = \max(n_{1,3}, n_C) + \max(n_{2,4}, n_C), \text{ and} \tag{4}$$

$$u_5 = \begin{cases} \max(n_{1,3} + n_{2,4}, n_{1,4} + n_{2,3}), & \text{if } n_{1,3} - n_{2,3} \neq n_{1,4} - n_{2,4}, \\ \max(n_{1,3}, n_{2,4}, n_{1,4}, n_{2,3}), & \text{otherwise.} \end{cases} \tag{5}$$

The achievability of the above theorem in proved in Appendix B, and the upperbounds are derived in Appendix D.

Theorem 2 Gaussian case. *The sum-capacity of the Gaussian channel with source cooperation is at most the minimum of the following five quantities and a sum-rate can be achieved to within a constant (18 bits) of this minimum.*

$$u_1 = \log\left(1 + \left(\frac{|h_{1,3}|}{\max(1, |h_{1,4}|)} + \frac{|h_{2,3}|}{\max(1, |h_C|)}\right)^2\right)(1 + |h_C|^2) \tag{6}$$

$$+ \log\left(1 + \left(\frac{|h_{2,4}|}{\max(1, |h_{2,3}|)} + \frac{|h_{1,4}|}{\max(1, |h_C|)}\right)^2\right)(1 + |h_C|^2), \tag{7}$$

$$u_2 = \log 2\left(1 + (|h_{1,3}| + |h_{2,3}|)^2\right)\left(1 + \frac{\max(|h_{2,4}|^2, |h_{2,3}|^2, |h_C|^2)}{\max(1, |h_{2,3}|^2)}\right), \tag{8}$$

$$u_3 = \log 2\left(1 + (|h_{2,4}| + |h_{1,4}|)^2\right)\left(1 + \frac{\max(|h_{1,3}|^2, |h_{1,4}|^2, |h_C|^2)}{\max(1, |h_{1,4}|^2)}\right), \tag{9}$$

$$u_4 = \log\left(1 + |h_{1,3}|^2 + |h_C|^2\right) + \log\left(1 + |h_{2,4}|^2 + |h_C|^2\right), \text{ and} \tag{10}$$

$$u_5 = \log\left(1 + 2\left(|h_{1,3}|^2 + |h_{2,4}|^2 + |h_{1,4}|^2 + |h_{2,3}|^2\right)\right.$$
$$\left. + 4\left(|h_{1,3}h_{2,4}|^2 + |h_{1,4}h_{2,3}|^2 - 2|h_{1,3}h_{2,4}h_{1,4}h_{2,3}|\cos\theta\right)\right). \tag{11}$$

The achievability proof is presented in Appendix C and the converse in Appendix D.

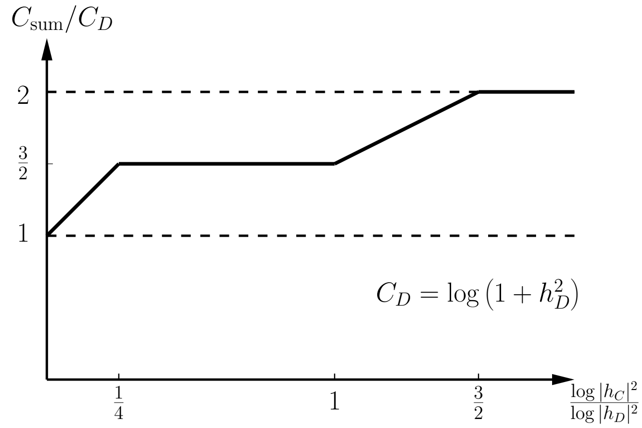


Figure 2: Normalized sum-capacity of the symmetric interference channel with $h_I = \sqrt{h_D}$ under source cooperation in the limit of $h_D \to \infty$ keeping $\log|h_I|^2/\log|h_D|^2$ and $\log|h_C|^2/\log|h_D|^2$ fixed.

3.2 Gains from cooperation

To illustrate the gains from cooperation, in this section we will explore a specific instance of the symmetric interference channel: $|h_{1,3}| = |h_{2,4}| = h_D$, $|h_{1,4}| = |h_{2,3}| = h_I = \sqrt{h_D}$, and $\theta = 0$. In appendix F we prove the following proposition.

Proposition 3 *Let C be the minimum of the following*

$$2\log 2h_D\left(1 + h_C^2\right),$$
$$\log 2h_D^2\left(1 + \frac{\max(h_D^2, h_C^2)}{h_D}\right),$$
$$\log\left(4h_D^4\right).$$

For the symmetric channel described above, for any $\epsilon > 0$, the sum-capacity lies in $(C - 6 - \epsilon, C + \epsilon)$, for sufficiently large h_D.

We plot in Fig. 2, as a function of $\log|h_C|^2/\log|h_D|^2$, the sum-rate C normalized by the capacity of the direct link, in the limit of $|h_D| \to \infty$ while keeping the ratios $\log|h_C|^2/\log|h_D|^2$ and $\log|h_I|^2/\log|h_D|^2$ constants. Since C is achievable within a constant gap (6 bits), this plot is also that of the sum-capacity in this limit. This reveals three regimes of operation:

- $\log|h_C|^2/\log|h_D|^2 \leq 1/2$. In this regime, the plot shows that the capacity increases linearly with the strength of the cooperation link (measured in the dB scale). For every 3dB increase in link strength the sum-capacity increases by 2 bits. Note that without cooperation, the sum-capacity is essentially achieved by time-sharing. Thus, cooperation can be seen to deliver significant gains.

- $1/2 < \log|h_C|^2/\log|h_D|^2 \leq 1$. The linear gain saturates when the cooperation link strength is half the direct link strength. No further gains are accrued until the cooperation link is as strong as the direct link.

- $1 < \log|h_C|^2/\log|h_D|^2 \leq 3/2$. The capacity again increases linearly with the cooperation link strength, but here an increase in capacity by 2 bits requires a 6dB increase in the cooperation channel strength. This linear increase continues until the cooperation capacity is approached when the cooperation link is 3/2 times as strong as the direct link, after which the capacity is flat.

4 Coding Scheme: Illustrative Examples

In this section, our cooperative coding scheme is illustrated through a few simple examples. These examples have been hand-picked so that uncoded (signal processing) strategies are enough to achieve the sum-capacity. However, they cover the key novel aspects of our coding strategy. In particular, the first example shows how the sources cooperate in conveying cooperative-public messages, which are messages decoded by both sources and which benefit from cooperation. Example 1 also involves the use of private messages which are decoded only by the destinations to which they are intended and get conveyed without the benefits of cooperation. Example 2 shows how cooperative-public messages can occur together with the two types of messages in Han-Kobayashi's scheme, namely, private messages and public messages (public messages are decoded by both the destinations and do not benefit from cooperation). Example 3 illustrates a second kind of cooperative message: cooperative-private message which benefits from cooperation and is decoded only by the destination which is interested in it.

Example 1: For the linear deterministic case, let us consider the symmetric channel with direct links $n_{1,3} = n_{2,4} = n_D$, say, and interference links $n_{1,4} = n_{2,3} = n_I$, say, such that $n_D = 2n_I$. When source cooperation is absent, *i.e.*, $n_C = 0$, the sum capacity turns out to be n_D and it can be achieved simply by time-sharing. Now, let us consider $n_C = n_D/4$, and in particular, $n_D = 4, n_I = 2, n_C = 1$. The sources transmit

$$x_1(t) = \begin{pmatrix} v_1(t) + v_1(t-1) \\ v_2(t-1) \\ z_{1a}(t) \\ z_{1b}(t) \end{pmatrix} \text{ and } x_2(t) = \begin{pmatrix} v_2(t) + v_2(t-1) \\ v_1(t-1) \\ z_{2a}(t) \\ z_{2b}(t) \end{pmatrix}.$$

Note that this is possible because

$$y_1(t) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ v_2(t) + v_2(t-1) \end{pmatrix} \text{ and } y_2(t) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ v_1(t) + v_1(t-1) \end{pmatrix},$$

which means that before time t, source 1 knows $v_2(t-1)$ and source 2 knows $v_1(t-1)$. The transmissions are over a long block of length T with the signals at time T such that

$z_{ka}(T) = z_{kb}(T) = 0,\ k = 1, 2$. Also, we interpret $v_k(0) = 0$. Then, destination 3 receives

$$y_3(t) = \begin{pmatrix} v_1(t) + v_1(t-1) \\ v_2(t-1) \\ z_{1a}(t) + v_2(t) + v_2(t-1) \\ z_{1b}(t) + v_1(t-1) \end{pmatrix},\ t = 1, 2, \ldots, T-1, \text{ and}$$

$$y_3(T) = \begin{pmatrix} v_1(T) + v_1(T-1) \\ v_2(T-1) \\ v_2(T) + v_2(T-1) \\ v_1(T-1) \end{pmatrix}.$$

At the end of time T, destination 3 starts reading off the signals backwards starting from what it received at time T. From $y_3(T)$ it can recover $v_1(T), v_2(T), v_1(T-1), v_2(T-1)$. Making use of the latter two, *i.e.*, $v_1(T-1), v_2(T-1)$, it can recover $z_{1a}(T-1), z_{1b}(T-1), v_1(T-2), v_2(T-2)$ from $y_3(T-1)$. Then, employing its knowledge of $v_1(T-2), v_2(T-2)$, it recovers $z_{1a}(T-2), z_{1b}(T-2), v_1(T-3), v_2(T-3)$ from $y_3(T-2)$, and so on. Thus, destination 3 can recover $\{(v_1(t), z_{1a}(t), z_{1b}(t), v_2(t)) : t = 1, 2, \ldots, T\}$. By symmetry, destination 4 can also recover its messages. Thus, a rate of $R_1 = 3,\ R_2 = 3$ can be achieved (asymptotically as $T \to \infty$). Thus we obtain a sum-rate of 6 which is in fact the sum-capacity of this channel with cooperation.

The above scheme has two kinds of signals:

- Private signals: $z_{1a}, z_{1b}, z_{2a}, z_{2b}$ are recovered only by the destination to which it is intended. Note that these signals occupy the lower levels of the transmitted vector such that they do not appear at the destination where they may act as interference.

- Cooperative-public signals: v_1, v_2. These signals are read off by the other source at the end of each time t and then incorporated into the transmission by both the sources at the time $t + 1$. Thus, the transmission of these signals exploits the possibility of cooperation among the two sources. Destinations perform "backwards decoding." They recover the cooperative-public signals sent cooperatively starting from the final received vector and proceeding backwards. Hence, the initial transmission used by the sources to convey the signals to each other, being already available, does not act as interference at the destinations. In order to facilitate the recovery of these signals at the sources, they occupy the higher levels of transmitted vector.

Example 2: Let us consider the following asymmetric linear deterministic case, $n_{1,3} = 6, n_{1,4} = 3, n_{2,4} = 4, n_{2,3} = 3$, and $n_C = 1$. The capacity as given by Theorem 1 is 7. To achieve this, the sources transmit

$$x_1(t) = \begin{pmatrix} v_1(t) + v_1(t-1) \\ u_1(t) + v_2(t-1) \\ u_1(t) \\ z_{1a}(t) \\ z_{1b}(t) \\ z_{1c}(t) \end{pmatrix} \text{ and } x_2(t) = \begin{pmatrix} v_2(t) + v_2(t-1) \\ v_1(t-1) \\ 0 \\ z_2(t) \\ 0 \\ 0 \end{pmatrix}.$$

Note that this is again possible because

$$y_1(t) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ v_2(t) + v_2(t-1) \end{pmatrix} \text{ and } y_2(t) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ v_1(t) + v_1(t-1) \end{pmatrix}.$$

and hence, source 1 knows $v_2(t-1)$ and source 2 knows $v_1(t-1)$ before time t. The transmissions are again over a long block of length T with the signals at time T such that at time T all the u, v, z signals are 0. Also, we interpret $v_k(0) = 0,\ k = 1, 2$. Then, destination 3 receives

$$y_3(t) = \begin{pmatrix} v_1(t) + v_1(t-1) \\ u_1(t) + v_2(t-1) \\ u_1(t) \\ z_{1a}(t) + v_2(t) + v_2(t-1) \\ z_{1b}(t) + v_1(t-1) \\ z_{1c}(t) \end{pmatrix},\ t = 1, 2, \ldots, T-1, \text{ and } y_3(T) = \begin{pmatrix} v_1(T-1) \\ v_2(T-1) \\ 0 \\ v_2(T-1) \\ v_1(T-1) \\ 0 \end{pmatrix},$$

and destination 4 receives

$$y_4(t) = \begin{pmatrix} 0 \\ 0 \\ v_2(t) + v_2(t-1) \\ v_1(t) + 2v_1(t-1) \\ u_1(t) + v_2(t-1) \\ u_1(t) + z_2(t) \end{pmatrix},\ t = 1, 2, \ldots, T-1, \text{ and } y_3(T) = \begin{pmatrix} 0 \\ 0 \\ v_2(T-1) \\ 2v_1(T-1) \\ v_2(T-1) \\ 0 \end{pmatrix},$$

Now it is easy to verify that if the destinations read off the signals starting from the vectors they received at time T and proceeding backwards as in the previous examples, destination 3 can recover $\{(v_1(t), u_1(t), z_{1a}(t), z_{1b}(t), z_{1c}(t), v_2(t)) : t = 1, 2, \ldots, T\}$, and destination 4 $\{(v_2(t), z_2(t), v_1(t), u_1(t)) : t = 1, 2, \ldots, T\}$, and hence achieve a sum-rate of 7.

This example involved a new type of signal apart from the private and cooperative-public types of the previous example.

- Public signal: u_1. This signal is decoded by both the destinations. However, note that unlike the cooperative-public signal, the other source does not participate in its transmission. Indeed, this signal is transmitted in such a way that it is not visible to the other source.

Both the examples above involved cooperative links which are weaker than the direct and interfering links. The mode of cooperation involved the sources cooperating in aiding the destinations recover an interfering signal. When the cooperative link is strong, yet another possible form of cooperation becomes feasible. The next example illustrates this.

Example 3: Let us again consider the symmetric case, but now with $n_D = 4, n_I = 3, n_C = 5$. Note that, the cooperation link is now stronger than both the direct and interference links.

Without cooperation, the sum capacity is 5; but with cooperation, we will show that a sum rate of 6 can be achieved. The sources transmit

$$x_1(t) = \begin{pmatrix} v_1(t) + v_1(t-1) \\ v_2(t-1) \\ s_1(t) \\ z_1(t) - s_2(t) \\ s_1(t+1) \end{pmatrix} \text{ and } x_2(t) = \begin{pmatrix} v_2(t) + v_2(t-1) \\ v_1(t-1) \\ s_2(t) \\ z_2(t) - s_1(t) \\ s_2(t+1) \end{pmatrix}.$$

Note that this transmission scheme is possible since $y_1(t) = x_2(t),\ y_2(t) = x_1(t)$. This allows the sources to exchange their $s(t)$ signals one time step in advance over the lowest level. Also, we set $s_k(1) = 0, v_k(T) = 0$, and interpret $v_k(0) = 0, s_k(T+1) = 0,\ k = 1, 2$. Destinations now receive

$$y_3(t) = \begin{pmatrix} 0 \\ v_1(t) + v_1(t-1) \\ v_2(t) + 2v_2(t-1) \\ s_1(t) + v_1(t-1) \\ z_1(t) \end{pmatrix},\ t = 1, 2, \ldots, T-1, \text{ and}$$

$$y_3(T) = \begin{pmatrix} 0 \\ v_1(T-1) \\ 2v_2(T-1) \\ s_1(T) + v_1(T-1) \\ z_1(T) \end{pmatrix}.$$

Recovery of signals proceeds backwards from the last received vector as in Example 1. Assuming that the characteristic of the field $\mathbb{F}$ is not 2, destination 3 can recover the signals $\{(v_1(t), s_1(t), z_1(t), v_2(t)) : t = 1, 2, \ldots, T\}$. This gives a rate of $R_1 = 3$, and by symmetry a sum-rate of 6.

In addition to the previously encountered private and cooperative-public types of signals, another kind of signal plays an important role in this example. The strong cooperation link allows the sources to share with each other signals which are eventually only recovered by the destination it is intended for. *i.e.,*

- Cooperative-private signals: s_1, s_2. The sources learn these signals from each other one time step ahead, and in the next time step they cooperate with each other to convey these signals only to the destination it is intended for. Note that the collaboration between the sources in this example is a rudimentary form of precoding. The two sources can be thought of as two antennas of a broadcast transmitter when they cooperate to transmit these signals. Only the destination to which the signal is intended for recovers it and the precoding ensures that no interference is caused at the other destination. The main differences from cooperative-public signals are: (1) the sources convey the cooperative-private signals to each other through the lowest levels of their transmission vector in such a way that it is not visible to the destinations, whereas the cooperative-public signals are conveyed over the top most levels of the transmission vector, and (2) the source collaborate in sending the cooperative-private signals by pre-coding the signals to ensure that no interference results at the destinations, while cooperative-public signals are visible to both the destinations which end up recovering them. Thus, while the role of cooperation is to aid both the destinations in recovering the cooperative-public signals, it aims to conceal the cooperative-private from the destination it is not intended for.

In general, such uncoded schemes are not sufficient to cover all possible linear deterministic channels (indeed, even in the above example, we relied on the characteristic of the field not being 2), and more importantly the Gaussian channels. But the basic intuition can be used to build coding schemes which do. These schemes are presented in the next section.

5 Coding schemes

We first present our key coding theorem (in Theorem 4). It is a block-Markov (in the sense of [14]) coding scheme which builds on Han and Kobayashi's classical superposition coding scheme [5] for the two-user interference channel and has elements of decode-and-forward strategy [14] and backwards-decoding [15] for relay channels. The schemes are generic in the sense that they apply to any memoryless interference channel with source cooperation $p_{Y_1,Y_2,Y_3,Y_4|X_1,X_2}$. Then, we apply these schemes to the problems at hand to obtain the achievability part of Theorems 1 and 2. We would like to point out that Theorem 4(a) is identical to the one which appears in [6].

Theorem 4 *(a) Given a joint distribution* $p_W\ p_{V_1,U_1,X_1|W}\ p_{V_2,U_2,X_2|W}$*, the rate pair* (R_1, R_2) *is achievable if there are non-negative* $r_{V_1}, r_{V_2}, r_{U_1}, r_{U_2}, r_{X_1}, r_{X_2}$ *such that* $R_1 = r_{V_1} + r_{U_1} + r_{X_1}$, $R_2 = r_{V_2} + r_{U_2} + r_{X_2}$*, and*

$$r_{V_1} \leq I(V_1; Y_2|W)$$

$$\begin{aligned}
r_{X_1} &\leq I(X_1; Y_3|V_1, V_2, W, U_1, U_2),\\
r_{U_1} + r_{X_1} &\leq I(U_1, X_1; Y_3|V_1, V_2, W, U_2),\\
r_{U_2} + r_{X_1} &\leq I(U_2, X_1; Y_3|V_1, V_2, W, U_1),\\
r_{U_1} + r_{U_2} + r_{X_1} &\leq I(U_1, U_2, X_1; Y_3|V_1, V_2, W),\\
(r_{V_1} + r_{V_2}) + r_{U_1} + r_{U_2} + r_{X_1} &\leq I(W, V_1, V_2, U_1, U_2, X_1; Y_3),
\end{aligned}$$

and the corresponding inequalities with subscripts 1 and 2 exchanged, and 3 replaced with 4.
(b) Given a joint distribution $p_W\ p_{V_1,U_1|W} p_{V_2,U_2|W}\ p_{S_1|W} p_{S_2|W}\ p_{Z_1|W,V_1,U_1,S_1} p_{Z_2|W,V_2,U_2,S_2}\ p_{X_1|W,V_1,U_1,Z_1,S_1,S_2}$ $p_{X_2|W,V_2,U_2,Z_2,S_1,S_2}$*, the rate pair* (R_1, R_2) *is achievable if there are non-negative* $r_{V_k}, r_{U_k}, r_{Z_k}, r_{S_k},\ k =$

