Interference Channels With Rate-Limited Feedback

Alireza Vahid, Changho Suh, and A. Salman Avestimehr

Abstract—We consider the two-user interference channel with *rate-limited* feedback. Related prior works focus on the case where feedback links have *infinite* capacity, while no research has been done for the rate-limited feedback problem. Several new challenges arise due to the capacity limitations of the feedback links, both in deriving inner bounds and outer bounds. We study this problem under three different interference models: the El Gamal–Costa deterministic model, the linear deterministic model, and the Gaussian model. For the first two models, we develop an achievable scheme that employs three techniques: Han-Kobayashi message splitting, quantize-and-binning, and decode-and-forward. We also derive new outer bounds for all three models and we show the optimality of our scheme under the linear deterministic model. In the Gaussian case, we propose a transmission strategy that incorporates lattice codes, inspired by the ideas developed in the first two models. For symmetric channel gains, we prove that the gap between the achievable sum rate of the proposed scheme and our new outer bounds is bounded by a constant number of bits, independent of the channel gains.

Index Terms—El Gamal–Costa deterministic model, Gaussian interference channel (IC), linear deterministic model, multiuser information theory, rate-limited feedback.

I. INTRODUCTION

THE history of feedback in communication systems traces back to Shannon. It is well known that feedback does not increase the capacity of discrete memoryless point-to-point channels [1]. However, feedback can enlarge the capacity region of multiuser networks, even in the most basic case of the two-user memoryless multiple-access channel [2], [3]. Hence, there has been a growing interest in developing feedback strategies and understanding the fundamental limits of communication over multiuser networks with feedback, in particular the two-user interference channel (IC). See [4]–[10] for example.

Especially in [9], the infinite-rate feedback capacity of the two-user Gaussian IC has been characterized to within a two-bit gap. One consequence of this result is that interestingly feedback can provide an unbounded capacity increase. This is in contrast to point-to-point and multiple-access channels where feedback provides no gain and bounded gain respectively.

While the feedback links are assumed to have *infinite* capacity in [9], a more realistic feedback model is one where feedback links are *rate limited*. In this paper, we study the impact of the rate-limited feedback in the context of the two-user IC. We focus on two fundamental questions: 1) what is the maximum capacity gain that can be obtained with access to feedback links at a specific rate of C_{FB}? 2) what are the transmission strategies that exploit the available feedback links efficiently? Specifically, we address these questions under three channel models: the El Gamal–Costa deterministic model [11], the linear deterministic model of [12], and the Gaussian model.

Under the El Gamal–Costa deterministic model, we derive inner bounds and outer bounds on the capacity region. As a result, we show that the capacity region can be enlarged using feedback by at most the amount of available feedback, *i.e.*, "one bit of feedback is at most worth one bit of capacity." Our achievable scheme employs three techniques: 1) Han–Kobayashi message splitting; 2) quantize-and-binning; and 3) decode-and-forward. Unlike the infinite-rate feedback case [9], in the rate-limited feedback case, a receiver cannot provide its *exact* received signal to its corresponding transmitter; therefore, the main challenge is how to smartly decide what to send back through the available rate-limited feedback links. We overcome this challenge as follows. We first split each transmitter's message into three parts: the cooperative common, the noncooperative common, and the private message. Next, each receiver quantizes its received signal and then generates a binning index so as to capture part of the other user's common message (that we call the cooperative common message) which causes interference to its own message. The receiver will then send back this binning index to its intended transmitter through the rate-limited feedback links. With this feedback, each transmitter decodes the other user's cooperative common message by exploiting its own message as side information. This way transmitters will be able to cooperate by means of the feedback links, thereby enhancing the achievable rates. This result will be described in Section IV.

We then study the problem under the linear deterministic model [12] which captures the key properties of the wireless channel, and thus provides insights that can lead to an approximate capacity of Gaussian networks [9], [12], [13]–[15]. We show that our inner bounds and outer bounds match under this linear deterministic model, thus establishing the capacity region. While this model is a special case of the El Gamal–Costa model, it has a significant role to play in motivating a generic achievable scheme for the El Gamal–Costa model. Moreover, the explicit achievable scheme in this model provides a concrete guideline to the Gaussian case. We will explain this result in Section V.

Inspired by the results in the deterministic models, we develop an achievable scheme and also derive new outer-bounds

for the Gaussian channel. In order to translate the main ideas in our achievability strategy for the deterministic models into the Gaussian case, we employ lattice coding which enables receivers to decode superposition of codewords. Specifically at each transmitter, we employ lattice codes for cooperative messages. By appropriate power assignment of the codewords, we make the desired lattice codes arrive at the same power level, hence, receivers being able to decode the superposition of codewords. Each receiver will, then, decode the index of the lattice code corresponding to the superposition and sends it back to its corresponding transmitter where the cooperative common message of the other user will be decoded. For symmetric channel gains, we show that the gap between the achievable sum rate and the outer bounds can be bounded by a constant, independent of the channel gains. This will be explained in Section VI.

Related Work: ICs with infinite-rate feedback have received previous attention [4]–[10]. Kramer [4], [5] developed a feedback strategy in the Gaussian IC; In [6], Gastpar and Kramer established an outer-bound on the usefulness of noisy feedback for the two-user IC. Tandon and Ulukus in [7] derived an outer bound using the dependence balance bound technique [16]. However, the gap between the inner bounds and the outer bounds becomes arbitrarily large with the increase of signal to noise ratio (SNR) and interference to noise ratio (INR). Jiang-Xin-Garg [8] derived an achievable region in the discrete memoryless IC with feedback, based on block Markov encoding [17] and binning. However, no outer bounds are provided. Suh and Tse in [9] developed new inner bounds and outer bounds to characterize the feedback capacity of the Gaussian IC to within two bits. Sahai *et al.* in [10], have shown that in order to achieve the infinite-rate feedback capacity, it is sufficient to have only one feedback link of infinite rate from one receiver to either of the two transmitters.

While no research has been done for the rate-limited feedback problem, some works have been done for the different yet related problem—the conferencing encoder problem [18]–[23]. Tuninetti in [18] has proposed a coding strategy for the two-user IC that results in higher achievable rates by exploiting overheard information by the transmitters; backward decoding is incorporated at the receivers which we also employ in our achievability scheme. This result was improved in [19] by incorporating Gelfand–Pinsker coding [24], [25] to send cooperatively the private messages. Prabhakaran and Viswanath [21] have made a connection between the feedback problem and the conferencing encoder problem However, the connection is loose especially when the feedback link is rate limited, although it can be strong for the infinite-rate feedback case. It turns out this distinction between the two problems leads to developing a new lattice-code-based achievable scheme in our problem.

The rest of this paper is organized as follows. In Section II, we formulate our problem and give a brief overview of the channel models. In Section III, we provide a motivating example which forms inspiration of our achievable scheme. In Section IV, we will provide our main results under the El Gamal–Costa deterministic model. We will then present the capacity theorem for the linear deterministic model in Section V. Next, in Section VI, we describe our main results for the Gaussian channel. Section VII concludes the paper.
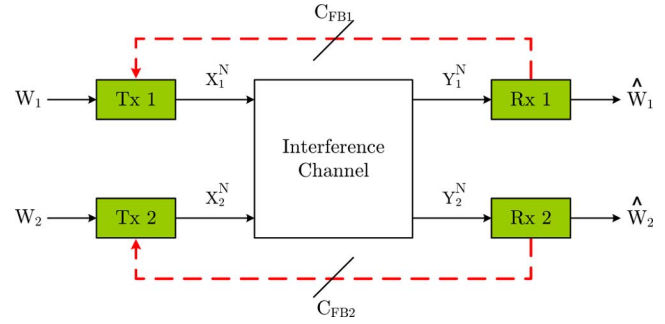

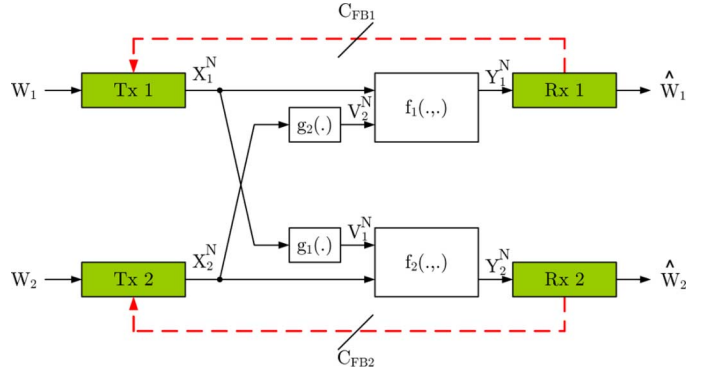Fig. 1. Two-user IC with rate-limited feedback.

Fig. 2. El Gamal–Costa deterministic IC with rate-limited feedback.
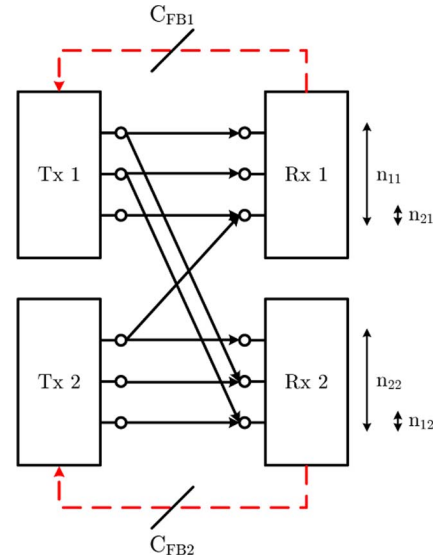
Fig. 3. Example of a linear deterministic IC with rate-limited feedback, where $n_{11} = n_{22} = 3$, $n_{12} = 2$, $n_{21} = 1$, and $q = 3$.

II. PROBLEM FORMULATION AND NETWORK MODEL

We consider a two-user IC where a noiseless rate-limited feedback link is available from each receiver to its corresponding transmitter (see Fig. 1). The feedback link from receiver k to transmitter k is assumed to have a capacity of $C_{\mathsf{FB}k}$, $k = 1, 2$. on.

Transmitters 1 and 2 wish to reliably communicate independent and uniformly distributed messages $W_1 \in \{1, 2, \ldots, 2^{NR_1}\}$ and $W_2 \in \{1, 2, \ldots, 2^{NR_2}\}$ to receivers 1 and 2 respectively, during N uses of the channel. The transmitted signal of transmitter k, $k = 1, 2$, at time i, $1 \leq i \leq N$, and the received signal of receiver k, $k = 1, 2$, at time i,

$1 \leq i \leq N$, are, respectively, denoted by $X_{k,i}$ and $Y_{k,i}$. There are two feedback encoders at the receivers that create the feedback signals from the received signals

$$\tilde{Y}_{k,i} = \tilde{e}_{k,i}(Y_{k,1}, \ldots, Y_{k,i-1}) = \tilde{e}_{k,i}(Y_k^{(i-1)}), \quad k = 1, 2. \tag{1}$$

where we use shorthand notation to indicate the sequence up to $i - 1$.

Due to the presence of feedback, the encoded signal $X_{k,i}$ of user k at time i is a function of both its own message and previous outputs of the corresponding feedback encoder:

$$X_{k,i} = e_{k,i}(W_k, \tilde{Y}_k^{(i-1)}), \quad k = 1, 2. \tag{2}$$

Each receiver k, $k = 1, 2$, uses a decoding function $d_{k,N}$ to get the estimate $\hat{W}_k$, from the channel outputs $\{Y_{k,i} : i = 1, \ldots, N\}$. An error occurs whenever $\hat{W}_k \neq W_k$. The average probability of error is given by

$$\lambda_{k,N} = \mathbb{E}[P(\hat{W}_k \neq W_k)], \quad k = 1, 2 \tag{3}$$

where the expectation is taken with respect to the random choice of the transmitted messages W_1 and W_2.

We say that a rate pair (R_1, R_2) is achievable, if there exists a block encoder at each transmitter, a block encoder at each receiver that creates the feedback signals, and a block decoder at each receiver as described earlier, such that the average error probability of decoding the desired message at each receiver goes to zero as the block length N goes to infinity. The capacity region $\mathcal{C}$ is the closure of the set of the achievable rate pairs.

We will consider the following three channel models to investigate this problem.

1) El Gamal–Costa Deterministic IC: Fig. 2 illustrates the El Gamal–Costa deterministic IC [11] with rate-limited feedback. In this model, the outputs Y_1 and Y_2 and the interferences V_1 and V_2 are (deterministic) functions of inputs X_1 and X_2[11]

$$\begin{aligned} Y_{1,i} &= f_1(X_{1,i}, V_{2,i}) \\ Y_{2,i} &= f_2(X_{2,i}, V_{1,i}) \\ V_{1,i} &= g_1(X_{1,i}) \\ V_{2,i} &= g_2(X_{2,i}) \end{aligned} \tag{4}$$

where $f_1(\cdot, \cdot)$ and $f_2(\cdot, \cdot)$ are such that

$$\begin{aligned} H(V_{2,i}|Y_{1,i}, X_{1,i}) &= 0 \\ H(V_{1,i}|Y_{2,i}, X_{2,i}) &= 0. \end{aligned} \tag{5}$$

Here, V_k is a part of $X_k (k = 1, 2)$, visible to the unintended receiver. This implies that in any system where each decoder can decode its message with arbitrary small error probability, V_1 and V_2 are completely determined at receivers 2 and 1, respectively, i.e., these are common signals.

2) Linear Deterministic IC: This model, which was introduced in [12], captures the effect of broadcast and superposition in wireless networks. We study this model to bridge from general deterministic networks into Gaussian networks. In this model, there is a nonnegative integer n_{kj} representing channel gain from transmitter k to receiver j, $k = 1, 2$, and $j = 1, 2$. In the linear deterministic IC, we can write the channel input to the

transmitter k at time i as $X_{k,i} = [X_{k,i}^1 \ X_{k,i}^2 \ldots X_{k,i}^q]^T \in \mathbb{F}_2^q$, $k = 1, 2$, such that $X_{k,i}^1$ and $X_{k,i}^q$ represent the most and the least significant bits of the transmitted signal, respectively. Also, q is the maximum of the channel gains in the network, i.e., $q = \max_{k,j}(n_{kj})$. At each time i, the received signals are given by

$$\begin{aligned} Y_{1,i} &= \mathbf{S}^{q-n_{11}} X_{1,i} \oplus \mathbf{S}^{q-n_{21}} X_{2,i} \\ Y_{2,i} &= \mathbf{S}^{q-n_{12}} X_{1,i} \oplus \mathbf{S}^{q-n_{22}} X_{2,i} \end{aligned} \tag{6}$$

where $\mathbf{S}$ is the $q \times q$ shift matrix and operations are performed in $\mathbb{F}_2$ (i.e., modulo two). See Fig. 3 for an example.

It is easy to see that this model also satisfies the conditions of (5); hence, it is a special class of the El Gamal–Costa deterministic IC.

3) Gaussian IC: In this model, there is a complex number h_{kj} representing the channel from transmitter k to receiver j, $k = 1, 2$, and $j = 1, 2$. The received signals are

$$\begin{aligned} Y_{1,i} &= h_{11} X_{1,i} + h_{21} X_{2,i} + Z_{1,i} \\ Y_{2,i} &= h_{12} X_{1,i} + h_{22} X_{2,i} + Z_{2,i} \end{aligned} \tag{7}$$

where $\{Z_{j,i}\}_{i=1}^N$ is the additive white complex Gaussian noise process with zero mean and unit variance at receiver j, $j = 1$, 2. Without loss of generality, we assume a power constraint of 1 at all nodes, i.e.

$$\frac{1}{N} \mathbb{E}\left(\sum_{i=1}^N |X_{k,i}|^2\right) \leq 1 \qquad k = 1, 2 \tag{8}$$

where N is the block length. We will use the following notations:

$$\begin{aligned} \mathsf{SNR}_1 &= |h_{11}|^2, \quad \mathsf{SNR}_2 = |h_{22}|^2, \\ \mathsf{INR}_{12} &= |h_{12}|^2, \quad \mathsf{INR}_{21} = |h_{21}|^2. \end{aligned} \tag{9}$$

III. MOTIVATING EXAMPLE

We start by analyzing a motivating example. Consider the linear deterministic IC with rate-limited feedback as depicted in Fig. 4(a). As we will see in Section V, the capacity region of this network is given by the region shown in Fig. 4(b). Our goal in this section is to demonstrate how feedback can help increase the capacity. In particular, we describe the achievability strategy for one of the corner points, i.e., $(R_1, R_2) = (4, 1)$. From this example, we will make important observations that will later provide insights into the achievable scheme.

The achievability strategy works as follows. In the first time slot, transmitter 1 sends four bits $a_1, \ldots, a_4$ and transmitter 2 sends only one bit b_1 at the third level [see Fig. 4(a)]. This way receiver 1 can decode its intended four bits interference free, while receiver 2 has access to only $a_1 \oplus b_1$ and a_2. In the second time slot, through the feedback link, receiver 2 feeds $a_1 \oplus b_1$ back to transmitter 2 who can remove b_1 from it to decode a_1. Also during the second time slot, transmitter 1 sends four fresh bits $a_5, \ldots, a_8$, whereas transmitter 2 sends one new bit b_2. In the third time slot, through the feedback link, receiver 2 feeds $b_2 \oplus a_5$ back to transmitter 2 who can remove b_2 from it to decode a_5. Moreover, during the third time slot, transmitter 1 sends
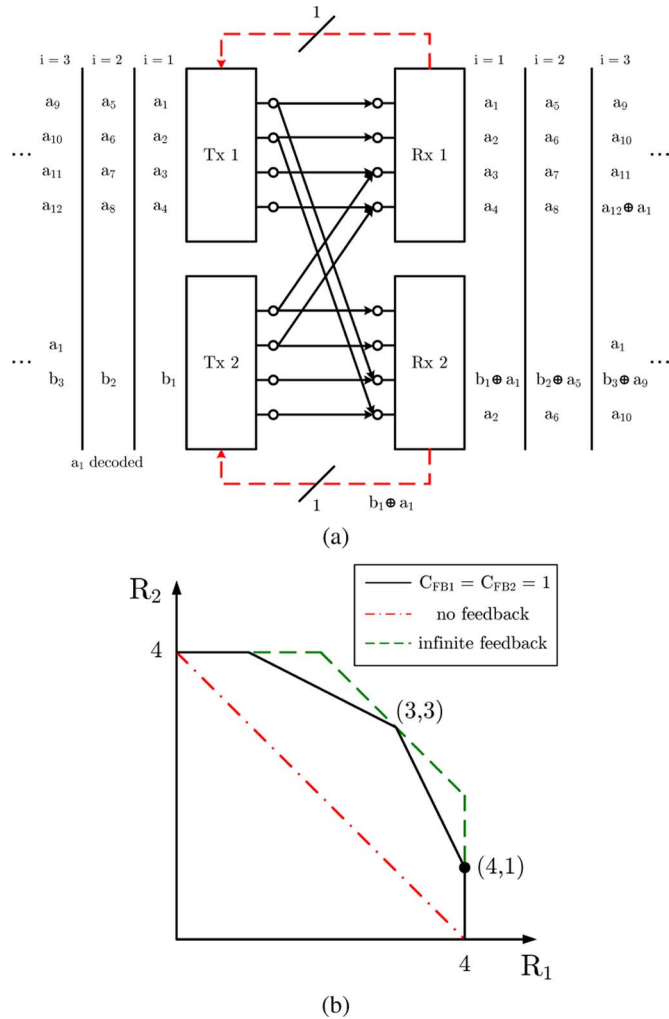
(a)

(b)

Fig. 4. (a) Two-user linear deterministic IC with channel gains $n_{11} = n_{22} = 4$, $n_{12} = n_{21} = 2$ and feedback rates $C_{\mathsf{FB1}} = C_{\mathsf{FB2}} = 1$, and (b) its capacity region.

four new bits $a_9, \ldots, a_{12}$, while transmitter 2 sends one new bit b_3 and at the level shown in Fig. 4(a), sends the other user's information bit a_1 decoded in the second time slot with the help of feedback. With this strategy, receiver 2 has now access to a_1 and can use it to decode b_1. Note that receiver 1 already knows a_1 and, hence, can decode a_{12}. This procedure will be repeated over the next time slots. During the last two time slots, only transmitter 2 sends the other user's information decoded before, while transmitter 1 sends nothing. Therefore, after B time slots, we achieve a rate of $(R_1, R_2) = \frac{B-2}{B}(4, 1)$, which converges to $(4, 1)$ as B goes to infinity.

