Interference Channels with Destination Cooperation

Vinod M. Prabhakaran and Pramod Viswanath

Abstract

Interference is a fundamental feature of the wireless channel. To better understand the role of cooperation in interference management, the two-user Gaussian interference channel where the destination nodes can cooperate by virtue of being able to *both* transmit and receive is studied. The sum-capacity of this channel is characterized up to a constant number of bits. The coding scheme employed builds up on the superposition scheme of Han and Kobayashi for two-user interference channels without cooperation. New upperbounds to the sum-capacity are also derived.

Index Terms

Cooperation, interference channel, relay channel, sum-capacity.

I. Introduction

Orthogonalization and treating interference as noise are the two most common ways of handling interference in practical wireless communication systems. However, it is well-known that better rates of operation may be achieved when the interfering systems are designed jointly as modelled in interference channels [4], [8]. Superposition coding and interference alignment have been shown to perform well for interference channels (where the sources only transmit and destinations only receive) [8], [3].

A further degree of cooperation is possible when the radios can both receive and transmit. Understanding the gains from this form of cooperation is the goal of this paper. We study a two-user Gaussian interference channel where the destinations are also equipped with transmit capabilities. This sets up the possibility of cooperation among the destination nodes. The cooperative links are over the same frequency band as the rest of the links. In this paper, we study the sum-rate when the destination nodes operate in full-duplex. The main result is a characterization of the sum-capacity within a constant number of bits. The constant we obtain is 43 bits, but we discuss how this gap could be improved. The two-user interference channel where the source radios have receive capabilities is explored in a companion paper [12]. A reversibility property exists between the two results [12, Section 7.1]. As we discuss there, one setting can be viewed as being obtained from the other by (a) reversing the roles of sources and destinations and (b) changing the directions of the links while preserving the channel coefficients. The sum-capacities of the two settings connected by this transformation are within a constant gap.

A scheme based on superposition coding of Cover [5] was proposed by Han and Kobayashi [9] for the two-user interference channel. Recently, Etkin, Tse, and Wang [8] showed that the scheme of Han and Kobayashi achieves the capacity of the two-user Gaussian interference channels to within one bit. The scheme of Han and Kobayashi 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 employs two additional types of messages which take advantage of cooperation:

The authors are with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801. Email: {vinodmp,pramodv}@illinois.edu

- Cooperative private messages are decoded by the destination to which is intended, but unlike private messages, they benefit from cooperation. The effect of cooperation is to ensure that these messages do not appear as interference at the destination to which they are not intended. This is achieved using a nulling scheme.
- Cooperative public messages are decoded by both destinations, and unlike public messages, they benefit from cooperation. The form of cooperation involves destinations exchanging messages with each other which carry information on their past observations. This has similarities to *compress-and-forward* schemes used in relay channels [6].

We also derive upperbounds on the sum-rate to show that these modes of cooperation are optimal up to a constant gap.

Related works include [10], [11], [17], [14]. The same model was studied in [10], but a constant-gap result was not obtained there. A two-stage, two-source interference network is studied in [11]. Two-user Gaussian interference channels with conferencing decoders (where the decoders communicate over an orthogonal conferencing channel) are studied in [17], [14]. One-sided interference channel with unidirectional conferencing between decoders is considered in [17], while [14] derives the capacity region of the two-user Gaussian interference channel with conferencing decoders within a gap of two bits.

II. Problem Statement

We consider the following model for destination cooperation (see Figure 1). At each discrete-time instant – indexed by $t = 1, 2, \dots$ – the source nodes 1 and 2 send out, respectively, $X_1[t]$ and $X_2[t]$ which belong to the set $\mathbb{C}$ of complex numbers. The destination nodes 3 and 4 can not only receive, but also transmit over the same channel. Let $X_3[t]$ and $X_4[t] \in \mathbb{C}$, respectively, denote what nodes 3 and 4 transmit at time t. Then the destination nodes receive

$$\begin{aligned} Y_3[t] &= g_{1,3}X_1[t] + g_{2,3}X_2[t] + g_{4,3}X_4[t] + N_3[t], \\ Y_4[t] &= g_{2,4}X_2[t] + g_{1,4}X_1[t] + g_{3,4}X_3[t] + N_4[t], \end{aligned}$$

where $N_3[t]$ and $N_4[t]$ are i.i.d. (over t), circularly symmetric, zero-mean, unit variance, complex Gaussian random variables which are independent of each other. The g's are constant, complex channel coefficients which are assumed to be known to all the nodes. We impose a natural causality constraint on the transmissions from the destination nodes – the transmissions from each destination is a deterministic function of everything it has received up to the previous time instant. *i.e.*,

$$X_k[t] = f_{k,t}(Y_k^{t-1}),\ k = 3, 4.$$

The source nodes 1 and 2 map their messages (which are assumed to be uniformly distributed over their alphabets and denoted by M_1 and M_2, respectively) to their channel inputs using deterministic encoding functions.

$$X_k[t] = f_{k,t}(M_k),\ k = 1, 2,$$

It is easy to see that, without loss of generality, we may consider a channel where the channel coefficients $g_{1,3}, g_{2,4}, g_{3,4}, g_{4,3}$ are replaced by their magnitudes $|g_{1,3}|, |g_{2,4}|, |g_{3,4}|, |g_{4,3}|$, and the channel coefficient $g_{1,4}$ is replaced by $|g_{1,4}|e^{j\theta/2}$ and $g_{2,3}$ is replaced by $|g_{2,3}|e^{j\theta/2}$, where $\theta \stackrel{\text{def}}{=} \arg(g_{1,4}) + \arg(g_{2,3}) - \arg(g_{1,3}) - \arg(g_{2,4})$. We will consider this channel. We will also assume that $|g_{3,4}| = |g_{4,3}| = g_C$, say, which models the reciprocity of the link between nodes 3 and 4. Further, we will consider unit power constraints which is without loss of generality when both destinations have the same power constraint. Thus, a blocklength-T codebook of rate (R_1, R_2) is a sequence of encoding functions, $f_{k,t},\ t = 1, 2, \dots, T$ as described above such that

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}|X_k[t]|^2\right] \leq 1,\ k = 1, 2, 3, 4,$$

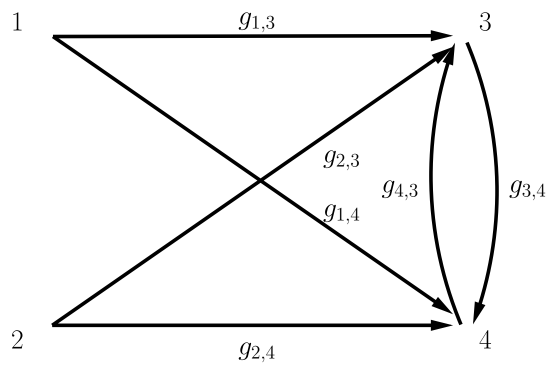


Fig. 1: Problem Setup

with message alphabets $\mathcal{M}_k = \{1, 2, \ldots, 2^{TR_k}\},\ k = 1, 2$ over which the messages M_k are uniformly distributed, and decoding functions $\tilde{f}_{k+2} : \mathcal{C}^T \to \mathcal{M}_k$. 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 $T \to \infty$,

$$\mathbb{P}\left(\tilde{f}_{k+2}(Y_{k+2}^T) \neq M_k\right) \to 0,\ k = 1, 2.$$

In this paper, we are interested in the largest $R_1 + R_2$ such that (R_1, R_2) is achievable.

We would also like to consider a linear deterministic model [1] for the above channel. In order to treat both models together, we will adopt the following notation: The destination nodes receive

$$Y_3[t] = g_{1,3}(X_1[t]) + g^*_{2,3}(X_2[t]) + g_{4,3}(X_4[t]),$$
$$Y_4[t] = g_{2,4}(X_2[t]) + g^*_{1,4}(X_1[t]) + g_{3,4}(X_3[t]),$$

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

$$X_k[t] = f_{k,t}(M_k),\ k = 1, 2,$$

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

$$X_k[t] = f_{k,t}(Y_k^{t-1}),\ k = 3, 4.$$

Gaussian case:

$$g_{1,3}(X_1) = g_{1,3}X_1,$$
$$g_{2,4}(X_2) = g_{2,4}X_2,$$
$$g^*_{2,3}(X_1) = g_{2,3}X_4 + N_3,$$
$$g^*_{1,4}(X_1) = g_{1,4}X_1 + N_4,$$
$$g_{3,4}(X_3) = g_{3,4}X_3,$$
$$g_{4,3}(X_4) = g_{4,3}X_4,$$

Note that g^*'s denote randomized maps while g's are deterministic.

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

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

[1]In the sequel, we will also (ab)use notation to denote $\mathbf{S}^0 \stackrel{\text{def}}{=} I$, the $n \times n$ identity matrix, and $\mathbf{S}^{-k} \stackrel{\text{def}}{=} \text{transpose}\left(\mathbf{S}^k\right), k \geq 0$.

We define

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

Further, to model the reciprocity of the links between the two receivers, we set $n_{3,4} = n_{4,3} = n_C$. The set of achievable (R_1, R_2) are defined as in the Gaussian case.

III. Main Results

We will first state our main result on the sum-rates of the channels presented in the previous section. Then we illustrate the gains resulting from cooperation using an example.

A. Sum-rate Characterization

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

$$u_1 = \max(n_{1,3} - n_{1,4} + n_C, n_{2,3}, n_C) + \max(n_{2,4} - n_{2,3} + n_C, n_{1,4}, n_C), \tag{1}$$
$$u_2 = \max(n_{2,4}, n_{2,3}) + \left(\max(n_{1,3}, n_{2,3}, n_C) - n_{2,3}\right), \tag{2}$$
$$u_3 = \max(n_{1,3}, n_{1,4}) + \left(\max(n_{2,4}, n_{1,4}, n_C) - n_{1,4}\right), \tag{3}$$
$$u_4 = \max(n_{1,3}, n_C) + \max(n_{2,4}, n_C), \text{ and} \tag{4}$$
$$u_5 = \begin{cases} \max(n_{1,3} + n_{2,4}, n_{1,4} + n_{2,3}), & \text{if } n_{1,3} - n_{2,3} \neq n_{1,4} - n_{2,4}, \\ \max(n_{1,3}, n_{2,4}, n_{1,4}, n_{2,3}), & \text{otherwise.} \end{cases} \tag{5}$$

The condition $n_{1,3} - n_{2,3} = n_{1,4} - n_{2,4}$ refers to a degenerate case where one of the destinations is degraded with respect to the other, *i.e.*, the signal that one of the destinations receives is part of what the other receives. We prove the achievability in appendix C and the upperbound in appendix E.

Theorem 2: Gaussian case. The sum-capacity of the Gaussian channel with destination cooperation is at most the minimum of the following five quantities and a sum-rate can be achieved within a gap of at most 43 bits of this minimum[2].

$$\begin{aligned} u_1 = &\begin{cases} \log\left(1 + \left(|g_{2,3}| + |g_C| + \left|\frac{g_{1,3}g_C}{g_{1,4}}\right|\right)^2 + \left|\frac{g_{1,3}}{g_{1,4}}\right|^2\right), & \text{if } |g_{1,4}| > \max(1, |g_C|) \\ \log\left(1 + (|g_{2,3}| + |g_C| + |g_{1,3}|)^2\right), & \text{otherwise} \end{cases} \\ &+ \begin{cases} \log\left(1 + \left(|g_{1,4}| + |g_C| + \left|\frac{g_{2,4}g_C}{g_{2,3}}\right|\right)^2 + \left|\frac{g_{2,4}}{g_{2,3}}\right|^2\right), & \text{if } |g_{2,3}| > \max(1, |g_C|) \\ \log\left(1 + (|g_{1,4}| + |g_C| + |g_{2,4}|)^2\right), & \text{otherwise,} \end{cases} \end{aligned} \tag{6}$$

$$u_2 = \log(1 + (|g_{1,3}| + |g_{2,3}| + |g_C|)^2) + \log\left(1 + \frac{|g_{2,4}|^2}{\max(1, |g_{2,3}|^2)}\right), \tag{7}$$

$$u_3 = \log(1 + (|g_{2,4}| + |g_{1,4}| + |g_C|)^2) + \log\left(1 + \frac{|g_{1,3}|^2}{\max(1, |g_{1,4}|^2)}\right), \tag{8}$$

$$u_4 = \log\left(1 + (|g_{1,3}| + |g_C|)^2\right) + \log\left(1 + (|g_{2,4}| + |g_C|)^2\right), \tag{9}$$

[2]All logarithms in this paper are to the base 2.

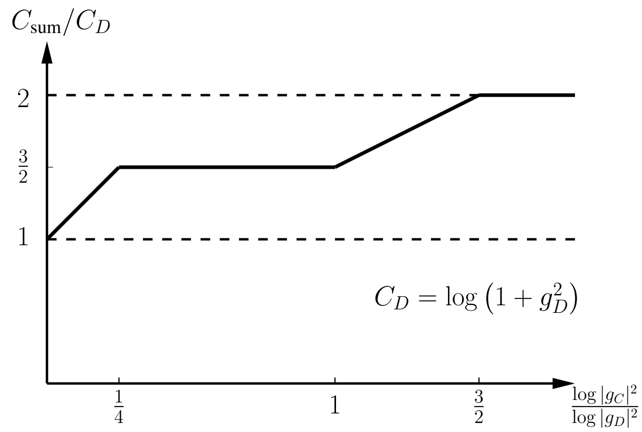


Fig. 2: Normalized sum-capacity of the symmetric interference channel with $g_I = \sqrt{g_D}$ under source cooperation in the limit of $g_D \to \infty$ keeping $\log|g_I|^2/\log|g_D|^2$ and $\log|g_C|^2/\log|g_D|^2$ fixed.

$$u_5 = \log\Bigg(1 + 2\left(|g_{1,3}|^2 + |g_{2,4}|^2 + |g_{1,4}|^2 + |g_{2,3}|^2\right) + 4\left(|g_{1,3}g_{2,4}|^2 + |g_{1,4}g_{2,3}|^2 - 2|g_{1,3}g_{2,4}g_{1,4}g_{2,3}|\cos\theta\right)\Bigg). \tag{10}$$

The achievability is proved in appendix D and the upperbound in appendix E.

B. *Gains from cooperation*

Let us consider the following two-user Gaussian interference channel to see the gains from cooperation: $|g_{1,3}| = |g_{2,4}| = g_D$, $|g_{1,4}| = |g_{2,3}| = g_I = \sqrt{g_D}$, and arbitrary θ. In Fig. 2, we plot the upperbound on the sum-rate from Theorem 2 normalized by the capacity of the direct link, as a function of $\log|g_C|^2/\log|g_D|^2$, in the limit of $|g_D| \to \infty$ while keeping the ratios $\log|g_C|^2/\log|g_D|^2$ and $\log|g_I|^2/\log|g_D|^2$ constant. Since this upperbound is achievable within a constant gap, this plot is also that of the sum-capacity in this limit. There are three distinct regimes.

- $\log|g_C|^2/\log|g_D|^2 \leq 1/2$. 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. As will become clear in the sequel, the benefits of cooperation in this regime come from messaging using cooperative private messages.
- $1/2 < \log|g_C|^2/\log|g_D|^2 \leq 1$. The linear gain saturates when the cooperation link strength is half the direct link strength. No further gains are available until the cooperation link is as strong as the direct link.
- $1 < \log|g_C|^2/\log|g_D|^2 \leq 3/2$. 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. The cooperative gains in this regime result from using cooperative public messages.

IV. Coding schemes: Illustrative examples

In this section we will present two linear deterministic channel examples to illustrate our achievable scheme. The examples have been chosen such that essentially signal processing schemes (*i.e.*, schemes which involve no coding) can achieve the sum-capacity. Our achievable scheme for the general problem relies on the basic intuition illustrated here.

Han and Kobayashi's achievable scheme for interference channels without cooperation involves two kinds of signals: (a) public signals which are decoded by all the destinations and (b) private signals which are decoded only by the destinations to which they are intended and treated as noise by all the other destinations. Our coding scheme involves two other forms of signals. The first example introduces a third type of signal called the cooperative private signal. It also shows that the private signal itself comes in two varieties now. Example 2 introduces a fourth type of signal called the cooperative public signal.

A. Example 1

Consider the symmetric linear deterministic 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 $n_D = 5$ and $n_I = 2$. When source cooperation is absent, *i.e.*, $n_C = 0$, the sum capacity is 6. With a cooperative link of $n_C = 1$, the sum-capacity turns out to be 8. A scheme which achieves this is as follows: The sources transmit

$$x_1[t] = \begin{pmatrix} u_1[t] \\ s_1[t] \\ z_{\uparrow 1}[t] \\ s_1[t+1] \\ z_{\downarrow 1}[t] \end{pmatrix} \text{ and } x_2[t] = \begin{pmatrix} u_2[t] \\ s_2[t] \\ z_{\uparrow 2}[t] \\ s_2[t+1] \\ z_{\downarrow 2}[t] \end{pmatrix},$$

where $s_k(1) = s_k(T+1) = 0,\ k = 1, 2$. The destination nodes will receive the following signals

$$y_3[t] = \begin{pmatrix} u_1[t] \\ s_1[t] \\ z_{\uparrow 1}[t] \\ s_1[t+1] \\ z_{\downarrow 1}[t] \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ u_2[t] \\ s_2[t] \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -(s_2[t] + u_1[t-1]) \end{pmatrix}$$

$$= \begin{pmatrix} u_1[t] \\ s_1[t] \\ z_{\uparrow 1}[t] \\ s_1[t+1] + u_2[t] \\ z_{\downarrow 1}[t] - u_1[t-1] \end{pmatrix}, \text{ and similarly}$$

$$y_4[t] = \begin{pmatrix} u_2[t] \\ s_2[t] \\ z_{\uparrow 2}[t] \\ s_2[t+1] + u_1[t] \\ z_{\downarrow 2}[t] - u_2[t-1] \end{pmatrix},$$

if they transmit at time t

$$x_3[t] = \begin{pmatrix} -(s_1[t] + u_2[t-1]) \\ -(z_{\downarrow 1}[t-1] - u_1[t-1]) \\ 0 \\ 0 \\ 0 \end{pmatrix} \text{ and } x_4[t] = \begin{pmatrix} -(s_2[t] + u_1[t-1]) \\ -(z_{\downarrow 2}[t-1] - u_2[t-1]) \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Clearly, the destinations nodes may transmit the above signals in this example and destination 3 can recover the signals $\{(u_1[t], s_1[t], z_{\uparrow 1}[t], z_{\downarrow 1}[t], , u_2[t]) : t = 1, 2, \ldots T\}$ while destination 4 can recover the signals $\{u_2[t], s_2[t], z_{\uparrow 2}[t], z_{\downarrow 2}[t], u_1[t] : t = 1, 2, \ldots T\}$. This implies that a rate $R_1 = 4$ (and similarly $R_2 = 4$) is achievable.