$1, 2$ *such that* $R_k = r_{V_k} + r_{U_k} + r_{Z_k} + r_{S_k}$, $k = 1, 2$ *and*

$$\begin{aligned} r_{S_1} &\le I(X_1; Y_2|W, S_1, S_2, Z_1, U_1, V_1) \\ r_{Z_1} + r_{S_1} &\le I(Z_1, X_1; Y_2|W, S_1, S_2, U_1, V_1) \\ r_{U_1} + r_{Z_1} + r_{S_1} &\le I(U_1, Z_1, X_1; Y_2|W, S_1, S_2, V_1) \\ r_{V_1} + r_{U_1} + r_{Z_1} + r_{S_1} &\le I(V_1, U_1, Z_1, X_1; Y_2|W, S_1, S_2), \end{aligned}$$

$$\begin{aligned} r_{Z_1} &\le I(Z_1; Y_3|V_1, V_2, W, U_1, U_2, S_1), \\ r_{U_1} + r_{Z_1} &\le I(U_1, Z_1; Y_3|V_1, V_2, W, U_2, S_1), \\ r_{S_1} + r_{Z_1} &\le I(S_1, Z_1; Y_3|V_1, V_2, W, U_1, U_2), \\ r_{S_1} + r_{U_1} + r_{Z_1} &\le I(S_1, U_1, Z_1; Y_3|V_1, V_2, W, U_2), \\ r_{U_2} + r_{Z_1} &\le I(U_2, Z_1; Y_3|V_1, V_2, W, U_1, S_1), \\ r_{U_2} + r_{U_1} + r_{Z_1} &\le I(U_2, U_1, Z_1; Y_3|V_1, V_2, W, S_1), \\ r_{U_2} + r_{S_1} + r_{Z_1} &\le I(U_2, S_1, Z_1; Y_3|V_1, V_2, W, U_1), \\ r_{U_2} + r_{S_1} + r_{U_1} + r_{Z_1} &\le I(U_2, S_1, U_1, Z_1; Y_3|V_1, V_2, W), \\ (r_{V_1} + r_{V_2}) + r_{U_1} + r_{U_2} + r_{S_1} + r_{Z_1} &\le I(W, V_1, V_2, U_1, U_2, S_1, Z_1; Y_3), \end{aligned}$$

and the corresponding inequalities with subscripts 1 and 2 exchanged, and 3 replaced by 4.
(c) Given a joint distribution $p_W\, p_{V_1,U_1|W} p_{V_2,U_2|W}\, p_{S_1|W}\, p_{Z_1|W,V_1,U_1,S_1} p_{Z_2|W,V_2,U_2}\, p_{X_1|W,V_1,U_1,Z_1,S_1}$ $p_{X_2|W,V_2,U_2,Z_2,S_1}$, *the rate pair* (R_1, R_2) *is achievable if there are non-negative* $r_{V_k}, r_{U_k}, r_{Z_k}$, $k = 1, 2$, *and* r_{S_1}, *such that* $R_1 = r_{V_1} + r_{U_1} + r_{Z_1} + r_{S_1}$, $R_2 = r_{V_1} + r_{U_1} + r_{Z_1}$, *and*

$$\begin{aligned} r_{S_1} &\le I(X_1; Y_2|W, S_1, Z_1, U_1, V_1) \\ r_{Z_1} + r_{S_1} &\le I(Z_1, X_1; Y_2|W, S_1, U_1, V_1) \\ r_{U_1} + r_{Z_1} + r_{S_1} &\le I(U_1, Z_1, X_1; Y_2|W, S_1, V_1) \\ r_{V_1} + r_{U_1} + r_{Z_1} + r_{S_1} &\le I(V_1, U_1, Z_1, X_1; Y_2|W, S_1), \end{aligned}$$

$$r_{V_2} \le I(V_2; Y_1|W, S_1),$$

$$\begin{aligned} r_{Z_1} &\le I(Z_1; Y_3|V_1, V_2, W, U_1, U_2, S_1), \\ r_{U_1} + r_{Z_1} &\le I(U_1, Z_1; Y_3|V_1, V_2, W, U_2, S_1), \\ r_{S_1} + r_{Z_1} &\le I(S_1, Z_1; Y_3|V_1, V_2, W, U_1, U_2), \\ r_{S_1} + r_{U_1} + r_{Z_1} &\le I(S_1, U_1, Z_1; Y_3|V_1, V_2, W, U_2), \\ r_{U_2} + r_{Z_1} &\le I(U_2, Z_1; Y_3|V_1, V_2, W, U_1, S_1), \\ r_{U_2} + r_{U_1} + r_{Z_1} &\le I(U_2, U_1, Z_1; Y_3|V_1, V_2, W, S_1), \\ r_{U_2} + r_{S_1} + r_{Z_1} &\le I(U_2, S_1, Z_1; Y_3|V_1, V_2, W, U_1), \\ r_{U_2} + r_{S_1} + r_{U_1} + r_{Z_1} &\le I(U_2, S_1, U_1, Z_1; Y_3|V_1, V_2, W), \\ (r_{V_1} + r_{V_2}) + r_{U_1} + r_{U_2} + r_{S_1} + r_{Z_1} &\le I(W, V_1, V_2, U_1, U_2, S_1, Z_1; Y_3), \end{aligned}$$

$$
\begin{aligned}
r_{Z_2} &\le I(Z_2; Y_4|V_1, V_2, W, U_1, U_2), \\
r_{U_2} + r_{Z_2} &\le I(U_2, Z_2; Y_4|V_1, V_2, W, U_1), \\
r_{U_1} + r_{Z_2} &\le I(U_1, Z_2; Y_4|V_1, V_2, W, U_2), \\
r_{U_1} + r_{U_2} + r_{Z_2} &\le I(U_1, U_2, Z_2; Y_4|V_1, V_2, W), \\
(r_{V_1} + r_{V_2}) + r_{U_1} + r_{U_2} + r_{Z_2} &\le I(W, V_1, V_2, U_1, U_2, Z_2; Y_4).
\end{aligned}
$$

We prove this theorem in Appendix A. Here, we interpret these theorems in the context of the examples and discussion in the previous section:

Scheme (a) involves the source nodes aiding their respective destination nodes decode part of the interference (the cooperative-public message from the interfering source) by essentially retransmitting the part of the interference observed by the source. In order to facilitate this, the cooperative-public message is coded separately (in V_1 and V_2) which are decoded by the sources and retransmitted (in W). There are further public messages (U_1 and U_2) which are not aided by the other source, and private messages as well. This scheme is closely related to Example 1 of the previous section.

In scheme (b), in addition to the above, the sources collaborate in sending private messages by sharing with each other in advance the part of the message on which they want to collaborate. Thus, the connection is to Example 3 of the previous section. The auxiliary random variables in scheme (b) have the following interpretation:

Aux. RV	Decoding destinations	Remarks
V_1	3,4	cooperative-public message from source 1
V_2	3,4	cooperative-public message from source 2
W	3,4	used by the sources to cooperatively send the two cooperative-public messages
U_1	3,4	public message from source 1
U_2	3,4	public message from source 2
S_1	3	carries the cooperative-private message to destination 3
S_2	4	carries the cooperative-private message to destination 4
Z_1	3	private message from source 1
Z_2	4	private message from source 2

Note that scheme (a) is not a special case of (b) as it might appear. The key difference is that while in scheme (b), the sources perform a joint decoding of all the messages sent by their counterparts (including the public and private messages which they do not aid in the transmission of), in scheme (a), only the part of the cooperative-public message meant to used for collaboration is decoded treating all the other messages as noise. At low strengths for the cooperative link, scheme (a) can perform better than scheme (b) specialized to not include a cooperative-private message. On the other hand, for strong cooperative links, adopting a joint decoding scheme at the sources can lead to better overall rates for other messages.

Scheme (c) combines these two schemes in a limited manner. Only source node 1's transmission benefits from transmission of some cooperative-public and some cooperative-private messages while source node 2's transmission benefits only from collaborative transmission of some cooperative-public message. Source node 1 adopts a decoding strategy similar to that adopted by the sources in scheme (a) whereas source node 2's decoding strategy is similar to the one in scheme (b). The three schemes (a), (b), and (c) can be easily combined into a single scheme (at the expense of considerably more involved notation) which may improve the achievable region in general, but since the focus here is on the sum-rate of the Gaussian case (within a constant gap) and linear deterministic cases which are obtained by considering the schemes separately, we do not explore this here.

Intuitively, a source employs a cooperative-private message only when the cooperative link is stronger than the direct link from this source, and the source shares this message with the other source in advance by having it "ride below" the other messages it sends. Thus, one of the differences of our scheme from the proposals of [7, 8] is that when sharing the cooperative-private message in advance, the decoding source performs a joint decoding of all the messages including those messages it does not aid in the transmission of, rather than treat these other messages as noise. Note that in our schemes (b) and (c), the sources cooperate in sending the cooperative-private messages by employing a simple form of precoding along the lines of Example 3 in the last section. Use of dirty-paper coding [16, 17] instead may lead to an improved gap.

The achievability proofs presented in Appendix B and C of our main theorems (Theorems 1 and 2) make use of these schemes.

6 Interference Channel with Feedback

A closely related problem is that of the interference channel with feedback. Let us consider the symmetric Gaussian interference channel with noiseless feedback. As in the model we considered earlier,

$$\begin{aligned} Y_3(t) &= h_{1,3}X_1(t) + h_{2,3}X_2(t) + N_3(t), \\ Y_4(t) &= h_{2,4}X_2(t) + h_{1,4}X_1(t) + N_4(t). \end{aligned}$$

However, instead of receiving signals through the cooperation link, the sources now receive noiseless feedback from their respective destinations. *i.e.*,

$$\begin{aligned} Y_1(t) &= Y_3(t), \\ Y_2(t) &= Y_4(t). \end{aligned}$$

As before, the transmissions from the sources are deterministic functions of their messages and their observations (here, the feedback received) in the past. Since the sources have access to the symbols they transmitted in the past, it is clear that the above problem is equivalent to one where the sources observe

$$\begin{aligned} Y_1(t) &= h_{2,3}X_2(t) + N_3(t), \\ Y_2(t) &= h_{1,4}X_1(t) + N_4(t). \end{aligned}$$

We can rewrite these as

$$Y_1(t) = h_{2,1}X_2(t) + N_1(t),$$
$$Y_2(t) = h_{1,2}X_1(t) + N_2(t).$$

Here $h_{2,1} = h_{2,3}$, $h_{1,2} = h_{1,4}$, $N_1(t) = N_3(t)$, and $N_2(t) = N_4(t)$. This is identical to the channel model stated at the beginning of section 2 except for the fact that now N_1, N_2, N_3, N_4 are not independent, but N_1 and N_3 are identical, and so are N_2 and N_4.

The above difference notwithstanding, we will argue below that the results presented in Section 3.1 on the sum-capacity of the interference channel with source cooperation have a bearing on this model as well. Note, however, that our source cooperation result was proved under the restriction that $|h_{1,2}| = |h_{2,1}|$. In general, this may not hold true for interference channels with feedback, but a range of channels including, most importantly, the symmetric interference channel is covered. We will focus our attention on the symmetric channel, where $|h_{1,3}| = |h_{2,4}| = h_D$, and $|h_{1,4}| = |h_{2,3}| = h_I$.

Let us begin by noting that a linear deterministic formulation of the above feedback problem is identical to the one with source cooperation, and hence, Theorem 1 applies directly. Turning to the Gaussian case, let us note that the achievability proof for the source cooperation case depended only on the marginal distributions of the noises and not on their correlation. Hence the achievability proof also holds directly. We only need to argue that the converse also applies. In appendix E, we will show that the biting upperbound for the symmetric channel with noiseless feedback indeed holds to give us the following proposition:

Proposition 5 *The sum-capacity of the symmetric, Gaussian interference channel with output feedback is within a constant (19 bits) of the minimum of the following terms:*

$$\log 2\left(1 + (|h_D| + |h_I|)^2\right)\left(1 + \frac{\max\left(|h_D|^2, |h_I|^2\right)}{\max\left(1, |h_I|^2\right)}\right). \tag{12}$$

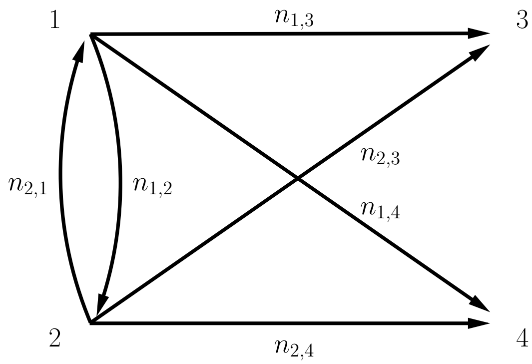
This case was studied independently in [10] which also characterizes the sum-capacity and obtains a better constant than we do here.

7 Discussion

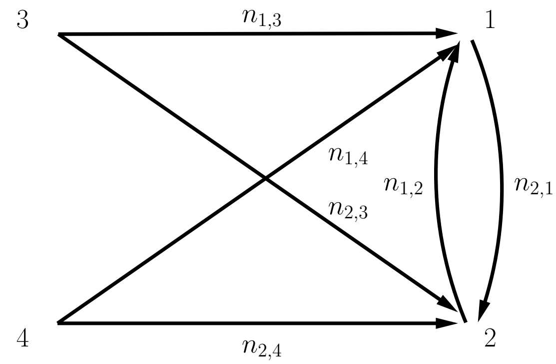
7.1 Reversibility

A related setting to the one studied in this paper is the interference channel with destination cooperation. This case will be presented in a companion paper [18]. An interesting *reversibility* property connects the two settings. We briefly discuss it here. For brevity, the linear deterministic model will be employed.

In the destination cooperation case, the destinations can not only receive, but they can also transmit. But these transmissions have to satisfy a causality constraint – the transmissions from each destination is a function of everything it has received up to the previous time



(a) Source Cooperation



(b) Destination Cooperation

Figure 3: Source Cooperation and Destination Cooperation. Note that the destination cooperation channel (b) is obtained from the source cooperation channel (a) by: (i) reversing the roles of the sources and destinations, *i.e.*, nodes 1 and 2 are sources in (a) and destinations in (b); nodes 3 and 4 are destinations in (a) and sources in (b), and (ii) reversing the directions of all the links while maintaining the link strengths to be the same. Both channels have the same sum-capacity.

instant. In order to illustrate the reversibility between destination and source cooperation scenarios, we will number the nodes as shown in the Fig. 3(b): 3 and 4 are the source nodes now which want to communicate to destination nodes 1 and 2 respectively, and there is a cooperation link between the destination nodes. The destination nodes receive

$$
\begin{aligned}
Y_1(t) &= h_{1,3}(X_3(t)) + h^*_{1,4}(X_4(t)) + h_{1,2}(X_2(t)), \\
Y_2(t) &= h_{2,4}(X_4(t)) + h^*_{2,3}(X_3(t)) + h_{2,1}(X_1(t)).
\end{aligned}
$$

where the (deterministic) encoding functions at the sources are of the form

$$X_k(t) = f_{k,t}(M_k),\ k = 3, 4,$$

and the (deterministic) relaying functions at the destinations are of the form

$$X_k(t) = f_{k,t}(Y_k^{t-1}),\ k = 2, 1.$$

Let $n_{1,3}, n_{2,3}, n_{1,4}, n_{2,4}, n_{2,1}, n_{1,2}$ be non-negative integers and $n \stackrel{\text{def}}{=} \max(n_{1,3}, n_{2,3}, n_{1,4}, n_{2,4}, n_{2,1}, n_{1,2})$. The inputs to the channel X_3 and X_4 are n-length vectors over a finite field $\mathbb{F}$. We define

$$
\begin{aligned}
h_{1,3}(X_3) &= \mathbf{S}^{n-n_{1,3}} X_3, \\
h_{2,4}(X_4) &= \mathbf{S}^{n-n_{2,4}} X_4, \\
h^*_{1,4}(X_4) &= \mathbf{S}^{n-n_{1,4}} X_4, \\
h^*_{2,3}(X_3) &= \mathbf{S}^{n-n_{2,3}} X_3, \\
h_{1,2}(X_2) &= \mathbf{S}^{n-n_{1,2}} X_2, \\
h_{2,1}(X_1) &= \mathbf{S}^{n-n_{2,1}} X_1.
\end{aligned}
$$

Further, to model the reciprocity of the links between the two receivers, we set $n_{2,1} = n_{1,2} = n_C$.

The reversibility property is that the sum-capacity expression in Theorem 1 is also the sum-capacity of the above channel with destination cooperation. This turns out to be a feature of our achievable strategy which holds in more general cases (larger networks, more number of sources-destination pairs etc.) as discussed further in [19]. It would be interesting to investigate whether the optimality results presented here also extend to more general settings.

7.2 Dependence on channel state information

Throughout the paper we assumed that full channel state information is available at both the sources and the destinations. However, this can be relaxed. The application of the scheme in Theorem 4(a) only requires the sources to know the channel strengths to the two destinations, and not their phases. Note that Theorem 4(a) caters to the case where the strength of the cooperative link is weaker than that of all the links to the two destinations. But, to apply the schemes in Theorem 4(b) and (c), we do require the sources to have full channel state information. This is not surprising since the analogous setting of a multiantenna broadcast channel also requires full channel state information.

A Proof of Theorem 4

(a) We present a block-Markov scheme with backwards decoding. Given $p_W p_{V_1,U_1,X_1|W} p_{V_2,U_2,X_2|W}$, we construct the following blocklength-T codebooks:

- W codebook: We create a W-codebook $\mathcal{C}_W$ of size 2^{Tr_W} with codewords of length n by choosing the elements independently according to the distribution p_W. We will denote the codewords by $c_W(m_W)$ where $m_W \in \{1, \ldots, 2^{T(r_W - \epsilon)}\}$, where $\epsilon > 0$.

- V codebooks: For each codeword $c_W(m_W) \in \mathcal{C}_W$, and for each $k = 1, 2$, we create V_k-codebook $\mathcal{C}_{V_k}(m_W)$ of size $2^{T(r_{V_k} - \epsilon)}$ respectively, by choosing elements independently according to $p_{V_k|W}(.|w)$ where w is set to the respective element of the $c_W(m_W)$ codeword. These codewords will be denoted by $c_{V_k}(m_{V_k}, m_W)$ where $m_{V_k} \in \{1, \ldots, 2^{T(r_{V_k} - \epsilon)}\}$. Moreover, we set
$$r_W = r_{V_1} + r_{V_2}.$$

- U codebooks: For each codeword $c_{V_k}(m_{V_k}, m_W)$, we create a U_k-codebook $\mathcal{C}_{U_k}(m_{V_k}, m_W)$ of size $2^{T(r_{U_k} - \epsilon)}$ by choosing elements according to $p_{U_k|V_k,W}(.|v_k, w)$ by setting v_k and w to be the respective elements of the $c_{V_k}(m_{V_k}, m_W)$ and $c_W(m_W)$ codewords.

- X codebooks: Finally, for each codeword $c_{U_k}(m_{U_k}, m_{V_k}, m_W)$, we similarly create a X_k-codebook $\mathcal{C}_{X_k}(m_{U_k}, m_{V_k}, m_W)$ of size $2^{T(r_{X_k} - \epsilon)}$ using $p_{X_k|U_k,V_k,W}$.

Encoding: For block-j, we will assume for the moment that the source nodes have successfully decoded the messages $m_{V_k}(j-1)$. Then the encoding proceeds as follows. Both encoders set $m_W(j) = (m_{V_1}(j-1), m_{V_2}(j-1))$. They then proceed to choose the codewords $c_W(m_W(j))$, $c_{V_k}(m_{V_k}(j), m_W(j))$, $c_{U_k}(m_{U_k}(j), m_{V_k}(j), m_W(j))$, and $c_{X_k}(m_{X_k}(j), m_{U_k}(j), m_{V_k}(j), m_W(j))$. The X-codewords are transmitted. For the first block, $j=1$, we set $m_W(1) = 1$ and for the last block J, we set $m_{V_1}(J) = m_{V_2}(J) = 1$.
Decoding at the sources: At the end of block-j, source 1 will try to decode the message $m_{V_2}(j)$ from source 2 and *vice versa*. Using standard arguments, we can show that for joint-typical decoding to succeed (with probability approaching 1 as the blocklength n approaches ∞), it is enough to ensure that

$$\begin{aligned} r_{V_1} &\leq I(V_1; Y_2|W), \text{ and} \\ r_{V_2} &\leq I(V_2; Y_1|W). \end{aligned}$$

When this decoding fails, we will say that "encoding at the sources has failed at block-j," and declare an error.
Decoding at the destinations: Destinations perform *backwards decoding*. We will assume that before destination 3 processes block-j, it has already successfully decoded $m_W(j+1) = (m_{V_1}(j), m_{V_2}(j))$. This is true for $j = J$ since $m_{V_1}(J) = m_{V_2}(J) = 1$. And, for each j, we will ensure that from block-j, destination 3 decodes $m_W(j)$ successfully thereby ensuring that the above assumption holds true. Assuming that $m_{V_1}(j), m_{V_2}(j)$ is available at destination 3, we will try to ensure that from block-j, the messages $m_W(j)$, $m_{U_1}(j)$, $m_{X_1}(j)$ are successfully decoded. In trying to decode these messages, destination 3 will also try to jointly decode the message $m_{U_2}(j)$. It can be shown that a high probability of decoding success can be ensured (*i.e.*, the probability of failure in decoding the messages $m_W(j)$, $m_{U_1}(j)$, and $m_{X_1}(j)$ from what destination 3 receives in block-j assuming $m_W(j+1)$ is available, goes to 0 as blocklength n goes to ∞) if the following conditions are met.

$$\begin{aligned} r_{X_1} &\leq I(X_1; Y_3|V_1, V_2, W, U_1, U_2), \\ r_{U_1} + r_{X_1} &\leq I(U_1, X_1; Y_3|V_1, V_2, W, U_2), \\ r_{U_2} + r_{X_1} &\leq I(U_2, X_1; Y_3|V_1, V_2, W, U_1), \\ r_{U_1} + r_{U_2} + r_{X_1} &\leq I(U_1, U_2, X_1; Y_3|V_1, V_2, W), \text{ and} \\ (r_{V_1} + r_{V_2}) + r_{U_1} + r_{U_2} + r_{X_1} &\leq I(W, V_1, V_2, U_1, U_2, X_1; Y_3). \end{aligned}$$

A similar set of conditions ensure success of decoding at destination 4. If decoding fails for block-j for either of the destinations, we will say that "decoding failed at block-j" and declare an error.