Based on this simple capacity-achieving strategy, we can now make several observations.

1) The messages coming from transmitter 1 can be split into three parts: 1) "cooperative common": this message is visible to both receivers, while interfering with the other user's signals (e.g., a_1 at transmitter 1 in the first time slot). This should be fed back to the transmitter so that it can be used later in refining the desired signals corrupted by the interfering signal; 2) "noncooperative common": this message is visible to both receivers; however, it does not cause any interference (e.g., a_2 at transmitter 1 in the

first time slot); 3) "private": this message is visible only to the intended receiver (e.g., a_3 and a_4 at transmitter 1 in the first time sot). Denote these messages by w_{kcc}, w_{knc}, and w_{kp}, respectively, where $k = 1, 2$ is the transmitter index.

2) To refine the desired signal corrupted by the cooperative common signal of transmitter 1 (i.e., a_1), receiver 2 utilizes the feedback link to send the interfered signal (i.e., $a_1 \oplus b_1$) back to transmitter 2. Transmitter 2 then employs a partial decode-and-forward to help receiver 1 decode its messages, i.e., the cooperative message of transmitter 1 is decoded at transmitter 2 and it will be forwarded to receiver 2 during another time slot.

3) As we can see in this example, encoding operations at each time slot depend on previous operations, thereby motivating us to employ block Markov encoding. As for the decoding, we implement backward decoding at receivers. Each receiver waits until the last time B and we use the last received signal to decode the message received at time $B - 2$. We then decode the message received at time $B - 3$ and all the way down to the message received at time 1.

These observations will form the basis for our achievable schemes in the following sections.

IV. DETERMINISTIC IC

In this section, we consider the El Gamal–Costa deterministic IC with rate-limited feedback, described in Section II. The motivating example in the previous section leads us to develop a generic achievable scheme based on three ideas: 1) Han–Kobayashi message splitting [26]; 2) quantize-and-binning; and 3) decode-and-forward [17]. As mentioned earlier, we split the message into three parts: the cooperative common message, noncooperative common message, and private message. We employ quantize-and-binning to feed back part of the interfered signals through the rate-limited feedback link. With feedback, each transmitter decodes part of the other user's common information (cooperative common) that interfered with its desired signals. We accomplish this by using the partial decode-and-forward scheme. We also derive a new outer bound based on the genie-aided argument [11] and the dependence-balance-bound technique [7], [27], [28].

A. Achievable Rate Region

Theorem 1: The capacity region of the El Gamal–Costa deterministic IC with rate-limited feedback includes the set $\mathcal{R}$ of (R_1, R_2) satisfying inequalities (10a)–(10m), as shown at the bottom of the next page.

Proof: We first provide an outline of our achievable scheme. We employ block Markov encoding with a total size B of blocks. In block 1, transmitter 1 splits its own message into cooperative common, noncooperative common, and private parts and then sends a codeword superimposing all of these messages. The cooperative common message is sent via the codeword $u_1^{N,(1)}$. The noncooperative common message is added to this, being sent via $v_1^{N,(1)}$. The private message is, then, superimposed on top of the previous messages, being sent via $x_1^{N,(1)}$. Similarly, transmitter 2 sends $x_2^{N,(1)}$. In block 2, receiver 1 quantizes its received signal $y_1^{N,(1)}$ into $\hat{y}_1^{N,(1)}$ with the rate of $\hat{R}_1$. Next, it generates a bin index by considering the

capacity of its feedback link and then feeds the bin index back to its corresponding transmitter. Similarly, receiver 2 feeds back the corresponding bin index. In block 3, with feedback, each transmitter decodes the other user's cooperative common message (sent in block 1) that interfered with its desired signals. The following messages are, then, available at the transmitter: 1) its own message; and 2) the other user's cooperative common message decoded with the help of feedback. Using its own cooperative common message as well as the other user's counterpart, each transmitter generates the codeword $u^{N,(3)}$. This captures the correlation between the two transmitters that might induce the cooperative gain. Conditioned on these previous cooperative common messages, each transmitter generates new cooperative common, noncooperative common, and private messages. It then sends the corresponding codeword. This procedure is repeated until block $B-2$. In the last two blocks $B-1$ and B, to facilitate backward decoding, each transmitter sends the predetermined common messages and a new private message. Each receiver waits until total B blocks have been received and then performs backward decoding.

Codebook Generation: Fix a joint distribution $p(u)p(u_1|u)p(u_2|u)p(x_1|u_1,u)p(x_2|u_2,u)p(\hat{y}_1|y_1)p(\hat{y}_2|y_2)$.

We will first show that $p(x_1|u,u_1,v_1)$ and $p(x_2|u,u_2,v_2)$ are functions of the aforementioned distributions. To see this, let us write a joint distribution $p(u,u_1,u_2,v_1,v_2,x_1,x_2)$ in two different ways

$$
\begin{aligned}
&p(u,u_1,u_2,v_1,v_2,x_1,x_2) \\
&= p(u)p(u_1|u)p(u_2|u)p(x_1|u,u_1)p(x_2|u,u_2)\delta(v_1-g_1(x_1)) \\
&\quad \times \delta(v_2-g_2(x_2)) \\
&= p(u)p(u_1|u)p(u_2|u)p(v_1|u,u_1)p(v_2|u,u_2)p(x_1|u,u_1,v_1) \\
&\quad \times p(x_2|u,u_2,v_2)
\end{aligned} \tag{11}
$$

where $\delta(\cdot)$ indicates the Kronecker delta function. Notice that by the El Gamal–Costa model assumption (4), $p(v_1|u,u_1,x_1) = \delta(v_1-g_1(x_1))$ and $p(v_2|u,u_2,x_2) = \delta(v_2-g_2(x_2))$. From this, we can easily see that

$$
\begin{aligned}
p(x_1|u,u_1,v_1) &= \frac{p(x_1|u,u_1)\delta(v_1-g_1(x_1))}{p(v_1|u,u_1)} \\
p(x_2|u,u_2,v_2) &= \frac{p(x_2|u,u_2)\delta(v_2-g_2(x_1))}{p(v_2|u,u_2)}.
\end{aligned} \tag{12}
$$

We now generate codewords as follows. Transmitter 1 generates $2^{N(R_{1cc}+R_{2cc})}$ independent code-

$$
R_1 \le I(U,V_2,X_1;Y_1) \tag{10a}
$$

$$
R_1 \le I(X_1;Y_1|U,U_1,V_2) + \min(I(U_1;Y_2|U,X_2),C_{\mathsf{FB2}}-\delta_2) \tag{10b}
$$

$$
R_2 \le I(U,V_1,X_2;Y_2) \tag{10c}
$$

$$
R_2 \le I(X_2;Y_2|U,U_2,V_1) + \min(I(U_2;Y_1|U,X_1),C_{\mathsf{FB1}}-\delta_1) \tag{10d}
$$

$$
R_1+R_2 \le I(X_1;Y_1|U,V_1,V_2) + I(U,V_1,X_2;Y_2) \tag{10e}
$$

$$
R_1+R_2 \le I(X_2;Y_2|U,V_1,V_2) + I(U,V_2,X_1;Y_1) \tag{10f}
$$

$$
\begin{aligned}
R_1+R_2 \le{}& \min(I(U_2;Y_1|U,X_1),C_{\mathsf{FB1}}-\delta_1) + \min(I(U_1;Y_2|U,X_2),C_{\mathsf{FB2}}-\delta_2) \\
&+ I(X_1,V_2;Y_1|U,U_1,U_2) + I(X_2;Y_2|U,V_1,V_2)
\end{aligned} \tag{10g}
$$

$$
\begin{aligned}
R_1+R_2 \le{}& \min(I(U_2;Y_1|U,X_1),C_{\mathsf{FB1}}-\delta_1) + \min(I(U_1;Y_2|U,X_2),C_{\mathsf{FB2}}-\delta_2) + I(X_2,V_1;Y_2|U,U_1,U_2) \\
&+ I(X_1;Y_1|U,V_1,V_2)
\end{aligned} \tag{10h}
$$

$$
\begin{aligned}
R_1+R_2 \le{}& \min(I(U_2;Y_1|U,X_1),C_{\mathsf{FB1}}-\delta_1) + \min(I(U_1;Y_2|U,X_2),C_{\mathsf{FB2}}-\delta_2) + I(X_1,V_2;Y_1|U,V_1,U_2) \\
&+ I(X_2,V_1;Y_2|U,V_2,U_1)
\end{aligned} \tag{10i}
$$

$$
\begin{aligned}
2R_1+R_2 \le{}& I(U,V_2,X_1;Y_1) + I(X_1;Y_1|U,V_1,V_2) + I(X_2,V_1;Y_2|U,U_1,V_2) \\
&+ \min(I(U_1;Y_2|U,X_2),C_{\mathsf{FB2}}-\delta_2)
\end{aligned} \tag{10j}
$$

$$
\begin{aligned}
2R_1+R_2 \le{}& 2\min(I(U_1;Y_2|U,X_2),C_{\mathsf{FB2}}-\delta_2) + \min(I(U_2;Y_1|U,X_1),C_{\mathsf{FB1}}-\delta_1) + I(X_1,V_2;Y_1|U,U_1,U_2) \\
&+ I(X_1;Y_1|U,V_1,V_2) + I(X_2,V_1;Y_2|U,U_1,V_2)
\end{aligned} \tag{10k}
$$

$$
\begin{aligned}
R_1+2R_2 \le{}& I(U,V_1,X_2;Y_2) + I(X_2;Y_2|U,V_2,V_1) + I(X_1,V_2;Y_1|U,U_2,V_1) \\
&+ \min(I(U_2;Y_1|U,X_1),C_{\mathsf{FB1}}-\delta_1)
\end{aligned} \tag{10l}
$$

$$
\begin{aligned}
R_1+2R_2 \le{}& 2\min(I(U_2;Y_1|U,X_1),C_{\mathsf{FB1}}-\delta_1) + \min(I(U_1;Y_2|U,X_2),C_{\mathsf{FB2}}-\delta_2) + I(X_2,V_1;Y_2|U,U_1,U_2) \\
&+ I(X_2;Y_2|U,V_1,V_2) + I(X_1,V_2;Y_1|U,U_2,V_1)
\end{aligned} \tag{10m}
$$

over all jointdistributions

$$
p(u)p(u_1|u)p(u_2|u)p(v_1|u,u_1)p(v_2|u,u_2)p(x_1|u,u_1,v_1)p(x_2|u,u_2,v_2)p(\hat{y}_1|y_1)p(\hat{y}_2|y_2),
$$

where

$$
\begin{aligned}
\delta_1 &= I(\hat{Y}_1;Y_1|U,U_2,X_1) \\
\delta_2 &= I(\hat{Y}_2;Y_2|U,U_1,X_2).
\end{aligned}
$$

words $u^N(i,j)$ according to $\prod_{i=1}^{N} p(u_i)^1$, where $i \in \{1,\dots,2^{NR_{1cc}}\}$ and $j \in \{1,\dots,2^{NR_{2cc}}\}$. For each codeword $u^N(i,j)$, it generates $2^{NR_{1cc}}$ independent codewords $u_1^N((i,j),k)$ according to $\prod_{i=1}^{N} p(u_{1i}|u_i)$ where $k \in \{1,\dots,2^{NR_{1cc}}\}$. Subsequently, for each pair of codewords $\left(u^N(i,j),u_1^N((i,j),k)\right)$, generate $2^{NR_{1nc}}$ independent codewords $v_1^N((i,j),k,l)$ according to $\prod_{i=1}^{N} p(v_{1i}|u_i,u_{1i})$ where $l \in \{1,\dots,2^{NR_{1nc}}\}$. Finally for each triple of codewords $\left(u^N(i,j),u_1^N((i,j),k),v_1^N((i,j),k,l)\right)$, generate $2^{NR_{1p}}$ independent codewords $x_1^N((i,j),k,l,m)$ according to $\prod_{i=1}^{N} p(x_{1i}|u_i,u_{1i},v_{1i})$ where $m \in \{1,\dots,2^{NR_{1p}}\}$. On the other hand, receiver 1 generates $2^{N\hat{R}_1}$ sequences $\hat{y}_1^N(q)$ according to $\prod_{i=1}^{N} p(\hat{y}_{1i})$ where $q \in \{1,\dots,2^{N\hat{R}_1}\}$. In the feedback strategy (to be described shortly), we will see how this codebook generation leads to the joint distribution $p(\hat{y}_1|y_1)$. Similarly, receiver 2 generates $\hat{y}_2^N$.

As it will be clarified later, for a given block b, indices i and j in $u^N(i,j)$ correspond to the cooperative common message of transmitter 1 and transmitter 2 sent during block $(b-2)$, respectively. Then, $\prod_{i=1}^{N} p(u_{1i}|u_i)$ is used to create $2^{NR_{1cc}}$ independent codewords corresponding to the cooperative common message of transmitter 1 in $\{1,\dots,2^{NR_{1cc}}\}$. Similarly, $2^{NR_{1nc}}$ independent codewords are created according to $\prod_{i=1}^{N} p(v_{1i}|u_i,u_{1i})$, corresponding to the noncooperative common message of transmitter 1 in $\{1,\dots,2^{NR_{1nc}}\}$. Finally, $\prod_{i=1}^{N} p(x_{1i}|u_i,u_{1i},v_{1i})$ is used to create $2^{NR_{1p}}$ independent codewords corresponding to the private message of transmitter 1 in $\{1,\dots,2^{NR_{1p}}\}$.

Notation: Notations are independently used only for this section. The index k indicates the cooperative common message of user 1 instead of user index. The index i is used for both purposes: 1) indicating the previous cooperative common message of user 1; 2) indicating time index. It could be easily differentiated from contexts.

Feedback Strategy (Quantize-and-Binning): Focus on the bth transmission block. First receiver 1 quantizes its received signal $y_1^{N,(b)}$ into $\hat{y}_1^{N,(b+1)}$ with the rate of $\hat{R}$. Next, it finds an index q such that $\left(\hat{y}_1^{N,(b+1)}(q),y_1^{N,(b)}\right) \in \mathcal{T}_\epsilon^{(N)}$, where $q \in [1:2^{N\hat{R}_1}]$ and $\mathcal{T}_\epsilon^{(N)}$ indicates a jointly typical set. The quantization rate $\hat{R}_1$ is chosen so as to ensure the existence of such an index with probability 1. The covering lemma in [29] guides us to choose $\hat{R}_1$ such that

$$\hat{R}_1 \geq I(\hat{Y}_1;Y_1) \tag{13}$$

since under the aforementioned constraint, the probability that there is no such an index becomes arbitrarily small as N goes to infinity. Notice that with this choice of $\hat{R}_1$, the codebook $\hat{y}_1^N(q)$ according to $\prod_{i=1}^{N} p(\hat{y}_{1i})$ would match the codeword according to $\prod_{i=1}^{N} p(\hat{y}_{1i}|y_{1i})$.

We then partition the set of indices $q \in [1:2^{N\hat{R}_1}]$ into the number $2^{NC_{FB1}}$ of equal-size subsets (that we call *bins*)

$$\mathcal{B}(r) = \left[(r-1)2^{N(\hat{R}_1-C_{FB1})}+1 : r2^{N(\hat{R}_1-C_{FB1})}\right]$$
$$r \in [1:2^{NC_{FB1}}].$$

---

[1] With a slight abuse of notation, we use the same index i to represent time.

Now, the idea is to feed back the bin index r such that $q \in \mathcal{B}(r)$. This helps transmitter 1 to decode the quantized signal $\hat{y}_1^{N,(b)}$. Specifically, using the bin index r, transmitter 1 finds a unique index $q \in \mathcal{B}(r)$ such that $\left(\hat{y}_1^{N,(b+1)}(q),x_1^{N,(b)},u^{N,(b)}\right) \in \mathcal{T}_\epsilon^{(N)}$. Notice that by the packing lemma in [29], the decoding error probability goes to zero if

$$\hat{R}_1 - C_{FB1} \leq I(\hat{Y}_1;X_1,U). \tag{14}$$

Using (13) and (14), transmitter 1 can now decode the quantized signal as long as

$$I(\hat{Y}_1;Y_1|X_1,U) \leq C_{FB1}. \tag{15}$$

Similarly, transmitter 2 can decode $\hat{y}_2^{N,(b+1)}(q)$ if

$$I(\hat{Y}_2;Y_2|X_2,U) \leq C_{FB2}. \tag{16}$$

Encoding: Given $\hat{y}_1^{N,(b-1)}$ (decoded with the help of feedback), transmitter 1 finds a unique index $\hat{w}_{2cc}^{(b-2)} = \hat{k}$ (sent from transmitter 2 in the $(b-2)$th block) such that

$$\left(u^N(\cdot),u_1^N(\cdot),v_1^N(\cdot),x_1^N(\cdot),u_2^N(\cdot,\hat{k}),\hat{y}_1^{N,(b-1)}\right) \in \mathcal{T}_\epsilon^{(N)}$$

where $(\cdot)$ indicates the known messages $(w_{1cc}^{(b-4)},\hat{w}_{2cc}^{(b-4)},w_{1nc}^{(b-2)},w_{1p}^{(b-2)})$. Notice that due to the feedback delay, the feedback signal contains information of the $(b-2)$th block. We assume that $\hat{w}_{2cc}^{(b-2)}$ is correctly decoded from the previous block. By the packing lemma [29], the decoding error probability becomes arbitrarily small (as N goes to infinity) if

$$\begin{aligned}
R_{2cc} &\leq I(U_2;\hat{Y}_1|X_1,U) \\
&= I(\hat{Y}_1;Y_1|U,X_1) - I(\hat{Y}_1;Y_1|U,U_2,X_1) \\
&\leq \min(C_{FB1}-\delta_1, I(U_2;Y_1|X_1,U)) \tag{17}
\end{aligned}$$

where the last inequality follows from (15), $\delta_1 := I(\hat{Y}_1;Y_1|U,U_2,X_1)$ and $I(U_2;\hat{Y}_1|X_1,U) \leq I(U_2;Y_1|X_1,U)$.