But we now interpret the steps involved at the destinations with an additional restriction that they can access lower levels of their observations only if the signals contributing to the higher levels have been recovered. As we will see in the next section, this restriction allows us to extend this scheme to a more general scheme which also works in the Gaussian case. The rough intuition is that the scheme will treat the signals in the lower levels, which represent lower power levels in the Gaussian context, as noise while processing the higher levels. And if a higher level has not been decoded, then the lower levels are essentially "drowned out" by the higher power of the undecoded signals.

At the end of time t, destination 3 performs a preliminary decoding (*phase-1 decoding*) where it recovers $u_1[t]$, $s_1[t]$, $z_{\uparrow 1}[t]$ in that order reading them off from the top levels of $y_3[t]$. Then, it removes the effect of these signals from $y_3[t]$ to obtain the residual signal

$$\tilde{y}_3[t] = \begin{pmatrix} 0 \\ 0 \\ 0 \\ s_1[t+1] + u_2[t] \\ z_{\downarrow 1}[t] - u_1[t-1] \end{pmatrix}.$$

This residual signal is multiplied by -1 and shifted upwards (equivalent to a scaling in the Gaussian case) and transmitted at time $t+1$ as $x_3[t+1]$. Note that destination 3 at this point has not decoded all the signals from $y[t]$. But phase-1 decoding of $y_3[t]$ allows it to construct $x_3[t+1]$ at the end of time t. Destination 3 has to wait till at least after time $t+1$ in order to recover $u_2[t]$ from the fourth level of its observation at time t, namely $s_1[t+1] + u_2[t]$. This is because neither $u_2[t]$ nor $s_1[t+1]$ is available separately until time $t+1$ when the latter is recovered. The destination then removes the effect of this level. This allows it to recover $z_{\downarrow 1}[t]$ since $u_1[t-1]$ is already recovered. Recovering In generalizing this example, our achievable scheme will adopt a similar approach as explained in the next section.

The above scheme involved three types of signals (Figure 3).

- Public signals. u_1 and u_2 in the scheme above are recovered by both the destinations. This signal is similar to the public message of Han and Kobayashi's superposition scheme for two-user interference channels.
- Private signals. $z_{\uparrow 1}$, $z_{\downarrow 1}$, $z_{\uparrow 2}$, and $z_{\downarrow 2}$ are recovered only by the destinations to which they are intended. This signal is again similar to the private message of Han and Kobayashi's scheme. We want to further divide these signals into two types in the context of our scheme:
  - $\uparrow$ private signals. $z_{\uparrow 1}[t]$ and $z_{\uparrow 2}[t]$ are recovered by the respective destinations on observing $y_3[t]$ and $y_4[t]$ respectively. The destinations remove the effects of these signals from the observations in order to prepare the residual signal which is transmitted at time $t+1$. In this sense these signals are treated differently by the destinations compared to the signals $z_{\downarrow 1}[t]$ and $z_{\downarrow 2}[t]$.
  - $\downarrow$ private signals. $z_{\downarrow 1}[t]$ and $z_{\downarrow 2}[t]$, unlike the above $\uparrow$ private signals, are recovered by the respective destinations with a certain time-lag.
- Cooperative private signals. s_1 and s_2 are also recovered only by the destinations to which they are intended. However, their effect at the destination where they could act as interference is nulled out by the actions of the other destination. In this sense, these signals benefit cooperation.

In this example, sources made use of the relaying capabilities of destination nodes to beamform and null-out part of the interference. This idea has similarities to a technique independently arrived at in [11] which the authors call *interference neutralization*.

B. Example 2

Consider the symmetric linear deterministic channel with $n_D = 2, n_I = 1, n_C = 3$.

$$x_1[t] = \begin{pmatrix} u_1[t] \\ z_1[t] \\ 0 \end{pmatrix} \text{ and } x_2[t] = \begin{pmatrix} u_2[t] \\ z_2[t] \\ 0 \end{pmatrix},$$

where $u_k(T) = u_k(0) = 0,\ k = 1, 2$. The destinations transmit

$$x_3[t] = \begin{pmatrix} u_1[t-1] \\ 0 \\ 0 \end{pmatrix} \text{ and } x_4[t] = \begin{pmatrix} u_2[t-1] \\ 0 \\ 0 \end{pmatrix}$$

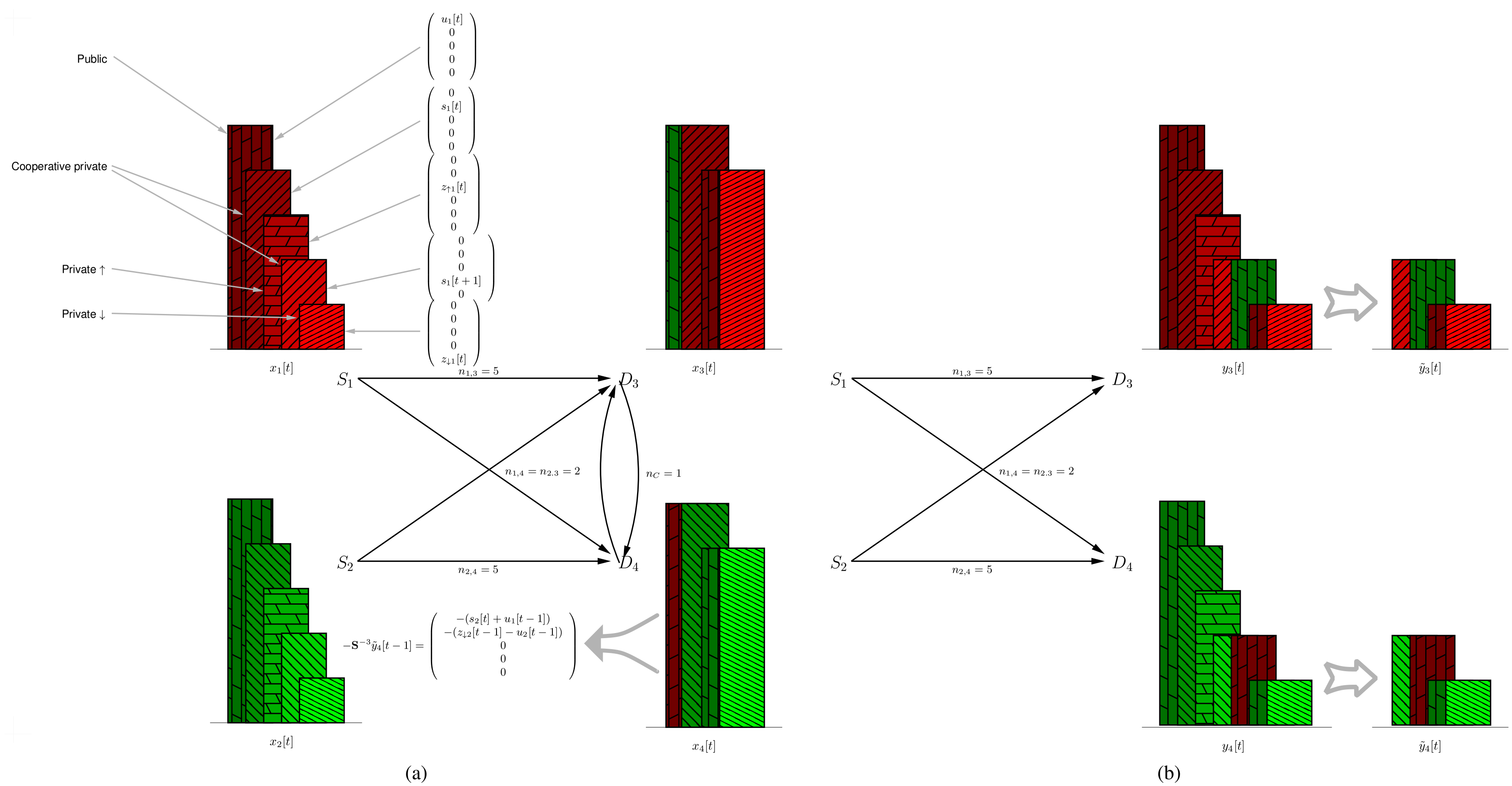


Fig. 3: Example 1. (a) Transmitted signals. (b) Received signals and residual signals after phase-1 decoding. Notice that the destinations transmit at time t a scaled up version of their residual signals from time $t-1$. This results in a nulling-out of the interference caused by the cooperative private signal from the sources.

which is possible since the destinations receive

$$y_3[t] = \begin{pmatrix} u_2[t-1] \\ u_1[t] \\ z_1[t] + u_2[t] \end{pmatrix} \text{ and } y_4[t] = \begin{pmatrix} u_1[t-1] \\ u_2[t] \\ z_2[t] + u_1[t] \end{pmatrix}.$$

Thus, destination 3 now recovers $\{(u_1[t], z_1[t], u_2[t]) : t = 1, 2, \ldots, T\}$ and destination 4 $\{u_2[t], z_2[t], u_1[t] : t = 1, 2, \ldots, T\}$ leading to rates $R_1 = 2, R_2 = 2$. Here the destinations helped each other decode part of the interference.

We have two types of signals in this example (Figure 4). Signals z_1 and z_2 are private signals in that they are decoded only by the destination to which it is intended and they did not benefit from cooperation. Signals u_1 and u_2 constitute a new kind of signal.

- Cooperative public signals. These signals are decoded by both destinations. Their transmission benefited from cooperation. In this example, destination 3 aided destination 4 in recovering the signal u_1. It should be noted that, if destination 3 at time t, instead of sending $u_1[t-1]$, sent an appropriate linear projection of its received vector truncated at the top two levels (so as not include any of the private signal in this linear projection), the scheme would continue to work. In generalizing this example, our achievable scheme will adopt a similar approach.

V. Coding schemes

We use two coding schemes to prove the achievability of our main results. These schemes cater to different regimes of the cooperation link. When the cooperation link is weaker than the other links, we use a scheme which extends the intuition from Example 1 of the previous section. The basic intuition is that sources may make use of the relaying capabilities of destination nodes to beamform and null-out part of the interference. When the cooperation link is stronger than the direct links, we use a coding scheme similar to Example 2. This is a form of compress-and-forward scheme where the destinations quantize their observations and convey these to each other over the cooperation link. We roughly sketch the schemes in this section and leave the formal proofs to the appendices. In the appendix (C and D), we also show that achievability in these two regimes imply achievability for the entire range of cooperative link strengths.

A. *Cooperation link weaker than other links*

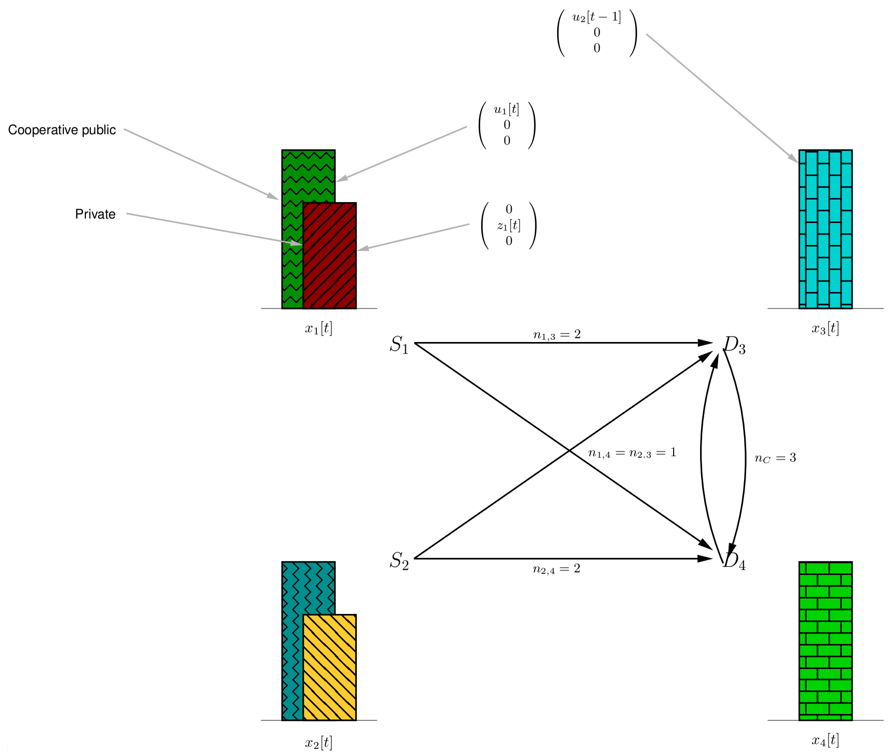
This is a block-Markov coding scheme which generalizes Example 1. The exact details are provided in Appendices C and D, we only provide a rough sketch here in the context of the Gaussian channel. For each block-j, the different types of messages, namely, public, $\uparrow$ private, $\downarrow$ private, and cooperative-private messages are coded using independent codebooks. For source-k, where $k = 1, 2$, let us denote these codewords, respectively, by $c_{u_k}^{(j)}$, $c_{\uparrow z_k}^{(j)}$, $c_{\downarrow z_k}^{(j)}$, and $c_{s_k}^{(j)}$, respectively. Source-k transmits a superposition of the $c_{u_k}^{(j)}$, $c_{\uparrow z_k}^{(j)}$, $c_{\downarrow z_k}^{(j)}$ codewords and a signal $f_{s_k}^{(j)}$ (defined below) which depends on the current and future c_{u_k} codewords, *i.e.*, on $\left\{ c_{u_k}^{(i)} : \ i = j, j+1, \ldots, J \right\}$, where J is the total number of blocks.

$$f_{s_k}^{(j)} \stackrel{\text{def}}{=} c_{s_k}^{(j)} + A_k c_{s_k}^{(j+1)} + A_k^2 c_{s_k}^{(j+2)} + A_k^{J-j} c_{s_k}^{(J)},$$

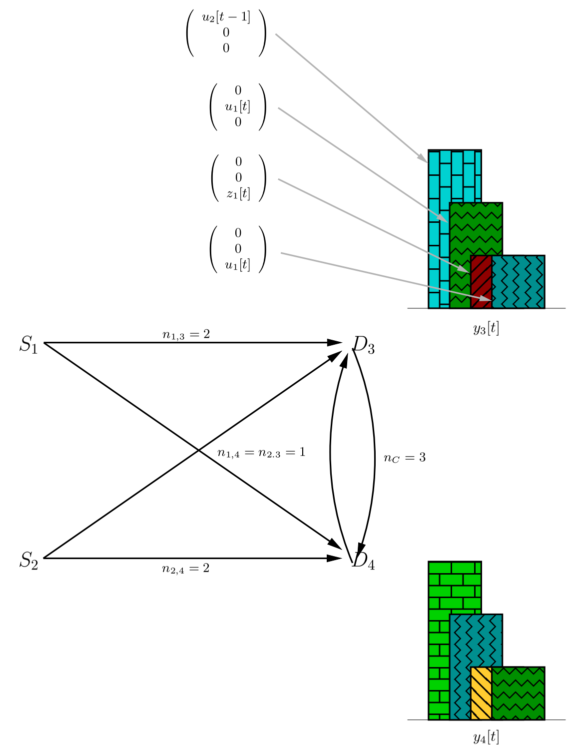
where A_k will be defined later. Note that f_{s_k} can be thought of as the effect of passing the signal c_{s_k} through an anti-causal filter (which acts on blocks rather than individual samples) with transfer function

$$\frac{1}{1 - A_k z^{-1}}.$$

Destinations decode in two phases. At the end of the j-th block, destination 3 decodes using successive cancellation decoding, the codewords $c_{u_1}^{(j)}$, $c_{s_1}^{(j)}$, and $c_{\uparrow z_1}^{(j)}$ in that order while treating all the other undecoded codewords and interference as noise. This constitutes the first phase of decoding. The residual signal from



(a)



(b)

Fig. 4: Example 2. (a) Transmitted signals. (b) Received signals.

which the contribution of the decoded signals is removed is scaled by a factor A_3 and transmitted as the X_3 signal in block-$j+1$. Destination 4 also performs its first phase of decoding in a similar manner and obtains its transmit signal X_4 for block-$j+1$. Assuming that all the decoding in previous blocks were successful, the signal at destination 3 in block-j is a linear combination of signals involving $\{c_{u_1}^{(i)} :\ i = j-1, j\}$, $\{c_{u_2}^{(j)}\}$, $\{c_{s_1}^{(i)} :\ i = j-1, j, \ldots, J\}$, $\{c_{s_2}^{(i)} :\ i = j, j+1, \ldots, J\}$, $\{c_{z_{\uparrow 1}}^{(i)} :\ i = j-1, j\}$, $\{c_{z_{\downarrow 1}}^{(i)} :\ i = 1, 2, \ldots, j\}$, $\{c_{z_{\uparrow 2}}^{(j)}\}$, and $\{c_{z_{\downarrow 2}}^{(i)} :\ i = 1, 2, \ldots, j\}$. Of these, let us consider the contribution of c_{s_2}. We first define the notation $\{f\}_{j-1}$ to denote f from which all terms which depend on $\{c_{s_2}^{(i)} :\ i = 1, 2, \ldots, j-1\}$ have been removed. For instance, for $i \geq 0$,

$$\begin{aligned}\left\{f_{s_2}^{(j-i)}\right\}_{j-1} &= \left\{c_{s_2}^{(j-i)} + A_2 c_{s_2}^{(j-i+1)} + A_2^2 c_{s_2}^{(j-i+2)} + \ldots + A_2^{J-j+i} c_{s_2}^{(J)}\right\}_{j-1} \\ &= A_2^i c_{s_2}^{(j)} + A_2^{i+1} c_{s_2}^{(j+1)} + \ldots + A_2^{J-j+i} c_{s_2}^{(J)} \\ &= A_2^i f_{s_2}^{(j)}.\end{aligned}$$

Then, the contribution of c_{s_2} in the signal received at destination 3 in block-j is[3]

$$\begin{aligned}&g_{2,3} f_{s_2}^{(j)} + g_C A_4 \Big\{ g_{2,4} f_{s_2}^{(j-1)} + g_C A_3 \left(g_{2,3} f_{s_2}^{(j-2)} + g_C A_4 g_{2,4} f_{s_2}^{(j-3)}\right) \\ &\qquad\qquad + g_C A_3 g_C A_4 g_C A_3 \left(g_{2,3} f_{s_2}^{(j-4)} + g_C A_4 g_{2,4} f_{s_2}^{(j-5)}\right) + \ldots \Big\}_{j-1} \\ &= \left(g_{2,3} f_{s_2}^{(j)} + g_C A_4 g_{2,4} \left\{f_{s_2}^{(j-1)}\right\}_{j-1}\right) + (g_c A_4 g_C A_3)\left(g_{2,3}\left\{f_{s_2}^{(j-2)}\right\}_{j-1} + g_C A_4 g_{2,4}\left\{f_{s_2}^{(j-3)}\right\}_{j-1}\right) \\ &\qquad + (g_C A_4 g_C A_3)^2 \left(g_{2,3}\left\{f_{s_2}^{(j-4)}\right\}_{j-1} + g_C A_4 g_{2,4}\left\{f_{s_2}^{(j-5)}\right\}_{j-1}\right) + \ldots \\ &= (g_{2,3} + g_C A_4 g_{2,4} A_2)\left(1 + (g_c A_4 g_c A_3) A_2^2 + (g_c A_4 g_c A_3)^2 A_2^4 + \ldots\right) f_{s_2}^{(j)}.\end{aligned}$$

Hence, if we choose A_2, A_4 such that

$$g_{2,3} = -g_C g_{2,4} A_4 A_2,$$

we can ensure that c_{s_2} causes no interference at destination 3. Similarly, choosing A_1, A_3 to satisfy

$$g_{1,4} = -g_C g_{1,3} A_3 A_1$$

ensures that destination 4 receives no interference from c_{s_1}. In appendix D we show that there are choices for A's such that these conditions and the power constraints at the transmitters can be satisfied in the regime of interest. The upshot of this is that the cooperative private messages do not cause interference at the destinations. Appendix C employs a similar scheme in the context of linear deterministic channels.