Overall, an error results if for at least one block-j, either encoding fails or decoding fails. Since there are a finite number J of blocks, by union bound, the above discussion implies that the probability of error goes to 0 as the blocklength goes to ∞ when the above conditions are met. This completes the random coding argument.
(b) We present a block-Markov scheme with backwards decoding at the destinations. Given $p_W\, p_{V_1,U_1|W}\, p_{V_2,U_2|W}\, p_{S_1|W} p_{S_2|W}\, p_{Z_1|W,V_1,U_1,S_1}\, p_{Z_2|W,V_2,U_2,S_2}\, p_{X_1|W,V_1,U_1,Z_1,S_1,S_2}\, p_{X_2|W,V_2,U_2,Z_2,S_1,S_2}$, we construct the following blocklength-n codebooks:

- W, V, and U codebooks: These codebooks are constructed as in scheme (a) above. We create a W-codebook $\mathcal{C}_W$ of size 2^{Tr_W} with codewords of length n by choosing the elements independently according to the distribution p_W. We will denote the codewords by $c_W(m_W)$, where $m_W \in \{1, \ldots, 2^{T(r_W-\epsilon)}\}$, where $\epsilon > 0$.

  For each codeword $c_W(m_W) \in \mathcal{C}_W$, and for each $k = 1, 2$, we create V_k-codebook $\mathcal{C}_{V_k}(m_W)$ of size $2^{T(r_{V_k}-\epsilon)}$ respectively, by choosing elements independently according to $p_{V_k|W}(.|w)$ where w is set to the respective element of the $c_W(m_W)$ codeword. These codewords will be denoted by $c_{V_k}(m_{V_k}, m_W)$ where $m_{V_k} \in \{1, \ldots, 2^{T(r_{V_k}-2\epsilon)}\}$. Moreover, we set

  $$r_W = r_{V_1} + r_{V_2}.$$

  For each codeword $c_{V_k}(m_{V_k}, m_W)$, we create a U_k-codebook $\mathcal{C}_{U_k}(m_{V_k}, m_W)$ of size $2^{T(r_{U_k}-\epsilon)}$ by choosing elements according to $p_{U_k|V_k,W}(.|v_k, w)$ by setting v_k and w to be the respective elements of the $c_{V_k}(m_{V_k}, m_W)$ and $c_W(m_W)$ codewords. These codewords will be denoted by $c_{U_k}(m_{U_k}, i_{V_k}, m_W)$ where $m_{U_k} \in \{1, \ldots, 2^{T(r_{U_k}-\epsilon)}\}$.

- S codebooks: For each $c_W(m_W) \in \mathcal{C}_W$, and for each $k = 1, 2$, we create an S_k-codebook $\mathcal{C}_{S_k}(m_W)$ of size $2^{T(r_{S_k}-\epsilon)}$ respectively, by choosing elements independently according to $p_{S_k|W}(.|w)$ where w is set to the respective element of the $c_W(m_W)$ codeword. These codewords will be denoted by $c_{S_k}(m_{S_k}, m_W)$ where $m_{S_k} \in \{1, \ldots, 2^{T(r_{S_k}-2\epsilon)}\}$.

- Z codebooks: For each pair of codewords $(c_{U_k}(m_{U_k}, m_{V_k}, m_W), c_{S_k}(m_{S_k}, m_W)) \in \mathcal{C}_{U_k}(m_{V_k}, m_W) \times \mathcal{C}_{S_k}(m_W)$, and for each $k = 1, 2$, we create a Z_k-codebook $\mathcal{C}_{Z_k}(m_{U_k}, m_{V_k}, m_W, i_{S_k})$ of size $2^{T(r_{Z_k}-\epsilon)}$ respectively, by choosing elements independently according to $p_{Z_k|W,V_k,U_k,S_k}(.|w, v_k, u_k, s_k)$ where w, v_k, u_k and s_k are set to the respective elements of the $c_W(m_W)$, $c_{V_k}(m_{V_k}, m_W)$, $c_{U_k}(m_{U_k}, m_{V_k}, m_W)$, and $c_{S_k}(m_{S_k}, m_W)$ codewords, respectively. The codewords so generated will be denoted by $c_{Z_k}(m_{Z_k}, m_{U_k}, m_{V_k}, m_W, m_{S_k})$ where $i_{Z_k} \in \{1, \ldots, 2^{T(r_{Z_k}-\epsilon)}\}$.

- X codebooks: Finally, consider pairs of codewords $(c_{Z_k}(m_{Z_k}, m_{U_k}, m_{V_k}, m_W, m_{S_k}), c_{S_{\bar{k}}}(m_{S_{\bar{k}}}, m_W)) \in \mathcal{C}_{Z_k}(m_{U_k}, m_{V_k}, m_W, m_{S_k}) \times \mathcal{C}_{S_{\bar{k}}}(m_W)$, for each $k = 1, 2$, where $\bar{k} = 2$, if $k = 1$, and $\bar{k} = 1$, if $k = 2$. For each pair, we create a X_k-codebook $\mathcal{C}_{X_k}(m_{Z_k}, m_{U_k}, m_{V_k}, m_W, i_{S_k}, i_{S_{\bar{k}}})$ of size $2^{T(r_{S_k}-\epsilon)}$ by choosing elements independently according to $p_{X_k|W,V_k,U_k,Z_k,S_k,S_{\bar{k}}}(.|w, v_k, u_k, s_k, s_{\bar{k}})$ where w, v_k, u_k, z_k, s_k, and $s_{\bar{k}}$ are set to the respective elements of the $c_W(m_W)$, $c_{V_k}(i_{V_k}, m_W)$, $c_{U_k}(i_{U_k}, i_{V_k}, m_W)$, $c_{Z_k}(i_{Z_k}, i_{U_k}, i_{V_k}, m_W)$, $c_{S_k}(i_{S_k}, m_W)$, and $c_{S_{\bar{k}}}(i_{S_{\bar{k}}}, m_W)$ codewords, respectively. The codewords so generated will be denoted by $c_X(m_{\text{PVT-COOP}_k}, m_{Z_k}, m_{U_k}, m_{V_k}, m_W, i_{S_{\bar{k}}}, i_{S_k})$ where $m_{\text{PVT-COOP}_k} \in \{1, \ldots, 2^{T(r'_{S_k}-\epsilon)}\}$.

Encoding: For block-j, we will assume for the moment that the source nodes have successfully decoded the messages $m_{V_k}(j-1), m_{\text{PVT-COOP}_k}(j-1)$. Then the encoding proceeds as follows. Both encoders set $m_W(j) = (m_{V_1}(j-1), m_{V_2}(j-1))$. Then, encoder-$k$ proceeds to select the codewords $c_W(m_W(j))$, $c_{V_k}(m_{V_k}(j), m_W(j))$, $c_{U_k}(m_{U_k}(j), m_{V_k}(j), m_W(j))$,

$c_{S_k}(m_{S_k}(j), m_W(j))$, and $c_{Z_k}(m_{Z_k}(j), m_{U_k}(j), m_{V_k}(j), m_W(j), i_{S_k}(j))$. It transmits the X-codeword
$c_X(m_{\text{PVT-COOP}_k}(j), m_{Z_k}(j), m_{U_k}(j), m_{V_k}(j), m_W(j), m_{S_{\bar{k}}}(j), m_{S_k}(j))$. For the first block, $j = 1$, we set $m_W(1) = m_{S_1}(1) = m_{S_2}(1) = 1$ and for the last block J, we set $m_{V_1}(J) = m_{V_2}(J) = m_{\text{PVT-COOP}_1}(J) = m_{\text{PVT-COOP}_2}(J) = 1$.
Decoding at the sources: At the end of block-j, source 2 will try to jointly decode

$$m_{V_1}(j), m_{U_1}(j), m_{Z_1}(j), m_{\text{PVT-COOP}_1}(j)$$

from source 1 and *vice versa*. Note that both sources have access to the W, S_1, and S_2-codewords. For joint-typical decoding to succeed (with probability approaching 1 as the blocklength T approaches ∞), we can show that it is enough to ensure that

$$\begin{aligned}
r_{S_1} &\leq I(X_1; Y_2 | W, S_1, S_2, Z_1, U_1, V_1) \\
r_{Z_1} + r_{S_1} &\leq I(Z_1, X_1; Y_2 | W, S_1, S_2, U_1, V_1) \\
r_{U_1} + r_{Z_1} + r_{S_1} &\leq I(U_1, Z_1, X_1; Y_2 | W, S_1, S_2, V_1) \\
r_{V_1} + r_{U_1} + r_{Z_1} + r_{S_1} &\leq I(V_1, U_1, Z_1, X_1; Y_2 | W, S_1, S_2).
\end{aligned}$$

When this decoding fails, we will say that "encoding at the sources has failed at block-j," and declare an error.
Decoding at the destinations: Destinations perform backwards decoding. We will assume that before destination 3 processes block-j, it has already successfully decoded $m_W(j+1) = (m_{V_1}(j), m_{V_2}(j))$. This is true for $j = J$ since $m_{V_1}(J) = m_{V_2}(J) = 1$. And, for each j, we will ensure that from block-j, destination 3 decodes $m_W(j)$ successfully thereby ensuring that the above assumption holds true. Assuming that $m_{V_1}(j), m_{V_2}(j)$ is available at destination 3, we will try to ensure that from block-j, the messages $m_W(j)$, $m_{U_1}(j)$, $m_{Z_1}(j)$, and $m_{S_1}(j)$ are successfully decoded. In trying to decode these messages, destination 3 will also try to jointly decode the message $m_{U_2}(j)$. The decoding is performed by looking for a unique tuple of $W, V_1, V_2, U_2, U_1, Z_1, S_1$ codewords consistent with the information already known (namely, $m_{V_1}(j), m_{V_2}(j)$) and which are jointly (strongly) typical with the (T-length) block of signal Y_3 received by destination 3 corresponding to the block-j. Using standard arguments, a high probability of decoding success can be ensured (*i.e.*, the probability of failure in decoding the messages $m_W(j)$, $m_{U_1}(j)$, and $m_{Z_1}(j)$ and $i_{S_1}(j)$ from what destination 3 receives in block-j assuming $m_W(j+1)$ is available, goes to 0 as blocklength T goes to ∞) if the following conditions are met.

$$\begin{aligned}
r_{Z_1} &\leq I(Z_1; Y_3 | V_1, V_2, W, U_1, U_2, S_1), \\
r_{U_1} + r_{Z_1} &\leq I(U_1, Z_1; Y_3 | V_1, V_2, W, U_2, S_1), \\
r_{S_1} + r_{Z_1} &\leq I(S_1, Z_1; Y_3 | V_1, V_2, W, U_1, U_2), \\
r_{S_1} + r_{U_1} + r_{Z_1} &\leq I(S_1, U_1, Z_1; Y_3 | V_1, V_2, W, U_2), \\
r_{U_2} + r_{Z_1} &\leq I(U_2, Z_1; Y_3 | V_1, V_2, W, U_1, S_1), \\
r_{U_2} + r_{U_1} + r_{Z_1} &\leq I(U_2, U_1, Z_1; Y_3 | V_1, V_2, W, S_1), \\
r_{U_2} + r_{S_1} + r_{Z_1} &\leq I(U_2, S_1, Z_1; Y_3 | V_1, V_2, W, U_1),
\end{aligned}$$

$$
\begin{aligned}
r_{U_2} + r_{S_1} + r_{U_1} + r_{Z_1} &\leq I(U_2, S_1, U_1, Z_1; Y_3 | V_1, V_2, W), \text{ and} \\
(r_{V_1} + r_{V_2}) + r_{U_1} + r_{U_2} + r_{S_1} + r_{Z_1} &\leq I(W, V_1, V_2, U_1, U_2, S_1, Z_1; Y_3).
\end{aligned}
$$

A similar set of conditions ensure success of decoding at destination 4. If decoding fails for block-j for either of the destinations, we will say that "decoding failed at block-j" and declare an error.

For the purposes of illustration, let us see how one of the above conditions is arrived at. One of the possible error-events under which decoding at destination 3 fails is when, for block-j, only the following decoding errors occur: $\widehat{m}_{Z_1} \neq m_{Z_1}$, $\widehat{m}_{U_2} \neq m_{U_2}$, and $\widehat{m}_{S_1} \neq m_{S_1}$, where the $\widehat{\ }$'s indicate the decoded values. The probability of this (under random coding as described above) is

$$
\sum_{\widehat{m}_{Z_1}, \widehat{m}_{U_2}, \widehat{m}_{S_1}} \mathbb{P}\Bigg((c_W, c_{V_1}, c_{V_2}, c_{U_1}, \widehat{c}_{Z_1}, \widehat{c}_{S_1}, \widehat{c}_{U_2}, Y_3^T) \in \mathcal{T}^\delta \Bigg| m_W, m_{V_1}, m_{V_2}, m_{U_1}, m_{U_2}, m_{Z_1}, m_{Z_2}, m_{S_1}, i_{S_2}, m_{\text{PVT-COOP}_1}, m_{\text{PVT-COOP}_2} \Bigg),
$$

where we suppressed the indices for the codewords, and the time-index j for the codeword indices of the conditioning event. The unhatted codewords have indices from the messages of the conditioning event, while the hatted codewords are short-hand notations for the codewords with the corresponding indices replaced by their hatted forms:

$$
\begin{aligned}
\widehat{c}_{Z_1} &= c_{Z_1}(\widehat{m}_{Z_1}, m_{U_1}, m_{V_1}, m_W, \widehat{m}_{S_1}), \\
\widehat{c}_{S_1} &= c_{S_1}(\widehat{m}_{S_1}, m_W), \text{ and} \\
\widehat{c}_{U_2} &= c_{U_2}(\widehat{m}_{U_2}, m_{V_2}, m_W).
\end{aligned}
$$

We also suppressed the subscript for the δ-typical set $\mathcal{T}_{W,V_1,V_2,U_1,U_2,Z_1,S_1,Y_3}$. We will continue to do that in the sequel; the appropriate subscripts will be clear from the context. Below, we will also suppress the conditioning event. Then,

$$
\begin{aligned}
&\mathbb{P}\Bigg((c_W, c_{V_1}, c_{V_2}, c_{U_1}, \widehat{c}_{Z_1}, \widehat{c}_{S_1}, \widehat{c}_{U_2}, Y_3^T) \in \mathcal{T}^\delta \Bigg) \\
&\leq \sum_{(\widetilde{c}_W, \widetilde{c}_{V_1}, \widetilde{c}_{V_2}, \widetilde{c}_{U_1}) \in \mathcal{T}^\delta} \mathbb{P}\left((c_W, c_{V_1}, c_{V_2}, c_{U_1}) = (\widetilde{c}_W, \widetilde{c}_{V_1}, \widetilde{c}_{V_2}, \widetilde{c}_{U_1}) \right) \\
&\qquad \cdot \mathbb{P}\Bigg((c_W, c_{V_1}, c_{V_2}, c_{U_1}, \widehat{c}_{Z_1}, \widehat{c}_{S_1}, \widehat{c}_{U_2}, Y_3^T) \in \mathcal{T}^\delta \Bigg| (c_W, c_{V_1}, c_{V_2}, c_{U_1}) = (\widetilde{c}_W, \widetilde{c}_{V_1}, \widetilde{c}_{V_2}, \widetilde{c}_{U_1}). \Bigg)
\end{aligned}
$$

Further,

$$
\begin{aligned}
&\mathbb{P}\left((c_W, c_{V_1}, c_{V_2}, c_{U_1}, \widehat{c}_{Z_1}, \widehat{c}_{S_1}, \widehat{c}_{U_2}, Y_3^T) \in \mathcal{T}^\delta \middle| (c_W, c_{V_1}, c_{V_2}, c_{U_1}) = (\widetilde{c}_W, \widetilde{c}_{V_1}, \widetilde{c}_{V_2}, \widetilde{c}_{U_1}) \right) \\
&= \sum_{(\widetilde{c}_{Z_1}, \widetilde{c}_{S_1}, \widetilde{x}_{U_2}, \widetilde{Y}_3^T) \in \mathcal{T}^\delta} \mathbb{P}\left((\widehat{c}_{Z_1}, \widehat{c}_{S_1}, \widehat{c}_{U_2}, Y_3^T) = (\widetilde{c}_{Z_1}, \widetilde{c}_{S_1}, \widetilde{c}_{U_2}, \widetilde{Y}_3^T) \middle| (c_W, c_{V_1}, c_{V_2}, c_{U_1}) = (\widetilde{c}_W, \widetilde{c}_{V_1}, \widetilde{c}_{V_2}, \widetilde{c}_{U_1}) \right),
\end{aligned}
$$

where $\mathcal{T}^\delta$ in the summation index is the set of conditionally δ-typical Z_1, S_1, U_2, Y_3 sequences conditioned on the (W, V_1, V_2, U_1)-typical sequence $(\widetilde{c}_W, \widetilde{c}_{V_1}, \widetilde{c}_{V_2}, \widetilde{c}_{U_1})$. Note that the cardinality of this set is upperbounded by $2^{T(H(Z_1,S_1,U_2,Y_3|W,V_1,V_2,U_1)+\delta)}$.

$$\begin{aligned}
&\mathbb{P}\left((\widehat{c}_{Z_1}, \widehat{c}_{S_1}, \widehat{c}_{U_2}, Y_3^T) = (\widetilde{c}_{Z_1}, \widetilde{c}_{S_1}, \widetilde{c}_{U_2}, \widetilde{Y}_3^T)\middle|(c_W, c_{V_1}, c_{V_2}, c_{U_1}) = (\widetilde{c}_W, \widetilde{c}_{V_1}, \widetilde{c}_{V_2}, \widetilde{c}_{U_1})\right)\\
&= \mathbb{P}\left((\widehat{c}_{Z_1}, \widehat{c}_{S_1}, \widehat{c}_{U_2}) = (\widetilde{c}_{Z_1}, \widetilde{c}_{S_1}, \widetilde{c}_{U_2})|(c_W, c_{V_1}, c_{V_2}, c_{U_1}) = (\widetilde{c}_W, \widetilde{c}_{V_1}, \widetilde{c}_{V_2}, \widetilde{c}_{U_1})\right) \cdot\\
&\qquad\qquad \mathbb{P}\left(Y_3^T = \widetilde{Y}_3^T\middle|(c_W, c_{V_1}, c_{V_2}, c_{U_1}) = (\widetilde{c}_W, \widetilde{c}_{V_1}, \widetilde{c}_{V_2}, \widetilde{c}_{U_1})\right)\\
&\leq 2^{-T(H(Z_1,S_1,U_2|W,V_1,V_2,U_1)-\delta)} 2^{-T(H(Y_3|W,V_1,V_2,U_1)-\delta)},
\end{aligned}$$

where the first step follows from the independence of the hatted-codewords and what destination 3 receives conditioned on the unhatted-codewords. Combining everything, we can conclude that the probability of the error-event under consideration is less than or equal to

$$2^{T(r_{U_2}+r_{S_1}+r_{Z_1}-I(U_2,S_1,Z_1;Y_3|W,V_1,V_2,U_1)-3\epsilon+3\delta)},$$

which goes to zero as the blocklength T goes to ∞ if we choose $0 < \delta < \epsilon$, and the rates satisfy the condition

$$r_{U_2} + r_{S_1} + r_{Z_1} \leq I(U_2, S_1, Z_1; Y_3|W, V_1, V_2, U_1).$$

Overall, an error results if for at least one block-j, either encoding fails or decoding fails. Since there are a finite number J of blocks, by union bound, the above discussion implies that the probability of error goes to 0 as the blocklength goes to ∞ when the above conditions are met. This completes the random coding argument.

(c) This scheme is a combination of the two schemes above. Now, only destination 3 receives a private message sent cooperatively by the two sources. Hence, only an S_1 codebook is present and there is no S_2 codebook. The W, V and U codebooks are exactly as in (a) and (b). The S_1 codebook is identical to scheme (b). The Z and X codebooks are also similar and differ only in that the distribution used to construct them has no S_2, and there is no $m_{\text{PVT-COOP}_2}$. Hence, we may set $Z_2 = X_2$, and set X_2 codeword to be identical to the Z_2 codeword.

The encoding at node 1 proceeds exactly as in scheme (b) above except that, since there is no S_2 codeword, node 1 need only decode the V_2-codeword. Unlike in scheme (b), node 1 treats all the other codewords from node 2 as noise when decoding the V_2 codeword. Thus, the only condition imposed by decoding at node 1 is

$$r_{V_2} \leq I(V_2; Y_1|W, S_1).$$

Encoding at node 2 is exactly as in scheme (b) except for the fact that there is no S_2 codeword. Decoding at the destination 3 is identical to that in scheme (b) while that at destination 4 is identical to that in scheme (a).