Based on $(w_{1cc}^{(b-2)},\hat{w}_{2cc}^{(b-2)})$, transmitter 1 generates a new cooperative-common message $w_{1cc}^{(b)}$, a noncooperative-common message $w_{1nc}^{(b)}$, and a private message $w_{1p}^{(b)}$. It then sends x_1^N. Similarly, transmitter 2 decodes $\hat{w}_{1cc}^{(b-2)}$, generates $(w_{2cc}^{(b)},w_{2nc}^{(b)},w_{2p}^{(b)})$, and then sends x_2^N.

Decoding: Each receiver waits until total B blocks have been received and then does backward decoding. Notice that a block index b starts from the last B and ends to 1. For block b, receiver 1 finds the unique indices $(\hat{i},\hat{j},\hat{k},\hat{l})$ such that for some $m \in [1:2^{NR_{2nc}}]$

$$\begin{aligned}
&\Big(u^N(\hat{i},\hat{j}),u_1^N((\hat{i},\hat{j}),\hat{w}_{1cc}^{(b)}),v_1^N((\hat{i},\hat{j}),\hat{w}_{1cc}^{(b)},\hat{k}),x_1^N((\hat{i},\hat{j}),\hat{w}_{1cc}^{(b)}\\
&,\hat{k},\hat{l}),u_2^N((\hat{i},\hat{j}),\hat{w}_{2cc}^{(b)}),v_2^N((\hat{i},\hat{j}),\hat{w}_{2cc}^{(b)},m),y_1^{N,(b)}\Big) \in \mathcal{T}_\epsilon^{(N)}
\end{aligned}$$

where we assumed that a pair of messages $(\hat{w}_{1cc}^{(b)},\hat{w}_{2cc}^{(b)})$ was successfully decoded from the future blocks. Similarly, receiver 2 decodes $(\hat{w}_{1cc}^{(b-2)},\hat{w}_{2cc}^{(b-2)},\hat{w}_{2nc}^{(b)},\hat{w}_{2p}^{(b)})$.

Analysis of Probability of Error: By symmetry, we consider the probability of error only for block b and

for a pair of transmitter 1 and receiver 1. We assume that $(w_{1cc}^{(b-2)}, w_{2cc}^{(b-2)}, w_{1nc}^{(b)}, w_{1p}^{(b)}) = (1,1,1,1)$ was sent through the blocks; and there was no backward decoding error from the future blocks, i.e., $(\hat{w}_{1cc}^{(b)}, \hat{w}_{2cc}^{(b)})$ are successfully decoded.

Define an event

$$E_{ijklm} = \Big\{ \big(u^N(i,j), u_1^N((i,j), \hat{w}_{1cc}^{(b)}), v_1^N((i,j), \hat{w}_{1cc}^{(b)}, k) \Big.$$
$$, x_1^N((i,j), \hat{w}_{1cc}^{(b)}, k, l), u_2^N((i,j), \hat{w}_{2cc}^{(b)}), v_2^N((i,j), \hat{w}_{2cc}^{(b)}, m)$$
$$\Big. , y_1^{N,(b)} \big) \in \mathcal{T}_\epsilon^{(N)} \Big\}.$$

Let E_{1111m}^c be the complement of the set E_{1111m}. Then, by AEP, $\Pr(E_{1111m}^c) \to 0$ as N goes to infinity. Hence, we focus only on the following error event:

$$\Pr\left(\bigcup_{(i,j,k,l) \neq (1,1,1,1),m} E_{ijklm} \right)$$
$$\leq \underbrace{\sum_{(i,j) \neq (1,1)} \Pr\left(\bigcup_{k,l,m} E_{ijklm} \right)}_{\triangleq \Pr(E_1)}$$
$$+ \underbrace{\sum_{k \neq 1} \Pr\left(\bigcup_l E_{11kl1} \right)}_{\triangleq \Pr(E_2)} + \underbrace{\sum_{k \neq 1, m \neq 1} \Pr\left(\bigcup_l E_{11klm} \right)}_{\triangleq \Pr(E_3)}$$
$$+ \underbrace{\sum_{l \neq 1} \Pr(E_{111l1})}_{\triangleq \Pr(E_4)} + \underbrace{\sum_{l \neq 1, m \neq 1} \Pr(E_{111lm})}_{\triangleq \Pr(E_5)}. \quad (18)$$

Here, we hav

$$\Pr(E_1) \leq 24$$
$$\times 2^{N(R_{1cc}+R_{2cc}+R_{1nc}+R_{2nc}+R_{1p}-I(U,X_1,V_2;Y_1)+5\epsilon)}$$
$$\Pr(E_2) \leq 2 \times 2^{N(R_{1nc}+R_{1p}-I(X_1;Y_1|U,U_1,V_2)+2\epsilon)}$$
$$\Pr(E_3) \leq 2 \times 2^{N(R_{1nc}+R_{2nc}+R_{1p}-I(X_1,V_2;Y_1|U,U_1,U_2)+3\epsilon)}$$
$$\Pr(E_4) \leq 2^{N(R_{1p}-I(X_1;Y_1|U,V_1,V_2)+\epsilon)}$$
$$\Pr(E_5) \leq 2^{N(R_{2nc}+R_{1p}-I(X_1,V_2;Y_1|U,V_1,U_2)+2\epsilon)}.$$

Notice in $\Pr(E_1)$ that as long as $(i,j) \neq (1,1)$, all of the cases (decided depending on whether or not $k \neq 1$, $l \neq 1$, and $m \neq 1$) are dominated by the worst case bound that occurs when $k \neq 1$, $l \neq 1$, and $m \neq 1$. Since $(i,j) \neq (1,1)$ covers three different cases and we have eight different cases depending on the values of (k,l,m), we have 24 cases in total. This number reflects the constant 24 in the earlier first inequality. Similarly, we get the other four inequalities as earlier.

Hence, the probability of error can be made arbitrarily small if

$$\begin{cases} R_{2cc} \leq \min(I(U_2;Y_1|U,X_1), C_{\mathsf{FB}1} - \delta_1) \\ R_{1p} \leq I(X_1;Y_1|U,V_1,V_2) \\ R_{1p} + R_{2nc} \leq I(X_1,V_2;Y_1|U,V_1,U_2) \\ R_{1p} + R_{1nc} \leq I(X_1;Y_1|U,U_1,V_2) \\ R_{1p} + R_{1nc} + R_{2nc} \leq I(X_1,V_2;Y_1|U,U_1,U_2) \\ R_{1p} + R_{1cc} + R_{2cc} + R_{1nc} + R_{2nc} \leq I(U,V_2,X_1;Y_1) \end{cases} \quad (19)$$

$$\begin{cases} R_{1cc} \leq \min(I(U_1;Y_2|U,X_2), C_{\mathsf{FB}2} - \delta_2) \\ R_{2p} \leq I(X_2;Y_2|U,V_1,V_2) \\ R_{2p} + R_{1nc} \leq I(X_2,V_1;Y_2|U,V_2,U_1) \\ R_{2p} + R_{2nc} \leq I(X_2;Y_2|U,U_2,V_1) \\ R_{2p} + R_{2nc} + R_{1nc} \leq I(X_2,V_1;Y_2|U,U_1,U_2) \\ R_{2p} + R_{2cc} + R_{1cc} + R_{1nc} + R_{2nc} \leq I(U,V_1,X_2;Y_2). \end{cases} \quad (20)$$

Employing Fourier–Motzkin elimination, we finally get the bounds of (10a)–(10m). ∎

Remark 1 (Connection to Related Work [18], [21]): The three-fold message splitting in our achievable scheme is a special case of the more-than-three-fold message splitting introduced in [18] and [21]. Also our scheme has similarity to the schemes in [18] and [21] in a sense that the three techniques (message-splitting, block Markov encoding, and backward decoding) are jointly employed. However, due to a fundamental difference between our rate-limited feedback problem and the conferencing encoder problem in [18] and [21], a new scheme is required for feedback strategy and this is reflected as the quantize-and-binning scheme in the El Gamal–Costa deterministic model. It turns out this distinction leads to a new lattice-code-based scheme in the Gaussian case, as will be explained in Section VI.

B. Outer Bound

Theorem 2: The capacity region of the two-user El Gamal–Costa deterministic IC with rate-limited feedback (as described in Section II) is included by the set $\bar{\mathcal{C}}$ of (R_1, R_2) such that

$$R_1 \leq \min\{H(Y_1), H(Y_1|V_1,V_2,U_1) \quad (21a)$$
$$+ H(Y_2|X_2,U_1)\}$$
$$R_1 \leq H(Y_1|X_2,U_1) + C_{\mathsf{FB}2} \quad (21b)$$
$$R_2 \leq \min\{H(Y_2), H(Y_2|V_1,V_2,U_2) \quad (21c)$$
$$+ H(Y_1|X_1,U_2)\}$$
$$R_2 \leq H(Y_2|X_1,U_2) + C_{\mathsf{FB}1} \quad (21d)$$
$$R_1 + R_2 \leq H(Y_1|V_1,V_2,U_1) + H(Y_2) \quad (21e)$$
$$R_1 + R_2 \leq H(Y_2|V_1,V_2,U_2) + H(Y_1) \quad (21f)$$
$$R_1 + R_2 \leq H(Y_1|V_1) + H(Y_2|V_2) \quad (21g)$$
$$+ C_{\mathsf{FB}1} + C_{\mathsf{FB}2}$$
$$2R_1 + R_2 \leq H(Y_1) + H(Y_1|V_1,V_2,U_1) \quad (21h)$$
$$+ H(Y_2|V_2) + C_{\mathsf{FB}2}$$
$$R_1 + 2R_2 \leq H(Y_2) + H(Y_2|V_1,V_2,U_2) \quad (21i)$$
$$+ H(Y_1|V_1) + C_{\mathsf{FB}1}$$

for joint distributions $p(u_1, u_2)p(x_1|u_1, u_2)p(x_2|u_1, u_2)$. As depicted in Fig. 2, $C_{\mathsf{FB}1}$ and $C_{\mathsf{FB}2}$ indicate the capacity of each feedback link.

Remark 2: In the nonfeedback case, i.e., $C_{\mathsf{FB}1} = C_{\mathsf{FB}2} = 0$, by setting $U_1 = U_2 = \emptyset$, we recover the outer bounds in [11, Th. 1]. Note that in this case, $H(Y_1|X_2) = H(Y_1|V_2)$ and $H(Y_2|X_1) = H(Y_2|V_1)$. In fact, our achievable region of Theorem 1 matches the outer bound under this model, thereby achieving the nonfeedback capacity region.

Remark 3 (Feedback Gain Under Symmetric Feedback Cost): Notice from (21g) that the sum-rate capacity can be at most in-

creased by the rate of available feedback, i.e., one bit of feedback provides a capacity increase of at most one bit. Therefore, if the cost of using the feedback link is the same as that of using the forward link, there is no feedback gain under the feedback cost. However, it turns out that there is indeed feedback gain when the costs are *asymmetric*. This will be discussed in more details in Remark 5 of Section V.

Proof: By symmetry, it suffices to prove the bounds of (21a), (21b), (21e), (21g), and (21h). The bounds of (21a) and (21b) are nothing but the cutset bounds (see Appendix A for details). Also, (21e) is the bound when the feedback link has infinite capacity [9]. Hence, proving the bounds of (21g) and (21h) is the main focus of the proof. We will present the proof of (21g) here and for completeness, the proof for all other bounds is provided in Appendix A.

Proof of (21g):

$$N(R_1 + R_2 - \epsilon_N) \leq I(W_1; Y_1^N) + I(W_2; Y_2^N)$$
$$= H(Y_1^N) - H(Y_1^N|W_1) + H(Y_2^N) - H(Y_2^N|W_2)$$
$$\overset{(a)}{=} H(Y_1^N) - H(V_1^N|W_2) + H(Y_2^N) - H(V_2^N|W_1)$$
$$= H(Y_1^N) - [H(V_1^N) - I(V_1^N; W_2)] + H(Y_2^N)$$
$$\quad - [H(V_2^N) - I(V_2^N; W_1)]$$
$$\overset{(b)}{\leq} I(V_1^N; W_2) + I(V_2^N; W_1) + H(Y_1^N, V_1^N) - H(V_1^N)$$
$$\quad + H(Y_2^N, V_2^N) - H(V_2^N)$$
$$= I(V_1^N; W_2) + I(V_2^N; W_1) + H(Y_1^N|V_1^N) + H(Y_2^N|V_2^N)$$
$$\overset{(c)}{\leq} I(V_1^N, W_1, \tilde{Y}_1^N; W_2) + I(V_2^N, W_2, \tilde{Y}_2^N; W_1)$$
$$\quad + H(Y_1^N|V_1^N) + H(Y_2^N|V_2^N)$$
$$\overset{(d)}{=} I(W_1, \tilde{Y}_1^N; W_2) + I(W_2, \tilde{Y}_2^N; W_1) + H(Y_1^N|V_1^N)$$
$$\quad + H(Y_2^N|V_2^N)$$
$$= I(\tilde{Y}_1^N; W_2|W_1) + I(\tilde{Y}_2^N; W_1|W_2) + H(Y_1^N|V_1^N) +$$
$$H(Y_2^N|V_2^N)$$
$$= H(\tilde{Y}_1^N|W_1) + H(\tilde{Y}_2^N|W_2) + H(Y_1^N|V_1^N) + H(Y_2^N|V_2^N)$$
$$\overset{(e)}{\leq} N(C_{\mathsf{FB1}} + C_{\mathsf{FB2}}) + \sum H(Y_{1i}|V_{1i}) + \sum H(Y_{2i}|V_{2i})$$

where (a) follows from $H(Y_1^N|W_1) = H(V_2^N|W_1)$ and $H(Y_2^N|W_2) = H(V_1^N|W_2)$ (see Claim 1 in the following); (b) follows from providing V_1^N and V_2^N to receiver 1 and 2, respectively; (c) follows from the fact that adding information increases mutual information; (d) follows from the fact that V_k^N is a function of $(W_k, \tilde{Y}_k^{N-1})$; (e) follows from the fact that $H(\tilde{Y}_k^N|W_k) \leq NC_{\mathsf{FB}k}$ and conditioning reduces entropy.

Claim 1: $H(Y_1^N|W_1) = H(V_2^N|W_1)$ and $H(Y_2^N|W_2) = H(V_1^N|W_2)$.

Proof: By symmetry, it suffices to prove the first one

$$H(Y_1^N|W_1) = \sum H(Y_{1i}|Y_1^{i-1}, W_1)$$
$$\overset{(a)}{=} \sum H(V_{2i}|Y_1^{i-1}, W_1)$$
$$\overset{(b)}{=} \sum H(V_{2i}|Y_1^{i-1}, W_1, X_1^i, V_2^{i-1})$$
$$\overset{(c)}{=} \sum H(V_{2i}|W_1, V_2^{i-1}) = H(V_2^N|W_1)$$

where (a) follows from the fact that Y_{1i} is a function of (X_{1i}, V_{2i}) and X_{1i} is a function of (W_1, Y_1^{i-1}); (b) follows

from the fact that X_1^i is a function of (W_1, Y_1^{i-1}) and V_2^{i-1} is a function of (Y_1^{i-1}, X_1^{i-1}); (c) follows from the fact that Y_1^{i-1} is a function of (X_1^{i-1}, V_2^{i-1}) and X_1^i is a function of (W_1, V_2^{i-1}) (see Claim 2 in the following). ∎

Claim 2: For $i \geq 1$, X_1^i is a function of (W_1, V_2^{i-1}). Similarly, X_2^i is a function of (W_2, V_1^{i-1}).

Proof: By symmetry, it is enough to prove the first one. Since the channel is deterministic, X_1^i is a function of (W_1, W_2). In Fig. 2, we see that information of W_2 to the first link pair must pass through V_{2i}. Also note that X_{1i} depends on the past output sequences until $i-1$. Therefore, X_1^i is a function of (W_1, V_2^{i-1}). ∎

V. LINEAR DETERMINISTIC IC

In this section, we consider the linear deterministic IC with rate-limited feedback described in Section II. Since this model is a special case of the El Gamal–Costa model, our inner and outer bounds derived in the previous section also apply to this model. We show that the inner bound and the outer bound derived in Theorems 1 and 2, respectively, coincide under this linear deterministic model, thus establishing the capacity region.

Theorem 3: The capacity region of the linear deterministic IC with rate-limited feedback is the set of non-negative (R_1, R_2) satisfying

$$R_1 \leq \min\{\max(n_{11}, n_{21}), \max(n_{11}, n_{12})\} \quad (22a)$$
$$R_1 \leq n_{11} + C_{\mathsf{FB2}} \quad (22b)$$
$$R_2 \leq \min\{\max(n_{22}, n_{12}), \max(n_{22}, n_{21})\} \quad (22c)$$
$$R_2 \leq n_{22} + C_{\mathsf{FB1}} \quad (22d)$$
$$R_1 + R_2 \leq (n_{11} - n_{12})^+ + \max(n_{22}, n_{12}) \quad (22e)$$
$$R_1 + R_2 \leq (n_{22} - n_{21})^+ + \max(n_{11}, n_{21}) \quad (22f)$$
$$R_1 + R_2 \leq \max\{n_{21}, (n_{11} - n_{12})^+\} \quad (22g)$$
$$\quad + \max\{n_{12}, (n_{22} - n_{21})^+\} + C_{\mathsf{FB1}} + C_{\mathsf{FB2}}$$
$$2R_1 + R_2 \leq (n_{11} - n_{12})^+ + \max(n_{11}, n_{21}) \quad (22h)$$
$$\quad + \max\{n_{12}, (n_{22} - n_{21})^+\} + C_{\mathsf{FB2}}$$
$$R_1 + 2R_2 \leq (n_{22} - n_{21})^+ + \max(n_{22}, n_{12}) \quad (22i)$$
$$\quad + \max\{n_{21}, (n_{11} - n_{12})^+\} + C_{\mathsf{FB1}}.$$

Remark 4: In the nonfeedback case, i.e., $C_{\mathsf{FB1}} = C_{\mathsf{FB2}} = 0$, this theorem recovers the result in [11] and [13]. In the infinite feedback case, i.e., $C_{\mathsf{FB1}} = C_{\mathsf{FB2}} = \infty$, this recovers the result in [9] and [30]. Considering the sum-rate capacity under symmetric setting, i.e., $n_{11} = n_{22} = n$, $n_{12} = n_{21} = m$, $C_{\mathsf{FB1}} = C_{\mathsf{FB2}}$, this recovers the result of [31].