The second phase of decoding is performed after the entire transmission is completed. At the end of the transmissions, the destinations have decoded their own public, cooperative private, and $\uparrow$ private messages for all blocks. They may cancel the effects of these messages from their observations. From the residual signal they successively decode the interferer's public message and their own $\downarrow$ private message in that order while treating the remaining interference as noise. The power allocations and the rates calculation are deferred to he appendices C and D.

B. Cooperation link stronger than direct links

The following theorem generalizes the scheme in Example 2. The coding scheme we use is a block-Markov coding scheme which has elements of compress-and-forward coding [6], which was originally proposed for relay channels, and superposition coding for interference channels [9]. The destinations perform a form of backwards decoding [15].

[3]Note that there are only a finite number of terms since the block indices start from $j = 1$.

Theorem 3: Given joint distributions $p_{U_1,X_1}\ p_{U_2,X_2}\ p_{X_3}p_{X_4}\ p_{Y_3|X_1,X_2,X_4}\ p_{Y_4|X_1,X_2,X_3}\ p_{V_3|Y_3}\ p_{V_4|Y_4}$ (where $p_{Y_3|X_1,X_2,X_4}$ and $p_{Y_4|X_1,X_2,X_4}$ are defined by the channel) the rate pair (R_1, R_2) is achievable if there are non-negative $r_{U_1}, r_{U_2}, r_{X_1}, r_{X_2}, r_3, r_4$ such that $R_1 = r_{U_1} + r_{X_1}$, $R_2 = r_{U_2} + r_{X_2}$, and

$$\begin{aligned}
r_{X_1} &\leq I(X_1; Y_3|X_4, U_1, U_2),\\
r_{U_2} &\leq I(U_2; Y_3, V_4|X_3, X_4, U_1),\\
r_{U_2} &\leq I(U_2; Y_3|X_3, X_4, U_1) + (r_4 - I(Y_4; V_4|X_3, X_4, U_1, U_2, Y_3)),\\
r_{U_1} &\leq I(U_1; Y_3, V_4|X_3, X_4, U_2),\\
r_{U_1} &\leq I(U_1; Y_3|X_3, X_4, U_2) + (r_4 - I(Y_4; V_4|X_3, X_4, U_1, U_2, Y_3)),\\
r_4 &\leq I(X_4; Y_3|U_1, U_2),\\
r_{U_2} + r_{U_1} &\leq I(U_2, U_1; Y_3, V_4|X_3, X_4),\\
r_{U_2} + r_{U_1} &\leq I(U_2, U_1; Y_3|X_3, X_4) + (r_4 - I(Y_4; V_4|X_3, X_4, U_1, U_2, Y_3)),\\
r_4 + r_{U_1} &\leq I(X_4, U_1; Y_3, V_4|X_3, U_2),\\
r_4 + r_{U_1} &\leq I(X_4, U_1; Y_3|X_3, U_2) + (r_4 - I(Y_4; V_4|X_3, X_4, U_1, U_2, Y_3)),\\
r_{U_2} + r_4 &\leq I(U_2, X_4; Y_3, V_4|X_3, U_1),\\
r_{U_2} + r_4 &\leq I(U_2, X_4; Y_3|X_3, U_1) + (r_4 - I(Y_4; V_4|X_3, X_4, U_1, U_2, Y_3)),\\
r_{U_2} + r_4 + r_{U_1} &\leq I(U_2, X_4, U_1; Y_3, V_4|X_3),\\
r_{U_2} + r_4 + r_{U_1} &\leq I(U_2, X_4, U_1; Y_3|X_3) + (r_4 - I(Y_4; V_4|X_3, X_4, U_1, U_2, Y_3)),
\end{aligned}$$

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

The theorem is proved in Appendix A. A rough interpretation follows. The auxiliary random variables and the information associated with them are described below:

- Cooperative public *messages*. U_1 and U_2 carry the cooperative public messages. These are decoded by both destinations with help from each other as will become evident.
- Private *messages*. The random variables X_1 conditioned on U_1 and X_2 conditioned on U_2 carry the private messages.
- *Quantized observations at the destinations*. The destinations quantize their observations over each block. The test channel for the quantizer employed by destination 3 is $p_{V_3|Y_3}$. Similarly, destination 4 quantizes its observation over a block using the test channel $p_{V_4|Y_4}$.
- *Messages from destinations to each other*. The quantization codebooks are binned and the bin-index of the quantized codewords are conveyed by the destinations to each other in the next block. Destination 3 sends the bin-index of its quantized codeword to destination 4 in the next block using X_3. Similarly destination 4 quantizes its observation over a block and conveys the bin-index to destination 3 in the next block using X_4.

The destinations start decoding from the last block and proceed backwards. For each block, destination 3 recovers (i) the two cooperative public messages for the current block, (ii) its private message for the current block and (iii) the message from destination 4 which conveys the bin-index of the quantized codeword at destination 4 for the previous block. In performing this decoding step, besides its observation for the current block, destination 3 may rely on the bin-index of the quantized codeword at destination 4 for the current block since this was recovered in the previous decoding step. This can be done by the destination jointly decoding the following: (i) the two cooperative public messages for the current block, (ii) its private message for the current block and (iii) the message from destination 4 which conveys the bin-index of the quantized codeword at destination 4 for the previous block, and (iv) the quantized codeword at destination 4 for the current block. Destination 3 makes use of its observation for the current block and the bin-index of the quantized codeword at destination 4 for the current block. The private message from the interfering source is treated as noise. Decoding at destination 4 also proceeds similarly.

As explained in detail in Appendices C and D where we prove the achievability part of our main results, this scheme is useful in the regime where the cooperative link is stronger than both the direct links from the sources to their respective destinations. In applying this scheme, we choose the power levels for the cooperative public and private signals in a manner similar to [8]. The private signals have a power level which ensures that they appear at or below the noise-level at the destinations where they act as interference. The quantizer test-channel we choose for our purposes is as follows: *At destination 3, the quantization-noise level is equal to the power level at which the* private *signal from source 1 is received at destination 3.* The intuition is that since destination 4 is not interested in decoding this private signal from source 1 and treats it as noise, quantizing the observation at destination 3 any finer than this will not help the joint decoding significantly. In fact, quantizing more finely could result in inferior performance since the destinations have to code the messages for each other at higher rates without producing significant benefits and this potentially results in overall lower rates for the other messages.

The decoding scheme in Theorem 3 deviates slightly from the earlier description. This is done primarily to simplify the evaluation of the achievable sum-rate for the Gaussian case. Instead of decoding the two cooperative public messages, the private message, the message from the other receiver, and the quantized codeword at the other destination for the current block, the destinations initially decode only (i) the two cooperative public messages, (ii) the message from the other destination, and (iii) the quantized codeword at the other destination for the current block. This is done treating both the private messages as noise. Once the above messages are recovered, (i) and (ii) are stripped off from the received signal and the private message is decoded from the residual signal. While this decoding scheme, in general, could lead to an inferior rate-region compared to the one described earlier, it is sufficient to obtain the achievability of sum-rate for the linear deterministic and the Gaussian cases (up to a constant gap for the Gaussian). For completeness, below we state a generic achievability theorem which implements the joint decoding described earlier. We sketch a proof in Appendix B. However, we do not use this scheme in proving Theorems 1 and 2.

Theorem 4: Given joint distributions $p_{U_1,X_1}\ p_{U_2,X_2}\ p_{X_3}p_{X_4}\ p_{Y_3|X_1,X_2,X_4}\ p_{Y_4|X_1,X_2,X_3}\ p_{V_3|Y_3}\ p_{V_4|Y_4}$ (where $p_{Y_3|X_1,X_2,X_4}$ and $p_{Y_4|X_1,X_2,X_4}$ are defined by the channel) the rate pair (R_1, R_2) is achievable if there are non-negative $r_{U_1}, r_{U_2}, r_{X_1}, r_{X_2}, r_3, r_4$ such that $R_1 = r_{U_1} + r_{X_1}$, $R_2 = r_{U_2} + r_{X_2}$, and

$$\begin{aligned}
r_{X_1} &\le I(X_1; Y_3, V_4|X_3, X_4, U_1, U_2),\\
r_{X_1} &\le I(X_1; Y_3|X_3, X_4, U_1, U_2)\\
&\qquad + (r_4 - I(Y_4; V_4|X_3, X_4, X_1, U_2, Y_3) + I(X_1; V_4)),\\
r_4 &\le I(X_4; Y_3|X_3, X_1, U_2),\\
r_{U_2} + r_{X_1} &\le I(U_2, X_1; Y_3, V_4|X_3, X_4, U_1),\\
r_{U_2} + r_{X_1} &\le I(U_2, X_1; Y_3|X_3, X_4, U_1)\\
&\qquad + (r_4 - I(Y_4; V_4|X_3, X_4, X_1, U_2, Y_3) + I(U_2, X_1; V_4)),\\
r_{U_1} + r_{X_1} &\le I(X_1; Y_3, V_4|X_3, X_4, U_2),\\
r_{U_1} + r_{X_1} &\le I(X_1; Y_3|X_3, X_4, U_2)\\
&\qquad + (r_4 - I(Y_4; V_4|X_3, X_4, X_1, U_2, Y_3) + I(X_1; V_4)),\\
r_4 + r_{X_1} &\le I(X_4, X_1; Y_3, V_4|X_3, U_2),\\
r_4 + r_{X_1} &\le I(X_4, X_1; Y_3|X_3, U_2)\\
&\qquad + (r_4 - I(Y_4; V_4|X_3, X_4, X_1, U_2, Y_3) + I(X_1; V_4)),\\
r_{U_2} + r_4 &\le I(U_2, X_4; Y_3, V_4|X_3, X_1),\\
r_{U_2} + r_4 &\le I(U_2, X_4; Y_3|X_3, X_1)\\
&\qquad + (r_4 - I(Y_4; V_4|X_3, X_4, X_1, U_2, Y_3) + I(U_2; V_4)),\\
r_{U_2} + r_{U_1} + r_{X_1} &\le I(U_2, X_1; Y_3, V_4|X_3, X_4),\\
r_{U_2} + r_{U_1} + r_{X_1} &\le I(U_2, X_1; Y_3|X_3, X_4)\\
&\qquad + (r_4 - I(Y_4; V_4|X_3, X_4, X_1, U_2, Y_3) + I(X_1, U_2; V_4)),
\end{aligned}$$

$$
\begin{aligned}
r_4 + r_{U_1} + r_{X_1} &\le I(X_4, X_1; Y_3, V_4|X_3, U_2),\\
r_4 + r_{U_1} + r_{X_1} &\le I(X_4, X_1; Y_3|X_3, U_2)\\
&\qquad + (r_4 - I(Y_4; V_4|X_3, X_4, X_1, U_2, Y_3) + I(X_1; V_4)),\\
r_{U_2} + r_4 + r_{X_1} &\le I(U_2, X_4, X_1; Y_3, V_4|X_3, U_1),\\
r_{U_2} + r_4 + r_{X_1} &\le I(U_2, X_4, X_1; Y_3|X_3, U_1)\\
&\qquad + (r_4 - I(Y_4; V_4|X_3, X_4, X_1, U_2, Y_3) + I(X_1, U_2; V_4)),\\
r_{U_2} + r_4 + r_{U_1} + r_{X_1} &\le I(U_2, X_4, X_1; Y_3, V_4|X_3),\\
r_{U_2} + r_4 + r_{U_1} + r_{X_1} &\le I(U_2, X_4, X_1; Y_3|X_3)\\
&\qquad + (r_4 - I(Y_4; V_4|X_3, X_4, X_1, U_2, Y_3) + I(X_1, U_2; V_4)),
\end{aligned}
$$

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

VI. Upperbounds

The upperbounds of Theorems 1 and 2 are derived in appendix E. The key ideas behind the upperbounds are as follows:

Upperbound 1: We simulate two dummy channels which are independent realizations of the original channel. In both channels, the same codebooks as in the original are used. In the first dummy channel, the message M_1 of sender 1 is replaced by a dummy message random variable M_1' which is also uniformly distributed over the alphabet $\mathcal{M}$, but is independent of both M_1 and M_2. Similarly, in the second dummy channel the message M_2 is replaced by a dummy message M_2''. A genie provides destination 3 with M_2'' and a certain signal $(g_{1,4}^*(X_1^T) + g_{3,4}(X_3^T))$ from the second dummy channel, and destination 4 with the symmetric counterparts from the first dummy channel. Applying Fano's inequality and conditions implied by the causality conditions on the sources, we derive upperbounds which imply (6) and (1).

Upperbound 2 and 3: To show upperbound 2, we consider a genie which provides $g_{2,3}^*(X_2^T)$ and M_1 to destination 4 and nothing at all to destination 3. Using Fano's inequality and using the causality conditions obeyed by the sources, we derive upperbounds which imply (7)-(8) and (2)-(3).

Upperbound 4: This is a simple cut-set upperbound with nodes 1 and 3 on one side of the cut and nodes 2 and 4 on the other side.

Upperbound 5: This is also a cut-set bound. The sources are on one side of the cut and the destinations on the other.

VII. Discussion

The gap in Theorem 2 can be easily improved by considering more elaborate schemes and further tightening the upperbound. We mention a couple of ideas to illustrate how this could be achieved. However, computing the best possible gap appears to be challenging and we do not pursue it here.

Even with the schemes we presented in the last section, we picked potentially sub-optimal power allocations for the different messages involved in order to simplify the calculations. Improvements in the gap can be achieved in specific instances simply by optimizing over these power allocations. But still further improvements can be achieved by considering other schemes. For instance, consider the case of a channel with a direct links which are weak compared to the interfering and cooperative links. Incorporating a form of decode-and-forward strategy can improve performance. To see this, let us consider this extreme case: both direct links are absent

$$g_{1,3} = g_{2,4} = 0,$$

the interfering links have the same strength

$$|g_{1,4}| = |g_{2,3}| = |g_I|,$$

and the cooperative links are such that the following condition is satisfied

$$\left(1+|g_C|^2\right)^2 \le \left(1+|g_I|^2+|g_C|^2\right).$$

Then, we can show that the sum-capacity is achieved by a simple decode-and-forward scheme. Both destinations decode the message from their interfering sources treating the signal in the cooperative-link as noise, and then in the next block they forward this decoded message to the other source over the cooperative link. Since the interference is decoded off first, the signal over the cooperative link can be decoded without any interference. The resulting rate, under the condition on the channel strengths mentioned above, is

$$2\log\left(1+|g_C|^2\right),$$

which is also what upperbound (9) works out to. However, no choice of power allocations in Theorems 3 or 4 can achieve this. This can be easily remedied by extending those schemes by incorporating a partial-decode-and-forward component. However, we do not pursue this direction since the gains are at most a constant and computing such gains to get an improved uniform bound appears to be involved.

The upperbounds could also be improved. Modifying the correlation of the Gaussian noise processes in the additional signals we provide to the destinations can lead to tighter upperbounds 1, 2, and 3. 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. Upperbound 5 can be easily improved by choosing the optimal input covariance matrix.

Appendix A
Proof of Theorem 3

We present a block-Markov scheme with backwards decoding. Given $p_{U_1,X_1}\ p_{U_2,X_2}\ p_{X_3}p_{X_4}\ p_{Y_3|X_1,X_2,X_4}\ p_{Y_4|X_1,X_2,X_3}\ p_{V_3|Y_3}\ p_{V_4|Y_4}$ (where $p_{Y_3|X_1,X_2,X_4}$ and $p_{Y_4|X_1,X_2,X_4}$ are defined by the channel), we construct the following blocklength-T codebooks:

- U codebooks: For $k=1,2$, we create U_k-codebooks $\mathcal{C}_{U_k}$ of size $2^{T(r_{U_k}-3\epsilon)}$ respectively, by choosing elements independently according to p_{U_k}. These codewords will be denoted by $c_{U_k}(m_{U_k})$ where $m_{U_k}\in\{1,\ldots,2^{T(r_{U_k}-3\epsilon)}\}$.
- X_1 and X_2 codebooks: For each codeword $c_{U_k}(m_{U_k})$, we create a X_k-codebook $\mathcal{C}_{X_k}(m_{U_k})$ of size $2^{T(r_{X_k}-3\epsilon)}$ by choosing elements i.i.d. according to $p_{X_k|U_k}(.|u_k)$ by setting u_k to be the respective element of the $c_{U_k}(m_{U_k})$ codeword. We denote these codewords by $c_{X_k}(m_{X_k},m_{U_k})$, where $m_{X_k}\in\{1,\ldots,2^{T(r_{X_k}-3\epsilon)}\}$.
- X_3 and X_4 codebooks: For $k=3,4$, we create X_k-codebooks $\mathcal{C}_{X_k}$ of size $2^{T(r_k-\epsilon)}$ by choosing the elements i.i.d. according to p_{X_k}. These codewords will be denoted by $c_{X_k}(m_{X_k})$ where $m_{X_k}\in\{1,\ldots,2^{T(r_k-\epsilon)}\}$.
- V_k codebooks: For $k=3,4$, we create V_k codebooks $\mathcal{C}_{V_k}$ of size $2^{T(I(Y_k;V_k)+\epsilon)}$ by choosing the elements i.i.d. according to the induced marginal distributions p_{V_k}. We bin these codebooks such that the number of bins is $2^{T(r_k-\epsilon)}$. The codewords will be denoted by $c_{V_k}(b_{V_k},i_{V_k})$ where the bin-indices are denoted by $b_{V_k}\in\{1,\ldots,2^{T(r_k-\epsilon)}\}$, and within each bin, the index of the codewords are denoted by $i_{V_k}\in\{1,\ldots,2^{T(I(Y_k;V_k)-r_k+2\epsilon)}\}$.