B Proof of achievability of Theorem 1

If we fix $n_{1,3}$, $n_{1,4}$, $n_{2,3}$, and $n_{2,4}$, and consider the u_i's in (1)-(4) as functions of n_C, the sum-rate expression in Theorem 1 (as a function of n_C) breaks up into four natural regimes. We use different strategies to achieve the sum-capacity in different regimes. The regimes are:

(i) $n_C \leq n_{\min} \stackrel{\text{def}}{=} \min(n_{1,3}, n_{1,4}, n_{2,3}, n_{2,4})$. It can be shown that for $n_C \geq n_{\min}$,

$$u_1(n_C) \geq \min(u_2(n_C), u_3(n_C), u_4(n_C), u_5).$$

Hence, we need consider u_1 only in the regime $n_C \leq n_{\min}$. Moreover, in this regime, $u_2(n_C)$ through $u_4(n_C)$ are constants (*i.e.*, they do not depend on n_C and their values are the same as when $n_C = 0$). Since $u_1(n_C)$ is monotonically increasing in n_C, this means that we need to employ cooperation only when $u_1(0) < \min(u_2(0), u_3(0), u_4(0), u_5)$, *i.e.*, when

$$\begin{aligned}\max(n_{1,3} - n_{1,4}, n_{2,3}) &+ \max(n_{2,4} - n_{1,4}, n_{1,4}) \\ &< \min(\max(n_{1,3}, n_{2,3}) + (\max(n_{2,4}, n_{2,3}) - n_{2,3})\,, \\ &\qquad \max(n_{2,4}, n_{1,4}) + (\max(n_{1,3}, n_{1,4}) - n_{1,4})\,, n_{1,3} + n_{2,4}). \qquad (13)\end{aligned}$$

When the above condition is not true, the sum-rate expression reduces to the sum-capacity without cooperation. We show below how Theorem 4(a) implies that the sum-rate expression is achievable in this region, both when cooperation is required and not.

(ii) $n_{\min} < n_C \leq \min(n_{1,3}, n_{2,4})$. In this regime, we can observe that the sum-rate expression takes on a constant value since $u_2(n_C)$, $u_3(n_C)$, and $u_4(n_C)$ are still constants. Hence, the achievability here is implied by the achievability in regime (i).

(iii) $\min(n_{1,3}, n_{2,4}) < n_C \leq \max(n_{1,3}, n_{2,4})$. In this regime, we use Theorem 4(c).

(iv) $\max(n_{1,3}, n_{2,4}) < n_C$. The sum-capacity is achieved in this regime by applying Theorem 4(b).

For integer q satisfying $1 \leq q \leq n$, we define

$$\mathcal{F}_q \stackrel{\text{def}}{=} \{x \in \mathbb{F}^n : x_i = 0,\ i \leq q\}\,,$$

i.e., all vectors in $\mathbb{F}^q$ such that their components in the range $1, \ldots, q$ are zeros. We take the indexing of the elements of vectors to start from the top as usual. For example, for binary field and $n = 4$,

$$\mathcal{F}_2 = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

Regime (i): When the condition (13) holds, we consider a restricted regime of n_C where

$$u_1(n_C) \leq \min(u_2(n_C), u_3(n_C), u_4(n_C), u_5).$$

Since $u_1(n_C)$ is monotonic in n_C, it is enough to prove achievability in this regime to obtain achievability in regime (i) when (13) holds. We make the following choices for the auxiliary random variables in Theorem 4(a): $W, V_1, U_1, Z_1, V_2, U_2, Z_2$ are independent of each other and uniformly distributed over their alphabets which are as follows – $V_1, V_2 \in \mathbb{F}^n$, $U_1, U_2 \in \mathcal{F}_{n_C}$, $Z_1 \in \mathcal{F}_{n_{1,4}}$, and $Z_2 \in \mathcal{F}_{n_{2,3}}$. W is independent of all these and has the same cardinality as (V_1, V_2). X_1 and X_2 are defined as

$$\begin{aligned} X_1 &= V_1 + U_1 + Z_1, \\ X_2 &= V_2 + U_2 + Z_2. \end{aligned}$$

This defines $p_W p_{V_1,U_1,X_1|W} p_{V_2,U_2,X_2|W}$. These choices are such that destination 4's observation does not depend on the "private" signal Z_1, and, similarly, destination 3's observation does not depend on Z_2. With these choices, the conditions on the non-negative rates $r_{V_1}, r_{V_2}, r_{U_1}, r_{U_2}, r_{Z_1}, r_{Z-2}$ after removing redundant conditions are

$$\begin{aligned} r_{V_1} &\leq n_C \\ r_{Z_1} &\leq [n_{1,3} - n_{1,4}], \\ r_{U_1} + r_{Z_1} &\leq n_{1,3} - n_C, \\ r_{U_2} + r_{Z_1} &\leq \max(n_{1,3} - n_{1,4}, n_{2,3} - n_C), \\ r_{U_1} + r_{U_2} + r_{Z_1} &\leq \max(n_{1,3} - n_C, n_{2,3} - n_C), \\ r_{V_1} + r_{V_2} + r_{U_1} + r_{U_2} + r_{Z_1} &\leq \max(n_{1,3}, n_{2,3}), \end{aligned}$$

and the corresponding inequalities with subscripts 1 and 2 exchanged, and 3 and 4 exchanged. Further, we make the following choices for the rates.

$$\begin{aligned} r_{Z_1} &= [n_{1,3} - n_{1,4}]_+, \\ r_{Z_2} &= [n_{2,4} - n_{2,3}]_+, \\ r_{V_1} = r_{V_2} &= n_C, \\ r_{U_1} &= \max(n_{2,4} - n_{2,3}, n_{1,4} - n_C) - [n_{2,4} - n_{2,3}]_+, \text{ and} \\ r_{U_2} &= \max(n_{1,3} - n_{1,4}, n_{2,3} - n_C) - [n_{1,3} - n_{1,4}]_+, \end{aligned}$$

where $[x]_+ = \max(x, 0)$. It can be shown that under the restricted regime of n_C, these choices satisfy all the conditions above. The resulting sum-rate is $u_1(n_C)$ as required.

When condition (13) does not hold, as we mentioned earlier, it is enough to prove that the sum-rate at $n_C = 0$ is achievable. We apply Theorem 4(a) where we set W, V_1, V_2 to be absent, and U_1, U_2, Z_1, Z_2 to be independent and uniformly distributed over their alphabets. $U_1, U_2 \in \mathbb{F}^n$, $Z_1 \in \mathcal{F}_{n_{1,4}}$, $Z_2 \in \mathcal{F}_{n_{2,3}}$, and X_1 and X_2 are defined as

$$\begin{aligned} X_1 &= V_1 + U_1 + Z_1, \\ X_2 &= V_2 + U_2 + Z_2. \end{aligned}$$

The conditions on the non-negative rates $r_{U_1}, r_{U_2}, r_{Z_1}, r_{Z-2}$ after removing redundant conditions are

$$
\begin{aligned}
r_{Z_1} &\leq [n_{1,3} - n_{1,4}], \\
r_{U_1} + r_{Z_1} &\leq n_{1,3}, \\
r_{U_2} + r_{Z_1} &\leq \max(n_{1,3} - n_{1,4}, n_{2,3}), \\
r_{U_1} + r_{U_2} + r_{Z_1} &\leq \max(n_{1,3}, n_{2,3}),
\end{aligned}
$$

and the corresponding inequalities with subscripts 1 and 2 exchanged, and 3 and 4 exchanged. Applying Fourier-Motzkin elimination, we can show that a sum-rate of $\min(u_1(0), u_2(0), u_3(0), u_4(0), u_5)$ is achievable.

Regime (iii): Without loss of generality, let us assume that $n_{1,3} \leq n_C \leq n_{2,4}$. We will apply Theorem 4(c) in two different ways to show achievability in this regime. The first application covers (1) $n_{2,4} \geq n_{1,4}$, while the second covers (2) $n_{2,4} < n_{1,4}$.

For case (1), $n_{2,4} \geq n_{1,4}$, we use the following choices for the auxiliary random variables: U_1 is a constants. $U_2, V_1, V_2, Z_1, Z_2, \widetilde{S}_{1,3}$, $S_{2,3}^{\perp}$, and S_1' are chosen to be independent and uniformly distributed over their alphabets. The alphabets are: $V_1, V_2 \in \mathbb{F}^n$, $Z_1 \in \mathcal{F}_{n_{1,4}}$, $Z_2 \in \mathcal{F}_{n_{2,3}}$, $\widetilde{S}_{1,3} \in \mathbb{F}^n$, $S_1' \in \mathcal{F}_{\max(n_{1,3}, n_{1,4})}$, $S_{2,3}^{\perp} \in \mathcal{F}_{n_{2,4}}$ and $U_2 \in \mathcal{F}_{n_C'}$, where n_C' ($n_C' \leq n_C$) is to be specified. W is independent of all these and has the same cardinality as (V_1, V_2). We define X_1 and X_2 as follows

$$
\begin{aligned}
X_1 &= V_1 + U_1 + Z_1 + \widetilde{S}_{1,3} + S_1', \\
X_2 &= V_2 + U_2 + Z_2 - \mathbf{S}^{n_{2,4} - n_{1,4}} \widetilde{S}_{1,3} + S_{2,3}^{\perp}
\end{aligned}
$$

with the result that

$$
\begin{aligned}
Y_3 &= \mathbf{S}^{n - n_{1,3}}(V_1 + Z_1) + \mathbf{S}^{n - n_{2,3}}(V_2 + U_2) + S_1 \\
Y_4 &= \mathbf{S}^{n - n_{2,4}}(V_2 + U_2 + Z_2) + \mathbf{S}^{n - n_{2,3}}(V_1),
\end{aligned}
$$

where we defined

$$
S_1 \stackrel{\text{def}}{=} (\mathbf{S}^{n - n_{1,3}} - \mathbf{S}^{n - n_{2,3} + n_{2,4} - n_{1,4}}) \widetilde{S}_{1,3} + \mathbf{S}^{n - n_{2,3}} S_{2,3}^{\perp}.
$$

We set $r_{U_1} = 0$ and the conditions on the other rates (after removing the redundant ones) are

$$
\begin{aligned}
r_{S_1} &\leq [n_C - \max(n_{1,3}, n_{1,4})]_+ \\
r_{V_1} + r_{U_1} + r_{Z_1} + r_{S_1} &\leq n_C,
\end{aligned}
$$

$$
r_{V_2} \leq n_C',
$$

$$
\begin{aligned}
r_{Z_1} &\leq [n_{1,3} - n_{1,4}]_+, \\
r_{U_2} + r_{Z_1} &\leq \max([n_{1,3} - n_{1,4}]_+, n_{2,3} - n_C'), \\
r_{S_1} + r_{Z_1} &\leq \begin{cases} \max([n_{1,3} - n_{1,4}]_+, n_{2,3} - (n_{2,4} - n_{1,4})), & \text{if } n_{1,3} + n_{2,4} \neq n_{1,4} + n_{2,3} \\ \max([n_{1,3} - n_{1,4}]_+, n_{2,3} - n_{2,4}), & \text{otherwise,} \end{cases}
\end{aligned}
$$

$$r_{U_2}+r_{S_1}+r_{Z_1} \leq \begin{cases} \max([n_{1,3}-n_{1,4}]_+, n_{2,3}-n'_C, n_{2,3}-(n_{2,4}-n_{1,4})), & \text{if } n_{1,3}+n_{2,4} \neq n_{1,4}+n_{2,3} \\ \max([n_{1,3}-n_{1,4}]_+, n_{2,3}-n'_C, n_{2,3}-n_{2,4}), & \text{otherwise,} \end{cases}$$

$$(r_{V_1}+r_{V_2})+r_{U_2}+r_{S_1}+r_{Z_1} \leq \max(n_{1,3}, n_{2,3}),$$

$$\begin{aligned} r_{Z_2} &\leq [n_{2,4}-n_{2,3}]_+, \\ r_{U_2}+r_{Z_2} &\leq \max([n_{2,4}-n_{2,3}]_+, n_{2,4}-n'_C), \\ (r_{V_1}+r_{V_2})+r_{U_2}+r_{Z_2} &\leq \max(n_{2,4}, n_{1,4}). \end{aligned}$$

By Fourier-Motzkin elimination, we may conclude that an achievable R_1+R_2 is given by the smaller of $u_2(n_C), u_3(n_C), u_4(n_C), u_5$ and

$$\frac{[n_C-\max(n_{1,3},n_{1,4})]_+ + \max(n_{2,4},n_{1,4}) + \max([n_{1,3}-n_{1,4}]_+, n_{2,3}-n'_C) + n_C + n'_C + [n_{2,4}-n_{2,3}]_+}{2}.$$

The last term above can be shown to be not smaller than the minimum of $u_2(n_C), u_3(n_C), u_4(n_C)$ and u_5 if n'_C is chosen to be such that $u_1(n'_C) = \min(u_2(n_C), u_3(n_C), u_4(n_C), u_5)$ if $u_1(0) < \min(u_2(0), u_3(0), u_4(0), u_5)$, and $n'_C = 0$ otherwise. Note that from earlier discussion, we know that this choice of n'_C must be less than or equal to n_C.

For (2) $n_{2,4} < n_{1,4}$, we apply Theorem 4(c) with the same auxiliary random variables $W, U_1, U_2, V_1, V_2, Z_1, Z_2, S_{2,3}^{\perp}, S'_1$ as in case (1) above. But instead of $\widetilde{S}_{1,3}$, we define $\widetilde{S}_{2,3}$ which is independent of all these random variables and distributed uniformly over $\mathbb{F}^n$. We define X_1, and X_2 as follows

$$\begin{aligned} X_1 &= V_1 + Z_1 + \mathbf{S}^{n_{1,4}-n_{2,4}}\widetilde{S}_{2,3} + S'_1, \\ X_2 &= V_2 + U_2 + Z_2 - \widetilde{S}_{2,3} + S_{2,3}^{\perp} \end{aligned}$$

with the result that

$$\begin{aligned} Y_3 &= \mathbf{S}^{n-n_{1,3}}(V_1+Z_1) + \mathbf{S}^{n-n_{2,3}}(V_2+U_2) + S_1, \\ Y_4 &= \mathbf{S}^{n-n_{2,4}}(V_2+U_2+Z_2) + \mathbf{S}^{n-n_{1,4}}(V_1), \end{aligned}$$

where we define S_1 as

$$S_1 \stackrel{\text{def}}{=} (\mathbf{S}^{n-n_{1,3}+n_{1,4}-n_{2,4}} - \mathbf{S}^{n-n_{2,3}})\widetilde{S}_{2,3} + \mathbf{S}^{n-n_{2,3}}S_{2,3}^{\perp}.$$

The conditions on the rates (after removing the redundant ones) are the same as in case (1) except for the following two

$$r_{S_1}+r_{Z_1} \leq \begin{cases} \max([n_{1,3}-n_{1,4}]_+, n_{1,3}-(n_{1,4}-n_{2,4}), n_{2,3}-n_{2,4}), & \text{if } n_{1,3}+n_{2,4} \neq n_{1,4}+n_{2,3}, \\ \max([n_{1,3}-n_{1,4}]_+, n_{2,3}-n_{2,4}), & \text{otherwise,} \end{cases}$$

$$r_{U_2}+r_{S_1}+r_{Z_1}$$
$$\leq \begin{cases} \max([n_{1,3}-n_{1,4}]_+, n_{2,3}-n'_C, n_{1,3}-(n_{1,4}-n_{2,4}), n_{2,3}-n_{2,4}), & \text{if } n_{1,3}+n_{2,4} \neq n_{1,4}+n_{2,3}, \\ \max([n_{1,3}-n_{1,4}]_+, n_{2,3}-n'_C, n_{2,3}-n_{2,4}), & \text{otherwise.} \end{cases}$$

By Fourier-Motzkin elimination, we may conclude that the achievable $R_1 + R_2$ is given by the smaller of $u_2(n_C), u_3(n_C), u_4(n_C), u_5$ and

$$\frac{[n_C - \max(n_{1,3}, n_{1,4})]_+ + \max(n_{2,4}, n_{1,4}) + \max([n_{1,3} - n_{1,4}]_+, n_{2,3} - n'_C) + n_C + n'_C + [n_{2,4} - n_{2,3}]_+}{2}.$$

The above term can be shown to be not smaller than the minimum of $u_2(n_C), u_3(n_C), u_4(n_C)$ and u_5 if n'_C is chosen as was done in case (1) above.

We can represent the two cases above together using the following notation:

$$\begin{aligned} X_1 &= V_1 + Z_1 + \bar{S}_{1,3} + S'_1, \\ X_2 &= V_2 + U_2 + Z_2 + \bar{S}_{2,3}, \end{aligned}$$

where

$$\bar{S}_{1,3} = \begin{cases} \widetilde{S}_{1,3}, & \text{if } n_{2,4} \geq n_{1,4}, \\ \mathbf{S}^{n_{1,4}-n_{2,4}} \widetilde{S}_{2,3}, & \text{otherwise,} \end{cases}$$

$$\bar{S}_{2,3} = \begin{cases} -\mathbf{S}^{n_{2,4}-n_{1,4}} \widetilde{S}_{1,3} + S^{\perp}_{2,3}, & \text{if } n_{2,4} \geq n_{1,4}, \\ -\widetilde{S}_{2,3} + S^{\perp}_{2,3}, & \text{otherwise,} \end{cases}$$

with the result that

$$\begin{aligned} Y_3 &= \mathbf{S}^{n-n_{1,3}}(V_1 + Z_1) + \mathbf{S}^{n-n_{2,3}}(V_2 + U_2) + S_1, \\ Y_4 &= \mathbf{S}^{n-n_{2,4}}(V_2 + U_2 + Z_2) + \mathbf{S}^{n-n_{1,4}}(V_1), \end{aligned}$$

where

$$S_1 \overset{\text{def}}{=} \begin{cases} (\mathbf{S}^{n-n_{1,3}} - \mathbf{S}^{n-n_{2,3}+n_{2,4}-n_{1,4}}) \widetilde{S}_{1,3} + \mathbf{S}^{n-n_{2,3}} S^{\perp}_{2,3}, & \text{if } n_{2,4} \geq n_{1,4}, \\ (\mathbf{S}^{n-n_{1,3}+n_{1,4}-n_{2,4}} - \mathbf{S}^{n-n_{2,3}}) \widetilde{S}_{2,3} + \mathbf{S}^{n-n_{2,3}} S^{\perp}_{2,3}, & \text{otherwise.} \end{cases}$$

We will use a symmetric form of this notation below for regime (iv).

Regime (iv): Application of Theorem 4(b) with the following auxiliary random variables covers this regime: U_1, U_2 are constants. $V_1, V_2, Z_1, Z_2, \widetilde{S}_{1,3}, \widetilde{S}_{2,3}, \widetilde{S}_{1,4}, \widetilde{S}_{2,4}, S^{\perp}_{2,3}, S^{\perp}_{1,4}, S'_1, S'_2$ are chosen to be independent and uniformly distributed over their alphabets. Also, we choose the alphabets to be $V_1, V_2, \widetilde{S}_{1,3}, \widetilde{S}_{2,3}, \widetilde{S}_{1,4}, \widetilde{S}_{2,4} \in \mathbb{F}^n$, $Z_1 \in \mathcal{F}_{n_{1,4}}$, $Z_2 \in \mathcal{F}_{n_{2,3}}$, $S^{\perp}_{2,3} \in \mathcal{F}_{n_{2,4}}$, $S^{\perp}_{1,4} \in \mathcal{F}_{n_{1,3}}$, $S'_1 \in \mathcal{F}_{\max(n_{1,3},n_{1,4})}$, and $S'_2 \in \mathcal{F}_{\max(n_{2,4},n_{2,3})}$. W is independent of all these and has the same cardinality as (V_1, V_2). Further, we define X_1, and X_2 as follows

$$\begin{aligned} X_1 &= V_1 + U_1 + Z_1 + \bar{S}_{1,3} + \bar{S}_{1,4} + S'_1 \\ X_2 &= V_2 + U_2 + Z_2 + \bar{S}_{2,3} + \bar{S}_{2,4} + S'_2, \end{aligned}$$

where

$$\bar{S}_{1,3} = \begin{cases} \widetilde{S}_{1,3}, & \text{if } n_{2,4} \geq n_{1,4}, \\ \mathbf{S}^{n_{1,4}-n_{2,4}} \widetilde{S}_{2,3}, & \text{otherwise}, \end{cases}$$

$$\bar{S}_{2,3} = \begin{cases} -\mathbf{S}^{n_{2,4}-n_{1,4}} \widetilde{S}_{1,3} + S_{2,3}^{\perp}, & \text{if } n_{2,4} \geq n_{1,4}, \\ -\widetilde{S}_{2,3} + S_{2,3}^{\perp}, & \text{otherwise}, \end{cases}$$

$$\bar{S}_{2,4} = \begin{cases} \widetilde{S}_{2,4}, & \text{if } n_{1,3} \geq n_{2,3}, \\ \mathbf{S}^{n_{2,3}-n_{1,3}} \widetilde{S}_{1,4}, & \text{otherwise}, \end{cases}$$

$$\bar{S}_{1,4} = \begin{cases} -\mathbf{S}^{n_{1,3}-n_{2,3}} \widetilde{S}_{2,4} + S_{1,4}^{\perp}, & \text{if } n_{1,3} \geq n_{2,3}, \\ -\widetilde{S}_{1,4} + S_{1,4}^{\perp}, & \text{otherwise}. \end{cases}$$

The upshot of this is that

$$\begin{aligned} Y_3 &= \mathbf{S}^{n-n_{1,3}}(V_1 + Z_1) + \mathbf{S}^{n-n_{2,3}} V_2 + S_1 \\ Y_4 &= \mathbf{S}^{n-n_{2,4}}(V_2 + Z_2) + \mathbf{S}^{n-n_{2,3}} V_1 + S_2, \end{aligned}$$

where

$$S_1 \stackrel{\text{def}}{=} \begin{cases} (\mathbf{S}^{n-n_{1,3}} - \mathbf{S}^{n-n_{2,3}+n_{2,4}-n_{1,4}}) \widetilde{S}_{1,3} + \mathbf{S}^{n-n_{2,3}} S_{2,3}^{\perp}, & \text{if } n_{2,4} \geq n_{1,4}, \\ (\mathbf{S}^{n-n_{1,3}+n_{1,4}-n_{2,4}} - \mathbf{S}^{n-n_{2,3}}) \widetilde{S}_{2,3} + \mathbf{S}^{n-n_{2,3}} S_{2,3}^{\perp}, & \text{otherwise}, \end{cases}$$

$$S_2 \stackrel{\text{def}}{=} \begin{cases} (\mathbf{S}^{n-n_{2,4}} - \mathbf{S}^{n-n_{1,4}+n_{1,3}-n_{2,3}}) \widetilde{S}_{2,4} + \mathbf{S}^{n-n_{1,4}} S_{1,4}^{\perp}, & \text{if } n_{1,3} \geq n_{2,3}, \\ (\mathbf{S}^{n-n_{2,4}+n_{2,3}-n_{1,3}} - \mathbf{S}^{n-n_{1,4}}) \widetilde{S}_{1,4} + \mathbf{S}^{n-n_{1,4}} S_{1,4}^{\perp}, & \text{otherwise}. \end{cases}$$

With these choices, the conditions on the non-negative rates $r_{V_1}, r_{V_2}, r_{S_1}, r_{S_2}, r_{Z_1}, r_{Z_2}$ after removing redundant conditions are

$$\begin{aligned} r_{S_1} &\leq [n_C - \max(n_{1,3}, n_{1,4})]_+, \\ r_{V_1} + r_{Z_1} + r_{S_1} &\leq n_C, \\ \\ r_{Z_1} &\leq [n_{1,3} - n_{1,4}]_+, \\ r_{S_1} + r_{Z_1} &\leq \max([n_{1,3} - n_{1,4}]_+, n_{S_1}), \\ (r_{V_1} + r_{V_2}) + r_{S_1} + r_{Z_1} &\leq \max(n_{1,3}, n_{2,3}), \end{aligned}$$

and the corresponding inequalities with subscripts 1 and 2 exchanged, and 3 replaced by 4, where

$$n_{S_1} = \begin{cases} \max(n_{1,3}, n_{2,3} - (n_{2,4} - n_{1,4})), & \text{if } n_{2,4} \geq n_{1,4} \text{ and } n_{1,3} + n_{2,4} \neq n_{1,4} + n_{2,3}, \\ \max(n_{2,3}, n_{1,3} - (n_{1,4} - n_{2,4})), & \text{if } n_{2,4} < n_{1,4} \text{ and } n_{1,3} + n_{2,4} \neq n_{1,4} + n_{2,3}, \\ (n_{2,3} - n_{2,4})_+, & \text{if } n_{1,3} + n_{2,4} = n_{1,4} + n_{2,3}. \end{cases}$$

We may apply Fourier-Motzkin elimination to obtain the sum-rate supported by this scheme. We get a sum-rate which is the minimum of $u_2(n_C), u_3(n_C), u_4(n_C), u_5,$. This completes the achievability proof.