Proof: The converse proof is trivial due to Theorem 2. For achievability, we will use the result in Theorem 1. By choosing the following input distribution, we will show the tightness of the outer bound. $\forall\, k \in \{1, 2\}$ and $j \neq k$, we choose

$$U = \varnothing$$
$$U_k = U \oplus X_{kcc}$$
$$V_k = U_k \oplus X_{knc} \quad (23)$$
$$X_k = V_k \oplus X_{kp}$$
$$\hat{Y}_k = \mathsf{LSB}_{\min(n_{kj}, C_{\mathsf{FB}j})}(Y_k)$$

where for any column vector A, $\mathsf{LSB}_n(A)$ takes the bottom $n(n \leq |A|)$ entries of A while returning zeros for the remaining part; and X_{kcc}, X_{knc}, and X_{kp} are independent random vectors of size $\max\{n_{kk}, n_{kj}\}$, such that

1) the random vector X_{kp} consists of $(n_{kk} - n_{kj})^+$ i.i.d. $\mathsf{Ber}\left(\frac{1}{2}\right)$ random variables at the bottom, denoted by $*$ in (24), corresponding to the number of private signal levels of transmitter k;
2) the random vector X_{kcc} consists of $(n_{kk} - n_{kj})^+$ i.i.d. $\mathsf{Ber}\left(\frac{1}{2}\right)$ random variables in the middle (above the private signal levels), denoted by $*$ in (24), corresponding to the number of common signal levels that will be resent cooperatively through the other communication link with the help of feedback;
3) the random vector X_{knc} consists of $(n_{kj} - C_{\mathsf{FBj}})^+$ i.i.d. $\mathsf{Ber}\left(\frac{1}{2}\right)$ random variables at the top, denoted by $*$ in (24), corresponding to the number of noncooperative common signal levels.

As we show in Appendix B, with this choice of random variables, the achievable region of Theorem 1 matches the outer bounds in Theorem 2

$$X_k = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \hline 0 \\ \hline \vdots \\ 0 \\ \hline * \\ \vdots \\ * \end{bmatrix} \oplus \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \hline * \\ \hline \vdots \\ * \\ \hline 0 \\ \vdots \\ 0 \end{bmatrix} \oplus \begin{bmatrix} * \\ \vdots \\ * \\ \hline 0 \\ \hline \vdots \\ 0 \\ \hline 0 \\ \vdots \\ 0 \end{bmatrix}, \quad k = 1, 2.$$

$$\underbrace{}_{X_{kp}} \quad \underbrace{}_{X_{kcc}} \quad \underbrace{}_{X_{knc}} \tag{24}$$

■

It is worth utilizing Theorem 3 to illustrate the impact of feedback on the sum-rate capacity of the linear deterministic IC. Consider a symmetric case where $n_{11} = n_{22} = n$, $n_{12} = n_{21} = \alpha n$, and $C_{\mathsf{FB1}} = C_{\mathsf{FB2}} = \beta n$. Using Theorem 3, we can derive the sum-rate capacity of this network (normalized by n)

$$\frac{C_{\mathsf{sum}}}{n} = \begin{cases} \min(2 - 2\alpha + 2\beta, 2 - \alpha) & \text{for } \alpha \in [0, 0.5] \\ \min(2\alpha + 2\beta, 2 - \alpha) & \text{for } \alpha \in [0.5, \frac{2}{3}] \\ 2 - \alpha & \text{for } \alpha \in [\frac{2}{3}, 1] \\ \alpha & \text{for } \alpha \in [1, 2 + 2\beta] \\ 2 + 2\beta & \text{for } \alpha \in [2 + 2\beta, \infty). \end{cases} \tag{25}$$

Fig. 5 illustrates the (normalized) sum-rate capacity as a function of α, for different values of $\beta = 0$ (i.e., no feedback), $\beta = \infty$ (i.e., infinite feedback), and $\beta = 0.125$. We note the following cases.

1) Case 1 $\left(\alpha \in \left[0, \frac{1}{2}\right]\right)$: In this regime, the sum-rate capacity is increased by the total amount of feedback rates and saturates at $2 - \alpha$ once the rate of each feedback link is larger than $\frac{\alpha n}{2}$.
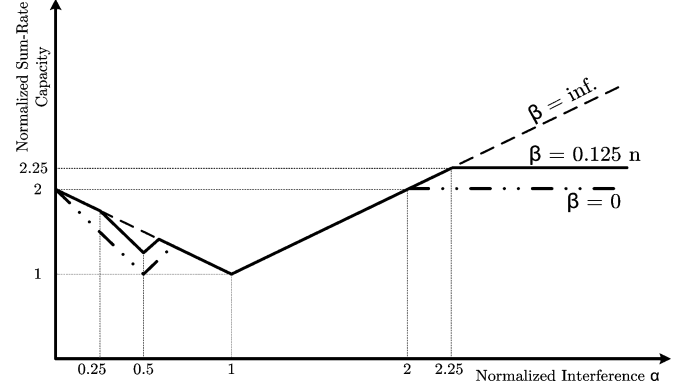

Fig. 5. Normalized sum-rate capacity for $\beta = 0$, $\beta = 0.125$, and $\beta = \infty$.

2) Case 2 $\left(\alpha \in \left[\frac{1}{2}, \frac{2}{3}\right]\right)$: In this regime, the sum-rate capacity is increased by the total amount of feedback rates and saturates at 2α once the rate of each feedback link is larger than $\frac{(2n - 3\alpha n)}{2}$.
3) Case 3 $\left(\alpha \in \left[\frac{2}{3}, 2\right]\right)$: In this regime, feedback does not increase the capacity.
4) Case 4 $(\alpha \in [2 + 2\beta, \infty))$: In this regime, the sum-rate capacity is increased by at most the total amount of feedback rates.

Remark 5 (Feedback Gain Under Asymmetric Feedback Cost): As it can be seen in (25), the sum-rate capacity is increased by at most the total amount of feedback rates. Let the *cost* be the amount of resources (e.g., time, frequency) paid for sending one bit. With this cost in mind, let us consider the effective gain of using feedback which counts the cost. Notice that by Cases 1, 2, and 4, there are many channel parameter scenarios where one bit of feedback can provide a capacity increase of exactly one bit. This implies that the effective feedback gain depends on the cost difference between feedback and forward links. So if the feedback cost is cheaper than that of using forward link, then there is indeed feedback gain. The cellular network may be this case. Suppose that downlink is used for feedback purpose, while uplink is used as a forward link. Then, this is the scenario where the feedback cost is cheaper than the cost of using the forward link, as downlink power is typically larger than uplink power, thus inducing cheaper feedback cost.

VI. GAUSSIAN IC

In this section, we consider the Gaussian IC with rate-limited feedback, described in Section II. We first derive an outer bound on the capacity region of this network. We then develop an achievability strategy based on the techniques discussed in the previous sections and then show that for symmetric channel gains it achieves a sum-rate within a constant gap to the optimality.

A. Outer Bound

Theorem 4: The capacity region of the Gaussian IC with rate-limited feedback is included by the set $\bar{\mathcal{C}}$ of (R_1, R_2) satisfying

(26a)–(26k), shown at the bottom of the page, for $0 \leq \rho \leq 1$ where $\rho = \mathbb{E}[X_1 X_2^*]^2$.

Proof: By symmetry, it suffices to prove the bounds of (26a)–(26c), (26g), (26h), and (26j). The bounds of (26a)–(26c) are nothing but cutset bounds. The bound of (26h) corresponds to the case of infinite feedback rate and was derived in [9]. Hence, proving the bounds of (26g) and (26j) is the main focus of this proof. We will present the proof of (26g) here, and defer the proof for remaining bounds to Appendix C.

Proof of (26g): The proof idea mostly follows the deterministic case proof of (21g). The only difference in the Gaussian case is that we define a noisy version of $g_{12} X_1^N$ corresponding to V_1^N in the deterministic case: $S_1^N := g_{12} X_1^N + Z_2^N$. Similarly, we define $S_2^N := g_{21} X_2^N + Z_1^N$ to mimic V_2^N. With this, we can now get

$$N(R_1 + R_2 - \epsilon_N) \leq I(W_1; Y_1^N) + I(W_2; Y_2^N)$$

$$\overset{(a)}{=} h(Y_1^N) - h(S_1^N|W_2) + h(Y_2^N) - h(S_2^N|W_1)$$

$$= h(Y_1^N) - [h(S_1^N) - I(S_1^N; W_2)] + h(Y_2^N)$$
$$- [h(S_2^N) - I(S_2^N; W_1)]$$

$$\overset{(b)}{\leq} I(S_1^N; W_2) + I(S_2^N; W_1)$$
$$+ h(Y_1^N) - [h(S_1^N) - h(S_1^N|Y_1^N) + h(S_1^N|Y_1^N, X_1^N)]$$
$$+ h(Y_2^N) - [h(S_2^N) - h(S_2^N|Y_2^N) + h(S_2^N|Y_2^N, X_2)]$$

[2]ρ captures the power gain that can be achieved by making the transmit signals correlated.

$$= I(S_1^N; W_2) + I(S_2^N; W_1) + h(Y_1^N) - I(S_1^N; Y_1^N)$$
$$- \sum h(Z_{1i}) + h(Y_2^N) - I(S_2^N; Y_1^N) - \sum h(Z_{2i})$$

$$= I(S_1^N; W_2) + I(S_2^N; W_1) + h(Y_1^N|S_1^N) - \sum h(Z_{1i})$$
$$+ h(Y_2^N|S_2^N) - \sum h(Z_{2i})$$

$$\overset{(c)}{\leq} I(S_1^N, W_1, \tilde{Y}_1^N; W_2) + I(S_2^N, W_2, \tilde{Y}_2^N; W_1)$$
$$+ h(Y_1^N|S_1^N) - \sum h(Z_{1i}) + h(Y_2^N|S_2^N) - \sum h(Z_{2i})$$

$$\overset{(d)}{=} I(W_1, \tilde{Y}_1^N; W_2) + I(W_2, \tilde{Y}_2^N; W_1) + h(Y_1^N|S_1^N)$$
$$- \sum h(Z_{1i}) + h(Y_2^N|S_2^N) - \sum h(Z_{2i})$$

$$= I(\tilde{Y}_1^N; W_2|W_1) + I(\tilde{Y}_2^N; W_1|W_2) + h(Y_1^N|S_1^N)$$
$$- \sum h(Z_{1i}) + h(Y_2^N|S_2^N) - \sum h(Z_{2i})$$

$$\leq H(\tilde{Y}_1^N|W_1) + H(\tilde{Y}_2^N|W_2) + h(Y_1^N|S_1^N)$$
$$- \sum h(Z_{1i}) + h(Y_2^N|S_2^N) - \sum h(Z_{2i})$$

$$= \sum H(\tilde{Y}_{1i}|W_1, \tilde{Y}_1^{i-1}, X_{1i})$$
$$+ \sum H(\tilde{Y}_{2i}|W_2, \tilde{Y}_2^{i-1}, X_{2i})$$
$$+ \sum [h(Y_{1i}|Y_1^{i-1}, S_1^N) - h(Z_{1i})]$$
$$+ \sum [h(Y_{2i}|Y_2^{i-1}, S_2^N) - h(Z_{2i})]$$

$$\overset{(e)}{\leq} \sum H(\tilde{Y}_{1i}|X_{1i}) + \sum H(\tilde{Y}_{2i}|X_{2i})$$

$$R_1 \leq \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_{21} + 2\rho\sqrt{\mathsf{SNR}_1 \cdot \mathsf{INR}_{21}}\right) \tag{26a}$$

$$R_1 \leq \log\left(1 + \frac{(1-\rho^2)\mathsf{SNR}_1}{1 + (1-\rho^2)\mathsf{INR}_{12}}\right) + \log\left(1 + (1-\rho^2)\mathsf{INR}_{12}\right) \tag{26b}$$

$$R_1 \leq \log\left(1 + (1-\rho^2)\mathsf{SNR}_1\right) + C_{\mathsf{FB2}} \tag{26c}$$

$$R_2 \leq \log\left(1 + \mathsf{SNR}_2 + \mathsf{INR}_{12} + 2\rho\sqrt{\mathsf{SNR}_2 \cdot \mathsf{INR}_{12}}\right) \tag{26d}$$

$$R_2 \leq \log\left(1 + \frac{(1-\rho^2)\mathsf{SNR}_2}{1 + (1-\rho^2)\mathsf{INR}_{21}}\right) + \log\left(1 + (1-\rho^2)\mathsf{INR}_{21}\right) \tag{26e}$$

$$R_2 \leq \log\left(1 + (1-\rho^2)\mathsf{SNR}_2\right) + C_{\mathsf{FB1}} \tag{26f}$$

$$R_1 + R_2 \leq \log\left(1 + (1-\rho^2)\mathsf{INR}_{21} + \frac{(1-\rho^2)\mathsf{SNR}_1}{1 + (1-\rho^2)\mathsf{INR}_{12}}\right) + \log\left(1 + (1-\rho^2)\mathsf{INR}_{12} + \frac{(1-\rho^2)\mathsf{SNR}_2}{1 + (1-\rho^2)\mathsf{INR}_{21}}\right)$$
$$+ \min\left\{C_{\mathsf{FB1}}, \log\left(1 + (1-\rho^2)\mathsf{INR}_{21}\right)\right\} + \min\left\{C_{\mathsf{FB2}}, \log\left(1 + (1-\rho^2)\mathsf{INR}_{12}\right)\right\} \tag{26g}$$

$$R_1 + R_2 \leq \log\left(1 + \frac{(1-\rho^2)\mathsf{SNR}_1}{1 + (1-\rho^2)\mathsf{INR}_{12}}\right) + \log\left(1 + \mathsf{SNR}_2 + \mathsf{INR}_{12} + 2\rho\sqrt{\mathsf{SNR}_2 \cdot \mathsf{INR}_{12}}\right) \tag{26h}$$

$$R_1 + R_2 \leq \log\left(1 + \frac{(1-\rho^2)\mathsf{SNR}_2}{1 + (1-\rho^2)\mathsf{INR}_{21}}\right) + \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_{21} + 2\rho\sqrt{\mathsf{SNR}_1 \cdot \mathsf{INR}_{21}}\right) \tag{26i}$$

$$2R_1 + R_2 \leq \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_{21} + 2\rho\sqrt{\mathsf{SNR}_1 \cdot \mathsf{INR}_{21}}\right) + \log\left(1 + \frac{(1-\rho^2)\mathsf{SNR}_1}{1 + (1-\rho^2)\mathsf{INR}_{12}}\right)$$
$$+ \log\left(1 + (1-\rho^2)\mathsf{INR}_{12} + \frac{(1-\rho^2)\mathsf{SNR}_2}{1 + (1-\rho^2)\mathsf{INR}_{21}}\right) + \min\left\{C_{\mathsf{FB2}}, \log\left(1 + (1-\rho^2)\mathsf{INR}_{12}\right)\right\} \tag{26j}$$

$$R_1 + 2R_2 \leq \log\left(1 + \mathsf{SNR}_2 + \mathsf{INR}_{12} + 2\rho\sqrt{\mathsf{SNR}_2 \cdot \mathsf{INR}_{12}}\right) + \log\left(1 + \frac{(1-\rho^2)\mathsf{SNR}_2}{1 + (1-\rho^2)\mathsf{INR}_{21}}\right)$$
$$+ \log\left(1 + (1-\rho^2)\mathsf{INR}_{21} + \frac{(1-\rho^2)\mathsf{SNR}_1}{1 + (1-\rho^2)\mathsf{INR}_{12}}\right) + \min\left\{C_{\mathsf{FB1}}, \log\left(1 + (1-\rho^2)\mathsf{INR}_{21}\right)\right\} \tag{26k}$$

$$+ \sum [h(Y_{1i}|S_{1i}) - h(Z_{1i})] + \sum [h(Y_{2i}|S_{2i}) - h(Z_{2i})]$$

where (a) follows from $h(Y_1^N|W_1) = h(S_2^N|W_1)$ and $h(Y_2^N|W_2) = h(S_1^N|W_2)$ (see Claim 3 in the following); (b) follows from providing S_1^N and S_2^N to receivers 1 and 2, respectively; (c) follows from the fact that adding information increases mutual information; (d) follows from the fact that S_k^N is a function of $(W_k, \tilde{Y}_k^{N-1})$; and (e) follows from the fact that $H(\tilde{Y}_{ki}|X_{ki}) \leq NC_{\mathsf{FB}k}$ and conditioning reduces entropy.

Claim 3: $h(S_1^N|W_2) = h(Y_2^N|W_2)$.

Proof:

$$\begin{aligned} h(Y_2^N|W_2) &= \sum h(Y_{2i}|Y_2^{i-1}, W_2) \\ &\overset{(a)}{=} \sum h(S_{1i}|Y_2^{i-1}, W_2) \\ &\overset{(b)}{=} \sum h(S_{1i}|Y_2^{i-1}, W_2, X_2^i, S_1^{i-1}) \\ &\overset{(c)}{=} \sum h(S_{1i}|W_2, S_1^{i-1}) = h(S_1^N|W_2) \end{aligned}$$

where (a) follows from the fact that Y_{2i} is a function of (X_{2i}, S_{1i}) and X_{2i} is a function of (W_2, Y_2^{i-1}); (b) follows from the fact that X_2^i is a function of (W_2, Y_2^{i-1}) and S_1^{i-1} is a function of (Y_2^{i-1}, X_2^{i-1}); (c) follows from Claim 4 (see in the following). ∎

Claim 4: For all $i \geq 1$, X_1^i is a function of (W_1, S_2^{i-1}) and X_2^i is a function of (W_2, S_1^{i-1}).

Proof: By symmetry, it is enough to prove only one. Notice that X_2^i is a function of (W_2, Y_2^{i-1}) and Y_2^{i-1} is a function of (X_2^{i-1}, S_1^{i-1}). Hence, X_2^i is a function of $(W_2, X_2^{i-1}, S_1^{i-1})$. Iterating the same argument, we conclude that X_2^i is a function of (W_2, X_{21}, S_1^{i-1}). Since X_{21} depends only on W_2, we complete the proof. ∎

From the earlier, we get

$$\begin{aligned} R_1 + R_2 \leq &\, h(Y_1|S_1) - h(Z_1) + h(Y_2|S_2) - h(Z_2) \\ &+ C_{\mathsf{FB}1} + C_{\mathsf{FB}2}. \end{aligned} \tag{27}$$

To incorporate the case of infinite feedback rate, we can slightly modify the last few steps in the previous equation to obtain

$$\begin{aligned} R_1 + R_2 \leq &\, h(Y_1|S_1) - h(Z_1) + h(Y_2|S_2) - h(Z_2) \\ &+ \min\{C_{\mathsf{FB}1}, h(Y_1|X_1) - h(Z_1)\} \\ &+ \min\{C_{\mathsf{FB}2}, h(Y_2|X_2) - h(Z_2)\}. \end{aligned} \tag{28}$$

Therefore, we get the desired upper bound. ∎

If we consider the symmetric channel gains, *i.e.*

$$\begin{aligned} |h_{11}| &= |h_{22}| = |h_d| \\ |h_{12}| &= |h_{21}| = |h_c| \end{aligned} \tag{29}$$

and

$$\begin{aligned} \mathsf{SNR}_1 &= \mathsf{SNR}_2 = \mathsf{SNR} = |h_d|^2 \\ \mathsf{INR}_{12} &= \mathsf{INR}_{21} = \mathsf{INR} = |h_c|^2 \end{aligned} \tag{30}$$

we get the following outer-bound result.