Encoding at the sources: For block-j, $j=1,2,\ldots,J-1$, the encoders at the sources choose the codewords $c_{U_k}(m_{U_k}(j))$, and $c_{X_k}(m_{X_k}(j),m_{U_k}(j))$. The X-codewords are transmitted. For the last block J, we set $m_{U_1}(J)=m_{X_1}(J)=m_{U_2}(J)=m_{X_2}(J)=1$.

Encoding at the destinations: At the end of block-j, $j=1,2,\ldots,J-1$, the destination 3 quantizes its (T-length) block of observations Y_3^T using the V_3 codebook by finding a codeword $c_{V_3}(b_{V_3}(j),i_{V_3}(j))$ which is jointly (strongly) typical[4] [7, Chapter 13] with its observation. If no such codeword exists, we

[4]In the sequel, we denote the set of strongly δ-typical sequences by $\mathcal{T}_T^\delta$.

will say that "encoding failed at block-j" and declare an error. However, encoding succeeds with high probability since the V_3 codebook has rate $I(V_3; Y_3)$ [16]. Then destination 3 sets $m_{X_3}(j+1) = b_{V_3}(j)$ and for block-$j+1$, destination 3 sends $c_{X_3}(m_{X_3}(j+1))$. The encoding at destination 4 proceeds similarly.

Decoding at the destinations: Destinations perform *backwards decoding* [15]. We will assume that before destination 3 processes block-j, it has already successfully decoded $m_{X_4}(j+1)$. This is true with high probability[5] for $j = J$ if

$$r_4 \leq I(X_4; Y_3 | X_1, X_2).$$

For each $j = 1, 2, \ldots, J-1$, we will ensure that from block-j, destination 3 decodes $m_{X_4}(j)$ successfully with high probability thereby ensuring that the above assumption holds true. Assuming that $m_{X_4}(j+1)$, which is equal to $b_{V_4}(j)$, is available at destination 3, we will ensure that from the observation Y_3 made by destination 3 in block-j, the messages $m_{U_1}(j)$, $m_{X_1}(j)$, and $m_{X_4}(j)$ can be successfully decoded with high probability. The decoding will proceed in two steps. In the first step, destination 3 will attempt to decode the messages $m_{U_1}(j)$ and $m_{X_4}(j)$ along with the message $m_{U_2}(j)$. Then, conditioned on these messages, it will try to decode the message $m_{X_1}(j)$. Concretely, in the first step the decoder looks for a unique collection of codewords such that they are jointly typical with its observation $Y_3^T(j)$ and the information it already has, namely $m_{X_3}(j)$ and $b_{V_4}(j)$. In other words, destination 3 searches for a unique $(\hat{m}_{U_1}(j), \hat{m}_{U_2}(j), \hat{m}_{X_4}(j))$ such that

$$(c_{U_1}(\hat{m}_{U_1}(j)), c_{U_2}(\hat{m}_{U_2}(j)), c_{X_4}(\hat{m}_{X_4}(j)), c_{X_3}(m_{X_3}(j)), c_{V_4}(b_{V_4}(j), \hat{i}_{V_4}(j)), Y_3^T(j)) \in \mathcal{T}_T^\delta,$$

for some $\hat{i}_{V_4}(j)$. We will argue below that this decoding succeeds, *i.e.*, $(\hat{m}_{U_1}(j), \hat{m}_{U_2}(j), \hat{m}_{X_4}(j)) = (m_{U_1}(j), m_{U_2}(j), m_{X_4}(j))$ with high probability if the following conditions are met.

$$\begin{aligned}
r_{U_2} &\leq I(U_2; Y_3, V_4 | X_3, X_4, U_1), \\
r_{U_2} &\leq I(U_2; Y_3 | X_3, X_4, U_1) + (r_4 - I(Y_4; V_4 | X_3, X_4, U_1, U_2, Y_3)), \\
r_{U_1} &\leq I(U_1; Y_3, V_4 | X_3, X_4, U_2), \\
r_{U_1} &\leq I(U_1; Y_3 | X_3, X_4, U_2) + (r_4 - I(Y_4; V_4 | X_3, X_4, U_1, U_2, Y_3)), \\
r_4 &\leq I(X_4; Y_3 | U_1, U_2), \\
r_{U_2} + r_{U_1} &\leq I(U_2, U_1; Y_3, V_4 | X_3, X_4), \\
r_{U_2} + r_{U_1} &\leq I(U_2, U_1; Y_3 | X_3, X_4) + (r_4 - I(Y_4; V_4 | X_3, X_4, U_1, U_2, Y_3)), \\
r_4 + r_{U_1} &\leq I(X_4, U_1; Y_3, V_4 | X_3, U_2), \\
r_4 + r_{U_1} &\leq I(X_4, U_1; Y_3 | X_3, U_2) + (r_4 - I(Y_4; V_4 | X_3, X_4, U_1, U_2, Y_3)), \\
r_{U_2} + r_4 &\leq I(U_2, X_4; Y_3, V_4 | X_3, U_1), \\
r_{U_2} + r_4 &\leq I(U_2, X_4; Y_3 | X_3, U_1) + (r_4 - I(Y_4; V_4 | X_3, X_4, U_1, U_2, Y_3)), \\
r_{U_2} + r_4 + r_{U_1} &\leq I(U_2, X_4, U_1; Y_3, V_4 | X_3), \\
r_{U_2} + r_4 + r_{U_1} &\leq I(U_2, X_4, U_1; Y_3 | X_3) + (r_4 - I(Y_4; V_4 | X_3, X_4, U_1, U_2, Y_3)).
\end{aligned}$$

In the second step, destination 3 decodes $m_{X_1}(j)$ using its observation $Y_3^T(j)$ and what it decoded in the previous step, namely, $\hat{m}_{U_1}(j)$, $\hat{m}_{U_2}(j)$, and $\hat{m}_{X_4}(j)$. It looks for a unique $\hat{m}_{X_1}(j)$ such that

$$(c_{X_1}(\hat{m}_{X_1}(j)), c_{U_1}(\hat{m}_{U_1}(j)), c_{U_2}(\hat{m}_{U_2}(j)), c_{X_4}(\hat{m}_{X_4}(j)), Y_3^T(j)) \in \mathcal{T}_T^\delta.$$

Assuming that the first step succeeded, it can be shown that the second step succeeds in decoding the correct message with a high probability if

$$r_{X_1} \leq I(X_1; Y_3 | U_1, U_2, X_4).$$

[5] *i.e.*, with probability approaching 1 as the blocklength T goes to ∞.

We will now argue that the probability of error in the first step is vanishingly small. We first note that the correct choice of messages will result in a jointly typical set of codewords with high probability. *i.e.*, when $T \to \infty$,

$$\mathbb{P}\bigg(\big(c_{U_1}(m_{U_1}(j)), c_{U_2}(m_{U_2}(j)), c_{X_4}(m_{X_4}(j)), c_{X_3}(m_{X_3}(j)), c_{V_4}(b_{V_4}(j), i_{V_4}(j)), Y_3^T(j)\big) \in \mathcal{T}_T^{\delta} \bigg) \to 1.$$

This is essentially a statement of Markov Lemma [2, Lemma 4.1] (also see [13, Corollary 3.2.3.1]). We need to show that the probability of the event E that there is some $(\hat{m}_{U_1}(j), \hat{m}_{U_2}(j), \hat{m}_{X_4}(j)) \neq (m_{U_1}(j), m_{U_2}(j), m_{X_4}(j))$ and some $\hat{i}_{V_4}$ such that

$$\Big(c_{U_1}(\hat{m}_{U_1}(j)), c_{U_2}(\hat{m}_{U_2}(j)), c_{X_4}(\hat{m}_{X_4}(j)), c_{X_3}(m_{X_3}(j)), c_{V_4}(b_{V_4}(j), \hat{i}_{V_4}(j)), Y_3^T(j)\Big) \in \mathcal{T}_T^{\delta}$$

is small. This event is the union of the following two events: (1) where $\hat{i}_{V_4}$ takes on its correct value i_{V_4}, and (2) where $\hat{i}_{V_4} \neq i_{V_4}$. Further, each of these events are unions of events where some or none (but not all) of the messages take on their correct value. We apply union bound to upperbound $\mathbb{P}(\mathsf{E})$. To illustrate, let us consider two events:

$$\mathsf{E}_{1,2,4} = \Big\{ \hat{m}_{U_1}(j) \neq m_{U_1}(j), \hat{m}_{U_2}(j) \neq m_{U_2}(j), \hat{m}_{X_4}(j) \neq m_{X_4}(j), \hat{i}_{V_4}(j) = i_{V_4}(j) \Big\},$$
$$\mathsf{E}^*_{1,2,4} = \Big\{ \hat{m}_{U_1}(j) \neq m_{U_1}(j), \hat{m}_{U_2}(j) \neq m_{U_2}(j), \hat{m}_{X_4}(j) \neq m_{X_4}(j), \hat{i}_{V_4}(j) \neq i_{V_4}(j) \Big\}.$$

We have

$$\begin{aligned}
&\mathbb{P}(\mathsf{E}_{1,2,4}) \\
&= \sum_{\substack{\hat{m}_{U_1} \neq m_{U_1}(j),\\ \hat{m}_{U_2} \neq m_{U_2}(j),\\ \hat{m}_{X_4} \neq m_{X_4}(j)}} \mathbb{P}\left((c_{U_1}(\hat{m}_{U_1}), c_{U_2}(\hat{m}_{U_2}), c_{X_4}(\hat{m}_{X_4}), c_{X_3}(m_{X_3}(j)), c_{V_4}(b_{V_4}(j), i_{V_4}(j)), Y_3^T(j)) \in \mathcal{T}_T^{\delta} \right) \\
&\leq 2^{T(r_{U_1}+r_{U_2}+r_4-7\epsilon)} 2^{T(-I(U_1,U_2,X_4;X_3,V_4,Y_3)+\delta)} \\
&= 2^{T(r_{U_1}+r_{U_2}+r_4-I(U_1,U_2,X_4;V_4,Y_3|X_3)-7\epsilon+\delta)}.
\end{aligned}$$

If the rates satisfy the condition

$$r_{U_1} + r_{U_2} + r_4 \leq I(U_1, U_2, X_4; V_4, Y_3|X_3),$$

the probability of the error event $\mathsf{E}_{1,2,4}$ can be made vanishingly small. Similarly,

$$\begin{aligned}
&\mathbb{P}(\mathsf{E}^*_{1,2,4}) \\
&= \sum_{\substack{\hat{m}_{U_1} \neq m_{U_1}(j),\\ \hat{m}_{U_2} \neq m_{U_2}(j),\\ \hat{m}_{X_4} \neq m_{X_4}(j),\\ \hat{i}_{V_4} \neq i_{V_4}(j)}} \mathbb{P}\left((c_{U_1}(\hat{m}_{U_1}), c_{U_2}(\hat{m}_{U_2}), c_{X_4}(\hat{m}_{X_4}), c_{X_3}(m_{X_3}(j)), c_{V_4}(b_{V_4}(j), \hat{i}_{V_4}), Y_3^T(j)) \in \mathcal{T}_T^{\delta} \right) \\
&= 2^{T(r_{U_1}+r_{U_2}+r_4+(I(Y_4;V_4)-r_4)-5\epsilon)} 2^{T(-I(U_1,U_2,X_4,V_4;X_3,Y_3)+\delta)} \\
&= 2^{T((r_{U_1}+r_{U_2}+r_4-I(U_1,U_2,X_4;X_3,Y_3))-(r_4-I(V_4;Y_4)+I(V_4;X_3,Y_3|U_1,U_2,X_4))-5\epsilon+\delta)}
\end{aligned}$$

To drive the probability of this error event ($\mathsf{E}^*_{1,2,4}$) to zero, it is enough to ensure that

$$r_{U_1} + r_{U_2} + r_4 \leq I(U_1, U_2, X_4; X_3, Y_3) + (r_4 - I(V_4; Y_4) + I(V_4; X_3, Y_3|U_1, U_2, X_4)).$$

Note that

$$\begin{aligned} I(V_4;Y_4)-I(V_4;X_3,Y_3|U_1,U_2,X_4)) &= I(V_4;Y_4,U_1,U_2,X_4,X_3,Y_3)-I(V_4;X_3,Y_3|U_1,U_2,X_4)) \\ &= I(V_4;U_1,U_2,X_4)+I(V_4;Y_4|X_3,X_4,U_1,U_2,Y_3) \\ &\geq I(V_4;Y_4|X_3,X_4,U_1,U_2,Y_3), \end{aligned}$$

where the first equality follows from the fact that $V_4 - Y_4 - (U_1, U_2, X_4, X_3, Y_3)$ is a Markov chain. Hence, $\mathbb{P}(\mathsf{E}^*_{1,2,4})$ can be made small if

$$r_{U_1} + r_{U_2} + r_4 \leq I(U_1, U_2, X_4; Y_3|X_3) + (r_4 - I(V_4; Y_4|X_3, X_4, U_1, U_2, Y_3)).$$

Similarly, considering the other possible error events results in the rest of the conditions. 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.

Appendix B
Proof sketch of Theorem 4

The codebook construction and encoding at the sources and destinations are identical to the one in Appendix A. The only difference is in how the messages are decoded by the destinations. The destinations again follow a backwards decoding procedure similar to the one there. However, instead of carrying it out in two steps, the destinations attempt to decode the same set of codewords as there, but in a single step. In particular, destination 3 while decoding block-j looks for a unique set of $(\hat{m}_{U_1}(j), \hat{m}_{U_2}(j), \hat{m}_{X_4}(j), \hat{m}_{X_1}(j))$ such that

$$(c_{U_1}(\hat{m}_{U_1}(j)), c_{U_2}(\hat{m}_{U_2}(j)), c_{X_4}(\hat{m}_{X_4}(j)), c_{X_1}(\hat{m}_{U_1}(j)), c_{X_3}(m_{X_3}(j)), c_{V_4}(b_{V_4}(j), \hat{i}_{V_4}(j)), Y_3^T(j)) \in \mathcal{T}_T^{\delta},$$

for some $\hat{i}_{V_4}(j)$. Note that, as in Appendix A, in performing this decoding step, destinations 3 makes uses of the bin-index $b_{V_4}(j)$ which was recovered from processing block-$j+1$. The conditions on the rates in Theorem 4 ensure that the probability of all relevant error events are small for sufficiently large values of T. The analysis is along the same lines as in Appendix A and is omitted.

Appendix C
Proof of the 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 three natural regimes. We use different strategies to achieve the sum-capacity in different regimes. The regimes are:

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

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

Hence, we need consider u_1 only in the regime $n_C \leq n_{\text{min}}$. 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

$$\begin{aligned} &\max(n_{1,3} - n_{1,4}, n_{2,3}) + \max(n_{2,4} - n_{1,4}, n_{1,4}) \\ &\quad < \min\Bigg(\max(n_{1,3}, n_{2,3}) + (\max(n_{2,4}, n_{2,3}) - n_{2,3}), \\ &\qquad\qquad \max(n_{2,4}, n_{1,4}) + (\max(n_{1,3}, n_{1,4}) - n_{1,4}), n_{1,3} + n_{2,4} \Bigg). \end{aligned} \tag{11}$$

When the above condition is not true, the sum-rate expression reduces to the sum-capacity without cooperation.

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

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 $\mathbb{F}^n$ 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): $n_C \leq n_{i,j},\ i \in \{1,2\}, j \in \{3,4\}$.

First of all, we note that if

$$n_{1,4} + n_{2,3} \geq n_{1,3} + n_C \text{ and } n_{1,4} + n_{2,3} \geq n_{2,4} + n_C,$$

then,

$$u_1(n_C) = u_1(0).$$

Putting this together with the fact that $u_2(n_C), u_3(n_C)$, and $u_4(n_C)$ are independent of n_C in regime (i) (as discussed above), we may conclude that it is enough to show achievability under no cooperative link (*i.e.*, under $n_C = 0$). But this achievability is already known (see [12], for instance). Hence, we will assume that at least one of the following two conditions is true.

$$n_{1,4} + n_{2,3} > n_{1,3} + n_C \text{ and } n_{1,4} + n_{2,3} > n_{2,4} + n_C.$$

We will first consider case (a) where the following conditions are satisfied.

$$n_{1,4} + n_{2,3} < n_{1,3} + n_C \text{ and } n_{1,4} + n_{2,3} < n_{2,4} + n_C.$$

Following that, we will consider case (b) where

$$n_{1,4} + n_{2,3} < n_{1,3} + n_C \text{ and } n_{1,4} + n_{2,3} \geq n_{2,4} + n_C.$$

By symmetry, this would also cover the third possibility of

$$n_{1,4} + n_{2,3} \geq n_{1,3} + n_C \text{ and } n_{1,4} + n_{2,3} < n_{2,4} + n_C.$$

For case (a), let us consider the following block-Markov scheme with superposition coding. Let $U_k, S_k, Z_{\uparrow k}, Z_{\downarrow k},\ k$ $1,2$ be independent auxiliary random variables with marginal distributions $p_{U_k}, p_{S_k}, p_{Z_{\uparrow k}}, p_{Z_{\downarrow k}}$. The alphabet for these random variables is $\mathbb{F}^n$. Corresponding to these random variables, random codebooks of blocklength-T and rates $r_{U_k}, r_{S_k}, r_{Z_{\uparrow k}}, r_{Z_{\downarrow k}}$, respectively, are defined as usual. For instance, the U_1-codebook, denoted by $\mathcal{C}_{U_1}$, is of size $2^{T(r_{U_1}-\epsilon)}$ is generated by choosing the T elements of each of the codewords independently according to p_{U_1}. These codewords will be denoted by $c_{U_1}(m_{U_1})$ where $m_{U_1} \in \{1, \ldots, 2^{T(r_{U_1}-\epsilon)}\}$. The blocks will be indexed by $j = 1, 2, \ldots, J$. The message transmitted by source 1 using the U_1-codebook in block-j will be denoted by $m_{U_1}(j)$, and the corresponding codeword by $c_{U_1}^{(j)} = c_{U_1}(m_{U_1}(j))$. For block-$j$, sources transmit the following blocklength-T vectors

$$X_1^{(j)} = c_{U_1}^{(j)} + c_{Z_{\uparrow 1}}^{(j)} + c_{Z_{\downarrow 1}}^{(j)} + f_{S_1}^{(j)},$$
$$X_2^{(j)} = c_{U_2}^{(j)} + c_{Z_{\uparrow 2}}^{(j)} + c_{Z_{\downarrow 2}}^{(j)} + f_{S_2}^{(j)},$$

where

$$f_{S_1}^{(j)} = c_{S_1}^{(j)} + \mathbf{A}_1 c_{S_1}^{(j+1)} + \ldots + \mathbf{A}_1^{J-j} c_{S_1}^{(J)},$$
$$f_{S_2}^{(j)} = c_{S_2}^{(j)} + \mathbf{A}_2 c_{S_2}^{(j+1)} + \ldots + \mathbf{A}_2^{J-j} c_{S_2}^{(J)},$$

where $\mathbf{A}_1$ and $\mathbf{A}_2$ are matrices given below. The addition is vector addition.

$$\mathbf{A}_1 = \mathbf{S}^{n_{1,3}+n_C-n_{1,4}-n_{2,3}}, \qquad \mathbf{A}_2 = \mathbf{S}^{n_{2,4}+n_C-n_{1,4}-n_{2,3}}.$$

In the sequel we will ensure that the rates of the codebooks are such that from observing $Y_3^{(j)}$, the j-th block observed by destination 3, it (destination 3) can decode with a high probability of success the codewords $c_{U_1}^{(j)}$, $c_{S_1}^{(j)}$, and $c_{Z_{\uparrow 1}}^{(j)}$, for all j. Similarly, we will make sure that destination 4 will successfully decode $c_{U_2}^{(j)}$, $c_{S_2}^{(j)}$, and $c_{Z_{\uparrow 2}}^{(j)}$ from $Y_4^{(j)}$. Then, the destinations will transmit, respectively, for $j = 1, 2, \ldots, J-1$,

$$X_3^{(j+1)} = \mathbf{A}_3 \left(Y_3^{(j)} - \mathbf{G}_{1,3} \left(c_{U_1}^{(j)} + c_{S_1}^{(j)} + c_{U_{\uparrow 1}}^{(j)} \right) - \mathbf{G}_C \mathbf{A}_4 \mathbf{G}_{1,4} \left(c_{U_1}^{(j-1)} + c_{Z_{\uparrow 1}}^{(j-1)} \right) \right),$$
$$X_4^{(j+1)} = \mathbf{A}_4 \left(Y_4^{(j)} - \mathbf{G}_{2,4} \left(c_{U_2}^{(j)} + c_{S_2}^{(j)} + c_{U_{\uparrow 2}}^{(j)} \right) - \mathbf{G}_C \mathbf{A}_3 \mathbf{G}_{2,3} \left(c_{U_2}^{(j-1)} + c_{Z_{\uparrow 2}}^{(j-1)} \right) \right),$$

where, $\mathbf{G}_{k_1,k_2}$ is a short-hand notation for $\mathbf{S}^{n-n_{k_1,k_2}}$, and $\mathbf{G}_C$ for $\mathbf{S}^{n-n_C}$. Also, $\mathbf{A}_3$ and $\mathbf{A}_4$ are matrices defined below.

$$\mathbf{A}_4 = -\mathbf{S}^{-(n-n_{1,4})},$$
$$\mathbf{A}_3 = -\mathbf{S}^{-(n-n_{2,3})}.$$

The choices for the distributions of the auxiliary random variables used to create the codebooks will be taken up in the sequel.