C Proof of achievability of Theorem 2

We prove Theorem 2 using Theorem 4. Note that we proved the latter for discrete alphabets, but the extension to the continuous alphabet case is standard and we will assume that version in this section. This proof will follow the proof of Theorem 1 closely. We first make the following definitions:

$$\begin{aligned} n_{k_1,k_2} &\stackrel{\text{def}}{=} [\log |h_{k_1,k_2}|^2]_+, \; k_1, k_2 \in \{1,2,3,4\}, \text{ and} \\ n_C &\stackrel{\text{def}}{=} [\log |h_C|^2]_+. \end{aligned}$$

First, we observe that the following four terms u_1', u_2', u_3', and u_4' are within a constant (5 bits) of the corresponding unprimed terms, u_1, u_2, u_3, and u_4, respectively

$$u_1' = \max(n_{1,3} - n_{1,4} + n_C, n_{2,3}, n_C) + \max(n_{2,4} - n_{2,3} + n_C, n_{1,4}, n_C), \tag{14}$$

$$u_2' = \max(n_{1,3}, n_{2,3}) + (\max(n_{2,4}, n_{2,3}, n_C) - n_{2,3}), \tag{15}$$

$$u_3' = \max(n_{2,4}, n_{1,4}) + (\max(n_{1,3}, n_{1,4}, n_C) - n_{1,4}), \tag{16}$$

$$u_4' = \max(n_{1,3}, n_C) + \max(n_{2,4}, n_C). \tag{17}$$

Hence, it is enough to show that the minimum of the four terms above and

$$\begin{aligned} u_5' = \log \Bigg(1 &+ \left(|h_{1,3}|^2 + |h_{2,4}|^2 + |h_{1,4}|^2 + |h_{2,3}|^2 \right) \\ &+ \left(|h_{1,3}h_{2,4}|^2 + |h_{1,4}h_{2,3}|^2 - 2|h_{1,3}h_{2,4}h_{1,4}h_{2,3}| \cos\theta \right) \Bigg), \end{aligned} \tag{18}$$

which is within a constant (3 bits) of u_5, is achievable. We again consider the same four regimes as in the proof of Theorem 1:

Regime (i): $n_C \le n_{\min} \stackrel{\text{def}}{=} \min(n_{1,3}, n_{1,4}, n_{2,3}, n_{2,4})$. The discussion for regime (i) in the linear deterministic case continues to hold here as well. Note that u_5 is such that

$$u_5 - 5 \le u_5'' \stackrel{\text{def}}{=} \begin{cases} \max(n_{1,3} + n_{2,4}, n_{1,4} + n_{2,3}), & \text{if } n_{1,3} - n_{2,3} \neq n_{1,4} - n_{2,4}, \\ \max(n_{1,3}, n_{2,4}, n_{1,4}, n_{2,3}), & \text{otherwise.} \end{cases}$$

Thus, when condition (13) does not hold, the achievability (within a gap of 7 bits from the upperbound) is implied by the results of Etkin-Tse-Wang [2] (where 2-bit gap comes from [2] and an additional 5 bits were incurred above). And, when condition (13) holds, we need only show achievability in the restricted regime of n_C where

$$u_1'(n_C) \le \min(u_2'(n_C), u_3'(n_C), u_4'(n_C), u_5'').$$

We employ Theorem 4(a) using the following auxiliary random variables $W, V_1, U_1, Z_1, V_2, U_2, Z_2$ are zero-mean Gaussian random variables and independent of each other with the following variances:

$$
\begin{aligned}
\sigma^2_{V_1} = \sigma^2_{V_2} &= 1/K, \\
\sigma^2_{U_1} = \sigma^2_{U_2} &= \frac{1/K}{\max(1, |h_C|)}, \\
\sigma^2_{Z_1} &= \frac{1/K}{\max(1, |h_{1,4}|^2)}, \text{ and} \\
\sigma^2_{Z_2} &= \frac{1/K}{\max(1, |h_{2,3}|^2)},
\end{aligned}
$$

where K is a constant which will be specified soon. W is independent of all these and has the same distribution as (V_1, V_2). X_1 and X_2 are defined as

$$
\begin{aligned}
X_1 &= V_1 + U_1 + Z_1, \\
X_2 &= V_2 + U_2 + Z_2.
\end{aligned}
$$

In order for the power constraint to be satisfied, it is enough to have $K < 3$. This defines $p_W p_{V_1,U_1,X_1|W} p_{V_2,U_2,X_2|W}$. These choices are such that the "private" signal Z_1 appears at destination 4 with less power than the noise, and, similarly, Z_2 appears at destination 3 with less power than the noise. With these choices, the conditions on the non-negative rates $r_{V_1}, r_{V_2}, r_{U_1}, r_{U_2}, r_{Z_1}, r_{Z_2}$ are

$$
\begin{aligned}
r_{V_1} &\leq \log\left(1 + \frac{|h_C|^2/K}{2/K + 1}\right) \\
r_{Z_1} &\leq \log\left(1 + \frac{|h_{1,3}|^2/(\max(1, |h_{1,4}|^2)K)}{1/K + 1}\right), \\
r_{U_1} + r_{Z_1} &\leq \log\left(1 + \frac{|h_{1,3}|^2/(\max(1, |h_C|^2)K) + |h_{1,3}|^2/(\max(1, |h_{1,4}|^2)K)}{1/K + 1}\right), \\
r_{U_2} + r_{Z_1} &\leq \log\left(1 + \frac{|h_{1,3}|^2/(\max(1, |h_{1,4}|^2)K) + |h_{2,3}|^2/(\max(1, |h_C|^2)K)}{1/K + 1}\right), \\
r_{U_1} + r_{U_2} + r_{Z_1} &\leq \log\left(1 + \frac{\frac{|h_{1,3}|^2}{\max(1,|h_C|^2)K} + \frac{|h_{1,3}|^2}{\max(1,|h_{1,4}|^2)K} + \frac{|h_{2,3}|^2}{\max(1,|h_C|^2)K}}{1/K + 1}\right), \\
r_{V_1} + r_{V_2} + r_{U_1} + r_{U_2} + r_{Z_1} &\leq \log\left(1 + \frac{\frac{|h_{1,3}|^2}{K} + \frac{|h_{1,3}|^2}{\max(1,|h_C|^2)K} + \frac{|h_{1,3}|^2}{\max(1,|h_{1,4}|^2)K} + \frac{|h_{2,3}|^2}{K} + \frac{|h_{2,3}|^2}{\max(1,|h_C|^2)K}}{1/K + 1}\right),
\end{aligned}
$$

Simplifying, we can show that these conditions imply that non-negative rates which satisfy the same conditions as in the linear deterministic case (up to a constant) are achievable.

$$
\begin{aligned}
r_{V_1} &\leq n_C - \log 5 \\
r_{Z_1} &\leq [n_{1,3} - n_{1,4}]_+ - \log 4,
\end{aligned}
$$

$$
\begin{aligned}
r_{U_1} + r_{Z_1} &\le n_{1,3} - n_C - \log 4,\\
r_{U_2} + r_{Z_1} &\le \max(n_{1,3} - n_{1,4}, n_{2,3} - n_C) - \log 4,\\
r_{U_1} + r_{U_2} + r_{Z_1} &\le \max(n_{1,3} - n_C, n_{2,3} - n_C) - \log 4,\\
r_{V_1} + r_{V_2} + r_{U_1} + r_{U_2} + r_{Z_1} &\le \max(n_{1,3}, n_{2,3}) - \log 4,
\end{aligned}
$$

and the corresponding inequalities with subscripts 1 and 2 exchanged, and 3 and 4 exchanged. Note that the right hand sides above should be interpreted as zero if they evaluate to less than zero. We will tacitly assume this for similar conditions in the sequel. Further, we make the following choices for the rates.

$$
\begin{aligned}
r_{Z_1} &= [n_{1,3} - n_{1,4}]_+ - \log 4,\\
r_{Z_2} &= [n_{2,4} - n_{2,3}]_+ - \log 4,\\
r_{V_1} = r_{V_2} &= n_C - \log 5,\\
r_{U_1} &= \max(n_{2,4} - n_{2,3}, n_{1,4} - n_C) - [n_{2,4} - n_{2,3}]_+ - \log 4, \text{ and}\\
r_{U_2} &= \max(n_{1,3} - n_{1,4}, n_{2,3} - n_C) - [n_{1,3} - n_{1,4}]_+ - \log 4,
\end{aligned}
$$

where we interpret the rates as zero if their values work out to less than zero. It can be shown that under the restricted regime of n_C, these choices satisfy all the conditions above. The resulting sum-rate is $u_1'(n_C)$ within a constant gap (of at most 13 bits) as required.

When condition (13) does not hold, as we mentioned earlier, it is enough to prove that the sum-rate at $n_C = 0$ is achievable. We may invoke the achievability proof of Etkin-Tse-Wang [2] to conclude that a sum-rate which matches our upperbounds (up to a constant) is achievable. Thus, overall, in regime (i), we may conclude that the upperbound is achievable within a constant gap of 18 bits.

Regime (ii): $n_{\min} < n_C \le \min(n_{1,3}, n_{2,4})$. As in the linear deterministic case, the achievability in this regime is implied by the achievability in regime (i).

Regime (iii): $\min(n_{1,3}, n_{2,4}) < n \le \max(n_{1,3}, n_{2,4})$. Without loss of generality, let us assume that $n_{1,3} \le n_C \le n_{2,4}$. We will apply Theorem 4(c) to establish the achievability. We consider two separate possibilities: (1) $|h_{2,4}| \ge |h_{1,4}|$ and (2) $|h_{2,4}| < |h_{1,4}|$.

(1) When $|h_{2,4}| \ge |h_{1,4}|$ (which implies that $n_{2,4} \ge n_{1,4}$), the auxiliary random variables are as follows: U_1 is set to a constant. $U_2, V_1, V_2, Z_1, Z_2, \widetilde{S}_{1,3}, S_{2,3}^{\perp}$ and S_1' are independent zero-mean Gaussian random variables. Their variances are as follows

$$
\begin{aligned}
\sigma_{V_1}^2 = \sigma_{V_2}^2 &= 1/K,\\
\sigma_{U_2}^2 &= \frac{1/K}{\max(1, |h_C'|^2)},\\
\sigma_{Z_1}^2 &= \frac{1/K}{\max(1, |h_{1,4}|^2)},\\
\sigma_{Z_2}^2 &= \frac{1/K}{\max(1, |h_{2,3}|^2)},\\
\sigma_{\widetilde{S}_{1,3}}^2 &= 1/K,
\end{aligned}
$$

$$\sigma^2_{S'_1} = \frac{1/K}{\max(1,|h_{1,3}|^2,|h_{1,4}|^2)},$$
$$\sigma^2_{S^\perp_{2,3}} = \frac{1/K}{\max(1,|h_{2,4}|^2)},$$

where K and h'_C ($0 < h'_C$) will be specified. Let us define $n'_C \stackrel{\text{def}}{=} [\log h'^2_C]_+$. We will pick a h'_C such that $n'_C \le \min(n_C, n_{2,3})$. We set W to be independent of all these and have the same distribution as (V_1, V_2). We define X_1, and X_2 as follows

$$X_1 = V_1 + Z_1 + \widetilde{S}_{1,3} + S'_1,$$
$$X_2 = V_2 + U_2 + Z_2 - \frac{|h_{1,4}|e^{j\theta/2}}{|h_{2,4}|}\widetilde{S}_{1,3} + S^\perp_{2,3}.$$

This satisfies the power constraint if $K < 5$ (where we used the fact that $|h_{2,4}| \ge |h_{1,4}|$). Let $S_1 = (\widetilde{S}_{1,3}, S^\perp_{2,3})$ and

$$\bar{S}_1 = \left(|h_{1,3}| - \frac{|h_{1,4}||h_{2,3}|e^{j\theta}}{|h_{2,4}|}\right)\widetilde{S}_{1,3} + |h_{2,3}|e^{j\theta/2}S^\perp_{2,3}.$$

Then

$$Y_3 = |h_{1,3}|(V_1 + Z_1 + S'_1) + \bar{S}_1 + N_3,$$
$$Y_4 = |h_{2,4}|(V_2 + U_2 + Z_2) + |h_{1,4}|e^{j\theta/2}(V_1 + Z_1 + S'_1) + |h_{2,4}|S^\perp_{2,3} + N_4.$$

The conditions on the rates are

$$r_{S_1} \le \log\left(1 + |h_C|^2/(\max(1,|h_{1,3}|^2,|h_{1,4}|^2)K)\right),$$
$$r_{Z_1} + r_{S_1} \le \log\left(1 + \frac{|h_C|^2}{\max(1,|h_{1,3}|^2,|h_{1,4}|^2)K} + \frac{|h_C|^2}{\max(1,|h_{1,4}|^2)K}\right),$$
$$r_{V_1} + r_{Z_1} + r_{S_1} \le \log\left(1 + \frac{|h_C|^2}{K} + \frac{|h_C|^2}{\max(1,|h_{1,3}|^2,|h_{1,4}|^2)K} + \frac{|h_C|^2}{\max(1,|h_{1,4}|^2)K}\right),$$

$$r_{V_2} \le \log\left(1 + \frac{|h_C|^2/K}{|h_C|^2/(\max(1,|h'_C|^2)K) + |h_C|^2/(\max(1,|h_{2,3}|^2)K) + 1}\right),$$

$$r_{Z_1} \le \log\left(1 + \frac{|h_{1,3}|^2/(\max(1,|h_{1,4}|^2)K)}{2/K+1}\right),$$
$$r_{U_2} + r_{Z_1} \le \log\left(1 + \frac{|h_{1,3}|^2/(\max(1,|h_{1,4}|^2)K) + |h_{2,3}|^2/(\max(1,|h'_C|^2)K)}{2/K+1}\right),$$
$$r_{S_1} + r_{Z_1} \le \log\left(1 + \frac{\frac{|h_{1,3}|^2}{\max(1,|h_{1,4}|^2)K} + \frac{\left||h_{1,3}||h_{2,4}| - |h_{1,4}||h_{2,3}|e^{j\theta}\right|^2}{|h_{2,4}|^2K} + \frac{|h_{2,3}|^2}{\max(1,|h_{2,4}|^2)K}}{2/K+1}\right),$$

$$r_{U_2}+r_{S_1}+r_{Z_1}$$
$$\leq \log\left(1+\frac{\frac{|h_{1,3}|^2}{\max(1,|h_{1,4}|^2)K}+\frac{\left||h_{1,3}||h_{2,4}|-|h_{1,4}||h_{2,3}|e^{j\theta}\right|^2}{|h_{2,4}|^2K}+\frac{|h_{2,3}|^2}{\max(1,|h_{2,4}|^2)K}+\frac{|h_{2,3}|^2}{\max(1,|h'_C|^2)K}}{2/K+1}\right),$$
$$(r_{V_1}+r_{V_2})+r_{U_2}+r_{S_1}+r_{Z_1}$$
$$\leq \log\left(1+\frac{\frac{|h_{1,3}|^2}{K}+\frac{|h_{1,3}|^2}{\max(1,|h_{1,4}|^2)K}+\frac{\left||h_{1,3}||h_{2,4}|-|h_{1,4}||h_{2,3}|e^{j\theta}\right|^2}{|h_{2,4}|^2K}+\frac{|h_{2,3}|^2}{K}+\frac{|h_{2,3}|^2}{\max(1,|h_{2,4}|^2)K}+\frac{|h_{2,3}|^2}{\max(1,|h'_C|^2)K}}{2/K+1}\right),$$

$$r_{Z_2} \leq \log\left(1+\frac{|h_{2,4}|^2/(\max(1,|h_{2,3}|^2)K}{3/K+1}\right),$$
$$r_{U_2}+r_{Z_2} \leq \log\left(1+\frac{|h_{2,4}|^2/(\max(1,|h'_C|^2)K)+|h_{2,4}|^2/(\max(1,|h_{2,3}|^2)K}{3/K+1}\right),$$
$$(r_{V_1}+r_{V_2})+r_{U_2}+r_{Z_2} \leq \log\left(1+\frac{\frac{|h_{2,4}|^2}{K}+\frac{|h_{2,4}|^2}{\max(1,|h'_C|^2)K}+\frac{|h_{2,4}|^2}{\max(1,|h_{2,3}|^2)K}+\frac{|h_{1,4}|^2}{K}}{3/K+1}\right).$$

Upon simplification, the above conditions imply that non-negative rates which satisfy the conditions below are achievable.

$$r_{S_1} \leq [n_C-\max(n_{1,3},n_{1,4})]_+ - \log 5,$$
$$r_{V_1}+r_{Z_1}+r_{S_1} \leq n_C - \log 5,$$

$$r_{V_2} \leq n'_C - \log 7,$$

$$r_{Z_1} \leq [n_{1,3}-n_{1,4}]_+ - \log 7,$$
$$r_{U_2}+r_{Z_1} \leq \max([n_{1,3}-n_{1,4}]_+, n_{2,3}-n'_C) - \log 7,$$

$$r_{S_1}+r_{Z_1}$$
$$\leq \log\left(1+\left|\frac{h_{1,3}}{\max(1,|h_{1,4}|)}\right|^2+\left|\frac{|h_{1,3}||h_{2,4}|-|h_{1,4}||h_{2,3}|e^{j\theta}}{h_{2,4}}\right|^2+\left|\frac{h_{2,3}}{\max(1,|h_{2,4}|)}\right|^2\right)-\log 7,$$
$$r_{U_2}+r_{S_1}+r_{Z_1}$$
$$\leq \log\left(1+\left|\frac{h_{2,3}}{\max(1,|h'_C|)}\right|^2+\left|\frac{h_{1,3}}{h_{1,4}}\right|^2+\left|\frac{|h_{1,3}||h_{2,4}|-|h_{1,4}||h_{2,3}|e^{j\theta}}{h_{2,4}}\right|^2+\left|\frac{h_{2,3}}{\max(1,|h_{2,4}|)}\right|^2\right)-\log 7,$$

$$(r_{V_1}+r_{V_2})+r_{U_2}+r_{S_1}+r_{Z_1} \le \max(n_{1,3},n_{2,3})-\log 7,$$

$$\begin{aligned} r_{Z_2} &\le [n_{2,4}-n_{2,3}]_+ - \log 8,\\ r_{U_2}+r_{Z_2} &\le \max([n_{2,4}-n_{2,3}]_+, n_{2,4}-n'_C)-\log 8,\\ (r_{V_1}+r_{V_2})+r_{U_2}+r_{Z_2} &\le \max(n_{2,4},n_{1,4})-\log 8. \end{aligned}$$

Note that the conditions on the rates are exactly as in the linear deterministic case up to a constant except for the constraints on $r_{S_1}+r_{Z_1}$ and $r_{U_2}+r_{S_1}+r_{Z_1}$. By Fourier-Motzkin elimination, we may conclude that a sum-rate R_1+R_2 within a constant (9 bits) of the smaller of $u'_2(n_C), u'_3(n_C), u'_4(n_C), u'_5$ and

$$\frac{[n_C-\max(n_{1,3},n_{1,4})]_+ + \max(n_{2,4},n_{1,4}) + \max([n_{1,3}-n_{1,4}]_+, n_{2,3}-n'_C) + n_C + n'_C + [n_{2,4}-n_{2,3}]_+}{2}$$

is achievable. The last term can be shown to be not smaller than the minimum of $u'_2(n_C), u'_3(n_C), u'_4(n_C)$ and u'_5, if n'_C is chosen to be such that $u_1(n'_C) = \min(u_2(n_C), u_3(n_C), u_4(n_C), u_5)$ when $u_1(0) < \min(u_2(0), u_3(0), u_4(0), u_5)$, and $n'_C = 0$, otherwise. Note that from earlier discussion, we know that this choice of n'_C must be less than or equal to n_C and all $n_{i,j}$, $i \in \{1,2\}$, $j \in \{3,4\}$, and in particular $n_{2,3}$.