Corollary 1: The sum-rate capacity of the symmetric Gaussian IC with rate-limited feedback is included by the set $\bar{\mathcal{C}}_{\mathsf{sym}}$ of $R_1 + R_2$ satisfying

$$R_1 + R_2 \leq 2\log(1 + \mathsf{SNR}) + C_{\mathsf{FB}1} + C_{\mathsf{FB}2} \tag{31a}$$

$$R_1 + R_2 \leq \log\left(1 + \frac{\mathsf{SNR}}{1 + \mathsf{INR}}\right) \tag{31b}$$

$$+ \log\left(1 + \mathsf{SNR} + \mathsf{INR} + 2\sqrt{\mathsf{SNR} \cdot \mathsf{INR}}\right)$$

$$R_1 + R_2 \leq 2\log\left(1 + \mathsf{INR} + \frac{\mathsf{SNR}}{1 + \mathsf{INR}}\right) \tag{31c}$$

$$+ \min\{C_{\mathsf{FB}1}, \log(1 + \mathsf{INR})\}$$

$$+ \min\{C_{\mathsf{FB}2}, \log(1 + \mathsf{INR})\}.$$

Proof: The proof is straightforward and is a direct consequence of the bounds in (26c), (26f)–(26h). For instance, (31a) is derived by combining (26c) and (26f) for $\rho = 0$. Note that $\rho = 0$ maximizes (26c) and (26f). ∎

B. Achievability Strategy

We first provide a brief outline of the achievability. Our achievable scheme is based on block Markov encoding with backward decoding where the scheme is implemented over B blocks. In each block (with the exception of the last two), new messages are transmitted. At the end of a block, each receiver creates a feedback signal and sends it back to its corresponding transmitter. This will provide each transmitter with part of the other user's information that caused interference. Each transmitter encodes this interfering message and transmits it to its receiver during a different block. Through this part of the transmitted signal, receivers will be able to complete the decoding of the previously received messages. During the last two blocks, no new messages will be transmitted and each transmitter provides its receiver with the interfering message coming from the other transmitter. Later, we let B go to infinity to get our desired result.

As we have seen in Section III, each receiver may need to decode the superposition of the two codewords (corresponding to the other user's cooperative common message and part of its own private message). In order to accomplish this in the Gaussian case, we employ lattice codes.

1) Lattice Coding Preliminaries: We briefly go over some preliminaries on lattice coding and summarize the results that will be used later. A lattice is a discrete additive subgroup of $\mathbb{R}^n$. The fundamental volume $V_f(\Lambda)$ of a lattice Λ is the reciprocal of the number of lattice points per unit volume.

Given integer p, denote the set of integers modulo p by $\mathbb{Z}_p$. Let $\mathbb{Z}^n \to \mathbb{Z}_p^n: v \mapsto \bar{v}$ be the componentwise modulo p operation over integer vectors. Also, let C be a linear (n, k) code over $\mathbb{Z}_p$. The lattice Λ_C defined as

$$\Lambda_C = \{v \in \mathbb{Z}^n : \bar{v} \in C\} \tag{32}$$

is generated with respect to the linear code C (see [32] for details). In [32], it has been shown that there exists good lattice codes for point-to-point communication channels, i.e., codes

that achieve a rate close to the capacity of the channel with arbitrary small decoding error probability. We summarize the result here.

Consider a point-to-point communication scenario over an additive noise channel

$$Y = X + Z \tag{33}$$

where X is the transmitted signal with power constraint P, Y is the received signal, and Z is the additive noise process with zero mean and variance σ^2.

A set $\mathcal{B}$ of linear codes over $\mathbb{Z}_p$ is called balanced if every nonzero element of $\mathbb{Z}_p^n$ is contained in the same number of codes in $\mathcal{B}$. Define $\mathcal{L}_\mathcal{B}$ as

$$\mathcal{L}_\mathcal{B} = \{\Lambda_C : C \in \mathcal{B}\}. \tag{34}$$

Lemma 1([32]): Consider a point-to-point additive noise channel described in (33). Let $\mathcal{B}$ be a balanced set of linear (n, k) codes over $\mathbb{Z}_p$. Averaged over all lattices from the set $\mathcal{L}_\mathcal{B}$ defined in (34), each scaled by $\gamma > 0$ and with a fundamental volume V, we have that for any $\delta > 0$, the average probability of decoding error is bounded by

$$\bar{P}_e < (1 + \delta) \frac{n \frac{1}{2} \log \left(2\pi e \sigma^2\right)}{V} \tag{35}$$

for sufficiently large p and small γ such that $\gamma^n p^{n-k} = V$.

See [32] for the proof. The next lemma describes the existence of a good lattice code for a point-to-point additive white Gaussian noise (AWGN) channel.

Lemma 2 ([32]): Consider a point-to-point additive noise channel described in (33) such that the transmitter satisfies a power constraint of P. Then, we can choose a lattice Λ generated using construction A, a shift s [3], and a shaping region S [4] such that the codebook $(\Lambda + s) \cap S$ achieves a rate R with arbitrarily small probability of error if

$$R \leq \frac{1}{2} \log \left(\frac{P}{\sigma^2} \right). \tag{36}$$

In other words, Lemma 2 describes the existence of a lattice code with sufficient codewords. See [32] for the proof. For a more comprehensive review of lattice codes, see [32]–[34].

Remark 6: In this paper, we consider complex AWGN channels. Similar to Lemma 2, one can show that using lattice codes, a rate of $\log \left(\frac{P}{\sigma^2}\right)$ is achievable in the complex channel setting.

2) Achievability Strategy for $\mathcal{C}_{FB1} = \mathcal{C}_{FB}$ and $\mathcal{C}_{FB2} = 0$: We describe our strategy for the extreme case where $C_{FB1} = C_{FB}$ and $C_{FB2} = 0$ (interchanging user IDs, one can get similar results for $C_{FB2} = C_{FB}$ and $C_{FB1} = 0$). Our strategy for any other feedback configuration will be based on a combination of the strategies for these extreme cases.

Codebook Generation and Encoding: The communication strategy consists of B blocks, each of length N channel

---

[3]Shift s is a vector in $\mathbb{R}^n$ and it is required in order to prove existence of good lattice codes; see [32] for more details.

[4]We need to consider the intersection of a lattice with some shaping region $S \subset \mathbb{R}^n$ to satisfy the power constraint.

uses. In block b, $b = 1, 2, \ldots, B - 2$, transmitter 1 has four messages $W_1^{(1,b)}$, $W_1^{(2,b)}$, $W_1^{(3,b)}$, and $W_1^{(4,b)}$, where $W_1^{(i,b)} \in \{1, 2, \ldots, 2^{NR_1^{(i)}}\}$. Out of these four messages, $W_1^{(1,b)}$, $W_1^{(2,b)}$, and $W_1^{(4,b)}$ are new messages and in particular $W_1^{(1,b)}$ and $W_1^{(2,b)}$ form the private message of transmitter 1 while $W_1^{(4,b)}$ is the noncooperative message (as it will be clarified shortly for the feedback strategy, the reason for splitting the private message of transmitter 1 into two parts is that in order to be able to use lattice codes, we would like the codeword corresponding to the cooperative common message of transmitter 2 to be received at the same power level as part of the codeword corresponding to the private message of transmitter 1). We will describe $W_1^{(3,b)}$ when we explain the feedback strategy. On the other hand, transmitter 2 has three new independent messages $W_2^{(1,b)}$, $W_2^{(2,b)}$, and $W_2^{(3,b)}$, the private, the cooperative common, and the noncooperative common message of transmitter 2, respectively.

At transmitter k, message $W_k^{(i,b)}$ is mapped to a Gaussian codeword $X_k^{(i,b)}$ picked from a codebook of size $2^{NR_k^{(i)}}$ and any element of this codebook is drawn i.i.d. from $\mathcal{CN}(0, P_k^{(i)})$, $(k, i) \in \{(1, 1), (1, 3), (1, 4), (2, 1), (2, 3)\}$. For notational simplicity, we have removed the superscript N.

Message $W_k^{(2,b)}$ is mapped to $X_k^{(2,b)}$ encoded by lattice $\Lambda_k^{(2,b)}$ with shift $s_k^{(2,b)}$ and spherical shaping region $S_k^{(2,b)}$. This gives a codebook of size $2^{NR_k^{(2)}}$ with power constraint of $P_k^{(2)}$, $k = 1$, 2. Denote this codebook by $(\Lambda_k^{(2,b)} + s_k^{(2,b)}) \cap S_k^{(2,b)}$.

Transmitter k will superimpose all of its transmitted signals to create $X_k^{(b)}$, its transmitted signal during block b, i.e., $X_1^{(b)} = X_1^{(1,b)} + X_1^{(2,b)} + X_1^{(3,b)} + X_1^{(4,b)}$ and $X_2^{(b)} = X_2^{(1,b)} + X_2^{(2,b)} + X_2^{(3,b)}$.

The power assignments should be such that they are nonnegative and satisfy the power constraint at each transmitter

$$P_1 = P_1^{(1)} + P_1^{(2)} + P_1^{(3)} + P_1^{(4)} \leq 1$$
$$P_2 = P_2^{(1)} + P_2^{(2)} + P_2^{(3)} \leq 1. \tag{37}$$

Feedback Strategy: Our feedback strategy is inspired by the motivating example in Section III. Remember that in this example, receiver 2 had to feed back the superposition of the two codewords (corresponding to transmitter 1's cooperative common message and part of its private message). To realize this in the Gaussian case, we incorporate lattice coding with appropriate power assignment as part of our strategy.

We set $\mathsf{SNR} P_1^{(2)} = \mathsf{INR} P_2^{(2)}$, so that $X_1^{(2,b)}$ and $X_2^{(2,b)}$ arrive at the same power level at receiver 1, and therefore, $h_d X_1^{(2,b)} + h_c X_2^{(2,b)}$ is a lattice point. We refer to this lattice index as $I_{\Lambda_{1,2}}^{(b)}$. Receiver 1 then feeds $\left(I_{\Lambda_{1,2}}^{(b)} \bmod 2^{NC_{FB}} \right)$ back to transmitter 1.

Given $\left(I_{\Lambda_{1,2}}^{(b)} \bmod 2^{NC_{FB}} \right)$, transmitter 1 removes $h_d X_1^{(2,b)}$ and decodes the message index of $W_2^{(2,b)}$. This can be done as long as the total number of lattice points for either of the two aligned messages is less than $2^{NC_{FB}}$, i.e., $R_1^{(2,b)}, R_2^{(2,b)} \leq C_{FB}$. Since the feedback transmission itself lasts a block, we set $W_1^{(3,b+2)} = W_2^{(2,b)}$.

Decoding: For notational simplicity, we ignore the block index and from our description it is clear whether the two signals belong to the same block or different ones. We also use the following shorthand notation:

$$P_k^{(1:j)} = P_k^{(1)} + P_k^{(2)} + \cdots + P_k^{(j)} \quad k = 1, 2. \quad (38)$$

Our achievable scheme employs different decoding orders depending on the channel gains. In other words, based on the channel gains, the number of required messages to achieve the desired sum-rate might vary. In fact, based on the channel gains, it might be sufficient to consider fewer messages than suggested earlier. In such cases, we assume the unnecessary messages to be deterministic (i.e., the corresponding rate to be zero). In particular, we have three different cases.

Case (a) $\log(\mathsf{INR}) \leq \frac{1}{2}\log(\mathsf{SNR})$: In this case, we set $R_1^{(4)} = R_2^{(3)} = 0$. In other words, $W_1^{(4)}$ and $W_2^{(3)}$ are deterministic messages. We then get

$$Y_1 = h_d \left(X_1^{(1)} + X_1^{(2)} + X_1^{(3)} \right) + h_c \left(X_2^{(1)} + X_2^{(2)} \right) + Z_1. \quad (39)$$

At the end of each block, receiver 1 first decodes $X_1^{(3)}$ by treating all other codewords as noise. $X_1^{(3)}$ can be decoded with small error probability if

$$R_1^{(3)} \leq \log \left(1 + \frac{\mathsf{SNR}P_1^{(3)}}{1 + \mathsf{INR}P_2 + \mathsf{SNR}P_1^{(1:2)}} \right). \quad (40)$$

It then removes $h_d X_1^{(3)}$ from the received signal and decodes $X_1^{(1)}$ by treating other codewords as noise. $X_1^{(1)}$ is decodable at receiver 1 with arbitrary small error probability if

$$R_1^{(1)} \leq \log \left(1 + \frac{\mathsf{SNR}P_1^{(1)}}{1 + \mathsf{INR}P_2 + \mathsf{SNR}P_1^{(2)}} \right). \quad (41)$$

After removing $h_d X_1^{(1)}$, receiver 1 has access to $h_d X_1^{(2)} + h_c X_2^{(2)} + h_c X_2^{(1)} + Z_1$. Since we have set $\mathsf{SNR}P_1^{(2)} = \mathsf{INR}P_2^{(2)}$, $h_d X_1^{(2)} + h_c X_2^{(2)}$ is a lattice point with some index $I_{\Lambda_{1,2}}^{(b)}$. Receiver 1 decodes $I_{\Lambda_{1,2}}^{(b)}$ by treating other codewords as noise, and sends back $\left(I_{\Lambda_{1,2}}^{(b)} \bmod 2^{NC_{\mathsf{FB}}} \right)$ to transmitter 1. From Lemma 2, decoding with arbitrary small error probability is feasible if

$$R_1^{(2)} \leq \left[\log \left(\frac{\mathsf{SNR}P_1^{(2)}}{1 + \mathsf{INR}P_2^{(1)}} \right) \right]^+$$
$$R_2^{(2)} \leq \left[\log \left(\frac{\mathsf{INR}P_2^{(2)}}{1 + \mathsf{INR}P_2^{(1)}} \right) \right]^+. \quad (42)$$

Here, $[\cdot]^+ = \max\{\cdot, 0\}$.

The decoding at receiver 2 proceeds as follows. At the end of each block, receiver 2 removes $h_c X_1^{(3)}$ from its received signal. Note that $X_1^{(3)}$ is in fact a function of $W_2^{(2,b-2)}$ and, thus, it is known to receiver 2 (assuming successful decoding in the previous blocks). Therefore, after removing $h_c X_1^{(3)}$, we get

$$Y_2 = h_d \left(X_2^{(1)} + X_2^{(2)} \right) + h_c \left(X_1^{(1)} + X_1^{(2)} \right) + Z_2. \quad (43)$$

Receiver 2 now decodes $X_2^{(2)}$ and $X_2^{(1)}$ by treating other codewords as noise. This can be done with arbitrary small error probability if

$$R_2^{(2)} \leq \left[\log \left(\frac{\mathsf{SNR}P_2^{(2)}}{1 + \mathsf{SNR}P_2^{(1)} + \mathsf{INR}P_1^{(1:2)}} \right) \right]^+$$
$$R_2^{(1)} \leq \log \left(1 + \frac{\mathsf{SNR}P_2^{(1)}}{1 + \mathsf{INR}P_1^{(1:2)}} \right). \quad (44)$$

The decoding strategy presented earlier describes a set of constraints on the rates, which is summarized as follows:

$$\begin{cases} R_1^{(1)} & \leq \log \left(1 + \frac{\mathsf{SNR}P_1^{(1)}}{1 + \mathsf{INR}P_2 + \mathsf{SNR}P_1^{(2)}} \right) \\ R_1^{(2)} & \leq \min \left\{ \log \left(\frac{\mathsf{SNR}P_1^{(2)}}{1 + \mathsf{INR}P_2^{(1)}} \right)^+, C_{\mathsf{FB}} \right\} \\ R_1^{(3)} & \leq \log \left(1 + \frac{\mathsf{SNR}P_1^{(3)}}{1 + \mathsf{INR}P_2 + \mathsf{SNR}P_1^{(1:2)}} \right) \\ R_2^{(1)} & \leq \log \left(1 + \frac{\mathsf{SNR}P_2^{(1)}}{1 + \mathsf{INR}P_1^{(1:2)}} \right) \\ R_2^{(2)} & \leq \min \left\{ \left[\log \left(\frac{\mathsf{INR}P_2^{(2)}}{1 + \mathsf{INR}P_2^{(1)}} \right) \right]^+, C_{\mathsf{FB}} \right\}. \end{cases} \quad (45)$$

Therefore, we can achieve a sum-rate $R_{\mathsf{SUM}}^{(a)} = R_1^{(1)} + R_1^{(2)} + R_2^{(1)} + R_2^{(2)}$, arbitrary close to [5]

$$R_{\mathsf{SUM}}^{(a)} = \log \left(1 + \frac{\mathsf{SNR}P_1^{(1)}}{1 + \mathsf{INR}P_2 + \mathsf{SNR}P_1^{(2)}} \right)$$
$$+ \min \left\{ \left[\log \left(\frac{\mathsf{SNR}P_1^{(2)}}{1 + \mathsf{INR}P_2^{(1)}} \right) \right]^+, C_{\mathsf{FB}} \right\}$$
$$+ \log \left(1 + \frac{\mathsf{SNR}P_2^{(1)}}{1 + \mathsf{INR}P_1^{(1:2)}} \right)$$
$$+ \min \left\{ \left[\log \left(\frac{\mathsf{INR}P_2^{(2)}}{1 + \mathsf{INR}P_2^{(1)}} \right) \right]^+, C_{\mathsf{FB}} \right\}. \quad (46)$$

Case (b) $\frac{1}{2}\log(\mathsf{SNR}) \leq \log(\mathsf{INR}) \leq \frac{2}{3}\log(\mathsf{SNR})$: In this case, we have

$$Y_1 = h_d \left(X_1^{(1)} + X_1^{(2)} + X_1^{(3)} + X_1^{(4)} \right)$$
$$+ h_c \left(X_2^{(1)} + X_2^{(2)} + X_2^{(3)} \right) + Z_1. \quad (47)$$

At the end of each block, receiver 1 first decodes $X_1^{(4)}$ by treating all other codewords as noise, and removes $h_d X_1^{(4)}$ from the received signal. This can be decoded with small error probability if

$$R_1^{(4)} \leq \log \left(1 + \frac{\mathsf{SNR}P_1^{(4)}}{1 + \mathsf{INR}P_2 + \mathsf{SNR}P_1^{(1:3)}} \right). \quad (48)$$

---

[5]Note that $X_1^{(3,b)}$ is a function of the cooperative common message of transmitter 2, i.e., $W_2^{(2,b-2)}$, hence, it does not contain any new information and it is not considered in the sum-rate.