With these, the received signals at the destinations are

$$Y_3^{(j)} = \overline{g_{1,3}}^{(j)}(c_{U_1} + c_{Z_{\uparrow 1}}) + \mathbf{G}_{1,3}(f_{S_1}^{(j)}) + \widetilde{g_{1,3}}^{(j)}(c_{Z_{\downarrow 1}}) + \mathbf{G}_{2,3}(c_{U_2}^{(j)} + c_{Z_{\uparrow 2}}^{(j)}) + \widetilde{g_{2,3}}^{(j)}(c_{Z_{\downarrow 2}}),$$
$$Y_4^{(j)} = \overline{g_{2,4}}^{(j)}(c_{U_2} + c_{Z_{\uparrow 2}}) + \mathbf{G}_{2,4}(f_{S_2}^{(j)}) + \widetilde{g_{2,4}}^{(j)}(c_{Z_{\downarrow 2}}) + \mathbf{G}_{1,4}(c_{U_1}^{(j)} + c_{Z_{\uparrow 1}}^{(j)}) + \widetilde{g_{1,4}}^{(j)}(c_{Z_{\downarrow 1}}),$$

where the functions are as defined below. Note that Y_3 does not have any terms which depend on S_2-codewords (and similarly, Y_4 does not involve any terms containing S_1-codewords). This was achieved by the appropriate choices above for $\mathbf{A}_1$ and $\mathbf{A}_4$ (respectively, $\mathbf{A}_2$ and $\mathbf{A}_3$).

$$\overline{g_{1,3}}^{(j)}(c_{U_1} + c_{Z_{\uparrow 1}}) = \mathbf{G}_{1,3}(c_{U_1}^{(j)} + c_{Z_{\uparrow 1}}^{(j)}) + \mathbf{G}_C \mathbf{A}_4 \mathbf{G}_{1,4}(c_{U_1}^{(j-1)} + c_{Z_{\uparrow 1}}^{(j-1)}),$$
$$\widetilde{g_{1,3}}^{(j)}(c_{Z_{\downarrow 1}}) = \sum_{i \in \{j, j-2, \ldots, 1\}} (\mathbf{G}_C \mathbf{A}_4 \mathbf{G}_C \mathbf{A}_3)^{\frac{j-i}{2}} \mathbf{G}_{1,3} c_{Z_{\downarrow 1}}^{(i)} + \sum_{i \in \{j-1, j-3, \ldots, 1\}} (\mathbf{G}_C \mathbf{A}_4 \mathbf{G}_C \mathbf{A}_3)^{\frac{j-1-i}{2}} \mathbf{G}_C \mathbf{A}_4 \mathbf{G}_{1,4} c_{Z_{\downarrow 1}}^{(i)},$$
$$\widetilde{g_{2,3}}^{(j)}(c_{Z_{\downarrow 2}}) = \sum_{i \in \{j, j-2, \ldots, 1\}} (\mathbf{G}_C \mathbf{A}_4 \mathbf{G}_C \mathbf{A}_3)^{\frac{j-i}{2}} \mathbf{G}_{2,3} c_{Z_{\downarrow 2}}^{(i)} + \sum_{i \in \{j-1, j-3, \ldots, 1\}} (\mathbf{G}_C \mathbf{A}_4 \mathbf{G}_C \mathbf{A}_3)^{\frac{j-1-i}{2}} \mathbf{G}_C \mathbf{A}_4 \mathbf{G}_{2,4} c_{Z_{\downarrow 2}}^{(i)},$$

and the functions $\overline{g_{2,4}}^{(j)}(c_{U_2} + c_{Z_{\uparrow 1}})$, $\widetilde{g_{2,4}}^{(j)}(c_{Z_{\downarrow 2}})$, and $\widetilde{g_{1,4}}^{(j)}(c_{Z_{\downarrow 1}})$ are defined similarly. Note that, as before, $\mathbf{G}_{k_1,k_2}$ is a short-hand notation for $\mathbf{S}^{n-n_{k_1,k_2}}$, and $\mathbf{G}_C$ for $\mathbf{S}^{n-n_C}$.

As mentioned earlier, destinations perform decoding in two phases. At the end of every block j, the destinations decode the $c_U^{(j)}$, $c_S^{(j)}$ and $c_{Z_\uparrow}^{(j)}$ intended for them (in that order). We call this *phase 1 decoding*.

At the end of block-J, the decoders perform a *phase 2 decoding* where it decodes the following codewords for all blocks, *i.e.*, for each $j = 1, 2, \ldots, J$, the destinations decode $c_U^{(j)}$ intended for the other user and $c_{Z_\downarrow}^{(j)}$ intended for itself (in that order). In both of the phases, the decodings are performed in the order mentioned above treating all the undecoded codewords and other interference as noise. Below, we will specify the distributions employed and evaluate the conditions on the rates to ensure successful decoding. This will establish achievability for regime (i).

The auxiliary random variables $U_1, S_1, Z_{\uparrow 1}$, and $Z_{\downarrow 1}$, respectively, are uniformly distributed over $\mathbb{F}^n$, $\mathcal{F}_{n_{1,4}-n_C}$, $\mathcal{F}_{n_{1,4}}$, and $\mathcal{F}_{n_{1,3}+n_C-n_{2,3}}$. Destination 3 transmits

$$\begin{aligned} X_3^{(j)} &= \mathbf{A}_3 \left(Y_3^{(j)} - \mathbf{G}_{1,3} \left(c_{U_1}^{(j)} + c_{S_1}^{(j)} + c_{U_{\uparrow 1}}^{(j)} \right) \right) \\ &= \mathbf{A}_3 \left(\widetilde{g_{1,3}}^{(j)}(c_{Z_{\downarrow 1}}) + \mathbf{G}_{2,3}(c_{U_2}^{(j)} + c_{Z_{\uparrow 2}}^{(j)}) + \widetilde{g_{2,3}}^{(j)}(c_{Z_{\downarrow 2}}) \right) \end{aligned}$$

The rates supported by the above scheme for source 1 are given by the following set of conditions.

$$\begin{aligned} r_{U_1} &\le [n_{1,4} - n_C], \\ r_{U_1} &\le n_C, \\ r_{S_1} &\le n_C - [n_{2,3} - (n_{1,3} - n_{1,4})]_+, \\ r_{Z_{\uparrow 1}} &\le [n_{1,3} - n_{1,4} - n_{2,3}]_+, \\ r_{Z_{\downarrow 1}} &\le [n_{2,3} - n_C]_+. \end{aligned}$$

The first constraint on r_{U_1} comes from the phase 1 decoding at destination 3, while the second condition is from the phase 2 decoding at destination 4. A similar set of constraints apply for the rates achievable by source 2.

Combining all these, an achievable sum-rate is given by

$$R_{\text{sum}} = \max(n_{1,3} - n_{1,4} + n_C, n_{1,3} - n_C) + \max(n_{2,4} - n_{2,3} + n_C, n_{2,4} - n_C).$$

This, combined with the fact that the achievability of a given sum-rate at a lower value of n_C implies its achievability for all larger values of n_C provided the rest of the channel coefficients remain the same allows us to conclude that the minimum of the following three terms is achievable.

$$\begin{aligned} &n_{1,3} - n_{1,4} + n_C + n_{2,4} - n_{2,3} + n_C, \\ &n_{1,3} + n_{2,4} - n_{1,4}, \text{ and} \\ &n_{2,4} + n_{1,3} - n_{2,3}. \end{aligned}$$

Under case (a), it is easy to verify that this is precisely what the upperbound evaluates to. Hence, we have shown achievability for case (a).

Let us now consider case (b) where

$$n_{1,4} + n_{2,3} < n_{1,3} + n_C \text{ and } n_{1,4} + n_{2,3} \ge n_{2,4} + n_C.$$

The achievable strategy we use for this case involves source 1 transmitting according to a scheme similar to the one above while source 2 employs a superposition coding scheme similar to that of Han and Kobayashi for the interference channel without a cooperative link. In particular, only node 3 uses its transmission capabilities.

The codebooks and the choice of distributions for source 1 are exactly as above, except for the choice of rates which will be presented in the sequel. Source 2 uses only the following codebooks: U_2 and $Z_{\downarrow 2}$. Moreover, $Z_{\downarrow 2}$ is now uniformly distributed over $\mathcal{F}_{n_{2,3}}$. Exactly as in the earlier scheme, destination 3 performs a two-phase decoding and transmits a shifted version of the residual signals after the first phase of decoding of the previous block. The shift matrix $\mathbf{A}_3$ is the same as above. Destination 4, on the other hand, performs only a single phase of decoding where U_2, U_1 and $Z_{\downarrow 2}$ codewords are decoded. As mentioned earlier, destination 4 does not transmit anything, *i.e.*, $\mathbf{A}_4 = 0$. The received signals can be seen to be

$$\begin{aligned}
Y_3^{(j)} &= \mathbf{G}_{1,3}(c_{U_1}^{(j)} + c_{Z_{\uparrow 1}}^{(j)} + f_{S_1}^{(j)} + c_{Z_{\downarrow 1}}^{(j)}) + \mathbf{G}_{2,3}(c_{U_2}^{(j)} + c_{Z_{\downarrow 2}}^{(j)}),\\
Y_4^{(j)} &= \mathbf{G}_{2,4}(c_{U_2}^{(j)} + c_{Z_{\downarrow 2}}^{(j)}) + \mathbf{G}_C\mathbf{A}_3\mathbf{G}_{2,3}(c_{U_2}^{(j-1)} + c_{Z_{\downarrow 2}}^{(j-1)}) + \mathbf{G}_{1,4}(c_{U_1}^{(j)} + c_{Z_{\uparrow 1}}^{(j)}) + \widetilde{g_{1,4}}^{(j)}(c_{Z_{\downarrow 1}}),
\end{aligned}$$

where, unlike earlier,

$$\widetilde{g_{1,4}}^{(j)}(c_{Z_{\downarrow 1}}) = \mathbf{G}_{1,4}c_{Z_{\downarrow 1}}^{(j)} + \mathbf{G}_C\mathbf{A}_3\mathbf{G}_{1,3}c_{Z_{\downarrow 1}}^{(j-1)}.$$

Note that, again the appropriate choice of $\mathbf{A}_1$ and $\mathbf{A}_3$ has ensured that no contribution from the S_1 codeword is observed at destination 4.

Decoding at destination 3 proceeds as in the above scheme. On the other hand, destination 4 decodes, at the end of each block, (i) first, the U_2 and U_1 codewords jointly treating all other signals and interference as noise, and then from the residual signal (ii) $Z_{\downarrow 2}$ codeword treating interference as noise. The conditions on the rates for successful decoding are given below.

$$\begin{aligned}
r_{U_1} &\le [n_{1,4} - n_C],\\
r_{U_2} &\le n_C,\\
r_{S_1} &\le n_C - [n_{2,3} - (n_{1,3} - n_{1,4})]_+,\\
r_{Z_{\uparrow 1}} &\le [n_{1,3} - n_{1,4} - n_{2,3}]_+,\\
r_{Z_{\downarrow 1}} &\le [n_{2,3} - n_C]_+,\\
r_{U_2} &\le \min(n_{2,3}, n_{2,4}),\\
r_{U_1} &\le [n_{1,4} - [n_{2,4} - n_{2,3}]_+],\\
r_{U_2} + r_{U_1} &\le \max(n_{2,4}, n_{1,4}) - [n_{2,4} - n_{2,3}]_+,\\
r_{Z_{\downarrow 2}} &\le [n_{2,4} - n_{2,3}]_+.
\end{aligned}$$

where the first five conditions ensure successful decoding at destination 3 and the rest of the three conditions does the same for decoding at destination 4. Upon simplifying, we may conclude that a sum-rate equal to the minimum of the following terms is achievable

$$\begin{aligned}
&n_{1,3} + [n_{2,4} - n_{2,3}]_+,\\
&n_{1,3} - n_{1,4} + \max(n_{1,4}, n_{2,4}), \text{and}\\
&n_{1,3} + n_C.
\end{aligned}$$

It is easy to check that this is what our upperbound evaluates to under case (b). Thus, we have also proved achievability under case (b).

Regime (iii): $\min(n_{1,3}, n_{2,4}) < n_C$. We employ Theorem 3 with the following choices for p_{U_1,X_1}, p_{U_2,X_2}, p_{X_3}, p_{X_4}, $p_{V_3|Y_3}$, and $p_{V_4|Y_4}$: U_1, U_2, X_3, X_4 are independent and identically distributed uniformly over $\mathbb{F}^n$. Z_1 and Z_2, respectively, are uniformly distributed over $\mathcal{F}_{n_{1,4}}$ and $\mathcal{F}_{n_{2,3}}$, respectively. They are independent of U_1, U_2, X_3, X_4 and of each other. We define

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

$p_{V_3|Y_3}$ and $p_{V_4|Y_4}$ are defined by the following deterministic *test-channels*

$$\begin{aligned}
V_3 &= \mathbf{S}^{[n_{1,3}-n_{1,4}]_+}Y_3,\\
V_4 &= \mathbf{S}^{[n_{2,4}-n_{2,3}]_+}Y_4.
\end{aligned}$$

Note that this amounts to the destinations truncating their observations to the level at which their own private-codewords (*e.g.*, Z_1 in the case of destination 3) are received. Thus, the quantized observations

V's only contain information on the public-codewords (the U's). This is consistent with the intuition that the information they forward on to the other destination is utilized to recover the public-codewords.

With these choices, the conditions on the non-negative rates for achievability from Theorem 3 work out to

$$
\begin{aligned}
r_{X_1} &\leq [n_{1,3}-n_{1,4}]_+,\\
r_{U_1} &\leq \max\left(n_{1,3}-[n_{1,3}-n_{1,4}]_+, n_{1,4}-[n_{2,4}-n_{2,3}]_+\right),\\
r_{U_1} &\leq n_{1,3}-[n_{1,3}-n_{1,4}]_+ + r_4,\\
r_{U_2} &\leq \max\left(n_{2,3}-[n_{1,3}-n_{1,4}]_+, n_{2,4}-[n_{2,4}-n_{2,3}]_+\right),\\
r_{U_2} &\leq [n_{2,3}-[n_{1,3}-n_{1,4}]_+]_+ + r_4,\\
r_4 &\leq [n_C-[n_{1,3}-n_{1,4}]_+]_+,\\
r_{U_1}+r_{U_2} &\leq \begin{cases}
\max\Big((n_{1,3}-[n_{1,3}-n_{1,4}]_+)+(n_{2,4}-[n_{2,4}-n_{2,3}]_+),\\
\qquad [n_{1,4}-[n_{2,4}-n_{2,3}]_+]_+ + [n_{2,3}-[n_{1,3}-n_{1,4}]_+]_+\Big),\\
\qquad\qquad \text{if } n_{1,3}+n_{2,4}\neq n_{1,4}+n_{2,3}\\
\max\Big((n_{1,3}-[n_{1,3}-n_{1,4}]_+),(n_{2,4}-[n_{2,4}-n_{2,3}]_+),\\
\qquad (n_{1,4}-[n_{2,4}-n_{2,3}]_+),(n_{2,3}-[n_{1,3}-n_{1,4}]_+)\Big),\\
\qquad\qquad \text{otherwise,}
\end{cases}\\
r_{U_1}+r_{U_2} &\leq \max(n_{1,3},n_{2,3})-[n_{1,3}-n_{1,4}]_+ + r_4,\\
r_{U_1}+r_4 &\leq \max\Big(\max(n_{1,3},n_C)-[n_{1,3}-n_{1,4}]_+,\\
&\qquad\qquad [n_{1,4}-[n_{2,4}-n_{2,3}]_+]_+ + [n_C-[n_{1,3}-n_{1,4}]_+]_+\Big),\\
r_{U_1}+r_4 &\leq n_{1,3}-[n_{1,3}-n_{1,4}]_+ + r_4,\\
r_{U_2}+r_4 &\leq \max\Big(\max(n_{2,3},n_C)-[n_{1,3}-n_{1,4}]_+,\\
&\qquad\qquad (n_{2,4}-[n_{2,4}-n_{2,3}]_+) + [n_C-[n_{1,3}-n_{1,4}]_+]_+\Big),\\
r_{U_2}+r_4 &\leq [n_{2,3}-[n_{1,3}-n_{1,4}]_+]_+ + r_4,\\
r_{U_2}+r_4+r_{U_1} &\leq \begin{cases}
\max\Big((n_{1,3}-[n_{1,3}-n_{1,4}]_+)+(n_{2,4}-[n_{2,4}-n_{2,3}]_+),\\
\qquad [n_{1,4}-[n_{2,4}-n_{2,3}]_+]_+ + [n_{2,3}-[n_{1,3}-n_{1,4}]_+]_+,\\
\qquad [n_C-[n_{1,3}-n_{1,4}]_+]_+ + (\max(n_{1,4},n_{2,4})-[n_{2,4}-n_{2,3}]_+)\Big),\\
\qquad\qquad \text{if } n_{1,3}+n_{2,4}\neq n_{1,4}+n_{2,3}\\
\max\Big((n_{1,3}-[n_{1,3}-n_{1,4}]_+),\ (n_{2,4}-[n_{2,4}-n_{2,3}]_+),\\
\qquad (n_{1,4}-[n_{2,4}-n_{2,3}]_+),(n_{2,3}-[n_{1,3}-n_{1,4}]_+),\\
\qquad [n_C-[n_{1,3}-n_{1,4}]_+]_+ + (\max(n_{1,4},n_{2,4})-[n_{2,4}-n_{2,3}]_+)\Big),\\
\qquad\qquad \text{otherwise,}
\end{cases}\\
r_{U_2}+r_4+r_{U_1} &\leq \max(n_{1,3},n_{2,3},n_C)-[n_{1,3}-n_{1,4}]_+ + r_4,
\end{aligned}
$$

and the corresponding inequalities with subscripts 1 and 2 exchanged, and 3 replaced with 4. Applying Fourier-Motzkin elimination to these conditions, we can show that a sum-rate equal to the minimum of the following terms is achievable in this regime.