(2) When $|h_{2,4}| < |h_{1,4}|$ (which implies that $n_{2,4} \le n_{1,4}$), we apply Theorem 4(c) as in case (1) above with the same choices for the auxiliary random variables $W, V_1, V_2, U_1, U_2, Z_1, Z_2, S^{\perp}_{2,3}, S'_1$. But, instead of $\widetilde{S}_{1,3}$ we now define an independent, zero-mean Gaussian random variable $\widetilde{S}_{2,3}$ with variance $\sigma^2_{\widetilde{S}_{2,3}} = 1/K$. The conditional distributions of X_1 and X_2 are defined through

$$\begin{aligned} X_1 &= V_1 + Z_1 - \frac{|h_{2,4}|}{|h_{1,4}|e^{j\theta/2}}\widetilde{S}_{2,3} + S'_1,\\ X_2 &= V_2 + U_2 + Z_2 + \widetilde{S}_{2,3} + S^{\perp}_{2,3}, \end{aligned}$$

which satisfies the power constraint if we set $K < 4$ (since $|h_{2,4}| < |h_{1,4}|$). We define $S_1 = (\widetilde{S}_{2,3}, S^{\perp}_{2,3})$ and

$$\bar{S}_1 = \left(|h_{2,3}|e^{j\theta/2} - \frac{|h_{1,3}||h_{2,4}|e^{-j\theta/2}}{|h_{1,4}|}\right)\widetilde{S}_{2,3} + |h_{2,3}|e^{j\theta/2}S^{\perp}_{2,3}.$$

The joint distribution of the signals received at the destinations is given by

$$\begin{aligned} Y_3 &= |h_{1,3}|(V_1+Z_1+S'_1) + |h_{2,3}|e^{j\theta}(V_2+U_2+Z_2) + \bar{S}_1 + N_3,\\ Y_4 &= |h_{2,4}|(V_2+U_2+Z_2) + |h_{1,4}|e^{j\theta}(V_1+Z_1+S'_1) + |h_{2,4}|S^{\perp}_{2,3} + N_4. \end{aligned}$$

The simplified conditions on the rates are identical to those in case (1), except for the following two

$$r_{S_1}+r_{Z_1} \le \log\left(1+\left|\frac{h_{1,3}}{h_{1,4}}\right|^2 + \left|\frac{|h_{1,3}||h_{2,4}|-|h_{1,4}||h_{2,3}|e^{j\theta}}{h_{1,4}}\right|^2 + \left|\frac{h_{2,3}}{h_{2,4}}\right|^2\right) - \log 6,$$

$$r_{U_2}+r_{S_1}+r_{Z_1} \le \log\left(1+\left|\frac{h_{2,4}}{h'_C}\right|^2 + \left|\frac{h_{1,3}}{h_{1,4}}\right|^2 + \left|\frac{|h_{1,3}||h_{2,4}|-|h_{1,4}||h_{2,3}|e^{j\theta}}{h_{1,4}}\right|^2 + \left|\frac{h_{2,3}}{h_{2,4}}\right|^2\right) - \log 6.$$

Applying Fourier-Motzkin elimination and choosing n'_C as in case (1) completes the achievability proof (to within 9 bits of the upperbound)

Thus, we may conclude that the upperbound is achievable with a gap of at most 9 bits in regime (iii). Note that we may represent the two cases together as follows:

$$\begin{aligned} X_1 &= V_1 + Z_1 + \bar{S}_{1,3} + S'_1, \\ X_2 &= V_2 + U_2 + Z_2 + \bar{S}_{2,4}, \end{aligned}$$

where

$$\bar{S}_{1,3} = \begin{cases} \widetilde{S}_{1,3}, & \text{if } |h_{2,4}| \geq |h_{1,4}|, \\ -\frac{|h_{2,4}|}{|h_{1,4}|e^{j\theta/2}}\widetilde{S}_{2,3}, & \text{otherwise.} \end{cases}$$

$$\bar{S}_{2,3} = \begin{cases} -\frac{|h_{1,4}|e^{j\theta/2}}{|h_{2,4}|}\widetilde{S}_{1,3} + S^{\perp}_{2,3}, & \text{if } |h_{2,4}| \geq |h_{1,4}|, \\ \widetilde{S}_{2,3} + S^{\perp}_{2,3}, & \text{otherwise.} \end{cases}$$

This gives

$$\begin{aligned} Y_3 &= |h_{1,3}|(V_1 + Z_1 + S'_1) + |h_{2,3}|e^{j\theta/2}(V_2 + U_2 + Z_2) + \bar{S}_1 + N_3, \\ Y_4 &= |h_{2,4}|(V_2 + U_2 + Z_2) + |h_{2,3}|e^{j\theta/2}(V_1 + Z_1 + S'_1) + |h_{2,4}|S^{\perp}_{2,3} + N_4, \end{aligned}$$

where, we define $S_1 = (\bar{S}_{2,3}, \bar{S}_{1,3}, S^{\perp}_{2,3})$ and

$$\bar{S}_1 = |h_{2,3}|e^{j\theta/2}\bar{S}_{2,3} + |h_{1,3}|\bar{S}_{1,3} + |h_{2,3}|e^{j\theta/2}S^{\perp}_{2,3}.$$

The distribution of S_1, S_2 we will employ in regime (iv) below is a symmetric generalization of this.

Regime (iv): $\max(n_{1,3}, n_{2,4}) < n_C$. In this regime, we employ Theorem 4(b) as we did for the linear deterministic case. U_1, U_2 are constants. $V_1, V_2, Z_1, Z_2, \widetilde{S}_{1,3}, \widetilde{S}_{2,3}, \widetilde{S}_{1,4}, \widetilde{S}_{2,4}, S^{\perp}_{2,3}, S^{\perp}_{1,4}, S'_1, S'_2$ are independent zero-mean Gaussian random variables. Their variances are as follows

$$\begin{aligned} \sigma^2_{V_1} = \sigma^2_{V_2} &= 1/K, \\ \sigma^2_{Z_1} &= \frac{1/K}{\max(1, |h_{1,4}|^2)}, \\ \sigma^2_{Z_2} &= \frac{1/K}{\max(1, |h_{2,3}|^2)}, \\ \sigma^2_{\widetilde{S}_{i,j}} &= 1/K, \;\; i \in \{1,2\}, \; j \in \{3,4\}, \\ \sigma^2_{S^{\perp}_{1,4}} &= \frac{1/K}{\max(1, |h_{1,3}|^2)}, \\ \sigma^2_{S^{\perp}_{2,3}} &= \frac{1/K}{\max(1, |h_{2,4}|^2)}, \\ \sigma^2_{S'_1} &= \frac{1/K}{\max(1, |h_{1,3}|^2, |h_{1,4}|^2)}, \\ \sigma^2_{S'_2} &= \frac{1/K}{\max(1, |h_{2,4}|^2, |h_{2,3}|^2)}. \end{aligned}$$

where K is to be specified. We set W to be independent of all these and have the same distribution as (V_1, V_2). We define X_1, and X_2 as follows

$$X_1 = V_1 + Z_1 + \bar{S}_{1,3} + \bar{S}_{1,4} + S'_1,$$
$$X_2 = V_2 + Z_2 + \bar{S}_{2,4} + \bar{S}_{1,3} + S'_2,$$

where

$$\bar{S}_{1,3} = \begin{cases} \widetilde{S}_{1,3}, & \text{if } |h_{2,4}| \geq |h_{1,4}|, \\ -\frac{|h_{2,4}|}{|h_{1,4}|e^{j\theta/2}} \widetilde{S}_{2,3}, & \text{otherwise.} \end{cases}$$
$$\bar{S}_{2,3} = \begin{cases} -\frac{|h_{1,4}|e^{j\theta/2}}{|h_{2,4}|} \widetilde{S}_{1,3} + S^{\perp}_{2,3}, & \text{if } |h_{2,4}| \geq |h_{1,4}|, \\ \widetilde{S}_{2,3} + S^{\perp}_{2,3}, & \text{otherwise.} \end{cases}$$
$$\bar{S}_{2,4} = \begin{cases} \widetilde{S}_{2,4}, & \text{if } |h_{1,3}| \geq |h_{2,3}|, \\ -\frac{|h_{1,3}|}{|h_{2,3}|e^{j\theta/2}} \widetilde{S}_{1,4}, & \text{otherwise.} \end{cases}$$
$$\bar{S}_{1,4} = \begin{cases} -\frac{|h_{2,3}|e^{j\theta/2}}{|h_{1,3}|} \widetilde{S}_{2,4} + S^{\perp}_{1,4}, & \text{if } |h_{1,3}| \geq |h_{2,3}|, \\ \widetilde{S}_{1,4} + S^{\perp}_{1,4}, & \text{otherwise.} \end{cases}$$

To satisfy the power constraint, it is enough to choose $K < 7$. Also, we define $S_1 = (\bar{S}_{2,3}, \bar{S}_{1,3}, S^{\perp}_{2,3})$, $S_2 = (\bar{S}_{1,4}, \bar{S}_{2,4}, S^{\perp}_{1,4})$, and

$$\bar{S}_1 = |h_{2,3}|e^{j\theta/2}\bar{S}_{2,3} + |h_{1,3}|\bar{S}_{1,3} + |h_{2,3}|e^{j\theta/2}S^{\perp}_{2,3},$$
$$\bar{S}_2 = |h_{1,4}|e^{j\theta/2}\bar{S}_{1,4} + |h_{2,4}|\bar{S}_{2,4} + |h_{1,4}|e^{j\theta/2}S^{\perp}_{1,4}.$$

Thus, S_1 and S_2 are independent of each other. Note that we defined $\bar{S}_{1,3}$ and $\bar{S}_{2,3}$ in the same way as we did in regime (iii). The destinations receive

$$Y_3 = |h_{1,3}|(V_1 + Z_1 + S'_1) + |h_{2,3}|e^{j\theta/2}(V_2 + Z_2 + S'_2) + \bar{S}_1 + |h_{1,3}|S^{\perp}_{1,4} + N_3,$$
$$Y_4 = |h_{2,4}|(V_2 + Z_2 + S'_2) + |h_{1,4}|e^{j\theta/2}(V_1 + Z_1 + S'_1) + \bar{S}_2 + |h_{2,4}|S^{\perp}_{2,3} + N_4.$$

It must be noted that $|h_{1,3}|S^{\perp}_{1,4}$ and $|h_{2,4}|S^{\perp}_{2,3}$ have variances at most unity (which is the variance of the noise). Since U_1 and U_2 are constants, we must set $r_{U_1} = r_{U_2} = 0$. The conditions on the non-negative rates are as follows.

$$r_{S_1} \leq \log\left(1 + |h_C|^2/(\max(1, |h_{1,3}|^2, |h_{1,4}|^2)K)\right),$$
$$r_{Z_1} + r_{S_1} \leq \log\left(1 + |h_C|^2/(\max(1, |h_{1,4}|^2)K) + |h_C|^2/(\max(1, |h_{1,3}|^2, |h_{1,4}|^2)K)\right)$$
$$r_{V_1} + r_{Z_1} + r_{S_1} \leq \log\left(1 + \frac{|h_C|^2}{K} + \frac{|h_C|^2}{\max(1, |h_{1,4}|^2)K} + \frac{|h_C|^2}{\max(1, |h_{1,3}|^2, |h_{1,4}|^2)K}\right),$$

$$r_{Z_1} \leq \log\left(1 + \frac{|h_{1,3}|^2/(\max(1, |h_{1,4}|^2)K)}{4/K + 1}\right),$$

$$r_{S_1} + r_{Z_1} \leq \begin{cases} \log\left(1 + \frac{\left(\left|\frac{|h_{1,3}|}{|h_{1,4}|}\right|^2 + \left|\frac{|h_{1,3}||h_{2,4}| - |h_{1,4}||h_{2,3}|e^{j\theta}}{|h_{2,4}|}\right|^2 + \left|\frac{|h_{2,3}|}{|h_{2,4}|}\right|^2\right)/K}{4/K+1}\right), & \text{if } |h_{2,4}| \geq |h_{1,4}| \\ \log\left(1 + \frac{\left(\left|\frac{|h_{1,3}|}{|h_{1,4}|}\right|^2 + \left|\frac{|h_{1,3}||h_{2,4}| - |h_{1,4}||h_{2,3}|e^{j\theta}}{|h_{1,4}|}\right|^2 + \left|\frac{|h_{2,3}|}{|h_{2,4}|}\right|^2\right)/K}{4/K+1}\right), & \text{otherwise} \end{cases}$$

$$\begin{aligned} &(r_{V_1} + r_{V_2}) + r_{S_1} + r_{Z_1} \\ &\leq \begin{cases} \log\left(1 + \frac{\left(|h_{1,3}|^2 + \left|\frac{|h_{1,3}|}{|h_{1,4}|}\right|^2 + |h_{2,3}|^2 + \left|\frac{|h_{1,3}||h_{2,4}| - |h_{1,4}||h_{2,3}|e^{j\theta}}{|h_{2,4}|}\right|^2 + \left|\frac{|h_{2,3}|}{|h_{2,4}|}\right|^2\right)/K}{4/K+1}\right), & \text{if } |h_{2,4}| \geq |h_{1,4}| \\ \log\left(1 + \frac{\left(|h_{1,3}|^2 + \left|\frac{|h_{1,3}|}{|h_{1,4}|}\right|^2 + |h_{2,3}|^2 + \left|\frac{|h_{1,3}||h_{2,4}| - |h_{1,4}||h_{2,3}|e^{j\theta}}{|h_{1,4}|}\right|^2 + \left|\frac{|h_{2,3}|}{|h_{2,4}|}\right|^2\right)/K}{4/K+1}\right), & \text{otherwise} \end{cases} \end{aligned}$$

and the corresponding conditions with 1 and 2 interchanged and 3 and 4 interchanged. On simplifictaion, this implies that non-negative rates which satisfy the following conditions are achievable.

$$\begin{aligned} r_{S_1} &\leq [n_C - \max(n_{1,3}, n_{1,4})]_+ - \log 7, \\ r_{V_1} + r_{Z_1} + r_{S_1} &\leq n_C - \log 7, \end{aligned}$$

$$r_{Z_1} \leq [n_{1,3} - n_{1,4}]_+ - \log 11,$$

$$r_{S_1} + r_{Z_1} \leq \begin{cases} \log\left(1 + \left|\frac{|h_{1,3}|}{|h_{1,4}|}\right|^2 + \left|\frac{|h_{1,3}||h_{2,4}| - |h_{1,4}||h_{2,3}|e^{j\theta}}{|h_{2,4}|}\right|^2 + \left|\frac{|h_{2,3}|}{|h_{2,4}|}\right|^2\right) - \log 11, & \text{if } |h_{2,4}| \geq |h_{1,4}| \\ \log\left(1 + \left|\frac{|h_{1,3}|}{|h_{1,4}|}\right|^2 + \left|\frac{|h_{1,3}||h_{2,4}| - |h_{1,4}||h_{2,3}|e^{j\theta}}{|h_{1,4}|}\right|^2 + \left|\frac{|h_{2,3}|}{|h_{2,4}|}\right|^2\right) - \log 11, & \text{otherwise,} \end{cases}$$

$$(r_{V_1} + r_{V_2}) + r_{S_1} + r_{Z_1} \leq \max(n_{1,3}, n_{2,3}) - \log 11,$$

and the corresponding conditions with 1 and 2 interchanged and 3 and 4 interchanged. Note that these conditions are identical to the ones for the linear deterministic case (up to a constant) except for the ones on $r_{S_1} + r_{Z_1}$ and $r_{S_2} + r_{Z_2}$. Fourier-Motzkin elimination shows that a sum-rate given by the minimum of $u_2'(n_C), u_3'(n_C), u_4'(n_C)$, and u_5' is achievable in regime (iv) up to a constant of 7 bits.

Thus, we can conclude that the upperbound is achievable within a constant gap (of 18 bits).

D Source cooperation: upperbounds

We prove four upperbounds to the sum-rate which will together imply the upperbounds in Theorems 1 and 2.

Upperbound 1: From Fano's inequality,

$$\begin{aligned}T(R_1+R_2-o(\epsilon)) &\leq I(M_1;Y_3^T)+I(M_2;Y_4^T)\\ &\leq I(M_1;Y_3^T,h_{1,4}^*(X_1^T),h_{1,2}^*(X_1^T),h_{2,1}^*(X_2^T))\\ &\quad + I(M_2;Y_4^T,h_{2,3}^*(X_2^T),h_{1,2}^*(X_1^T),h_{2,1}^*(X_2^T)).\end{aligned}$$

Note that we have provided additional signals to both the destinations – for instance, destination 4 now has access to $h_{1,4}^*(X_1^T)$, $h_{1,2}^*(X_1^T)$, and $h_{2,1}^*(X_2^T)$) in addition to its channel output Y_3^T. We will now upperbound the two symmetric mutual information terms above; only the first term is shown below.

$$\begin{aligned}&I(M_1;Y_3^T,h_{1,4}^*(X_1^T),h_{1,2}^*(X_1^T),h_{2,1}^*(X_2^T))\\ &\quad= H(h_{1,4}^*(X_1^T),h_{1,2}^*(X_1^T),h_{2,1}^*(X_2^T))+H(Y_3^T|h_{1,4}^*(X_1^T),h_{1,2}^*(X_1^T),h_{2,1}^*(X_2^T))\\ &\qquad - H(Y_3^T,h_{1,4}^*(X_1^T),h_{1,2}^*(X_1^T),h_{2,1}^*(X_2^T)|M_1)\\ &\quad\leq H(h_{1,4}^*(X_1^T),h_{1,2}^*(X_1^T),h_{2,1}^*(X_2^T))+H(Y_3^T|h_{1,4}^*(X_1^T),h_{2,1}^*(X_2^T))\\ &\qquad - H(Y_3^T,h_{1,4}^*(X_1^T),h_{1,2}^*(X_1^T),h_{2,1}^*(X_2^T)|M_1).\end{aligned}$$

We now derive an upperbound for the third term.