Next, it decodes $X_1^{(3)}$ by treating other codewords as noise and removes $h_d X_1^{(3)}$ from the received signal. This can be decoded with arbitrary small error probability if

$$R_1^{(3)} \le \log\left(1 + \frac{\mathsf{SNR}P_1^{(3)}}{1 + \mathsf{INR}P_2 + \mathsf{SNR}P_1^{(1:2)}}\right). \tag{49}$$

We proceed by decoding the noncooperative common message of transmitter 2, i.e., $X_2^{(3)}$ by treating other codewords as noise. This can be decoded with arbitrary small error probability if

$$R_2^{(3)} \le \log\left(1 + \frac{\mathsf{INR}P_2^{(3)}}{1 + \mathsf{INR}P_2^{(1:2)} + \mathsf{SNR}P_1^{(1:2)}}\right). \tag{50}$$

It then removes $h_c X_2^{(3)}$ from the received signal, having now access to $h_d X_1^{(2)} + h_c X_2^{(2)} + h_d X_1^{(1)} + h_c X_2^{(1)} + Z_1$. We decode the lattice index of $h_d X_1^{(2)} + h_c X_2^{(2)}$, i.e., $I_{\Lambda_{1,2}}^{(b)}$, by treating other codewords as noise. It then sends back $\left(I_{\Lambda_{1,2}}^{(b)} \bmod 2^{NC_{\mathsf{FB}}}\right)$ to transmitter 1. From Lemma 2, decoding with arbitrary small error probability is feasible if

$$R_1^{(2)} \le \left[\log\left(\frac{\mathsf{SNR}P_1^{(2)}}{1 + \mathsf{INR}P_2^{(1)} + \mathsf{SNR}P_1^{(1)}}\right)\right]^+$$
$$R_2^{(2)} \le \left[\log\left(\frac{\mathsf{INR}P_2^{(2)}}{1 + \mathsf{INR}P_2^{(1)} + \mathsf{SNR}P_1^{(1)}}\right)\right]^+. \tag{51}$$

After decoding and removing $h_d X_1^{(2)} + h_c X_2^{(2)}$, receiver 1 decodes $X_1^{(1)}$. This can be done with arbitrary small error probability if

$$R_1^{(1)} \le \log\left(1 + \frac{\mathsf{SNR}P_1^{(1)}}{1 + \mathsf{INR}P_2^{(1)}}\right). \tag{52}$$

Similar to the previous case, receiver 2 removes $X_1^{(3)}$ from its received signal. The decoding at receiver 2 proceeds as follows. Receiver 2 decodes $X_2^{(3)}$ by treating other codewords as noise and removes $h_d X_2^{(3)}$ from the received signal. Next, $X_2^{(2)}$, the noncooperative common message of transmitter 1, will be decoded while treating other codewords as noise. Receiver 2 removes $h_d X_2^{(2)}$ from the received signal and, then, decodes $X_1^{(4)}$ by treating other codewords as noise. After removing $h_c X_1^{(4)}$, we now decode the private message of transmitter 2, i.e., $X_2^{(1)}$. This can be done with arbitrary small error probability if

$$R_2^{(3)} \le \log\left(1 + \frac{\mathsf{SNR}P_2^{(3)}}{1 + \mathsf{SNR}P_2^{(1:2)} + \mathsf{INR}(P_1 - P_1^{(3)})}\right)$$
$$R_2^{(2)} \le \left[\log\left(\frac{\mathsf{SNR}P_2^{(2)}}{1 + \mathsf{SNR}P_2^{(1)} + \mathsf{INR}(P_1 - P_1^{(3)})}\right)\right]^+$$
$$R_1^{(4)} \le \log\left(1 + \frac{\mathsf{INR}P_1^{(4)}}{1 + \mathsf{INR}P_1^{(1:2)} + \mathsf{SNR}P_2^{(1)}}\right) \tag{53}$$
$$R_2^{(1)} \le \log\left(1 + \frac{\mathsf{SNR}P_2^{(1)}}{1 + \mathsf{INR}P_1^{(1:2)}}\right).$$

The decoding strategy presented earlier describes a set of constraints on the rates, which is summarized as follows:

$$\begin{cases} R_1^{(1)} & \le \log\left(1 + \frac{\mathsf{SNR}P_1^{(1)}}{1 + \mathsf{INR}P_2^{(1)}}\right) \\[2mm] R_1^{(2)} & \le \min\left\{\left[\log\left(\frac{\mathsf{SNR}P_1^{(2)}}{1 + \mathsf{INR}P_2^{(1)} + \mathsf{SNR}P_1^{(1)}}\right)\right]^+, C_{\mathsf{FB}}\right\} \\[2mm] R_1^{(3)} & \le \log\left(1 + \frac{\mathsf{SNR}P_1^{(3)}}{1 + \mathsf{INR}P_2 + \mathsf{SNR}P_1^{(1:2)}}\right) \\[2mm] R_1^{(4)} & \le \log\left(1 + \frac{\mathsf{INR}P_1^{(4)}}{1 + \mathsf{INR}P_1^{(1:2)} + \mathsf{SNR}P_2^{(1)}}\right) \\[2mm] R_2^{(1)} & \le \log\left(1 + \frac{\mathsf{SNR}P_2^{(1)}}{1 + \mathsf{INR}P_1^{(1:2)}}\right) \\[2mm] R_2^{(2)} & \le \min\left\{\left[\log\left(\frac{\mathsf{INR}P_2^{(2)}}{1 + \mathsf{INR}P_2^{(1)} + \mathsf{SNR}P_1^{(1)}}\right)\right]^+, C_{\mathsf{FB}}\right\} \\[2mm] R_2^{(3)} & \le \log\left(1 + \frac{\mathsf{INR}P_2^{(3)}}{1 + \mathsf{INR}P_2^{(1:2)} + \mathsf{SNR}P_1^{(1:2)}}\right). \end{cases} \tag{54}$$

Therefore, we can achieve a sum-rate $R_{\mathsf{SUM}}^{(b)} = R_1^{(1)} + R_1^{(2)} + R_1^{(4)} + R_2^{(1)} + R_2^{(2)} + R_2^{(3)}$, arbitrary close to [6]

$$R_{\mathsf{SUM}}^{(b)} = \log\left(1 + \frac{\mathsf{SNR}P_1^{(1)}}{1 + \mathsf{INR}P_2^{(1)}}\right)$$
$$+ \min\left\{\left[\log\left(\frac{\mathsf{SNR}P_1^{(2)}}{1 + \mathsf{INR}P_2^{(1)} + \mathsf{SNR}P_1^{(1)}}\right)\right]^+, C_{\mathsf{FB}}\right\}$$
$$+ \log\left(1 + \frac{\mathsf{INR}P_1^{(4)}}{1 + \mathsf{INR}P_1^{(1:2)} + \mathsf{SNR}P_2^{(1)}}\right)$$
$$+ \log\left(1 + \frac{\mathsf{SNR}P_2^{(1)}}{1 + \mathsf{INR}P_1^{(1:2)}}\right)$$
$$+ \min\left\{\left[\log\left(\frac{\mathsf{INR}P_2^{(2)}}{1 + \mathsf{INR}P_2^{(1)} + \mathsf{SNR}P_1^{(1)}}\right)\right]^+, C_{\mathsf{FB}}\right\}$$
$$+ \log\left(1 + \frac{\mathsf{INR}P_2^{(3)}}{1 + \mathsf{INR}P_2^{(1:2)} + \mathsf{SNR}P_1^{(1:2)}}\right). \tag{55}$$

Case (c) $2\log(\mathsf{SNR}) \le \log(\mathsf{INR})$: In this case, there is no need to decode the superposition of the two messages. So set $R_1^{(1)}$, $R_1^{(2)}$, and $R_2^{(1)}$ equal to zero. We then get

$$Y_1 = h_d\left(X_1^{(3)} + X_1^{(4)}\right) + h_c\left(X_2^{(2)} + X_2^{(3)}\right) + Z_1 \tag{56}$$
$$Y_2 = h_d\left(X_2^{(2)} + X_2^{(3)}\right) + h_c\left(X_1^{(3)} + X_1^{(4)}\right) + Z_2. \tag{57}$$

As for the feedback strategy, receiver 1 decodes $X_2^{(2)}$ by treating other codewords as noise, and sends the lattice index of $W_2^{(2)}$ back to transmitter 1 during the following block. Transmitter 1 later encodes this message as $X_1^{(3)}$ and transmits it. It is worth mentioning that in this case, it is in fact receiver 2 who wants to exploit the feedback link of user 1 to get part of its message. In other words, we have two paths for information flow from transmitter 2 to receiver 2; one through the direct link between them and the other one through receiver 1, feedback

---

[6]Note that $X_1^{(3,b)}$ is a function of the cooperative common message of transmitter 2, i.e., $W_2^{(2,b-2)}$; hence, it is not considered in the sum rate.
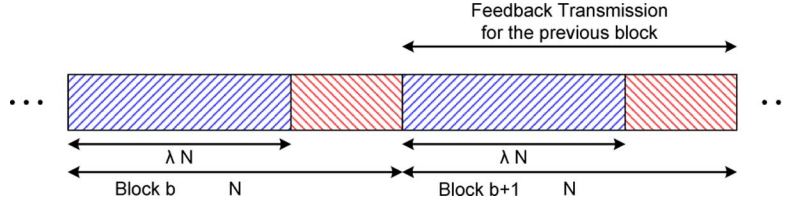
Fig. 6. Achievability strategy for $C_{FB1} = \lambda C_{FB}$ and $C_{FB2} = (1 - \lambda)C_{FB}$.

link and transmitter 1. The decoding works very similar to what we described earlier and we get the following set of constraints to guarantee small error probability at the decoders:

$$\begin{cases} R_1^{(3)} \le \log\left(1 + \frac{\mathsf{INR}P_1^{(3)}}{1 + \mathsf{SNR}P_2}\right) \\ R_1^{(4)} \le \log\left(1 + \frac{\mathsf{SNR}P_1^{(4)}}{1 + \mathsf{SNR}P_1^{(3)}}\right) \\ R_2^{(2)} \le \min\left\{\left[\log\left(\frac{\mathsf{INR}P_2^{(2)}}{1 + \mathsf{SNR}P_1}\right)\right]^+, C_{FB}\right\} \\ R_2^{(3)} \le \log\left(1 + \frac{\mathsf{SNR}P_2^{(3)}}{1 + \mathsf{SNR}P_2^{(2)}}\right). \end{cases} \quad (58)$$

As before $X_1^{(3,b)}$ is a function of $W_2^{(2,b-2)}$. Therefore, we can achieve a sum rate $R_{\mathsf{SUM}}^{(c)} = R_1^{(4)} + R_2^{(2)} + R_2^{(3)}$ arbitrary close to

$$\begin{aligned} R_{\mathsf{SUM}}^{(c)} = &\log\left(1 + \frac{\mathsf{SNR}P_1^{(4)}}{1 + \mathsf{SNR}P_1^{(3)}}\right) \\ &+ \min\left\{\left[\log\left(\frac{\mathsf{INR}P_2^{(2)}}{1 + \mathsf{SNR}P_1}\right)\right]^+, C_{FB}\right\} \\ &+ \log\left(1 + \frac{\mathsf{SNR}P_2^{(3)}}{1 + \mathsf{SNR}P_2^{(2)}}\right). \end{aligned} \quad (59)$$

Case (d) $\frac{2}{3}\log(\mathsf{SNR}) \le \log(\mathsf{INR}) \le 2\log(\mathsf{SNR})$: As we will show in Appendix D, in this regime, feedback can at most increase the sum-rate capacity by 4 bits/s/Hz. Hence, we ignore the feedback and use the nonfeedback transmission strategy in [35] (i.e., having only one private and one common message at each transmitter and jointly decoding at receivers).

3) General Feedback Assignment: We now describe our achievable scheme for general feedback capacity assignment based on a combination of the achievability schemes for the extreme cases. Let $C_{FB1} = \lambda C_{FB}$ and $C_{FB2} = (1 - \lambda)C_{FB}$, such that $0 \le \lambda \le 1$. We call the achievable sum rate of the extreme case $C_{FB1} = C_{FB}$ and $C_{FB2} = 0$ by $R_{\mathsf{SUM}}^{C_{FB2}=0}$, and similarly, we refer to the achievable sum-rate of the other extreme case by $R_{\mathsf{SUM}}^{C_{FB1}=0}$. We split any block b, $b = 1, 2, \ldots, B - 2$, of length N into two subblocks: b_1 of length λN and b_2 of length $(1 - \lambda)N$. See Fig. 6 for a depiction. During block b_1, we implement the transmission strategy of the extreme case $C_{FB1} = C_{FB}$ and $C_{FB2} = 0$, with a block length of λN; and during block b_2, the achievability scheme of the extreme case $C_{FB1} = 0$ and $C_{FB2} = C_{FB}$, with a block length of $(1 - \lambda)N$.

At the end of each subblock, receivers decode the messages as described earlier and create the feedback messages. During block $b + 1$, the feedback messages of sub-blocks b_1 and b_2 will

be sent back to corresponding transmitters, as shown in Fig. 6. Note that we use C_{FB1} during the entire length of block $b + 1$; hence, the effective feedback rate of user 1 (total feedback use divided by number of transmission time slots) would be

$$C_{FB1}^{\mathsf{eff}} = \frac{NC_{FB1}}{\lambda N} = \frac{\lambda N C_{FB}}{\lambda N} = C_{FB}. \quad (60)$$

Hence, we can implement the achievability strategy corresponding to the extreme case $C_{FB1} = C_{FB}$ and $C_{FB2} = 0$. Similar argument is valid for the other extreme case. With this achievability scheme, as N goes to infinity, we achieve a sum rate of $\lambda R_{\mathsf{SUM}}^{C_{FB2}=0} + (1 - \lambda) R_{\mathsf{SUM}}^{C_{FB1}=0}$.

4) Power Splitting: We have yet to specify the values of the powers associated with the codewords at the transmitters (i.e., $P_k^{(i)}: k \in \{1, 2\}, \mathrm{i} \in \{1, 2, 3, 4\}$). In general, one can solve an optimization problem to find the optimal choice of power level assignments that maximizes the achievable sum rate. We have performed numerical analysis for this optimization problem. Fig. 7 shows the gap between our proposed achievable scheme and the outer-bounds in Corollary 1 at (a) $\mathsf{SNR} = 20$ dB, (b) $\mathsf{SNR} = 40$ dB, and (c) $\mathsf{SNR} = 60$ dB, for $C_{FB} = 10$ bits. In fact, through our numerical analysis, we can see that the gap is at most 5.75, 5.25, and 5.15 bits/s/Hz for the given values of SNR, respectively. Note that sharp points in Fig. 7 are due to the change of achievability scheme for different values of INR as described earlier.

In Appendix D, we present an explicit choice of power assignments such that the gap between the achievability scheme and the outer bounds does not scale with SNR. As a result, we get the following Theorem.

Theorem 5: The sum-rate capacity of the Gaussian IC with rate-limited feedback is within at most 14.8 bits/s/Hz of the maximum $R_1 + R_2$ satisfying

$$0 \le R_1 + R_2 \le 2\log(1 + \mathsf{SNR}) + C_{FB1} + C_{FB2} \quad (61a)$$

$$0 \le R_1 + R_2 \le \log\left(1 + \frac{\mathsf{SNR}}{1 + \mathsf{INR}}\right) \quad (61b)$$
$$+ \log\left(1 + \mathsf{SNR} + \mathsf{INR} + 2\sqrt{\mathsf{SNR} \cdot \mathsf{INR}}\right)$$

$$0 \le R_1 + R_2 \le 2\log\left(1 + \mathsf{INR} + \frac{\mathsf{SNR}}{1 + \mathsf{INR}}\right) \quad (61c)$$
$$+ \min\{C_{FB1}, \log(1 + \mathsf{INR})\}$$
$$+ \min\{C_{FB2}, \log(1 + \mathsf{INR})\}.$$

Remark 7: Note that the given choice of power assignment in Appendix D is not necessarily optimal, and our analysis is pessimistic in the sense that we consider the worst case scenario, and we calculate the gap for the worst case.
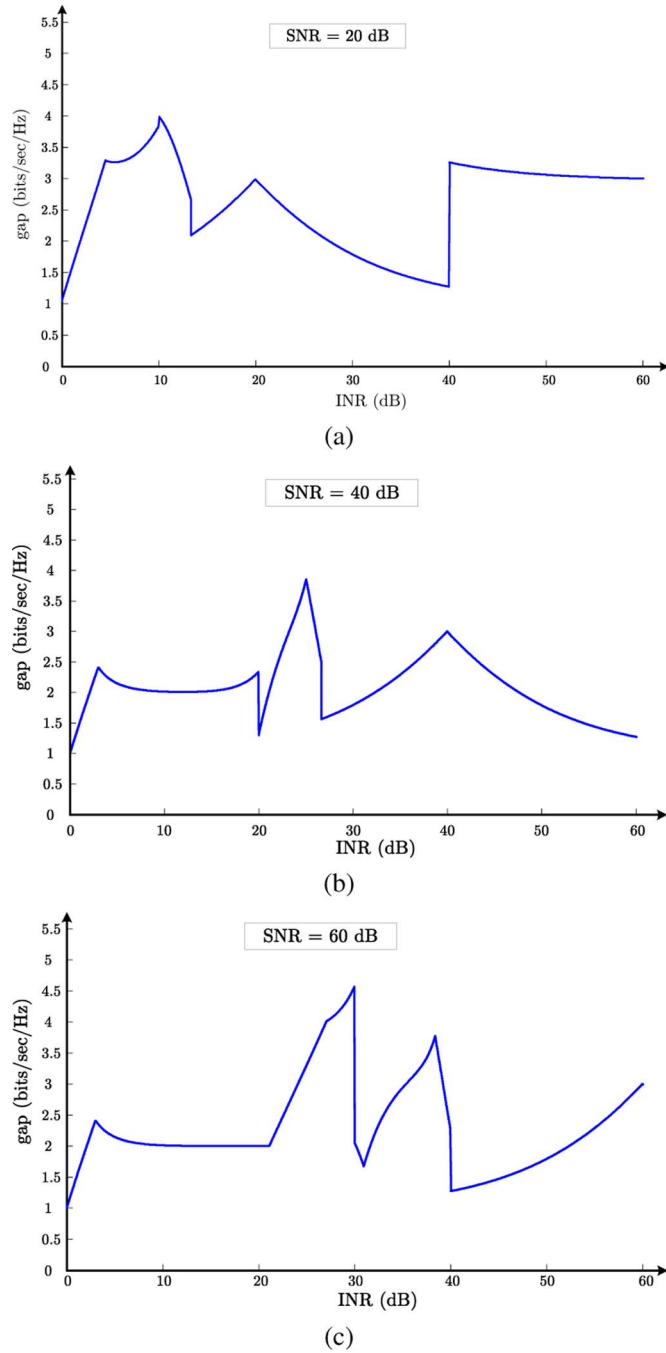
Fig. 7. Numerical analysis: gap between achievable scheme and the outer bounds in Corollary 1 at (a) $\mathsf{SNR} = 20$ dB, (b) $\mathsf{SNR} = 40$ dB, and (c) $\mathsf{SNR} = 60$ dB for $C_{FB} = 10$ bits.

As a corollary, we characterize the symmetric capacity of the two-user Gaussian IC with rate-limited feedback, as defined in the following, to within a constant number of bits.

Definition 1: The symmetric capacity is defined by

$$\mathcal{C}_{\mathsf{sym}} = \sup\{R : (R, R) \in \mathcal{C}\} \qquad (62)$$

where $\mathcal{C}$ is the capacity region.