$$\begin{aligned}
u_2 &= \max(n_{2,4}, n_{2,3}) + \left(\max(n_{1,3}, n_{2,3}, n_C) - n_{2,3}\right), \\
u_3 &= \max(n_{1,3}, n_{1,4}) + \left(\max(n_{2,4}, n_{1,4}, n_C) - n_{1,4}\right), \\
u_4 &= \max(n_{1,3}, n_C) + \max(n_{2,4}, n_C), \text{ and} \\
u_5 &= \begin{cases} \max(n_{1,3} + n_{2,4}, n_{1,4} + n_{2,3}), & \text{if } n_{1,3} - n_{2,3} \neq n_{1,4} - n_{2,4}, \\ \max(n_{1,3}, n_{2,4}, n_{1,4}, n_{2,3}), & \text{otherwise.} \end{cases}
\end{aligned}$$

Appendix D
Proof of the achievability of Theorem 2

The proof of Theorem 2 will follow the proof of Theorem 1 closely. We first make the following definitions:

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

Let us observe that the minimum of the following four terms u_1', u_2', u_3', and u_4' are within a constant (7 bits) of the minimum of the corresponding unprimed terms, u_1, u_2, u_3, and u_4.

$$\begin{aligned}
u_1' &= \max(n_{1,3} - n_{1,4} + n_C, n_{2,3}, n_C) + \max(n_{2,4} - n_{2,3} + n_C, n_{1,4}, n_C), \\
u_2' &= \max(n_{1,3}, n_{2,3}) + \left(\max(n_{2,4}, n_{2,3}, n_C) - n_{2,3}\right), \\
u_3' &= \max(n_{2,4}, n_{1,4}) + \left(\max(n_{1,3}, n_{1,4}, n_C) - n_{1,4}\right), \\
u_4' &= \max(n_{1,3}, n_C) + \max(n_{2,4}, n_C).
\end{aligned}$$

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

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

which is also within a constant (2 bits) of u_5, is achievable. We again consider the same three regimes as in Appendix C.

Regime (i): $|g_C| \leq |g_{i,j}|, \ i \in \{1,2\}, j \in \{3,4\}$.

We will assume that $|g_C| > 1$. If $|g_C| \leq 1$, it can be verified that our upperbound is not more than 4 bits away from the upperbound for the corresponding Gaussian interference channel without a cooperation link (*i.e.* $g_C = 0$) in [8] which itself is known to be achievable with a gap of at most 2 bits. Hence overall, if $|g_C| \leq 1$, the upperbound is achievable with a gap of 6 bits.

Further, we will show achievability only for the case where $2|g_C| \leq \min(|g_{1,4}|, |g_{2,3}|)$. Note that the upperbounds change by at most 2 bits if we do not impose this restriction (but still maintain the restriction that $|g_C| \leq |g_{i,j}|, \ i \in \{1,2\}, j \in \{3,4\}$), and achievability of a given sum-rate at a lower value of $|g_C|$ implies its achievability for all larger values of $|g_C|$ (*i.e.*, the sum-capacity is monotonic in $|g_C|$) provided the rest of the channel coefficients remain the same. Hence, showing achievability under this restricted $|g_C|$ regime implies a proof of achievability for the whole regime with a further gap of 2 bits from the upperbound.

Let us define

$$A_1 \overset{\text{def}}{=} \left| \frac{g_{1,4}g_{2,3}}{g_{1,3}g_C} \right|,$$
$$A_2 \overset{\text{def}}{=} \left| \frac{g_{1,4}g_{2,3}}{g_{2,4}g_C} \right|.$$

We first note that in regime (i) with the additional assumptions we made above, if both $A_1, A_2 \geq 1/2$, then our upperbound can be shown to be not more than 10 bits from the upperbound in [8] for the Gaussian interference channel without a cooperative link. Since that upperbound is known to be achievable within two bits, the gap to the upperbound is at most 12. Hence, we will only consider the other three possibilities: (a) $A_1 < 1/2, \quad A_2 < 1/2$, (b) $A_1 < 1/2, \quad A_2 \geq 1/2$, and (c) $A_1 \geq 1/2, \quad A_2 < 1/2$. We will show achievability for cases (a) and (b). Case (c) will follow from case (b) by symmetry.

The coding scheme will very closely resemble the one we used for regime (i) in Appendix C. We repeat all the details below for completeness.

For case (a), let us consider the following block-Markov scheme with superposition coding. Let $U_k, S_k, Z_{\uparrow k}, Z_{\downarrow k},\ k$ $1, 2$ be independent auxiliary random variables with marginal distributions $p_{U_k}, p_{S_k}, p_{Z_{\uparrow k}}, p_{Z_{\downarrow k}}$. The alphabet for these random variables is the set of complex numbers. Corresponding to these random variables, random codebooks of blocklength-T and rates $r_{U_k}, r_{S_k}, r_{Z_{\uparrow k}}, r_{Z_{\downarrow k}}$, respectively, are defined as usual. For instance, the U_1-codebook, denoted by $\mathcal{C}_{U_1}$, is of size $2^{T(r_{U_1}-\epsilon)}$ is generated by choosing the T elements of each of the codewords independently according to p_{U_1}. These codewords will be denoted by $c_{U_1}(m_{U_1})$ where $m_{U_1} \in \{1, \ldots, 2^{T(r_{U_1}-\epsilon)}\}$. The blocks will be indexed by $j = 1, 2, \ldots, J$. The message transmitted by source 1 using the U_1-codebook in block-j will be denoted by $m_{U_1}(j)$, and the corresponding codeword by $c_{U_1}^{(j)} = c_{U_1}(m_{U_1}(j))$.

For block-j, sources transmit the following blocklength-T vectors

$$X_1^{(j)} = c_{U_1}^{(j)} + c_{Z_{\uparrow 1}}^{(j)} + c_{Z_{\downarrow 1}}^{(j)} + f_{S_1}^{(j)},$$
$$X_2^{(j)} = c_{U_2}^{(j)} + c_{Z_{\uparrow 2}}^{(j)} + c_{Z_{\downarrow 2}}^{(j)} + f_{S_2}^{(j)},$$

where

$$f_{S_1}^{(j)} = c_{S_1}^{(j)} + A_1 c_{S_1}^{(j+1)} + \ldots + A_1^{J-j} c_{S_1}^{(J)},$$
$$f_{S_2}^{(j)} = c_{S_2}^{(j)} + A_2 c_{S_2}^{(j+1)} + \ldots + A_2^{J-j} c_{S_2}^{(J)}.$$

In the sequel we will ensure that the rates of the codebooks are such that from observing $Y_3^{(j)}$, the j-th block observed by destination 3, it (destination 3) can decode with a high probability of success the codewords $c_{U_1}^{(j)}$, $c_{S_1}^{(j)}$, and $c_{Z_{\uparrow 1}}^{(j)}$, for all j. Similarly, we will make sure that destination 4 will successfully decode $c_{U_2}^{(j)}$, $c_{S_2}^{(j)}$, and $c_{Z_{\uparrow 2}}^{(j)}$ from $Y_4^{(j)}$. Then, the destinations will transmit, respectively, for $j = 1, 2, \ldots, J-1$,

$$X_3^{(j+1)} = A_3 \left(Y_3^{(j)} - g_{1,3} \left(c_{U_1}^{(j)} + c_{S_1}^{(j)} + c_{U_{\uparrow 1}}^{(j)} \right) - g_C A_4 g_{1,4} \left(c_{U_1}^{(j-1)} + c_{Z_{\uparrow 1}}^{(j-1)} \right) \right),$$
$$X_4^{(j+1)} = A_4 \left(Y_4^{(j)} - g_{2,4} \left(c_{U_2}^{(j)} + c_{S_2}^{(j)} + c_{U_{\uparrow 2}}^{(j)} \right) - g_C A_3 g_{2,3} \left(c_{U_2}^{(j-1)} + c_{Z_{\uparrow 2}}^{(j-1)} \right) \right),$$

where, A_3 and A_4 are defined below

$$A_4 = -\frac{g_{2,3}}{g_C g_{2,4} A_2},$$
$$A_3 = -\frac{g_{1,4}}{g_C g_{1,3} A_1},$$

The choices for the distributions and the fact that the power constraints are satisfied will be taken up in the sequel.

With these, the received signals at the destinations are

$$Y_3^{(j)} = \overline{g_{1,3}}^{(j)}(c_{U_1} + c_{Z_{\uparrow 1}}) + g_{1,3}(f_{S_1}^{(j)}) + \widetilde{g_{1,3}}^{(j)}(c_{Z_{\downarrow 1}}) + g_{2,3}(c_{U_2}^{(j)} + c_{Z_{\uparrow 2}}^{(j)}) + \widetilde{g_{2,3}^*}^{(j)}(c_{Z_{\downarrow 2}}),$$
$$Y_4^{(j)} = \overline{g_{2,4}}^{(j)}(c_{U_2} + c_{Z_{\uparrow 2}}) + g_{2,4}(f_{S_2}^{(j)}) + \widetilde{g_{2,4}}^{(j)}(c_{Z_{\downarrow 2}}) + g_{1,4}(c_{U_1}^{(j)} + c_{Z_{\uparrow 1}}^{(j)}) + \widetilde{g_{1,4}^*}^{(j)}(c_{Z_{\downarrow 1}}),$$

where the functions are as defined below. Note that Y_3 does not have any terms which depend on S_2-codewords (and similarly, Y_4 does not involve any terms containing S_1-codewords). This was achieved by the appropriate choices above for A_1 and A_4 (respectively, A_2 and A_3). Here

$$\overline{g_{1,3}}^{(j)}(c_{U_1} + c_{Z_{\uparrow 1}}) = g_{1,3}(c_{U_1}^{(j)} + c_{Z_{\uparrow 1}}^{(j)}) + g_C A_4 g_{1,4}(c_{U_1}^{(j-1)} + c_{Z_{\uparrow 1}}^{(j-1)}),$$
$$g_{2,3}(c_{U_2}^{(j)} + c_{Z_{\uparrow 2}}^{(j)}) = g_{2,3}(c_{U_2}^{(j)} + c_{Z_{\uparrow 2}}^{(j)}),$$

$$\widetilde{g_{1,3}}^{(j)}(c_{Z_{\downarrow 1}}) = \sum_{i\in\{j,j-2,\ldots,1\}} (g_C A_4 g_C A_3)^{\frac{j-i}{2}} g_{1,3} c_{Z_{\downarrow 1}}^{(i)} + \sum_{i\in\{j-1,j-3,\ldots,1\}} (g_C A_4 g_C A_3)^{\frac{j-1-i}{2}} g_C A_4 g_{1,4} c_{Z_{\downarrow 1}}^{(i)},$$
$$\widetilde{g_{2,3}^*}^{(j)}(c_{Z_{\downarrow 2}}) = \sum_{i\in\{j,j-2,\ldots,1\}} (g_C A_4 g_C A_3)^{\frac{j-i}{2}} g_{2,3} c_{Z_{\downarrow 2}}^{(i)} + \sum_{i\in\{j-1,j-3,\ldots,1\}} (g_C A_4 g_C A_3)^{\frac{j-1-i}{2}} g_C A_4 g_{2,4} c_{Z_{\downarrow 2}}^{(i)}$$
$$+ \sum_{i\in\{j,j-2,\ldots,1\}} (g_C A_4 g_C A_3)^{\frac{j-i}{2}} N_1^{(i)} + \sum_{i\in\{j-1,j-3,\ldots,1\}} g_C A_4 (g_C A_4 g_C A_3)^{\frac{j-1-i}{2}} N_2^{(i)},$$

and the functions $\widetilde{g_{2,4}}^{(j)}(c_{Z_{\downarrow 2}})$, $g_{1,4}(c_{U_1}^{(j)} + c_{Z_{\uparrow 1}}^{(j)})$, and $\widetilde{g_{1,4}^*}^{(j)}(c_{Z_{\downarrow 1}})$ are defined similarly.

As before, destinations perform decoding in two phases. At the end of every block j, the destinations decode the $c_U^{(j)}$, $c_S^{(j)}$ and $c_{Z_\uparrow}^{(j)}$ intended for them (in that order). We call this *phase 1 decoding*. At the end of block-J, the decoders perform a *phase 2 decoding* where it decodes the following codewords for all blocks, *i.e.*, for each $j = 1, 2, \ldots, J$, the destinations decode $c_U^{(j)}$ intended for the other user and $c_{Z_\downarrow}^{(j)}$ intended for itself (in that order). In both of the phases, the decodings are performed in the order mentioned above treating all the undecoded codewords and other interference as noise. Below, we will specify the distributions employed and evaluate the conditions on the rates to ensure successful decoding. This will establish achievability for regime (i).

The auxiliary random variables are all Gaussian with the following powers:

$$\sigma_{U_1}^2 = 1/K,$$
$$\sigma_{S_1}^2 = \frac{1}{K}\left|\frac{g_C}{g_{1,4}}\right|^2,$$
$$\sigma_{Z_{\uparrow 1}}^2 = \frac{1}{K}\left|\frac{1}{g_{1,4}}\right|^2,$$
$$\sigma_{Z_{\downarrow 1}}^2 = \frac{1}{K}\left|\frac{g_{2,3}}{g_{1,3} g_C}\right|^2,$$

where K is a constant which will be chosen presently to satisfy the power constraint. The power allocation for the auxiliary random variables for source 2 is chosen similarly.

Destination 3 transmits

$$X_3^{(j)} = A_3\left(Y_3^{(j)} - g_{1,3}\left(c_{U_1}^{(j)} + c_{S_1}^{(j)} + c_{U_{\uparrow 1}}^{(j)}\right)\right)$$
$$= A_3\left(\widetilde{g_{1,3}}^{(j)}(c_{Z_{\downarrow 1}}) + g_{2,3}(c_{U_2}^{(j)} + c_{Z_{\uparrow 2}}^{(j)}) + \widetilde{g_{2,3}^*}^{(j)}(c_{Z_{\downarrow 2}})\right)$$

If $2|g_C| \leq |g_{1,4}|, |g_{2,3}|$ and $|g_C| \geq 1$, under the above power allocation, the average power of these terms can be shown to be

$$\begin{aligned}
(1/T)\|A_3\widetilde{g_{1,3}}^{(j)}(c_{Z_{\downarrow 1}})\|^2 &\leq \frac{|g_{2,3}|^2(2/K)}{|g_C|^2} \\
(1/T)\|A_3 g_{2,3}(c_{U_2}^{(j)} + c_{Z_{\uparrow 2}}^{(j)})\|^2 &\leq 2/K \\
(1/T)\mathbb{E}\left[\|A_3\widetilde{g_{2,3}^*}^{(j)}(c_{Z_{\downarrow 2}})\|^2\right] &\leq \frac{(2/K)+2}{|g_{2,3}|^2}.
\end{aligned}$$

We can easily verify from this that if we choose $K = 9$, the power constraints at all the transmitters are satisfied. The rates supported by the above scheme for source 1 are given by the following set of conditions

$$\begin{aligned}
r_{U_1} &\leq \log\left(1 + \frac{\frac{|g_{1,3}|^2}{K}}{\frac{2|g_{1,3}g_C|^2}{K|g_{1,4}|^2} + \frac{|g_{1,3}|^2}{K|g_{1,4}|^2} + \frac{2|g_{2,3}|^2}{K|g_C|^2} + \frac{2|g_{2,3}|^2}{K} + \frac{2}{K} + 2}\right), \\
r_{U_1} &\leq \log\left(1 + \frac{\frac{|g_{1,4}|^2}{K}}{\frac{|g_{1,4}|^2}{K|g_C|^2} + \frac{2}{K} + \frac{2}{K} + 2}\right), \\
r_{S1} &\leq \log\left(1 + \frac{\frac{|g_{1,3}g_C|^2}{K|g_{1,4}|^2}}{\frac{2|g_{2,3}|^2}{K} + \frac{|g_{1,3}|^2}{K|g_{1,4}|^2} + \frac{2|g_{2,3}|^2}{K|g_C|^2} + \frac{2|g_{2,3}|^2}{K} + \frac{2}{K} + 2}\right), \\
r_{Z_{\uparrow 1}} &\leq \log\left(1 + \frac{\frac{|g_{1,3}|^2}{K|g_{1,4}|^2}}{\frac{2|g_{2,3}|^2}{K} + \frac{2|g_{2,3}|^2}{K|g_C|^2} + \frac{2|g_{2,3}|^2}{K} + \frac{2}{K} + 2}\right), \\
r_{Z_{\downarrow 1}} &\leq \log\left(1 + \frac{\frac{|g_{2,3}|^2}{K|g_C|^2}}{\frac{2}{K} + 2}\right).
\end{aligned}$$

The first constraint on r_{U_1} comes from the phase 1 decoding at destination 3, while the second condition is from the phase 2 decoding at destination 4. A similar set of constraints apply for the rates achievable by source 2.