$$\begin{aligned}&H(Y_3^T,h_{1,4}^*(X_1^T),h_{1,2}^*(X_1^T),h_{2,1}^*(X_2^T)|M_1)\\ &=\sum_{t=1}^{T}H(Y_3(t),h_{1,4}^*(X_1(t)),h_{1,2}^*(X_1(t)),h_{2,1}^*(X_2(t))|Y_3^{t-1},h_{1,4}^*(X_1^{t-1}),h_{1,2}^*(X_1^{t-1}),h_{2,1}^*(X_2^{t-1}),M_1)\\ &\stackrel{(a)}{=}\sum_{t=1}^{T}H\Bigg(h_{2,3}^*(X_2(t)),h_{1,4}^*(X_1(t)),h_{1,2}^*(X_1(t)),h_{2,1}^*(X_2(t))\Bigg|X_1^t,h_{2,3}^*(X_2^{t-1}),h_{1,4}^*(X_1^{t-1}),\\ &\qquad\qquad h_{1,2}^*(X_1^{t-1}),h_{2,1}^*(X_2^{t-1}),M_1\Bigg)\\ &\stackrel{(b)}{=}\sum_{t=1}^{T}H(h_{1,4}^*(X_1(t))|X_1(t))+H(h_{1,2}^*(X_1(t))|X_1(t))\\ &\qquad + H(h_{2,3}^*(X_2(t)),h_{2,1}^*(X_2(t))|X_1^t,h_{2,3}^*(X_2^{t-1}),h_{1,4}^*(X_1^{t-1}),h_{1,2}^*(X_1^{t-1}),h_{2,1}^*(X_2^{t-1}),M_1)\\ &\stackrel{(c)}{=}\sum_{t=1}^{T}H(h_{1,4}^*(X_1(t))|X_1(t))+H(h_{1,2}^*(X_1(t))|X_1(t))\\ &\qquad + H(h_{2,3}^*(X_2(t)),h_{2,1}^*(X_2(t))|h_{2,3}^*(X_2^{t-1}),h_{1,2}^*(X_1^{t-1}),h_{2,1}^*(X_2^{t-1}))\\ &\stackrel{(d)}{=}\sum_{t=1}^{T}H(h_{1,4}^*(X_1(t))|X_1(t))+H(h_{1,2}^*(X_1(t))|X_1(t))\\ &\qquad + H(h_{2,3}^*(X_2(t)),h_{1,2}^*(X_1(t)),h_{2,1}^*(X_2(t))|h_{2,3}^*(X_2^{t-1}),h_{1,2}^*(X_1^{t-1}),h_{2,1}^*(X_2^{t-1}))\end{aligned}$$

$$
\begin{aligned}
&\quad - H(h^*_{1,2}(X_1(t))|h^*_{2,3}(X_2^t), h^*_{1,2}(X_1^{t-1}), h^*_{2,1}(X_2^t)) \\
&\overset{(e)}{\geq} \sum_{t=1}^{T} H(h^*_{1,4}(X_1(t))|X_1(t)) + H(h^*_{1,2}(X_1(t))|X_1(t)) \\
&\quad + H(h^*_{2,3}(X_2(t)), h^*_{1,2}(X_1(t)), h^*_{2,1}(X_2(t))|h^*_{2,3}(X_2^{t-1}), h^*_{1,2}(X_1^{t-1}), h^*_{2,1}(X_2^{t-1})) \\
&\quad - H(h^*_{1,2}(X_1(t))|h^*_{1,2}(X_1^{t-1})) \\
&= H(h^*_{2,3}(X_2^T), h^*_{1,2}(X_1^T), h^*_{2,1}(X_2^T)) - H(h^*_{1,2}(X_1^T)) \\
&\qquad + \left(\sum_{t=1}^{T} H(h^*_{1,4}(X_1(t))|X_1(t)) + H(h^*_{1,2}(X_1(t))|X_1(t))\right)
\end{aligned}
$$

where (a) follows from the following facts: (1) $Y_3(s) = h_{1,3}(X_1(s)) + h^*_{2,3}(X_2(s))$, (2) $h_{1,3}$ is a deterministic function, and (3) $X_1(s)$ is a deterministic function of $M_1, h^*_{2,1}(X_2^{t-1})$, for all $s \leq t$. Equality (b) can be seen from the channel model by which, conditioned on $X_1(t)$, the following three sets of random variables are independent: (1) $h^*_{1,4}(X_1(t))$, (2) $h^*_{1,2}(X_1(t))$, and (3) M_1 and all other quantities with indices up to and including t. The next equality (c) is a consequence of the fact that the following is a Markov chain

$$
(M_2, X_2^t, h^*_{2,3}(X_2^t)) - (h^*_{1,2}(X_1^{t-1}), h^*_{2,1}(X_2^{t-1})) - (M_1, X_1^t).
$$

This follows from (1) the channel $p_{h^*_{1,2}(X_1), h^*_{2,1}(X_2), h^*_{2,3}(X_2)|X_1,X_2} = p_{h^*_{1,2}(X_1)|X_1} p_{h^*_{2,1}(X_2)|X_1} p_{h^*_{2,3}(X_2)|X_2}$, (2) the independence of M_1 and M_2, and (3) the fact that the channel inputs depend deterministically on the messages at the respective sources and what these sources have received. Equality (d) is just the chain rule of entropy, and the inequality (e) follows from the non-negativity of mutual information.
Combining everything, we have

$$
\begin{aligned}
T(R_1 + R_2 - o(\epsilon)) \leq & \Bigg(\left\{ H(Y_3^T|h^*_{1,4}(X_1^T), h^*_{2,1}(X_2^T)) - \sum_{t=1}^{T} H(h^*_{2,3}(X_2(t))|X_2(t)) \right\} \\
& \qquad + \left\{ H(h^*_{1,2}(X_1^T)) - \sum_{t=1}^{T} H(h^*_{1,2}(X_1(t))|X_1(t)) \right\} \Bigg) \\
& + \Bigg(\left\{ H(Y_4^T|h^*_{2,3}(X_2^T), h^*_{1,2}(X_1^T)) - \sum_{t=1}^{T} H(h^*_{1,4}(X_1(t))|X_1(t)) \right\} \\
& \qquad + \left\{ H(h^*_{2,1}(X_2^T)) - \sum_{t=1}^{T} H(h^*_{2,1}(X_2(t))|X_2(t)) \right\} \Bigg).
\end{aligned}
$$

Linear deterministic model: Evaluating the bound directly gives us

$$
R_1 + R_2 \leq \max(n_{1,3} - n_{1,4}, n_{2,3} - n_C, 0) + n_C + \max(n_{2,4} - n_{2,3}, n_{1,4} - n_C, 0) + n_C.
$$

Gaussian model: Consider the first bracketed term.

$$
\begin{aligned}
\frac{1}{T}&\left\{H(Y_3^T|h_{1,4}^*(X_1^T),h_{2,1}^*(X_2^T))-\sum_{t=1}^{T}H(h_{2,3}^*(X_2(t))|X_2(t))\right\}\\
&\qquad\le\frac{1}{T}H(Y_3^T-h_{1,3}h_{1,4}^{-1}h_{1,4}^*(X_1^T)-h_{2,3}h_{2,1}^{-1}h_{2,3}^*(X_2^T))-H(N_3)\\
&\qquad=H(N_3-h_{1,3}h_{1,4}^{-1}N_4-h_{2,3}h_{2,1}^{-1}N_1)-H(N_3)\\
&\qquad\le\log\left(1+\left|\frac{h_{1,3}}{h_{1,4}}\right|^2+\left|\frac{h_{2,3}}{h_{2,1}}\right|^2\right).\\
\frac{1}{T}&\left\{H(h_{1,2}^*(X_1^T))-\sum_{t=1}^{T}H(h_{1,2}^*(X_1(t))|X_1(t))\right\}\\
&\qquad=\frac{1}{T}H(h_{1,2}X_1^T+N_2^T)-H(N_2)\\
&\qquad\le\log(1+|h_{1,2}|^2).
\end{aligned}
$$

If $|h_C|<1$, we do not subract the term $h_{2,3}h_{2,1}^{-1}h_{2,3}^*(X_2^T)$ while upperbounding the first term. This gives

$$
\begin{aligned}
\frac{1}{T}&\left\{H(Y_3^T|h_{1,4}^*(X_1^T),h_{2,1}^*(X_2^T))-\sum_{t=1}^{T}H(h_{2,3}^*(X_2(t))|X_2(t))\right\}\\
&\qquad\le\frac{1}{T}H(Y_3^T-h_{1,3}h_{1,4}^{-1}h_{1,4}^*(X_1^T))-H(N_3)\\
&\qquad=H(N_3^T-h_{1,3}h_{1,4}^{-1}N_4^T+h_{2,3}X_2^T)-H(N_3)\\
&\qquad\le\log\left(1+\left|\frac{h_{1,3}}{h_{1,4}}\right|^2+|h_{2,3}|^2\right).
\end{aligned}
$$

And similarly, we do not subtract the term $h_{1,3}h_{1,4}^{-1}h_{1,4}^*(X_1^T)$ if $|h_{1,4}|<1$ which gives the following upperbound for the first term.

$$
\begin{aligned}
\frac{1}{T}&\left\{H(Y_3^T|h_{1,4}^*(X_1^T),h_{2,1}^*(X_2^T))-\sum_{t=1}^{T}H(h_{2,3}^*(X_2(t))|X_2(t))\right\}\\
&\qquad\le\log\left(1+|h_{1,3}|^2+\left|\frac{h_{2,3}}{h_{2,1}}\right|^2\right).
\end{aligned}
$$

If both $|h_{1,4}|<1$ and $|h_C|<1$, we do not subtract either terms in which case we get the upperbound

$$
\begin{aligned}
\frac{1}{T}&\left\{H(Y_3^T|h_{1,4}^*(X_1^T),h_{2,1}^*(X_2^T))-\sum_{t=1}^{T}H(h_{2,3}^*(X_2(t))|X_2(t))\right\}\\
&\qquad\le\log\left(1+(|h_{1,3}|+|h_{2,3}|)^2\right).
\end{aligned}
$$

Similarly, upperbounding the second bracketed term and combining, we have

$$R_1 + R_2 \leq \log\left(1 + \left(\frac{|h_{1,3}|}{\max(1, |h_{1,4}|)} + \frac{|h_{2,3}|}{\max(1, |h_C|)}\right)^2\right)(1 + |h_C|^2)$$
$$+ \log\left(1 + \left(\frac{|h_{2,4}|}{\max(1, |h_{2,3}|)} + \frac{|h_{1,4}|}{\max(1, |h_C|)}\right)^2\right)(1 + |h_C|^2).$$

Upperbounds 2 and 3: We prove upperbound 2 below, the third one follows from a similar argument. From Fano's inequality,

$$\begin{aligned}
&T(R_1 + R_2 - o(\epsilon)) \\
&\leq I(M_1; Y_3^T) + I(M_2; Y_4^T) \\
&\leq I(M_1; Y_3^T) + I(M_2; Y_4^T, h_{2,3}^*(X_2^T), h_{1,2}^*(X_1^T), h_{2,1}^*(X_2^T), M_1) \\
&\leq I(M_1; Y_3^T) + I(M_2; Y_4^T, h_{2,3}^*(X_2^T), h_{1,2}^*(X_1^T), h_{2,1}^*(X_2^T)|M_1) \\
&\leq \sum_{t=1}^{T} I(M_1; Y_3(t)|Y_3^{t-1}) \\
&\qquad + I(M_2; Y_4(t), h_{2,3}^*(X_2(t)), h_{1,2}^*(X_1(t)), h_{2,1}^*(X_2(t))|Y_4^{t-1}, h_{2,3}^*(X_2^{t-1}), h_{1,2}^*(X_1^{t-1}), h_{2,1}^*(X_2^{t-1}), M_1).
\end{aligned}$$

Below, we upperbound these terms separately.

$$\begin{aligned}
I(M_1; Y_3(t)|Y_3^{t-1}) &= H(Y_3(t)|Y_3^{t-1}) - H(Y_3(t)|Y_3^{t-1}, M_1) \\
&\leq H(Y_3(t)) - H(Y_3(t)|Y_3^{t-1}, Y_4^{t-1}, h_{1,2}^*(X_1^{t-1}), h_{2,1}^*(X_2^{t-1}), M_1) \\
&\overset{\text{(a)}}{\leq} H(Y_3(t)) - H(h_{2,3}^*(X_2(t))|Y_4^{t-1}, h_{2,3}^*(X_2^{t-1}), h_{1,2}^*(X_1^{t-1}), h_{2,1}^*(X_2^{t-1}), M_1),
\end{aligned}$$

where (a) follows from the fact that $Y_3(t) = h_{1,3}(X_1(t)) + h_{2,3}^*(X_2(t))$, and, $h_{1,3}$ is a deterministic function and $X_1(t)$ is a deterministic function $f_{1,t}$ of $(M_1, h_{2,1}^*(X_2^{t-1}))$.

$$\begin{aligned}
I(M_2;&Y_4(t), h_{2,3}^*(X_2(t)), h_{1,2}^*(X_1(t)), h_{2,1}^*(X_2(t))|Y_4^{t-1}, h_{2,3}^*(X_2^{t-1}), h_{1,2}^*(X_1^{t-1}), h_{2,1}^*(X_2^{t-1}), M_1) \\
&= I(M_2; h_{2,3}^*(X_2(t))|Y_4^{t-1}, h_{2,3}^*(X_2^{t-1}), h_{1,2}^*(X_1^{t-1}), h_{2,1}^*(X_2^{t-1}), M_1) \\
&\quad + I(M_2; Y_4(t)|Y_4^{t-1}, h_{2,3}^*(X_2^{t}), h_{1,2}^*(X_1^{t-1}), h_{2,1}^*(X_2^{t-1}), M_1) \\
&\quad + I(M_2; h_{1,2}^*(X_1(t)), h_{2,1}^*(X_2(t))|Y_4^{t}, h_{2,3}^*(X_2^{t}), h_{1,2}^*(X_1^{t-1}), h_{2,1}^*(X_2^{t-1}), M_1)
\end{aligned}$$

We upperbound these three terms separately now.

$$\begin{aligned}
I(M_2; h_{2,3}^*(X_2(t))|&Y_4^{t-1}, h_{2,3}^*(X_2^{t-1}), h_{1,2}^*(X_1^{t-1}), h_{2,1}^*(X_2^{t-1}), M_1) \\
&= H(h_{2,3}^*(X_2(t))|Y_4^{t-1}, h_{2,3}^*(X_2^{t-1}), h_{1,2}^*(X_1^{t-1}), h_{2,1}^*(X_2^{t-1}), M_1) \\
&\quad - H(h_{2,3}^*(X_2(t))|Y_4^{t-1}, h_{2,3}^*(X_2^{t-1}), h_{1,2}^*(X_1^{t-1}), h_{2,1}^*(X_2^{t-1}), M_1, M_2). \\
I(M_2; Y_4(t)|Y_4^{t-1}, &h_{2,3}^*(X_2^{t}), h_{1,2}^*(X_1^{t-1}), h_{2,1}^*(X_2^{t-1}), M_1) \\
&= H(Y_4(t)|Y_4^{t-1}, h_{2,3}^*(X_2^{t}), h_{1,2}^*(X_1^{t-1}), h_{2,1}^*(X_2^{t-1}), M_1) \\
&\quad - H(Y_4(t)|Y_4^{t-1}, h_{2,3}^*(X_2^{t}), h_{1,2}^*(X_1^{t-1}), h_{2,1}^*(X_2^{t-1}), M_1, M_2) \\
&\overset{\text{(a)}}{\leq} H(Y_4(t)|X_1(t), h_{2,3}^*(X_2(t))) - H(Y_4(t)|X_1(t), X_2(t)),
\end{aligned}$$

where (a) follows from the channel model (memorylessness and independence of the noise processes at the different nodes) and the fact that $X_1(t)$ ($X_2(t)$, resp.) is a deterministic function $f_{1,t}$ ($f_{2,t}$, resp.) of $(M_1, h^*_{2,1}(X_2^{t-1}))$ ($(M_2, h^*_{1,2}(X_1^{t-1}))$, resp.).

$$\begin{aligned}
&I(M_2; h^*_{1,2}(X_1(t)), h^*_{2,1}(X_2(t))|Y_4^t, h^*_{2,3}(X_2^t), h^*_{1,2}(X_1^{t-1}), h^*_{2,1}(X_2^{t-1}), M_1) \\
&\quad \overset{(a)}{=} I(M_2; h^*_{2,1}(X_2(t))|Y_4^t, h^*_{2,3}(X_2^t), h^*_{1,2}(X_1^{t-1}), h^*_{2,1}(X_2^{t-1}), M_1) \\
&\quad = H(h^*_{2,1}(X_2(t))|Y_4^t, h^*_{2,3}(X_2^t), h^*_{1,2}(X_1^{t-1}), h^*_{2,1}(X_2^{t-1}), M_1) \\
&\qquad - H(h^*_{2,1}(X_2(t))|Y_4^t, h^*_{2,3}(X_2^t), h^*_{1,2}(X_1^{t-1}), h^*_{2,1}(X_2^{t-1}), M_1, M_2) \\
&\quad \overset{(b)}{\leq} H(h^*_{2,1}(X_2(t))|Y_4(t), h^*_{2,3}(X_2(t)), X_1(t)) - H(h^*_{2,1}(X_2(t))|X_2(t)),
\end{aligned}$$

where (a) follows from the channel model (memorylessness and the independence of the noise processes at the different nodes) and the fact that $X_1(t)$ is a deterministic function $f_{1,t}$ of $(M_1, h^*_{2,1}(X_2^{t-1}))$, and (b) follows from the facts above, its analogue for $X_2(t)$ and the channel model.
Combining everything, we have

$$\begin{aligned}
T(R_1 + R_2 - o(\epsilon)) \leq \sum_{t=1}^{T} \Big\{ & H(Y_3(t)|Y_3^{t-1}) \\
& - H(h^*_{2,3}(X_2(t))|Y_4^{t-1}, h^*_{2,3}(X_2^{t-1}), h^*_{1,2}(X_1^{t-1}), h^*_{2,1}(X_2^{t-1}), M_1, M_2) \Big\} \\
& + \Big\{ H(Y_4(t)|X_1(t), h^*_{2,3}(X_2(t))) - H(Y_4(t)|X_1(t), X_2(t)) \Big\} \\
& + \Big\{ H(h^*_{2,1}(X_2(t))|Y_4(t), h^*_{2,3}(X_2(t)), X_1(t)) - H(h^*_{2,1}(X_2(t))|X_2(t)) \Big\}.
\end{aligned}$$

Linear deterministic model: Suppose $n_C \leq \max(n_{2,4}, n_{2,3})$, then evaluating the upperbound,

$$R_1 + R_2 \leq \{\max(n_{1,3}, n_{2,3}) - 0\} + \{[n_{2,4} - n_{2,3}]_+ - 0\} + \{0 - 0\}.$$

Otherwise, if $n_C > \max(n_{2,4}, n_{2,3})$ we have either: (a) $n_C > n_{2,4} \geq n_{2,3}$ in which case, we lower the noise level at node 4 by $n_C - n_{2,4}$ level by defining $n'_{2,4} = n_{2,4} + (n_C - n_{2,4})$ and $n'_{1,4} = n_{1,4} + (n_C - n_{2,4})$. Now the upperbound evaluates to

$$R_1 + R_2 \leq \max(n_{1,3}, n_{2,3}) + n_C - n_{2,3} = n_C + [n_{1,3} - n_{2,3}]_+,$$

or (b) $n_C > n_{2,3} > n_{2,4}$ in which case, we lower the noise level at node 3 by $n_C - n_{2,3}$ by defining $n'_{1,3} = n_{1,3} + (n_C - n_{2,3})$ and $n'_{2,3} = n_{2,3} + (n_C - n_{2,3})$. Now the upperbound becomes

$$R_1 + R_2 \leq \max(n_{1,3}, n_{2,3}) + n_C - n_{2,3} = n_C + [n_{1,3} - n_{2,3}]_+.$$

Hence, without any conditions on $n_C, n_{2,4}, n_{2,3}$, we have

$$R_1 + R_2 \leq \max(n_{2,4}, n_{2,3}, n_C) + [n_{1,3} - n_{2,3}]_+.$$

By symmetry, we also have

$$R_1 + R_2 \le \max(n_{1,3}, n_{1,4}, n_C) + [n_{2,4} - n_{1,4}]_+.$$

Gaussian model: When $|h_C| \le \max(|h_{2,3}|, |h_{2,4}|)$,

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\Big\{H(Y_3(t)|Y_3^{t-1}) - H(h^*_{2,3}(X_2(t))|Y_4^{t-1}, h^*_{2,3}(X_2^{t-1}), h^*_{1,2}(X_1^{t-1}), h^*_{2,1}(X_2^{t-1}), M_1, M_2)\Big\} \\
&\qquad\qquad \le \frac{1}{T}\sum_{t=1}^{T} H(Y_3(t)) - H(N_3(t)) \\
&\qquad\qquad \le \log(1 + (|h_{1,3}| + |h_{2,3}|)^2).
\end{aligned}$$

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\Big\{H(Y_4(t)|X_1(t), h^*_{2,3}(X_2(t))) - H(Y_4(t)|X_1(t), X_2(t))\Big\} \\
&\qquad\qquad \le \frac{1}{T}\sum_{t=1}^{T} H(Y_4(t) - h_{1,4}X_1(t) - h_{2,4}h_{2,3}^{-1}h^*_{2,3}(X_2(t))) - H(N_4(t)) \\
&\qquad\qquad = \frac{1}{T}\sum_{t=1}^{T} H(N_4(t) - h_{2,4}h_{2,3}^{-1}N_3(t)) - H(N_4(t)) \\
&\qquad\qquad \le \log\left(1 + \left|\frac{h_{2,4}}{h_{2,3}}\right|^2\right).
\end{aligned}$$

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\Big\{H(h^*_{2,1}(X_2(t))|Y_4(t), h^*_{2,3}(X_2(t)), X_1(t)) - H(h^*_{2,1}(X_2(t))|X_2(t))\Big\} \\
&\qquad\qquad \le \min\Bigg(\frac{1}{T}\sum_{t=1}^{T} H(h^*_{2,1}(X_2(t)) - h_{2,1}h_{2,3}^{-1}h^*_{2,3}(X_2(t))) - H(N_1(t)), \\
&\qquad\qquad\qquad\qquad \frac{1}{T}\sum_{t=1}^{T} H(h^*_{2,1}(X_2(t)) - h_{2,1}h_{2,4}^{-1}(Y_4(t) - h_{1,4}X_1(t))) - H(N_1(t))\Bigg) \\
&\qquad\qquad = \min\Bigg(\frac{1}{T}\sum_{t=1}^{T} H(N_1(t) - h_{2,1}h_{2,3}^{-1}N_3(t)) - H(N_1(t)), \\
&\qquad\qquad\qquad\qquad \frac{1}{T}\sum_{t=1}^{T} H(N_1(t) - h_{2,1}h_{2,4}^{-1}N_4(t)) - H(N_1(t))\Bigg) \\
&\qquad\qquad \le \log\Big(1 + \min(|h_C|^2/|h_{2,3}|^2, |h_C|^2/|h_{2,4}|^2)\Big) \\
&\qquad\qquad \le 1.
\end{aligned}$$