Corollary 2: For the symmetric Gaussian IC with equal feedback link capacities, i.e., $C_{\mathsf{FB1}} = C_{\mathsf{FB2}}$, the presented achievability strategy achieves to within at most 7.4 bits/s/Hz/user

to the symmetric capacity $\mathcal{C}_{\mathsf{sym}}$ defined in (62), for all channel gains.

Proof: Theorem 5 says that we can achieve to within at most 14.8 bits/s/Hz of the outer bounds in Corollary 1 for any feedback assignment. Therefore, in symmetric IC with equal feedback link capacities $C_{\mathsf{FB1}} = C_{\mathsf{FB2}} = \frac{1}{2}C_{\mathsf{FB1}}$, the gap between the achievability and the symmetric capacity is at most 7.4 bits/s/Hz/user. ∎

VII. CONCLUDING REMARKS

We have addressed the two-user IC with rate-limited feedback under three different models: the El Gamal–Costa deterministic model [11], the linear deterministic model [12], and the Gaussian model. We developed new achievable schemes and new outer bounds for all of the three models. We showed the optimality of our scheme under the linear deterministic model. Under the Gaussian model, we established new outer bounds on the capacity region with rate-limited feedback, and we proposed a transmission strategy employing lattice codes and the ideas developed in the first two models. Furthermore, we proved that the gap between the achievable sum rate of the proposed scheme and the outer bound is bounded by a constant number of bits, independent of the channel gains.

One of the future directions would be to extend this result to the capacity region of the asymmetric two-user Gaussian IC with rate-limited feedback. The same achievability scheme can be applied there; however, the gap analysis will be cumbersome. Therefore, one interesting direction is to find out new techniques to bound the gap between the achievable region and the outer bounds on the capacity region of the asymmetric two-user Gaussian IC with rate-limited feedback.

APPENDIX A
PROOF OF THEOREM 2

Proof of (21a) (Cutset Bound): Starting with Fano's inequality, we get

$$N(R_1 - \epsilon_N) \le I(W_1; Y_1^N) \le \sum H(Y_{1i})$$

where the second inequality follows from the fact that conditioning reduces entropy. If (R_1, R_2) is achievable, then $\epsilon_N \to 0$ as $N \to \infty$. Thus, we obtain the left term of the bound. Notice that this is a cutset bound, as the bound is obtained assuming that the two transmitters fully collaborate.

To obtain the right term, we consider

$$N(R_1 - \epsilon_N) \le I(W_1; Y_1^N, Y_2^N, W_2)$$
$$\overset{(a)}{=} \sum H(Y_{1i}, Y_{2i} | W_2, Y_1^{i-1}, Y_2^{i-1}, X_2^i)$$
$$= \sum H(Y_{1i} | W_2, Y_1^{i-1}, Y_2^i, X_2^i)$$
$$+ \sum H(Y_{2i} | W_2, Y_1^{i-1}, Y_2^{i-1}, X_2^i)$$
$$\overset{(b)}{=} \sum H(Y_{1i} | W_2, Y_1^{i-1}, Y_{2i}, X_{2i}, U_{1i})$$
$$+ \sum H(Y_{2i} | W_2, Y_1^{i-1}, X_{2i}, U_{1i})$$
$$\overset{(c)}{\le} \sum H(Y_{1i} | X_{2i}, Y_{2i}, U_{1i}) + \sum H(Y_{2i} | X_{2i}, U_{1i})$$
$$\overset{(d)}{\le} \sum H(Y_{1i} | V_{1i}, V_{2i}, U_{1i}) + \sum H(Y_{2i} | X_{2i}, U_{1i}),$$

where (a) follows from the fact that W_1 is independent from W_2, and X_2^i is a function of (W_2, Y_2^{i-1}); (b) follows from the fact that $U_{1i} := (X_2^{i-1}, \tilde{Y}_2^{i-1})$ and $\tilde{Y}_2^{i-1}$ is a function of Y_2^{i-1}; (c) follows from the fact that conditioning reduces entropy; (d) follows from the fact that $(V_{1i}$ is a function of (X_{2i}, Y_{2i}), Y_{2i} is a function of (X_{2i}, V_{1i}), and V_{2i} is a function of X_{2i}. Thus, we get the right term of the bound. Notice that this is a cutset bound, as the bound is obtained assuming that the two receivers fully collaborate.

Proof of (21b) (Cutset Bound): Starting with Fano's inequality, we get

$$
\begin{aligned}
N(R_1 - \epsilon_N) &\leq I(W_1; Y_1^N, \tilde{Y}_2^N, W_2) \\
&\overset{(a)}{=} \sum H(Y_{1i}, \tilde{Y}_{2i}|W_2, Y_1^{i-1}, \tilde{Y}_2^{i-1}, X_2^i) \\
&= \sum H(Y_{1i}|W_2, Y_1^{i-1}, \tilde{Y}_2^{i-1}, X_2^i) \\
&\quad + \sum H(\tilde{Y}_{2i}|W_2, Y_1^i, \tilde{Y}_2^{i-1}, X_2^i) \\
&\overset{(b)}{\leq} \sum H(Y_{1i}|X_{2i}, U_{1i}) + \sum H(\tilde{Y}_{2i}|Y_{1i}, X_{2i}) \\
&\overset{(c)}{\leq} \sum H(Y_{1i}|X_{2i}, U_{1i}) + NC_{\mathsf{FB2}},
\end{aligned}
$$

where (a) follows from the fact that W_1 is independent from W_2, and X_2^i is a function of $(W_2, \tilde{Y}_2^{i-1})$; (b) follows from the fact that conditioning reduces entropy; (d) follows from $H(\tilde{Y}_{2i}|Y_{1i}, X_{2i}) \leq C_{\mathsf{FB2}}$. Therefore, we get the desired bound.

Proof of (21e): Starting with Fano's inequality, we get

$$
\begin{aligned}
N(R_1 + R_2 - \epsilon_N) &\leq I(W_1; Y_1^N|W_2) + I(W_2; Y_2^N) \\
&= H(Y_1^N|W_2) + I(W_2; Y_2^N) \\
&= H(Y_1^N|W_2) + H(Y_2^N) \\
&\quad - \left\{ H(Y_1^N, Y_2^N|W_2) - H(Y_1^N|Y_2^N, W_2) \right\} \\
&= H(Y_1^N|Y_2^N, W_2) - H(Y_2^N|Y_1^N, W_2) + H(Y_2^N) \\
&\overset{(a)}{=} \sum H(Y_{1i}|W_2, Y_1^{i-1}, Y_2^N, X_2^i, V_{1i}) + H(Y_2^N) \\
&\overset{(b)}{\leq} \sum [H(Y_{1i}|V_{2i}, V_{1i}, U_{1i}) + H(Y_{2i})],
\end{aligned}
$$

where (a) follows from the fact that X_2^i is a function of (W_2, Y_2^{i-1}) and V_{1i} is a function of (X_{2i}, Y_{2i}); (b) follows from the fact that U_{1i} is a function of (X_2^{i-1}, Y_2^{i-1}) and conditioning reduces entropy.

Proof of (21h):

$$
\begin{aligned}
N(2R_1 &+ R_2 - \epsilon_N) \\
&\leq I(W_1; Y_1^N) + I(W_1; Y_1^N|W_2) + I(W_2; Y_2^N) \\
&\overset{(a)}{\leq} [H(Y_1^N) - H(Y_1^N|W_1)] \\
&\quad + I(W_1; Y_1^N, V_1^N|W_2) + [H(Y_2^N) - H(Y_2^N|W_2)] \\
&\overset{(b)}{=} [H(Y_1^N) - H(V_2^N|W_1)] + H(V_1^N|W_2) \\
&\quad + H(Y_1^N|W_2, V_1^N) + [H(Y_2^N) - H(V_1^N|W_2)] \\
&= H(Y_1^N) + H(Y_1^N|W_2, V_1^N) \\
&\quad + H(Y_2^N) - [H(V_2^N) - I(W_1; V_2^N)] \\
&\overset{(c)}{\leq} I(W_1; V_2^N) + H(Y_1^N) + H(Y_1^N|W_2, V_1^N) \\
&\quad + H(Y_2^N, V_2^N) - H(V_2^N)
\end{aligned}
$$

$$
\begin{aligned}
&\overset{(d)}{\leq} I(W_1; V_2^N, W_2, \tilde{Y}_2^N) + H(Y_1^N) \\
&\quad + H(Y_1^N|W_2, V_1^N) + H(Y_2^N|V_2^N) \\
&\overset{(e)}{=} I(W_1; \tilde{Y}_2^N|W_2) + H(Y_1^N) \\
&\quad + H(Y_1^N|W_2, V_1^N, X_2^N, \tilde{Y}_2^N) + H(Y_2^N|V_2^N) \\
&\overset{(f)}{\leq} NC_{\mathsf{FB2}} + \sum H(Y_{1i}) + \sum H(Y_{2i}|V_{2i}) \\
&\quad + \sum H(Y_{1i}|V_{1i}, V_{2i}, U_{1i}),
\end{aligned}
$$

where (a) follows from the fact that adding information increases mutual information; (b) follows from Claim 1; (c) follows from providing V_2^N to receiver 2; (d) follows from the fact that adding information increases mutual information; follows from the fact that V_k^N is a function of $(W_k, \tilde{Y}_k^{N-1})$; (e) follows from the fact that X_2^N is a function of (W_2, V_1^{N-1}) (by Claim 2) and $\tilde{Y}_2^N$ is a function of (X_2^N, V_1^N); (f) follows from the fact that $U_{1i} := (X_2^{i-1}, \tilde{Y}_2^{i-1})$, $U_{2i} := (X_1^{i-1}, \tilde{Y}_1^{i-1})$, $H(\tilde{Y}_2^N|W_2) \leq NC_{\mathsf{FB2}}$ and conditioning reduces entropy.

To complete the proof, we will show that given $U_i := (U_{1i}, U_{2i})$, X_{1i} and X_{2i} are conditionally independent. Remember that our input distribution is of the form of $p(u_1, u_2)p(x_1|u_1, u_2)p(x_2|u_1, u_2)$.

Claim 5: Given $U_i := (U_{1i}, U_{2i}) = (X_2^{i-1}, \tilde{Y}_2^{i-1}, X_1^{i-1}, \tilde{Y}_1^{i-1})$, X_{1i} and X_{2i} are conditionally independent.

Proof: The proof is based on the dependence-balance-bound technique [27], [28]. For completeness, we describe details. We first show that $I(W_1; W_2|U_i) = 0$, implying that W_1 and W_2 are independent given U_i. We will then show that X_{1i} and X_{2i} are conditionally independent given U_i.

Consider

$$
\begin{aligned}
0 &\leq I(W_1; W_2|U_i) \overset{(a)}{=} I(W_1; W_2|U_i) - I(W_1; W_2) \\
&\overset{(b)}{=} -H(W_1) - H(W_2) - H(U_i) + H(W_1, W_2) \\
&\quad + H(W_1, U_i) + H(W_2, U_i) - H(W_1, W_2, U_i) \\
&\overset{(c)}{=} -H(U_i) + H(U_i|W_1) + H(U_i|W_2) \\
&\overset{(d)}{=} \sum_{j=1}^{i-1} \Big[-H(X_{1j}, X_{2j}|X_1^{j-1}, X_2^{j-1}) \\
&\quad + H(X_{1j}, X_{2j}|W_1, X_1^{j-1}, X_2^{j-1}) \\
&\quad + H(X_{1j}, X_{2j}|W_2, X_1^{j-1}, X_2^{j-1}) \Big] \\
&\overset{(e)}{=} \sum_{j=1}^{i-1} \Big[-H(X_{1j}, X_{2j}|X_1^{j-1}, X_2^{j-1}) \\
&\quad + H(X_{2j}|W_1, X_1^j, X_2^{j-1}) + H(X_{1j}|W_2, X_1^{j-1}, X_2^j) \Big] \\
&= \sum_{j=1}^{i-1} \Big[-H(X_{1j}|X_1^{j-1}, X_2^{j-1}) + H(X_{1j}|W_2, X_1^{j-1}, X_2^j) \\
&\quad - H(X_{2j}|X_1^j, X_2^{j-1}) + H(X_{2j}|W_1, X_1^j, X_2^{j-1}) \Big] \overset{(f)}{\leq} 0,
\end{aligned}
$$

where (a) follows from $I(W_1; W_2) = 0$; (b) follows from the chain rule; (c) follows from the chain rule and $H(U_i|W_1, W_2) = 0$; (d) follows from the fact that $(\tilde{Y}_{1j}, \tilde{Y}_{2j})$ is

a function of $(\tilde{X}_{1j}, \tilde{X}_{2j})$; (e) follows from the fact that X_{kj} is a function of $(W_k, X_1^{j-1}, X_2^{j-1})$; (f) follows from the fact that conditioning reduces entropy. Therefore, $I(W_1; W_2|U_i) = 0$, which shows the independence of W_1 and W_2 given U_i.

Notice that X_{ki} is a function of $(W_k, X_1^{i-1}, X_2^{i-1})$. Hence, it follows easily that $I(X_{1i}; X_{2i}|U_i) = I(X_{1i}; X_{2i}|X_1^{i-1}, X_2^{i-1}) = 0$. This proves the independence of X_{1i} and X_{2i} given U_i. ∎

APPENDIX B
ACHIEVABILITY PROOF OF THEOREM 3

With the choice of distribution given in (23), we have

$$\delta_1 = I(\hat{Y}_1; Y_1|U, U_2, X_1) = 0 \tag{63a}$$

$$\delta_2 = I(\hat{Y}_2; Y_2|U, U_1, X_2) = 0 \tag{63b}$$

$$I(U, V_2, X_1; Y_1) = \max(n_{11}, n_{21}) \tag{63c}$$

$$I(U, V_1, X_2; Y_2) = \max(n_{22}, n_{12}) \tag{63d}$$

$$I(X_1; Y_1|U, V_1, V_2) = (n_{11} - n_{12})^+ \tag{63e}$$

$$I(X_2; Y_2|U, V_1, V_2) = (n_{22} - n_{21})^+ \tag{63f}$$

$$I(U_2; Y_1|U, X_1) = \min(n_{21}, C_{\mathsf{FB1}}) \tag{63g}$$

$$I(U_1; Y_2|U, X_2) = \min(n_{12}, C_{\mathsf{FB2}}) \tag{63h}$$

$$I(X_1; Y_1|U, U_1, V_2) = (n_{11} - n_{12})^+ + \min\{n_{11}, (n_{12} - C_{\mathsf{FB2}})^+\} \tag{63i}$$

$$I(X_2; Y_2|U, U_2, V_1) = (n_{22} - n_{21})^+ + \min\{n_{22}, (n_{21} - C_{\mathsf{FB1}})^+\} \tag{63j}$$

$$I(X_1, V_2; Y_1|U, V_1, U_2) =$$
$$(n_{21} - C_{\mathsf{FB1}})^+ + [(n_{11} - n_{12})^+ - n_{21}]^+$$
$$+ \min\{(n_{11} - n_{12})^+, \min(n_{21}, C_{\mathsf{FB1}})\} \tag{63k}$$

$$I(X_2, V_1; Y_2|U, V_2, U_1) =$$
$$(n_{12} - C_{\mathsf{FB2}})^+ + [(n_{22} - n_{21})^+ - n_{12}]^+$$
$$+ \min\{(n_{22} - n_{21})^+, \min(n_{12}, C_{\mathsf{FB2}})\} \tag{63l}$$

$$I(X_1, V_2; Y_1|U, U_1, U_2) =$$
$$(n_{11} - n_{12})^+ + \min\{n_{11}, (n_{12} - C_{\mathsf{FB2}})^+\}$$
$$+ [n_{21} - \max\{n_{11}, \min(n_{21}, C_{\mathsf{FB1}})\}]^+$$
$$+ \left[\min\{n_{21}, [n_{11} - (n_{12} - C_{\mathsf{FB2}})^+]^+\}\right.$$
$$\left. - \max\{\min(n_{21}, C_{\mathsf{FB1}}), (n_{11} - n_{12})^+\}\right]^+ \tag{63m}$$

$$I(X_2, V_1; Y_2|U, U_1, U_2) =$$
$$(n_{22} - n_{21})^+ + \min\{n_{22}, (n_{21} - C_{\mathsf{FB1}})^+\}$$
$$+ [n_{12} - \max\{n_{22}, \min(n_{12}, C_{\mathsf{FB2}})\}]^+$$
$$+ \left[\min\{n_{12}, [n_{22} - (n_{21} - C_{\mathsf{FB1}})^+]^+\}\right.$$
$$\left. - \max\{\min(n_{12}, C_{\mathsf{FB2}}), (n_{22} - n_{21})^+\}\right]^+. \tag{63n}$$

Using this computation, one can show that the inequalities of (10g) and (10h) are implied by (10b), (10d)–(10f); the inequality (10k) is implied by (10b), (10d), and (10j); and the inequality

(10m) is implied by (10b), (10d), and (10l). We omit the tedious calculation. With further computation, we get

$$R_1 \le \max(n_{11}, n_{21}) \tag{64a}$$

$$R_1 \le (n_{11} - n_{12})^+ + \min\{n_{11}, (n_{12} - C_{\mathsf{FB2}})^+\} + \min(n_{12}, C_{\mathsf{FB2}}) \tag{64b}$$

$$R_2 \le \max(n_{22}, n_{12}) \tag{64c}$$

$$R_2 \le (n_{22} - n_{21})^+ + \min\{n_{22}, (n_{21} - C_{\mathsf{FB1}})^+\} + \min(n_{21}, C_{\mathsf{FB1}}) \tag{64d}$$

$$R_1 + R_2 \le (n_{11} - n_{12})^+ + \max(n_{22}, n_{12}) \tag{64e}$$

$$R_1 + R_2 \le (n_{22} - n_{21})^+ + \max(n_{11}, n_{21}) \tag{64f}$$

$$R_1 + R_2 \le \max\{(n_{11} - n_{12})^+, n_{21}\} + \max\{(n_{22} - n_{21})^+, n_{12}\} + \min\{(n_{11} - n_{12})^+, n_{21}, C_{\mathsf{FB1}}\} + \min\{(n_{22} - n_{21})^+, n_{12}, C_{\mathsf{FB2}}\} \tag{64g}$$

$$2R_1 + R_2 \le (n_{11} - n_{12})^+ + \max(n_{11}, n_{21}) + \max\{(n_{22} - n_{21})^+, n_{12}\} + \min\{(n_{22} - n_{21})^+, n_{12}, C_{\mathsf{FB2}}\} \tag{64h}$$

$$R_1 + 2R_2 \le (n_{22} - n_{21})^+ + \max(n_{22}, n_{12}) + \max\{(n_{11} - n_{12})^+, n_{21}\} + \min\{(n_{11} - n_{12})^+, n_{21}, C_{\mathsf{FB1}}\}. \tag{64i}$$

We will show that the inequalities developed previously are equivalent to the capacity region in Theorem 3. Note that (64b) can be written as

$$R_1 \le \begin{cases} n_{11} + C_{\mathsf{FB2}}, & n_{11} + C_{\mathsf{FB2}} \le n_{12}; \\ \max(n_{11}, n_{12}), & \text{otherwise.} \end{cases}$$