Simplifying, these constraints imply that rates which satisfy the following are also achievable

$$\begin{aligned}
r_{U_1} &\leq [n_{1,4} - n_C] - 4, \\
r_{U_1} &\leq n_C - 4, \\
r_{S_1} &\leq n_C - [n_{2,3} - (n_{1,3} - n_{1,4})]_+ - 4, \\
r_{Z_{\uparrow 1}} &\leq [n_{1,3} - n_{1,4} - n_{2,3}]_+ - 4, \\
r_{Z_{\downarrow 1}} &\leq [n_{2,3} - n_C]_+ - 5.
\end{aligned}$$

Combining all these, an achievable sum-rate is given by

$$R_{\text{sum}} = \max(n_{1,3} - n_{1,4} + n_C, n_{1,3} - n_C) + \max(n_{2,4} - n_{2,3} + n_C, n_{2,4} - n_C) - 34$$

This, combined with the fact that the achievability of a given sum-rate at a lower value of $|g_C|$ implies its achievability for all larger values of $|g_C|$ provided the rest of the channel coefficients remain the same allows us to conclude that the minimum of the following three terms is achievable with a gap of 34.

$$\begin{aligned}
&n_{1,3} - n_{1,4} + n_C + n_{2,4} - n_{2,3} + n_C, \\
&n_{1,3} + n_{2,4} - n_{1,4}, \ \text{and} \\
&n_{2,4} + n_{1,3} - n_{2,3}.
\end{aligned}$$

The above minimum is also the minimum of u_1', u_2', u_3' under case (a). This shows achievability in case (a).

Let us now consider case (b) where $A_1 < 1/2$, but $A_2 \geq 1/2$. The achievable strategy we use for this case involves source 1 transmitting according to a scheme similar to the one above while source 2 employs a superposition coding scheme similar to that of Han and Kobayashi for the interference channel without a cooperative link. In particular, only node 3 uses its transmission capabilities.

The codebooks and the choice of distributions (power allocations) for source 1 is exactly as above, except for the choice of rates which will be presented in the sequel. Source 2 uses only the following codebooks: U_2 and $Z_{\downarrow 2}$. Moreover, the choice of $\sigma^2_{Z_{\downarrow 2}}$ is different.

$$\sigma^2_{Z_{\downarrow 2}} = \frac{1}{K}\frac{1}{|g_{2,3}|^2}.$$

Exactly as in the above scheme, destination 3 performs a two-phase decoding and transmits a scaled version of the residual signals after the first phase of decoding of the previous block. The scaling factor A_3 is the same as above. Destination 4, on the other hand, performs only a single phase of decoding where U_2, U_1 and $Z_{\downarrow 2}$ codewords are decoded. As mentioned earlier, destination 4 does not transmit anything, *i.e.*, $A_4 = 0$. The received signals can be seen to be

$$\begin{aligned}
Y_3^{(j)} &= g_{1,3}(c_{U_1}^{(j)} + c_{Z_{\uparrow 1}}^{(j)} + f_{S_1}^{(j)} + c_{Z_{\downarrow 1}}^{(j)}) + g_{2,3}^*(c_{U_2}^{(j)} + c_{Z_{\downarrow 2}}^{(j)}),\\
Y_4^{(j)} &= g_{2,4}(c_{U_2}^{(j)} + c_{Z_{\downarrow 2}}^{(j)}) + g_C A_3 g_{2,3}(c_{U_2}^{(j-1)} + c_{Z_{\downarrow 2}}^{(j-1)}) + g_{1,4}(c_{U_1}^{(j)} + c_{Z_{\uparrow 1}}^{(j)}) + \widetilde{g_{1,4}^*}^{(j)}(c_{Z_{\downarrow 1}}),
\end{aligned}$$

where, unlike above,

$$\widetilde{g_{1,4}^*}^{(j)}(c_{Z_{\downarrow 1}}) = g_{1,4}c_{Z_{\downarrow 1}}^{(j)} + g_C A_3 g_{1,3} c_{Z_{\downarrow 1}}^{(j-1)} + N_2^{(j)} + g_C A_3 N_2^{(j-1)},$$

and the other functions are as before. Note that, again the appropriate choice of A_1 and A_3 has ensured that no contribution from the S_1 codeword is observed at destination 4. The fact that the power constraints are satisfied at all the transmitters under the earlier choice of $K = 9$ is easy to verify.

Decoding at destination 3 proceeds as in the above scheme. Whereas, destination 4 decodes, at the end of each block, (i) first, the U_2 and U_1 codewords jointly treating all other signals and interference as noise, and then from the residual signal (ii) $Z_{\downarrow 2}$ codeword treating interference as noise. The conditions on the rates for successful decoding are given below.

$$\begin{aligned}
r_{U_1} &\leq \log\left(1 + \frac{|g_{1,3}|^2/K}{2|g_{1,3}g_C|^2/(K|g_{1,4}|^2) + |g_{1,3}|^2/(K|g_{1,4}|^2) + |g_{2,3}|^2/(K|g_C|^2) + |g_{2,3}|^2 + 1}\right),\\
r_{S_1} &\leq \log\left(1 + \frac{|g_{1,3}g_C|^2/(K|g_{1,4}|^2)}{2|g_{2,3}|^2/K + |g_{1,3}|^2/(K|g_{1,4}|^2) + |g_{2,3}|^2/(K|g_C|^2) + |g_{2,3}|^2 + 1}\right),\\
r_{Z_{\uparrow 1}} &\leq \log\left(1 + \frac{|g_{1,3}|^2/(K|g_{1,4}|^2)}{2|g_{2,3}|^2/K + |g_{2,3}|^2/(K|g_C|^2) + |g_{2,3}|^2 + 1}\right),\\
r_{U_2} &\leq \log\left(1 + \frac{|g_{2,3}|^2/K}{|g_{2,3}|^2/(K|g_C|^2) + 1 + 1}\right),\\
r_{Z_{\downarrow 1}} &\leq \log\left(1 + \frac{|g_{2,3}|^2/(K|g_C|^2)}{1+1}\right),
\end{aligned}$$

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

$$r_{Z_{\downarrow 2}} \leq \log\left(1 + \frac{|g_{2,4}|^2/(K|g_{2,3}|^2)}{1/K + 3/2K + 3/2}\right),$$

where the first five conditions ensure successful decoding at destination 3 and the rest of the three conditions does the same for decoding at destination 4. We may simplify the terms to conclude that rates which satisfy all the conditions below are achievable.

$$\begin{aligned}
r_{U_1} &\leq [n_{1,4} - n_C] - 4,\\
r_{U_2} &\leq n_C - 4,\\
r_{S_1} &\leq n_C - [n_{2,3} - (n_{1,3} - n_{1,4})]_+ - 4,\\
r_{Z_{\uparrow 1}} &\leq [n_{1,3} - n_{1,4} - n_{2,3}]_+ - 4,\\
r_{Z_{\downarrow 1}} &\leq [n_{2,3} - n_C]_+ - 5,\\
r_{U_2} &\leq \min(n_{2,3}, n_{2,4}) - 5,\\
r_{U_1} &\leq [n_{1,4} - [n_{2,4} - n_{2,3}]_+] - 5,\\
r_{U_2} + r_{U_1} &\leq \max(n_{2,4}, n_{1,4}) - [n_{2,4} - n_{2,3}]_+ - 5,\\
r_{Z_{\downarrow 2}} &\leq [n_{2,4} - n_{2,3}]_+ - 4.
\end{aligned}$$

Upon simplifying, we may conclude that a sum-rate equal to the minimum of the following terms is achievable within 28 bits

$$\begin{aligned}
&n_{1,3} + [n_{2,4} - n_{2,3}]_+,\\
&n_{1,3} - n_{1,4} + \max(n_{1,4}, n_{2,4}), \text{and}\\
&n_{1,3} + n_C.
\end{aligned}$$

This is the minimum of u_1', u_2', and u_3' under case (b). Thus, we have shown achievability under case (b) as well.

Overall, we have shown achievability of the upperbound in regime (i) with a gap of at most 43 bits.

Regime (ii): As in Appendix C, achievability in regime (i) implies the achievability in regime (ii) as well since in this regime

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

and $u_2'(n_C)$, $u_3'(n_C)$, $u_4'(n_C)$ and u_5' are constants.

Regime (iii): Note that we proved Theorem 3 for discrete alphabets, but the extension to the continuous alphabet case is standard and we will assume that version for proving achievability here.

We apply Theorem 3 with the following choices for the auxiliary random variables. $Z_1, Z_2, U_1, U_2, X_3, X_4$ are independent and identically distributed zero-mean Gaussian random variables. The variances are, respectively

$$\begin{aligned}
\sigma_{Z_1}^2 &= \frac{1/K}{\max(1, |g_{1,4}|^2)},\\
\sigma_{Z_2}^2 &= \frac{1/K}{\max(1, |g_{2,3}|^2)},\\
\sigma_{U_1}^2 = \sigma_{U_2}^2 &= 1/K,\\
\sigma_{X_3}^2 = \sigma_{X_4}^2 &= 1,
\end{aligned}$$

where we set $K = 1/2$. Further,

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

It is easy to see that this satisfies the power constraint since $K = 1/2$. $p_{V_3|Y_3}$ and $p_{V_4|Y_4}$ are defined by the following *test-channels*

$$V_3 = Y_3 + Q_3, \text{ and}$$
$$V_4 = Y_4 + Q_4,$$

where Q_3 and Q_4 are independent, zero-mean Gaussian random variables which are also independent of Y_3, Y_4 and all the other auxiliary random variables. Their variances are, respectively

$$\sigma^2_{Q_3} = \max\left(1, \frac{\max(1, |g_{1,3}|^2)}{\max(1, |g_{1,4}|^2)}\right), \text{ and}$$
$$\sigma^2_{Q_4} = \max\left(1, \frac{\max(1, |g_{2,4}|^2)}{\max(1, |g_{2,3}|^2)}\right).$$

Note that this choice amounts to the destinations quantizing their observations with the quantization noise level set to the power level at which their own private-codewords (*e.g.*, Z_1 in the case of destination 3) is received. This is consistent with the intuition that the information they forward on to the other destination is used to recover only the public-codewords.

With these choices, it can be shown that Theorem 3 implies that the non-negative rates $r_{X_1}, r_{X_2}, r_{U_1}, r_{U_2}, r_3, r_4$ can be achieved if the following conditions are satisfied

$$
\begin{aligned}
r_{X_1} &\le [n_{1,3} - n_{1,4}]_+ - 2, \\
r_{U_1} &\le \max\left(n_{1,3} - [n_{1,3} - n_{1,4}]_+, n_{1,4} - [n_{2,4} - n_{2,3}]_+\right) - \log 36, \\
r_{U_1} &\le n_{1,3} - [n_{1,3} - n_{1,4}]_+ + r_4 - 3, \\
r_{U_2} &\le \max\left(n_{2,3} - [n_{1,3} - n_{1,4}]_+, n_{2,4} - [n_{2,4} - n_{2,3}]_+\right) - \log 36, \\
r_{U_2} &\le [n_{2,3} - [n_{1,3} - n_{1,4}]_+]_+ + r_4 - 3, \\
r_4 &\le [n_C - [n_{1,3} - n_{1,4}]_+]_+ - 1, \\
r_{U_1} + r_{U_2} &\le \log\Bigg(1 + \left|\frac{g_{1,3}}{\alpha_1}\right|^2 + \left|\frac{g_{2,4}}{\alpha_2}\right|^2 + \left|\frac{g_{1,4}}{\alpha_2}\right|^2 + \left|\frac{g_{2,3}}{\alpha_1}\right|^2 \\
&\qquad + \left|\frac{g_{1,3}g_{2,4}}{\alpha_1\alpha_2}\right|^2 + \left|\frac{g_{1,4}g_{2,3}}{\alpha_1\alpha_2}\right|^2 - 2\left|\frac{g_{1,3}g_{2,4}g_{1,4}g_{2,3}}{\alpha_1^2\alpha_2^2}\right|\cos\theta\Bigg) - \log 36, \\
r_{U_1} + r_{U_2} &\le \max(n_{1,3}, n_{2,3}) - [n_{1,3} - n_{1,4}]_+ + r_4 - 3, \\
r_{U_1} + r_4 &\le \max\Bigg(\max(n_{1,3}, n_C) - [n_{1,3} - n_{1,4}]_+, \\
&\qquad [n_{1,4} - [n_{2,4} - n_{2,3}]_+]_+ + [n_C - [n_{1,3} - n_{1,4}]_+]_+\Bigg) - \log 36, \\
r_{U_1} + r_4 &\le n_{1,3} - [n_{1,3} - n_{1,4}]_+ + r_4 - 3, \\
r_{U_2} + r_4 &\le \max\Bigg(\max(n_{2,3}, n_C) - [n_{1,3} - n_{1,4}]_+, \\
&\qquad (n_{2,4} - [n_{2,4} - n_{2,3}]_+) + [n_C - [n_{1,3} - n_{1,4}]_+]_+\Bigg) - \log 36, \\
r_{U_2} + r_4 &\le [n_{2,3} - [n_{1,3} - n_{1,4}]_+]_+ + r_4 - 3, \\
r_{U_2} + r_4 + r_{U_1} &\le \log\Bigg(1 + \left|\frac{g_{1,3}}{\alpha_1}\right|^2 + \left|\frac{g_{2,4}}{\alpha_2}\right|^2 + \left|\frac{g_{1,4}}{\alpha_2}\right|^2 + \left|\frac{g_{2,3}}{\alpha_1}\right|^2 + \left|\frac{g_C}{\alpha_1}\right|^2 + \left|\frac{g_C g_{1,4}}{\alpha_1\alpha_2}\right|^2 + \left|\frac{g_C g_{2,4}}{\alpha_1\alpha_2}\right|^2
\end{aligned}
$$

$$+\left|\frac{g_{1,3}g_{2,4}}{\alpha_1\alpha_2}\right|^2+\left|\frac{g_{1,4}g_{2,3}}{\alpha_1\alpha_2}\right|^2-2\left|\frac{g_{1,3}g_{2,4}g_{1,4}g_{2,3}}{\alpha_1^2\alpha_2^2}\right|\cos\theta\Bigg)-\log 36,\ \text{and}$$

$$r_{U_2}+r_4+r_{U_1}\leq \max(n_{1,3},n_{2,3},n_C)-[n_{1,3}-n_{1,4}]_+ +r_4-3,$$

and the corresponding inequalities with subscripts 1 and 2 exchanged, and 3 replaced with 4. Above, we used

$$\alpha_1=\sqrt{\max\left(1,\frac{\max(1,|g_{1,3}|^2)}{\max(1,|g_{1,4}|^2)}\right)},\ \text{and}$$
$$\alpha_2=\sqrt{\max\left(1,\frac{\max(1,|g_{2,4}|^2)}{\max(1,|g_{2,3}|^2)}\right)}.$$

To illustrate, we show how a couple of the above conditions are arrived at. The rest are also derived similarly. Two of the conditions on $r_{U_1}+r_{U_2}$ from Theorem 3 are

$$r_{U_1}+r_{U_2}\leq I(U_1,U_2;Y_3,V_4|X_3,X_4),\ \text{and} \tag{12}$$
$$r_{U_1}+r_{U_2}\leq I(U_1,U_2;Y_3|X_3,X_4)+r_4-I(Y_4;V_4|X_3,X_4,U_1,U_2,Y_3). \tag{13}$$

Below, we show the following:

$$\begin{aligned}I(U_1,U_2;Y_3,V_4|X_3,X_4)\geq \log\Bigg(1&+\left|\frac{g_{1,3}}{\alpha_1}\right|^2+\left|\frac{g_{2,4}}{\alpha_2}\right|^2+\left|\frac{g_{1,4}}{\alpha_2}\right|^2+\left|\frac{g_{2,3}}{\alpha_1}\right|^2\\
&+\left|\frac{g_{1,3}g_{2,4}}{\alpha_1\alpha_2}\right|^2+\left|\frac{g_{1,4}g_{2,3}}{\alpha_1\alpha_2}\right|^2-2\left|\frac{g_{1,3}g_{2,4}g_{1,4}g_{2,3}}{\alpha_1^2\alpha_2^2}\right|\cos\theta\Bigg)-\log 36,\\
I(U_1,U_2;Y_3|X_3,X_4)&\geq \max(n_{1,3},n_{2,3})-[n_{1,3}-n_{1,4}]_+-2,\ \text{and}\\
I(Y_4;V_4|X_3,X_4,U_1,U_2,Y_3)&\leq 1.\end{aligned}$$

This will allow us to conclude that in order for r_{U_1}, r_{U_2} to satisfy (12)-(13), it is enough if they satisfy

$$\begin{aligned}r_{U_1}+r_{U_2}\leq \log\Bigg(1&+\left|\frac{g_{1,3}}{\alpha_1}\right|^2+\left|\frac{g_{2,4}}{\alpha_2}\right|^2+\left|\frac{g_{1,4}}{\alpha_2}\right|^2+\left|\frac{g_{2,3}}{\alpha_1}\right|^2\\
&+\left|\frac{g_{1,3}g_{2,4}}{\alpha_1\alpha_2}\right|^2+\left|\frac{g_{1,4}g_{2,3}}{\alpha_1\alpha_2}\right|^2-2\left|\frac{g_{1,3}g_{2,4}g_{1,4}g_{2,3}}{\alpha_1^2\alpha_2^2}\right|\cos\theta\Bigg)-\log 36,\ \text{and}\\
r_{U_1}+r_{U_2}&\leq \max(n_{1,3},n_{2,3})-[n_{1,3}-n_{1,4}]_+ +r_4-3.\end{aligned}$$

From the choices for the auxiliary random variables we made,

$$\begin{aligned}I(U_1,U_2;Y_3,V_4|X_3,X_4)&=I\left(\begin{bmatrix}U_1\\U_2\end{bmatrix};\begin{bmatrix}\frac{Y_3}{\alpha_1}\\\frac{V_4}{\alpha_2}\end{bmatrix}\right)\\
&=I\left(\begin{bmatrix}U_1\\U_2\end{bmatrix};\begin{bmatrix}g_{1,3}&g_{2,3}\\g_{1,4}&g_{2,4}\end{bmatrix}\begin{bmatrix}U_1\\U_2\end{bmatrix}+\begin{bmatrix}\frac{g_{1,3}Z_1+g_{2,3}Z_2+N_3}{\alpha_1}\\\frac{g_{1,4}Z_1+g_{2,4}Z_2+N_4}{\alpha_2}\end{bmatrix}+\begin{bmatrix}0\\\frac{Q_4}{\alpha_2}\end{bmatrix}\right)\\
&\geq I\left(\begin{bmatrix}U_1\\U_2\end{bmatrix};\begin{bmatrix}g_{1,3}&g_{2,3}\\g_{1,4}&g_{2,4}\end{bmatrix}\begin{bmatrix}U_1\\U_2\end{bmatrix}+\begin{bmatrix}\frac{g_{1,3}Z_1+g_{2,3}Z_2+N_3}{\alpha_1}\\\frac{g_{1,4}Z_1+g_{2,4}Z_2+N_4}{\alpha_2}\end{bmatrix}+\begin{bmatrix}Q_3'\\\frac{Q_4}{\alpha_2}\end{bmatrix}\right)\\
&=I\left(\mathbf{U};\mathbf{HU}+\mathbf{N}+\mathbf{Q}\right),\end{aligned}$$

where Q_3' is a unit variance Gaussian random variable independent of everything else. In the last step, we defined the Gaussian vectors $\mathbf{U},\mathbf{H},\mathbf{N}$, and $\mathbf{Q}$. Note that $\frac{Q_4}{\alpha_2}$ is a unit variance Gaussian random variable which makes $\mathbf{Q}$ a Gaussian random vector whose covariance matrix is the identity matrix $\mathbf{I}$. Also, note that the covariance matrix of $\mathbf{U}$ is $\mathbf{K_U} = K^2\mathbf{I}$. Continuing,

$$\begin{aligned} I(U_1, U_2; Y_3, V_4|X_3, X_4) &\geq \log \frac{\left|\mathbf{H}\mathbf{K_U}\mathbf{H}^\dagger + \mathbf{K_N} + \mathbf{I}\right|}{|\mathbf{K_N} + \mathbf{I}|} \\ &\geq \log \frac{\left|\mathbf{H}\mathbf{K_U}\mathbf{H}^\dagger + \mathbf{I}\right|}{|\mathbf{K_N} + \mathbf{I}|} \end{aligned}$$