Also, (when $|h_{2,3}| < 1$) we may upperbound the second term without subtracting $h_{2,4}h_{2,3}^{-1}h^*_{2,3}(X_2(t))$

$$\frac{1}{T}\sum_{t=1}^{T}\Big\{H(Y_4(t)|X_1(t), h^*_{2,3}(X_2(t))) - H(Y_4(t)|X_1(t), X_2(t))\Big\} \le \log\Big(1 + |h_{2,4}|^2\Big).$$

Thus, we have

$$R_1 + R_2 \le \log 2(1 + (|h_{1,3}| + |h_{2,3}|)^2) + \log\left(1 + \frac{|h_{2,4}|^2}{\max(1, |h_{2,3}|^2)}\right).$$

If $|h_C| > |h_{2,4}| \geq |h_{2,3}|$, we consider the following "enhanced" channel with channel coefficients indicated by primed quantities defined by

$$\frac{h'_{2,4}}{h_{2,4}} = \frac{|h_C|}{|h_{2,4}|},$$
$$\frac{h'_{1,4}}{h_{1,4}} = \frac{|h_C|}{|h_{2,4}|}.$$

It is easy to see that this is equivalent to lowering the noise variance at destination 4 from unity to $|h_{2,4}|^2/|h_C|^2$. Hence, an upperbound on the "enhanced" channel is also an upperbound for the original. Thus, we have

$$\begin{aligned}R_1 + R_2 &\leq \log 2(1 + (|h_{1,3}| + |h_{2,3}|)^2) + \log\left(1 + \frac{|h'_{2,4}|^2}{\max(1, |h_{2,3}|^2)}\right)\\ &= \log 2(1 + (|h_{1,3}| + |h_{2,3}|)^2) + \log\left(1 + \frac{|h_C|^2}{\max(1, |h_{2,3}|^2)}\right)\end{aligned}$$

Similarly, if $|h_C| > |h_{2,3}| \geq |h_{2,4}|$, we consider the following "enhanced" channel

$$\frac{h'_{1,3}}{h_{1,3}} = \frac{|h_C|}{|h_{2,3}|},$$
$$\frac{h'_{2,3}}{h_{2,3}} = \frac{|h_C|}{|h_{2,3}|},$$

which is equivalent to lowering the noise variance at destination 3 from unity to $|h_{2,3}|^2/|h_C|^2$. Then, for $|h_{2,3}| \geq 1$,

$$\begin{aligned}R_1 + R_2 &\leq \log 2(1 + (|h'_{1,3}| + |h'_{2,3}|)^2) + \log\left(1 + \frac{|h_{2,4}|^2}{\max\left(1, |h'_{2,3}|^2\right)}\right)\\ &\leq \log 2(1 + (|h_{1,3}| + |h_{2,3}|)^2) + \log\left(1 + \frac{|h_{2,4}|^2}{|h_{2,3}|^2}\right).\end{aligned}$$

For $|h_{2,3}| < 1$, we upperbound the three terms directly

$$\begin{aligned}&\frac{1}{T}\sum_{t=1}^{T}\left\{H(Y_3(t)|Y_3^{t-1}) - H(h^*_{2,3}(X_2(t))|Y_4^{t-1}, h^*_{2,3}(X_2^{t-1}), h^*_{1,2}(X_1^{t-1}), h^*_{2,1}(X_2^{t-1}), M_1, M_2)\right\}\\ &\qquad \leq \frac{1}{T}\sum_{t=1}^{T} H(Y_3(t)) - H(N_3(t))\\ &\qquad \leq \log(1 + (|h_{1,3}| + |h_{2,3}|)^2).\\ &\frac{1}{T}\sum_{t=1}^{T}\left\{H(Y_4(t)|X_1(t), h^*_{2,3}(X_2(t))) - H(Y_4(t)|X_1(t), X_2(t))\right\}\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{T}\sum_{t=1}^{T} H(Y_4(t) - h_{1,4}X_1(t)) - H(N_4(t)) \\
&= \frac{1}{T}\sum_{t=1}^{T} H(N_4(t) + h_{2,4}X_2(t)) - H(N_4(t)) \\
&\leq \log\left(1 + |h_{2,4}|^2\right) \\
&\leq \log 2.
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}&\left\{H(h^*_{2,1}(X_2(t))|Y_4(t), h^*_{2,3}(X_2(t)), X_1(t)) - H(h^*_{2,1}(X_2(t))|X_2(t))\right\} \\
&\leq \frac{1}{T}\sum_{t=1}^{T} H(h^*_{2,1}(X_2(t))) - H(N_1(t)) \\
&\leq \log\left(1 + |h_C|^2\right).
\end{aligned}
$$

Thus, in general, we have

$$
R_1 + R_2 \leq \log 2(1 + (|h_{1,3}| + |h_{2,3}|)^2) + \log\left(1 + \frac{\max(|h_C|^2, |h_{2,3}|^2, |h_{2,4}|^2)}{\max(1, |h_{2,3}|^2)}\right).
$$

Upperbound 4:

This is a simple cut-set upperbound [20] with nodes 1 and 4 on one side of the cut and nodes 2 and 3 on the other. It is easy to verify that

$$
\begin{aligned}
R_1 &\leq \max_{p_{X_1}} I(X_1; Y_3, Y_2), \\
R_2 &\leq \max_{p_{X_2}} I(X_2; Y_4, Y_1).
\end{aligned}
$$

Under the linear deterministic model, this translates to an upperbound on the sum-rate of

$$
R_1 + R_2 \leq \max(n_{1,3}, n_C) + \max(n_{2,4}, n_C),
$$

and for the Gaussian case, we get an upperbound of

$$
R_1 + R_2 \leq \log\left(1 + |h_{1,3}|^2 + |h_C|^2\right) + \log\left(1 + |h_{2,4}|^2 + |h_C|^2\right).
$$

Upperbound 5:

This is also a simple cut-set upperbound. Nodes 1 and 2 are on one side of the cut and nodes 3 and 4 are on the other. The resulting upperbound on the sum-rate is

$$
R_1 + R_2 \leq \max_{p_{X_1, X_2}} I(X_1; X_2; Y_3, Y_4).
$$

For the linear deterministic case, this gives

$$
R_1 + R_2 \leq \begin{cases} \max(n_{1,3} + n_{2,4}, n_{1,4} + n_{2,3}), & \text{if } n_{1,3} - n_{2,3} \neq n_{1,4} - n_{2,4}, \\ \max(n_{1,3}, n_{2,4}, n_{1,4}, n_{2,3}), & \text{otherwise,} \end{cases}
$$

and for the Gaussian case, using the fact the eigenvalues of the input ($[X_1, X_2]$) covariance matrix cannot exceed 2, we may upperbound the sum-rate by

$$R_1 + R_2 \leq \log \Bigg(1 + 2\left(|h_{1,3}|^2 + |h_{2,4}|^2 + |h_{1,4}|^2 + |h_{2,3}|^2 \right) \\ + 4\left(|h_{1,3}h_{2,4}|^2 + |h_{1,4}h_{2,3}|^2 - 2|h_{1,3}h_{2,4}h_{1,4}h_{2,3}|\cos\theta \right) \Bigg).$$

Upperbounds 1, 2, and 3 can be further tightened to obtain a smaller constant gap in Theorem 2 by (1) considering the correlation between the noise processes at the destinations, as well as (2) modifying the correlation of the Gaussian noise processes in the additional signals we provide to the destinations. Also, the correlation between the input signals can be explicitly accounted for instead of assuming the worst-case correlation at different stages as we do here. Upperbound 5 can be easily improved by choosing the optimal input covariance matrix.

E Upperbounds for the Gaussian interference channel with feedback

Below, we will show that (12) is an upperbound on the sum-rate. To see that this is the only biting upperbound in this regime, let us consider the achievability proof. The achievability proof in appendix C only depends on the marginals of the noises and hence holds without change for the feedback case. Hence, the sum-rate achieved in appendix C is achievable for the feedback problem as well. Moreover, for the symmetric case, u'_4 of (17) in the achievability proof (appendix C) is strictly subsumed by u'_5 of (18). Also, since for noiseless feedback, n_C and $n_{\min}$ as defined in appendix C are equal, u'_1 of (14) is subsumed by $\min(u'_2, u'_3, u'_4, u'_5)$ as we argued in that appendix (where u'_2 through u'_5 are defined in (15)-(18). For the symmetric channel $u'_2 = u'_3$. Evaluating u'_5 and u'_2 for the symmetric channel with noiseless feedback reveals that u'_5 is never smaller than u'_2 by more than 1 bit. Thus, the achievability proof is appendix C when applied to the symmetric channel with noiseless feedback achieves a sum-rate of u'_2 within a gap of at most 13 bits. Also, u'_2 is within a gap of at most 5 bits from (12). Thus, overall, the achievability proof is appendix C achieves (12) with a gap of at most 19 bits.

It only remains to show that (12) is an upperbound to the sum-rate.The line of argument is similar to the one in the proof of upperbound 2 for the cooperation case; the main difference is that the the genie does not provide $h^*_{1,2}(X_1^T)$ to destination 4. Let us define $h^*_{2,3}(X_2) = h_{2,3}X_2 + N_1$ as before. Starting from Fano's inequality, we write

$$\begin{aligned} T(R_1 + R_2 - o(\epsilon)) &\leq I(M_1; Y_3^T) + I(M_2; Y_4^T) \\ &\leq I(M_1; Y_3^T) + I(M_2; Y_4^T, h^*_{2,3}(X_2^T), M_1) \\ &\leq I(M_1; Y_3^T) + I(M_2; Y_4^T, h^*_{2,3}(X_2^T)|M_1) \end{aligned}$$

$$\leq \sum_{t=1}^{T} I(M_1; Y_3(t)|Y_3^{t-1}) + I(M_2; Y_4(t), h^*_{2,3}(X_2(t))|Y_4^{t-1}, h^*_{2,3}(X_2^{t-1}), M_1).$$

Below, we upperbound these terms separately.

$$\begin{aligned}
I(M_1; Y_3(t)|Y_3^{t-1}) &= H(Y_3(t)|Y_3^{t-1}) - H(Y_3(t)|Y_3^{t-1}, M_1) \\
&\leq H(Y_3(t)|Y_3^{t-1}) - H(Y_3(t)|Y_3^{t-1}, Y_4^{t-1}, M_1) \\
&\overset{(a)}{\leq} H(Y_3(t)|Y_3^{t-1}) - H(h^*_{2,3}(X_2(t))|Y_4^{t-1}, h^*_{2,3}(X_2^{t-1}), M_1),
\end{aligned}$$

where (a) follows from the fact that $Y_3(t) = h_{1,3}X_1(t) + h^*_{2,3}(X_2(t))$, and $X_1(t)$ is a deterministic function $f_{1,t}$ of (M_1).

$$\begin{aligned}
&I(M_2; Y_4(t), h^*_{2,3}(X_2(t))|Y_4^{t-1}, h^*_{2,3}(X_2^{t-1}), M_1) \\
&\qquad = I(M_2; h^*_{2,3}(X_2(t))|Y_4^{t-1}, h^*_{2,3}(X_2^{t-1}), M_1) + I(M_2; Y_4(t)|Y_4^{t-1}, h^*_{2,3}(X_2^{t}), M_1).
\end{aligned}$$

We upperbound the above two terms separately now.

$$\begin{aligned}
&I(M_2; h^*_{2,3}(X_2(t))|Y_4^{t-1}, h^*_{2,3}(X_2^{t-1}), M_1) \\
&\qquad = H(h^*_{2,3}(X_2(t))|Y_4^{t-1}, h^*_{2,3}(X_2^{t-1}), M_1) - H(h^*_{2,3}(X_2(t))|Y_4^{t-1}, h^*_{2,3}(X_2^{t-1}), M_1, M_2) \\
&\qquad \overset{(a)}{=} H(h^*_{2,3}(X_2(t))|Y_4^{t-1}, h^*_{2,3}(X_2^{t-1}), M_1) - H(N_3(t)), \\
&I(M_2; Y_4(t)|Y_4^{t-1}, h^*_{2,3}(X_2^{t}), M_1) \\
&\qquad = H(Y_4(t)|Y_4^{t-1}, h^*_{2,3}(X_2^{t}), M_1) - H(Y_4(t)|Y_4^{t-1}, h^*_{2,3}(X_2^{t}), M_1, M_2) \\
&\qquad \overset{(b)}{\leq} H(Y_4(t)|X_1(t), h^*_{2,3}(X_2(t))) - H(N_4(t)),
\end{aligned}$$

where (a) and (b) follow from (1) the channel model (memorylessness and independence of the noise processes at the two destination nodes, and $Y_4(t) = h_{2,4}X_2(t) + h^*_{1,4}(X_1(t)) = h_{2,4}X_2(t) + h_{1,4}X_1(t) + N_4(t)$), and (2) the fact that $X_1(t)$ and $X_2(t)$, resp., are deterministic functions of $(M_1, h^*_{2,3}(X_2^{t-1}))$ and $(M_2, h^*_{1,4}(X_1^{t-1}))$, resp. Combining everything, we have

$$T(R_1 + R_2 - o(\epsilon)) \leq \sum_{t=1}^{T} \left\{ H(Y_3(t)|Y_3^{t-1}) - H(N_3(t)) \right\} + \left\{ H(Y_4(t)|X_1(t), h^*_{2,3}(X_2(t))) - H(N_4(t)) \right\}.$$

The rest of the argument is exactly as in Appendix D.

F A Gaussian example

The upperbounds follow from Appendix D. We can show that the sum-rate is upperbounded by all of the following for the symmetric channel with $h_I = \sqrt{h_D}$.

$$\begin{aligned}
U_1 &= 2 \log\left(1 + 2h_D\right)\left(1 + h_C^2\right), \\
u_2 = u_3 &= \log 2 \left(1 + \left(h_D + \sqrt{h_D}\right)^2\right)\left(1 + \frac{\max(h_D^2, h_C^2)}{h_D}\right), \\
u_5 &= \log\left(1 + 4(h_D^2 + h_D) + 4\left(h_D^2 - h_D\right)^2\right).
\end{aligned}$$

Note that u_1 above is slightly stronger than the one on Theorem 2, but follows directly from the proof in appendix D when specialized to the symmetric channel with $h_I = \sqrt{h_D}$. Also, we have left out u_4 since this upperbound is not important for this channel. The above upperbounds imply that for any $\epsilon > 0$, the sum-capacity is upperbounded by $C + \epsilon$, for sufficiently large h_D.

The achievability again depends on different schemes depending on the regime. For $h_C \leq \sqrt{h_D}$, we apply Theorem 4(a). We choose $Z_1, Z_2, U_1, U_2, V_1, V_2$ as independent,zero-mean Gaussian auxiliary random variables with the following variances.

$$\begin{aligned}
\sigma^2_{Z_1} &= \sigma^2_{Z_2} = \frac{1}{h_D}, \\
\sigma^2_{U_1} &= \sigma^2_{U_2} = \frac{1}{h_C^2}, \text{ and} \\
\sigma^2_{V_1} &= \sigma^2_{V_2} = 1 - \frac{1}{h_D 2} - \frac{1}{h_C^2}.
\end{aligned}$$

W is independent of all these and has the same distribution as (V_1, V_2). X_1 and X_2 are as follows.

$$\begin{aligned}
X_1 &= V_1 + U_1 + Z_1, \\
X_2 &= V_2 + U_2 + Z_2.
\end{aligned}$$

Evaluating the expressions in Theorem 4(a) and simplifying using Fourier-Motzkin elimination, it can be shown that the upperbound is achievable in the regime $h_C \leq \sqrt{h_D}$ within a gap of 6 bits for sufficiently large h_D^2.

In the range of $\sqrt{h_D} \leq h_C \leq h_D$, we find that, for sufficiently large h_D, $u_2 = \log 2\left(1 + \left(h_D + \sqrt{h_D}\right)^2\right)(1 + h_D)$, which is independent of h_C in this regime, dominates the other bounds. Thus, the achievability in the regime $h_C \leq \sqrt{h_D}$ implies achievability in this regime as well.

For $h_C > h_D$, we apply Theorem 4(b) with the following choices for the auxiliary random variables. $S_1', S_2', Z_1, Z_2, S_1, S_2, V_1, V_2$ are independent, zero-mean Gaussian auxiliary random variables with the following variances.

$$\begin{aligned}
\sigma^2_{S_1'} &= \sigma^2_{S_2'} = \frac{1}{h_D^2}, \\
\sigma^2_{Z_1} &= \sigma^2_{Z_2} = \frac{1/2}{h_D}, \\
\sigma^2_{S_1} &= \sigma^2_{S_2} = \frac{1}{2\left(1 + \frac{1}{h_D}\right)}\left(1 - \frac{1}{h_D^2} - \frac{1/2}{h_D}\right), \text{ and} \\
\sigma^2_{V_1} &= \sigma^2_{V_2} = \frac{1}{2}\left(1 - \frac{1}{h_D^2} - \frac{1/2}{h_D}\right).
\end{aligned}$$

W is independent of all these and has the same distribution as (V_1, V_2), and U_1, U_2 are constants. X_1 and X_2 are as follows.

$$X_1 = V_1 + S_1 - \frac{h_X}{h_D} S_2 + Z_1 + S_1',$$
$$X_2 = V_2 + S_2 - \frac{h_X}{h_D} S_1 + Z_2 + S_2'.$$

$$Y_3 = h_D \left(V_1 + Z_1 + S_1'\right) + h_X \left(V_2 + Z_2 + S_2'\right) + \left(h_D - \frac{h_X^2}{h_D}\right) S_1 + N_3,$$
$$Y_4 = h_D \left(V_2 + Z_2 + S_2'\right) + h_X \left(V_1 + Z_1 + S_1'\right) + \left(h_D - \frac{h_X^2}{h_D}\right) S_2 + N_4,$$

Evaluating the expressions in Theorem 4(b) and simplifying using Fourier-Motzkin elimination, it can be shown that the upperbound is achievable in the regime $h_C > h_D$ within a gap of 5 bits for sufficiently large h_D.

References

[1] A. B. Carleial. "Interference Channels," *IEEE Trans. on Inform. Theory*, 24(1):60–70, 1978.

[2] R. Etkin, D. Tse, and H. Wang. "Gaussian Interference Channel Capacity to Within One Bit," *IEEE Trans. on Inform. Theory*, 54:5534–5562, 2008.

[3] V. R. Cadambe, S. A. Jafar, "Interference Alignment and the Degrees of Freedom for the K User Interference Channel," *IEEE Trans. on Inform. Theory*, 54(8):3425–3441, 2008.

[4] A. Ozgur, O. Leveque, and D. Tse, "Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks," *IEEE Trans. on Inform. Theory*, 53(10):3549–3572, 2007.

[5] T. S. Han and K. Kobayashi. "A New Achievable Rate Region for the Interference Channel," *IEEE Transactions on Information Theory*, 27(1):49–60, 1981.

[6] D. Tuninetti. "On Interference Channels with Generalized Feedback." In *Proc. IEEE Int. Symp. on Inform. Theory*, June 2007.

[7] Y. Cao and B. Chen. "An Achievable Region for Interference Channel with Conferencing." In *Proc. IEEE Int. Symp. on Inform. Theory*, June 2007.

[8] S. Yang and D. Tuninetti. "A New Achievable Region for Interference Channel with Generalized Feedback." In *Proc. Annual Conf. on Inform. Sciences and Systems*, March 2008.

[9] A. Høst-Madsen. Capacity Bounds for Cooperative Diversity "*IEEE Transactions on Information Theory*," 52(4):1522–1544, 2006.

[10] C. Suh and D. Tse. "Feedback Capacity of the Gaussian Interference Channel to Within 1.7075 Bits: the Symmetric Case." `http://arxiv.org/abs/0901.3580v1` 2009

[11] S. Mohajer, S. N. Diggavi, C. Fragouli, and D. Tse. "Transmission Techniques for Relay-Interference Networks." In *Proc. of Forty-Sixth Annual Allerton Conf. on Comm., Control, and Comp.*, September 2008.

[12] C. Thejaswi, A. Bennatan, J. Zhang, R. Calderbank, D. Cochran. "Rate-Achievability Strategies for Two-Hop Interference Flows." In *Proc. of Forty-Sixth Annual Allerton Conf. on Comm., Control, and Comp.*, September 2008.

[13] S. Avestimehr, S. N. Diggavi, and D. Tse. "Wireless Network Information Flow." In *Proc. Forty-Fifth Annual Allerton Conf. on Comm., Control, and Comp.*, September 2007.

[14] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, 25:572–584, September 1979.

[15] F. M. J. Willems, *Information theoretical Results for the Discrete Memoryless Multiple Access Channels.* Ph.D. dissertation, Katholieke Universiteit Leuven, Leuven, Belgium, October 1982.

[16] S. I. Gel'fand and M. S. Pinsker. Coding for Channel with Random Parameters. *Problems of Control and Information Theory*, 9(1):19–31, 1980.

[17] M. H. M Costa. Writing on Dirty Paper. *IEEE Transactions on Information Theory*, 29(3):439–441, May 1983.

[18] V. Prabhakaran and P. Viswanath. "Interference Channels with Destination Cooperation." Under preparation.

[19] A. Raja, V. Prabhakaran, and P. Viswanath. "Reversibility in Linear Deterministic Networks under Linear Coding." In *Proc. IEEE Information Theory Workshop*, Volvos, 2009.

[20] T. M. Cover and J. A. Thomas. *Elements of Information Theory,* Wiley-Interscience, 1991.