From the choices made for the variances of Z_1, Z_2, we can find a uniform upperbound for the denominator for all possible channels:

$$|\mathbf{K_N} + \mathbf{I}| \leq 9.$$

Evaluating the lowerbound on $I(U_1, U_2; Y_3, V_4|X_3, X_4)$ using this and substituting $K = 1/2$, we can show that

$$\begin{aligned} I(U_1, U_2; Y_3, V_4|X_3, X_4) \geq \log \Bigg(1 &+ \left|\frac{g_{1,3}}{\alpha_1}\right|^2 + \left|\frac{g_{2,4}}{\alpha_2}\right|^2 + \left|\frac{g_{1,4}}{\alpha_2}\right|^2 + \left|\frac{g_{2,3}}{\alpha_1}\right|^2 \\ &+ \left|\frac{g_{1,3}g_{2,4}}{\alpha_1\alpha_2}\right|^2 + \left|\frac{g_{1,4}g_{2,3}}{\alpha_1\alpha_2}\right|^2 - 2\left|\frac{g_{1,3}g_{2,4}g_{1,4}g_{2,3}}{\alpha_1^2\alpha_2^2}\right|\cos\theta\Bigg) - \log 36. \end{aligned}$$

Similarly,

$$\begin{aligned} I(U_1, U_2; Y_3|X_3, X_4) &= I(U_1, U_2; g_{1,3}U_1 + g_{2,3}U_2 + g_{1,3}Z_1 + g_{2,3}Z_2 + N_3) \\ &\geq \log\left(1 + \frac{|g_{1,3}|^2/K + |g_{2,3}|^2/K}{1 + \max\left(1, \frac{\max(1,|g_{1,3}|^2)}{\max(1,|g_{1,4}|^2)}\right)}\right) \\ &\geq \max(n_{1,3}, n_{2,3}) - [n_{1,3} - n_{1,4}]_+ - \log 4, \end{aligned}$$

and

$$\begin{aligned} I(Y_4; V_4|X_3, X_4, U_1, U_2, Y_3) &\leq I(g_{1,4}Z_1 + g_{2,4}Z_2 + N_4; g_{1,4}Z_1 + g_{2,4}Z_2 + N_4 + Q_4) \\ &= \log\left(1 + \frac{\alpha_2^2}{1+\alpha_2^2}\right) \leq \log 2 = 1. \end{aligned}$$

Applying Fourier-Motzkin elimination to the set of all conditions on the rates, we can show that a sum-rate equal to the minimum of the following terms is achievable with a gap of at most 15 bits in this regime.

$$\begin{aligned} u'_2 &= \max(n_{2,4}, n_{2,3}) + (\max(n_{1,3}, n_{2,3}, n_C) - n_{2,3}), \\ u'_3 &= \max(n_{1,3}, n_{1,4}) + (\max(n_{2,4}, n_{1,4}, n_C) - n_{1,4}), \\ u'_4 &= \max(n_{1,3}, n_C) + \max(n_{2,4}, n_C), \text{ and} \\ u'_5 &= \log\Big(1 + \left(|g_{1,3}|^2 + |g_{2,4}|^2 + |g_{1,4}|^2 + |g_{2,3}|^2\right) \\ &\qquad + \left(|g_{1,3}g_{2,4}|^2 + |g_{1,4}g_{2,3}|^2 - 2|g_{1,3}g_{2,4}g_{1,4}g_{2,3}|\cos\theta\right)\Big). \end{aligned}$$

APPENDIX E
PROOF OF THE UPPERBOUNDS OF THEOREMS 1 AND 2

Upperbounds 1-3 are new, upperbounds 4 and 5 are cut-set upperbounds which also appeared in the two-user interference channel with source cooperation [12]. Below, we prove upperbounds 1-3 and, for completeness, repeat the proofs for upperbound 4 and 5.

Upperbound 1:

We create two dummy channels in both of which, all the noise processes are independent of those in the original channel, but have identical distributions to their counterparts in the original channel. All the nodes use the same strategies as in the original problem (*i.e.*, same codebooks at the nodes 1 and 2, and the same $f_{k,t}$'s at nodes 3 and 4), but the messages transmitted by the nodes are different from that in the original channel as explained below. In the first dummy channel (where all quantities are denoted by adding a prime $'$), the message at node 1 is identical to the message at node 1 in the original channel, *i.e.*, $M_1' = M_1$, however, the message M_2' at node 2 is independent of the messages M_1, M_2 and distributed uniformly over its alphabet. We note that $X_1' = X_1$, but X_2', X_3', and X_4' are, in general, different from their counterparts in the original channel. Similarly, $M_2'' = M_2$, but M_1'' is independent of M_1, M_2, M_2' and distributed uniformly over its alphabet. We start from Fano's inequality.

$$\begin{aligned}
&T(R_1 + R_2 - o(\epsilon))\\
&\qquad\le I(M_1; Y_3^T) + I(M_2; Y_4^T)\\
&\qquad\le I(M_1; Y_3^T, g_{1,4}^*(X_1'^T) + g_{3,4}(X_3'^T), M_2') + I(M_2; Y_4^T, g_{2,3}^*(X_2''^T) + g_{4,3}(X_4''^T), M_1'')\\
&\qquad\le I(M_1; Y_3^T, g_{1,4}^*(X_1'^T) + g_{3,4}(X_3'^T)|M_2') + I(M_2; Y_4^T, g_{2,3}^*(X_2''^T) + g_{4,3}(X_4''^T)|M_1'').
\end{aligned}$$

These two symmetric terms can be further upperbounded. Below we will show how the first is upperbounded; the second term can be similarly upperbounded.

$$\begin{aligned}
&I(M_1; Y_3^T, g_{1,4}^*(X_1'^T) + g_{3,4}(X_3'^T)|M_2')\\
&\qquad= H(g_{1,4}^*(X_1'^T) + g_{3,4}(X_3'^T)|M_2') + H(Y_3^T|g_{1,4}^*(X_1'^T) + g_{3,4}(X_3'^T), M_2')\\
&\qquad\quad - H(Y_3^T|M_1, M_2') - H(g_{1,4}^*(X_1'^T) + g_{3,4}(X_3'^T)|Y_3^T, M_1, M_2')\\
&\qquad\overset{\text{(a)}}{=} H(g_{1,4}^*(X_1'^T) + g_{3,4}(X_3'^T)|M_2') + H(Y_3^T|g_{1,4}^*(X_1'^T) + g_{3,4}(X_3'^T), M_2')\\
&\qquad\quad - H(g_{2,3}^*(X_2^T) + g_{4,3}(X_4^T)|M_1, M_2') - H(g_{1,4}^*(X_1'^T) + g_{3,4}(X_3'^T)|Y_3^T, M_1, M_2')\\
&\qquad\overset{\text{(b)}}{=} H(g_{1,4}^*(X_1'^T) + g_{3,4}(X_3'^T)|M_2') + H(Y_3^T|g_{1,4}^*(X_1'^T) + g_{3,4}(X_3'^T), M_2')\\
&\qquad\quad - H(g_{2,3}^*(X_2^T) + g_{4,3}(X_4^T)|M_1) - H(g_{1,4}^*(X_1'^T) + g_{3,4}(X_3'^T)|Y_3^T, M_1, M_2')\\
&\qquad\overset{\text{(c)}}{=} H(g_{1,4}^*(X_1'^T) + g_{3,4}(X_3'^T)|M_2') + H(Y_3^T|g_{1,4}^*(X_1'^T) + g_{3,4}(X_3'^T), M_2')\\
&\qquad\quad - H(g_{2,3}^*(X_2^T) + g_{4,3}(X_4^T)|M_1) - H(g_{1,4}^*(X_1'^T) + g_{3,4}(X_3'^T)|M_1, M_2').
\end{aligned}$$

where (a) follows from the fact that $Y_3[t] = g_{1,3}(X_1[t]) + g_{2,3}^*(X_2[t]) + g_{4,3}(X_4[t])$ and $g_{1,3}(X_1[t])$ is a deterministic function of M_1, and (b) is due to the independence of $(M_1, g_{2,3}^*(X_2^T), g_{4,3}(X_4^T))$ and M_2'. Equality (c) follows from the fact that conditioned on M_1, the primed quantities and the unprimed quantities are independent. Further, we can upperbound the second and fourth terms as follows

$$H(Y_3^n|g_{1,4}^*(X_1'^T) + g_{3,4}(X_3'^T), M_2') \le H(Y_3^n|g_{1,4}^*(X_1'^T) + g_{3,4}(X_3'^T)),$$

$$
\begin{aligned}
&-H(g^*_{1,4}(X'^T_1)+g_{3,4}(X'^T_3)|M_1,M'_2)\\
&\quad=\sum_{t=1}^{T}-H(g^*_{1,4}(X'_1[t])+g_{3,4}(X'_3[t])|g^*_{1,4}(X'^{t-1}_1)+g_{3,4}(X'^{t-1}_3),M_1,M'_2)\\
&\quad\leq\sum_{t=1}^{T}-H(g^*_{1,4}(X'_1[t])+g_{3,4}(X'_3[t])|g^*_{1,4}(X'^{t-1}_1)+g_{3,4}(X'^{t-1}_3),M_1,M'_2,X'_1[t],g_{3,4}(X'_3[t]))\\
&\quad=\sum_{t=1}^{T}-H(g^*_{1,4}(X'_1[t])|X'_1[t]).
\end{aligned}
$$

We combine these and use the following facts

$$
\begin{aligned}
H(g^*_{1,4}(X'^T_1)+g_{3,4}(X'^T_3)|M'_1)&=H(g^*_{1,4}(X^T_1)+g_{3,4}(X^T_3)|M_1),\\
H(g^*_{2,3}(X''^T_2)+g_{4,3}(X''^T_4)|M''_2)&=H(g^*_{2,3}(X^T_2)+g_{4,3}(X^T_4)|M_2).
\end{aligned}
$$

We arrive at

$$
\begin{aligned}
T(R_1+R_2-o(\epsilon))\leq{}&H(Y^T_3|g^*_{1,4}(X'^T_1)+g_{3,4}(X'^T_3))+H(Y^T_4|g^*_{2,3}(X''^T_2)+g_{4,3}(X''^T_4))\\
&-\left(\sum_{t=1}^{T}H(g^*_{1,4}(X'_1[t])|X'_1[t])+H(g^*_{2,3}(X''_2[t])|X''_2[t])\right).
\end{aligned}
$$

Linear deterministic case:
We have

$$
\begin{aligned}
R_1+R_2-o(\epsilon)&\leq\max(n_{2,3},n_{1,3}-n_{1,4}+n_{4,3},n_{4,3})-0+\max(n_{1,4},n_{2,4}-n_{2,3}+n_{3,4},n_{3,4})-0\\
&=\max(n_{2,3},n_{1,3}-n_{1,4}+n_C,n_C)+\max(n_{1,4},n_{2,4}-n_{2,3}+n_C,n_C).
\end{aligned}
$$

Gaussian case:
We have,

$$
\begin{aligned}
R_1+R_2-o(\epsilon)\leq{}&\begin{cases}\log\left(1+\left(|g_{2,3}|+|g_C|+\left|\frac{g_{1,3}g_C}{g_{1,4}}\right|\right)^2+\left|\frac{g_{1,3}}{g_{1,4}}\right|^2\right), & \text{if } |g_{1,4}|>\max(1,|g_C|)\\ \log\left(1+(|g_{2,3}|+|g_C|+|g_{1,3}|)^2\right), & \text{otherwise}\end{cases}\\
&+\begin{cases}\log\left(1+\left(|g_{1,4}|+|g_C|+\left|\frac{g_{2,4}g_C}{g_{2,3}}\right|\right)^2+\left|\frac{g_{2,4}}{g_{2,3}}\right|^2\right), & \text{if } |g_{2,3}|>\max(1,|g_C|)\\ \log\left(1+(|g_{1,4}|+|g_C|+|g_{2,4}|)^2\right), & \text{otherwise}\end{cases}
\end{aligned}
$$

Upperbounds 2 and 3:
We start from Fano's inequality.

$$
\begin{aligned}
T(R_1+R_2-o(\epsilon))&\leq I(M_1;Y^T_3)+I(M_2;Y^T_4)\\
&\leq I(M_1;Y^T_3)+I(M_2;Y^T_4,g^*_{2,3}(X^T_2),M_1)\\
&\leq I(M_1;Y^T_3)+I(M_2;Y^T_4,g^*_{2,3}(X^T_2),|M_1)\\
&\leq\sum_{t=1}^{T}I(M_1;Y_3[t]|Y^{t-1}_3)+I(M_2;Y_4[t],g^*_{2,3}(X_2[t]),|Y^{t-1}_4,g^*_{2,3}(X^{t-1}_2),M_1).
\end{aligned}
$$

Below, we upperbound these terms separately.

$$
\begin{aligned}
I(M_1;Y_3[t]|Y^{t-1}_3)&=H(Y_3[t]|Y^{t-1}_3)-H(Y_3[t]|Y^{t-1}_3,M_1)\\
&\leq H(Y_3[t])-H(Y_3[t]|Y^{t-1}_3,Y^{t-1}_4,M_1)\\
&\overset{\text{(a)}}{\leq}H(Y_3[t])-H(g^*_{2,3}(X_2[t])|Y^{t-1}_4,g^*_{2,3}(X^{t-1}_2),M_1),
\end{aligned}
$$

where (a) follows from the fact that $Y_3[t] = g_{1,3}(X_1[t]) + g^*_{2,3}(X_2[t]) + g_{4,3}(X_4[t])$, and $X_1[t]$ is a deterministic function of M_1, $X_4[t]$ is a deterministic function of Y_4^{t-1}, and $g_{1,3}$ and $g_{4,3}$ are deterministic maps.

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

We upperbound these two terms separately now.

$$\begin{aligned}I(M_2; &g^*_{2,3}(X_2[t])|Y_4^{t-1}, g^*_{2,3}(X_2^{t-1}), M_1)\\ &= H(g^*_{2,3}(X_2[t])|Y_4^{t-1}, g^*_{2,3}(X_2^{t-1}), M_1) - H(g^*_{2,3}(X_2[t])|X_2[t]),\end{aligned}$$

which follows from the channel model (memorylessness and independence of the noise processes at the different nodes).

$$\begin{aligned}I(M_2; &Y_4[t]|Y_4^{t-1}, g^*_{2,3}(X_2^{t}), M_1)\\ &\overset{\text{(a)}}{=} I(M_2; Y_4[t]|Y_4^{t-1}, g^*_{2,3}(X_2^{t}), M_1, X_3[t])\\ &= H(Y_4[t]|Y_4^{t-1}, g^*_{2,3}(X_2^{t}), M_1, X_3[t]) - H(Y_4[t]|Y_4^{t-1}, g^*_{2,3}(X_2^{t}), X_3[t], M_1, M_2)\\ &\leq H(X_2[t] + g^*_{1,4}(X_1[t])|g^*_{2,3}(X_2[t]), X_1[t]) - H(Y_4[t]|Y_4^{t-1}, g^*_{2,3}(X_2^{t}), X_3[t], M_1, M_2)\\ &\overset{\text{(b)}}{=} H(X_2[t] + g^*_{1,4}(X_1[t])|g^*_{2,3}(X_2[t]), X_1[t]) - H(g^*_{1,4}(X_1[t])|X_1[t]),\end{aligned}$$

where (a) can be seen by noting that: (1) $X_3[t]$ is a deterministic function $f_{3,t}$ of Y_3^{t-1}, (2) Y_3^{t-1} in turn is such that

$$Y_3(s) = g_{1,3}(X_1(s)) + g^*_{2,3}(X_2(s)) + g_{4,3}(X_4(s)),\ s \leq t-1,$$

and (3) for all $s \leq t-1$, $g_{1,3}(X_1(s))$ is a deterministic function of M_1, and $g_{4,3}(X_4(s))$ is a deterministic function of Y_4^{s-1}. Also, (b) 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 M_1.
Combining everything, we have

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

Linear deterministic channel: We have

$$\begin{aligned}R_1 + R_2 - o(\epsilon) &\leq \{\max(n_{1,3}, n_{2,3}, n_{4,3}) - 0\} + \{[n_{2,4} - n_{2,3}]_+ - 0\}\\ &= \max(n_{1,3}, n_{2,3}, n_C) + [n_{2,4} - n_{2,3}]_+.\end{aligned}$$

Gaussian channel: We have

$$R_1 + R_2 - o(\epsilon) \leq \log(1 + (|g_{1,3}| + |g_{2,3}| + |g_C|)^2) + \log\left(1 + \frac{|g_{2,4}|^2}{\max(1, |g_{2,3}|^2)}\right).$$

Upperbound 4:

This is a simple cut-set upperbound [7] 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

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

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

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

REFERENCES

[1] 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.

[2] T. Berger, "Multiterminal source coding," *The Inform. Theory Approach to Communications*, G. Longo,. Ed., New York: Springer-Verlag, 1977.

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

[5] T. M. Cover, "An achievable rate region for the broadcasting channel," *IEEE Trans. Inform. Theory*, 21:399–404, July 1979.

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

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

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

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

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

[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] V. Prabhakaran and P. Viswanath. "Interference Channels with Source Cooperation," Submitted to *IEEE Trans. Inform. Theory*, May 2009.

[13] S-Y. Tung. *Multiterminal Source Coding.* Ph.D. dissertation, Cornell University, Ithaca, May 1978.

[14] I-Hsiang Wang and D. Tse, "Interference mitigation through limited receiver cooperation," submitted to IEEE Trans. on Inform. Theory, Nov, 2009. `http://arxiv.org/abs/0911.2053`

[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] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, 22(1):1–10, January 1976.

[17] L. Zhou and W. Yu, "Gaussian Z-interference channel with a relay link: Achievability region and asymptotic sum capacity," submitted to *IEEE Trans. on Inform. Theory*, June 2010.