This shows that this inequality is implied by (22a) and (22b). Similarly, (64d) is implied by (22c) and (22d). Next consider (64g)

1) if $C_{\mathsf{FB1}} \le \min\{(n_{11} - n_{12})^+, n_{21}\}$, $C_{\mathsf{FB2}} \le \min\{(n_{22} - n_{21})^+, n_{12}\}$:

$$R_1 + R_2 \le \max\{(n_{11} - n_{12})^+, n_{21}\} + \max\{(n_{22} - n_{21})^+, n_{12}\} + C_{\mathsf{FB1}} + C_{\mathsf{FB2}} \tag{65}$$

2) if $C_{\mathsf{FB1}} > \min\{(n_{11} - n_{12})^+, n_{21}\}$, $C_{\mathsf{FB2}} > \min\{(n_{22} - n_{21})^+, n_{12}\}$:

$$R_1 + R_2 \le \max(n_{11}, n_{12}) + \max(n_{21}, n_{22}) \tag{66}$$

3) if $C_{\mathsf{FB1}} \le \min\{(n_{11} - n_{12})^+, n_{21}\}$, $C_{\mathsf{FB2}} > \min\{(n_{22} - n_{21})^+, n_{12}\}$:

$$R_1 + R_2 \le \max\{(n_{11} - n_{12})^+, n_{21}\} + C_{\mathsf{FB1}} + n_{12} + (n_{22} - n_{21})^+ \tag{67}$$

4) and finally, if $C_{\mathsf{FB1}} > \min\{(n_{11} - n_{12})^+, n_{21}\}$, $C_{\mathsf{FB2}} \le \min\{(n_{22} - n_{21})^+, n_{12}\}$:

$$R_1 + R_2 \le \max\{(n_{22} - n_{21})^+, n_{12}\} + C_{\mathsf{FB2}} + n_{21} + (n_{11} - n_{12})^+. \tag{68}$$

Note that the first case is implied by (22g); and the second case is implied by (22a) and (22c). Also notice that the third case is implied by (22c) and (22i); and the last case is implied by (22a) and (22h). Finally, we consider (64h)

$$2R_1 + R_2 \leq \begin{cases} (n_{11} - n_{12})^+ + \max(n_{11}, n_{21}) \\ \quad + \max\left\{(n_{22} - n_{21})^+, n_{12}\right\} + C_{\mathsf{FB2}} \\ \qquad \text{if } C_{\mathsf{FB2}} \leq \min\left\{(n_{22} - n_{21})^+, \ n_{12}\right\} \\ \max(n_{11}, n_{12}) + \max(n_{11}, n_{21}) \\ \quad + (n_{22} - n_{21})^+ \\ \qquad \text{if } C_{\mathsf{FB2}} > \min\left\{(n_{22} - n_{21})^+, \ n_{12}\right\}. \end{cases}$$

Note that the first case is implied by (22h); and the second case is implied by (22a) and (22f). Similarly, it can be shown that (64i) is implied by (22i), (22c), and (22e). Therefore, the inequalities of (64a)–(64i) are equivalent to those of (22a)–(22i), thus proving the achievability of Theorem 3.

APPENDIX C
PROOF OF THEOREM 4

Proof of (26a) and (26b): Starting with Fano's inequality, we get

$$N(R_1 - \epsilon_N) \leq I(W_1; Y_1^N)$$
$$\leq \sum [h(Y_{1i}) - h(Y_{1i}|W_1, Y_1^{i-1}, X_{1i})]$$
$$= \sum [h(Y_{1i}) - h(Z_{1i})],$$

where the second inequality follows from the fact that conditioning reduces entropy and X_{1i} is a function of (W_1, Y_1^{i-1}); and the third equality follows from the memoryless property of the channel. If (R_1, R_2) is achievable, then $\epsilon_N \to 0$ as $N \to \infty$. Assume that X_1 and X_2 have covariance ρ, i.e., $\rho = \mathbb{E}[X_1 X_2^*]$. We can, then, obtain (26a).

To obtain (26b), consider

$$N(R_1 - \epsilon_N) \leq I(W_1; Y_1^N, Y_2^N | W_2)$$
$$= \sum h(Y_{1i}, Y_{2i}|W_2, Y_1^{i-1}, Y_2^{i-1}) - h(Y_1^N, Y_2^N|W_1, W_2)$$
$$\overset{(a)}{=} \sum h(Y_{1i}, Y_{2i}|W_2, Y_1^{i-1}, Y_2^{i-1}, X_{2i})$$
$$\quad - \sum [h(Z_{1i}) + h(Z_{2i})],$$
$$\overset{(b)}{=} \sum h(Y_{2i}|W_2, Y_1^{i-1}, Y_2^{i-1}, X_{2i})$$
$$\quad + \sum h(Y_{1i}|W_2, Y_2^i, X_{2i}, S_{1i}) - \sum [h(Z_{1i}) + h(Z_{2i})],$$
$$\overset{(c)}{\leq} \sum [h(Y_{2i}|X_{2i}) - h(Z_{2i})]$$
$$\quad + \sum [h(Y_{1i}|X_{2i}, S_{1i}) - h(Z_{1i})],$$

where (a) follows from the fact that X_2^i is a function of (W_2, Y_2^{i-1}) and $h(Y_1^N, Y_2^N|W_1, W_2) = \sum[h(Z_{1i}) + h(Z_{2i})]$ (see Claim 6 in the following); (b) follows from the fact that $S_1^i := g_{12} X_1^i + Z_2^i$ is a function of (Y_2^i, X_2^i); (c) follows from the fact that conditioning reduces entropy. Hence, we get

$$R_1 \leq h(Y_2|X_2) - h(Z_2) + h(Y_1|X_2, S_1) - h(Z_1)$$
$$\overset{(a)}{\leq} \log\left(1 + (1 - |\rho|^2)\mathsf{INR}_{12}\right)$$
$$\quad + \log\left(1 + \frac{(1 - |\rho|^2)\mathsf{SNR}_1}{1 + (1 - |\rho|^2)\mathsf{INR}_{12}}\right),$$

where (a) follows from the fact that

$$h(Y_2|X_2) \leq \log 2\pi e \left(1 + (1 - |\rho|^2)\mathsf{INR}_{12}\right) \qquad (69)$$
$$h(Y_1|X_2, S_1) \leq \log 2\pi e \left(1 + \frac{(1 - |\rho|^2)\mathsf{SNR}_1}{1 + (1 - |\rho|^2)\mathsf{INR}_{12}}\right). (70)$$

The inequality of (70) is obtained as follows. Given (X_2, S_1), the variance of Y_1 is upper bounded by

$$\mathrm{Var}\,[Y_1|X_2, S_1] \leq K_{Y_1} - K_{Y_1(X_2,S_1)} K_{(X_2,S_1)}^{-1} K_{Y_1(X_2,S_1)}^*,$$

where

$$K_{Y_1} = E\left[|Y_1|^2\right]$$
$$= 1 + \mathsf{SNR}_1 + \mathsf{INR}_{21} + \rho g_{11}^* g_{21} + \rho^* g_{11} g_{21}^*$$
$$K_{Y_1(X_2,S_1)} = E\left[Y_1[X_2^*, S_1^*]\right]$$
$$= [\rho g_{11} + g_{21}, g_{12}^* g_{11} + \rho^* g_{21} g_{12}^*]$$
$$K_{(X_2,S_1)} = E\left[\begin{bmatrix} |X_2|^2 & X_2 S_1^* \\ X_2^* S_1 & |S_1|^2 \end{bmatrix}\right]$$
$$= \begin{bmatrix} 1 & \rho^* g_{12}^* \\ \rho g_{12} & 1 + \mathsf{INR}_{12} \end{bmatrix}. \qquad (71)$$

By further calculation, we can get (70).

Claim 6: $h(Y_1^N, Y_2^N|W_1, W_2) = h(Y_1^N, S_1^N|W_1, W_2) = \sum[h(Z_{1i}) + h(Z_{2i})]$.
Proof:

$$h(Y_1^N, Y_2^N|W_1, W_2) = \sum h(Y_{1i}, Y_{2i}|W_1, W_2, Y_1^{i-1}, Y_2^{i-1})$$
$$\overset{(a)}{=} \sum h(Y_{1i}, Y_{2i}|W_1, W_2, Y_1^{i-1}, Y_2^{i-1}, X_{1i}, X_{2i})$$
$$\overset{(b)}{=} \sum h(Z_{1i}, Z_{2i}|W_1, W_2, Y_1^{i-1}, Y_2^{i-1}, X_{1i}, X_{2i})$$
$$\overset{(c)}{=} \sum [h(Z_{1i}) + h(Z_{2i})],$$

where (a) follows from the fact that X_{1i} is a function of (W_1, Y_1^{i-1}) and X_{2i} is a function of (W_2, Y_2^{i-1}); (b) follows from the fact that $Y_{1i} = g_{11}X_{1i} + g_{21}X_{2i} + Z_{1i}$ and $S_{1i} := g_{12}X_{1i} + Z_{2i}$; (c) follows from the memoryless property of the channel and the independence assumption of Z_{1i} and Z_{2i}. Similarly, one can show that $h(Y_1^N, S_1^N|W_1, W_2) = \sum[h(Z_{1i}) + h(Z_{2i})]$. ∎

Proof of (26c): Starting with Fano's inequality, we get

$$N(R_1 - \epsilon_N) \leq I(W_1; Y_1^N, \tilde{Y}_2^N | W_2)$$
$$= \sum h(Y_{1i}, \tilde{Y}_{2i}|W_2, Y_1^{i-1}, \tilde{Y}_2^{i-1}) - h(Y_1^N, \tilde{Y}_2^N|W_1, W_2)$$
$$\overset{(a)}{\leq} \sum h(Y_{1i}, \tilde{Y}_{2i}|W_2, Y_1^{i-1}, \tilde{Y}_2^{i-1}, X_{2i}) - \sum h(Z_{1i})$$
$$= \sum H(\tilde{Y}_{2i}|W_2, Y_1^{i-1}, \tilde{Y}_2^{i-1}, X_{2i})$$
$$\quad + \sum h(Y_{1i}|W_2, Y_1^{i-1}, \tilde{Y}_2^i, X_{2i}) - \sum h(Z_{1i})$$
$$\overset{(b)}{\leq} N C_{\mathsf{FB2}} + \sum [h(Y_{1i}|X_{2i}) - h(Z_{1i})],$$

where (a) follows from the fact that $h(Y_1^N, \tilde{Y}_2^N|W_1, W_2) \geq \sum h(Z_{1i})$ (see Claim 7 in the following) and X_{2i} is a function

of $(W_2, \tilde{Y}_2^{i-1})$; (c) follows from the fact that $H(\tilde{Y}_{2i}) \leq C_{\mathsf{FB2}}$ and conditioning reduces entropy. So we get

$$R_1 \leq h(Y_1|X_2) - h(Z_1) + C_{\mathsf{FB2}}$$
$$\leq \log\left(1 + (1-|\rho|^2)\mathsf{SNR}_1\right) + C_{\mathsf{FB2}}.$$

Claim 7: $h(Y_1^N, \tilde{Y}_2^N|W_1, W_2) \geq \sum h(Z_{1i})$.
Proof:

$$h(Y_1^N, \tilde{Y}_2^N|W_1, W_2) = H(\tilde{Y}_2^N|W_1, W_2)$$
$$+ h(Y_1^N|W_1, W_2, \tilde{Y}_2^N)$$
$$\overset{(a)}{\geq} \sum h(Y_{1i}|W_1, W_2, Y_1^{i-1}, \tilde{Y}_2^{i-1}, X_{1i}, X_{2i})$$
$$= \sum h(Z_{1i}|W_1, W_2, Y_1^{i-1}, \tilde{Y}_2^{i-1}, X_{1i}, X_{2i})$$
$$\overset{(b)}{=} \sum h(Z_{1i}),$$

where (a) follows from the fact that entropy is nonnegative and X_{1i} is a function of (W_1, Y_1^{i-1}) and X_{2i} is a function of $(W_2, \tilde{Y}_2^{i-1})$; (b) follows from the memoryless property of the channel. ∎

Proof of (26h):

$$N(R_1 + R_2 - \epsilon_N) \leq I(W_1; Y_1^N) + I(W_2; Y_2^N)$$
$$\overset{(a)}{\leq} I(W_1; Y_1^N, S_1^N, W_2) + I(W_2; Y_2^N)$$
$$\overset{(b)}{=} h(Y_1^N, S_1^N|W_2) - h(Y_1^N, S_1^N|W_1, W_2) + I(W_2; Y_2^N)$$
$$\overset{(c)}{=} h(Y_1^N, S_1^N|W_2) - \sum [h(Z_{1i}) + h(Z_{2i})] + I(W_2; Y_2^N)$$
$$\overset{(d)}{=} h(Y_1^N|S_1^N, W_2) - \sum h(Z_{1i}) + h(Y_2^N) - \sum h(Z_{2i})$$
$$\overset{(e)}{=} h(Y_1^N|S_1^N, W_2, X_2^N) - \sum h(Z_{1i}) + h(Y_2^N) - \sum h(Z_{2i})$$
$$\overset{(f)}{\leq} \sum_{i=1}^{N} [h(Y_{1i}|S_{1i}, X_{2i}) - h(Z_{1i}) + h(Y_{2i}) - h(Z_{2i})],$$

where (a) follows from the fact that adding information increases mutual information (providing a *genie*); (b) follows from the independence of W_1 and W_2; (c) follows from $h(Y_1^N, S_1^N|W_1, W_2) = \sum[h(Z_{1i}) + h(Z_{2i})]$ (see Claim 6); (d) follows from $h(S_1^N|W_2) = h(Y_2^N|W_2)$ (see Claim 3); (e) follows from the fact that X_2^N is a function of (W_2, S_1^{N-1}) (see Claim 4); (f) follows from the fact that conditioning reduces entropy. Hence, we get

$$R_1 + R_2 \leq h(Y_1|S_1, X_2) - h(Z_1) + h(Y_2) - h(Z_2).$$

Note that

$$h(Y_2) \leq \log 2\pi e \left(1 + \mathsf{SNR}_2 + \mathsf{INR}_{12} + 2|\rho|\sqrt{\mathsf{SNR}_2 \cdot \mathsf{INR}_{12}}\right). \tag{72}$$

From (70) and (72), we get the desired upper bound.
Proof of (26j):

$$N(2R_1 + R_2 - \epsilon_N) \leq I(W_1; Y_1^N)$$
$$+ I(W_1; Y_1^N|W_2) + I(W_2; Y_2^N)$$
$$\overset{(a)}{\leq} [h(Y_1^N) - h(Y_1^N|W_1)] + I(W_1; Y_1^N, S_1^N|W_2)$$
$$+ [h(Y_2^N) - h(Y_2^N|W_2)]$$

$$\overset{(b)}{=} [h(Y_1^N) - h(Y_1^N|W_1)] + h(Y_1^N, S_1^N|W_2)$$
$$- \sum [h(Z_{1i}) + h(Z_{2i})] + [h(Y_2^N) - h(Y_2^N|W_2)]$$
$$\overset{(c)}{=} [h(Y_1^N) - h(S_2^N|W_1)] + h(Y_2^N) + h(Y_1^N|W_2, S_1^N)$$
$$- \sum [h(Z_{1i}) + h(Z_{2i})]$$
$$= h(Y_1^N) - [h(S_2^N) - I(W_1; S_2^N)] + h(Y_2^N)$$
$$+ h(Y_1^N|W_2, S_1^N) - \sum [h(Z_{1i}) + h(Z_{2i})]$$
$$\overset{(d)}{\leq} I(W_1; S_2^N) + h(Y_1^N) - [h(S_2^N)$$
$$- h(S_2^N|Y_2^N) + h(S_2^N|Y_2^N, X_1^N)]$$
$$+ h(Y_2^N) + h(Y_1^N|W_2, S_1^N) - \sum [h(Z_{1i}) + h(Z_{2i})]$$
$$= I(W_1; S_2^N) + h(Y_1^N) - I(S_2^N; Y_2^N) - \sum h(Z_{1i})$$
$$+ h(Y_2^N) + h(Y_1^N|W_2, S_1^N) - \sum [h(Z_{1i}) + h(Z_{2i})]$$
$$\overset{(e)}{\leq} I(W_1; S_2^N, W_2, \tilde{Y}_2^N) + h(Y_1^N) + h(Y_1^N|W_2, S_1^N)$$
$$+ h(Y_2^N|S_2^N) - \sum [2h(Z_{1i}) + h(Z_{2i})]$$
$$\overset{(f)}{=} I(W_1; \tilde{Y}_2^N|W_2) + h(Y_1^N) + h(Y_1^N|W_2, S_1^N, X_2^N, \tilde{Y}_2^N)$$
$$+ h(Y_2^N|S_2^N) - \sum [2h(Z_{1i}) + h(Z_{2i})]$$
$$\overset{(g)}{\leq} NC_{\mathsf{FB2}} + \sum [h(Y_{1i}) - h(Z_{1i})]$$
$$+ \sum [h(Y_{1i}|V_{1i}, V_{2i}, U_{1i}) - h(Z_{1i})]$$
$$+ \sum [h(Y_{2i}|S_{2i}) - h(Z_{2i})],$$

where (a) follows from the fact that adding information increases mutual information; (b) follows from $h(Y_1^N, S_1^N|W_1, W_2) = \sum[h(Z_{1i}) + h(Z_{2i})]$ (by Claim 6); (c) follows from Claim 3; (d) follows from providing S_2^N to receiver 2; (e) follows from the fact that adding information increases mutual information; follows from the fact that S_k^N is a function of $(W_k, \tilde{Y}_k^{N-1})$; (f) follows from the fact that X_2^N is a function of (W_2, S_1^{N-1}) (by Claim 4) and $\tilde{Y}_2^N$ is a function of (X_2^N, S_1^N); (g) follows from the fact that $h(\tilde{Y}_2^N|W_2) \leq NC_{\mathsf{FB2}}$ and conditioning reduces entropy.

Incorporating the infinite-feedback case, we get

$$2R_1 + R_2 \leq h(Y_1) - h(Z_1) + h(Y_1|S_1, X_2) - h(Z_1)$$
$$+ h(Y_2|S_2) - h(Z_2)$$
$$+ \min\{C_{\mathsf{FB2}}, h(Y_2|X_2) - h(Z_2)\}. \tag{73}$$

Therefore, we get the desired bound.

APPENDIX D
GAP ANALYSIS OF THEOREM 5

We show that our proposed achievability strategy in Section IV-B2 results in a sum rate to within at most 14.8 bits/s/Hz of the outer bounds in Corollary 1. It is sufficient to prove this for the extreme case of feedback capacity assignment, i.e., where $C_{\mathsf{FB1}} = C_{\mathsf{FB}}$ and $C_{\mathsf{FB2}} = 0$ (or symmetrically $C_{\mathsf{FB1}} = 0$ and $C_{\mathsf{FB2}} = C_{\mathsf{FB}}$). The reason is as follows. Consider our achievability strategy for the general feedback strategy described previously, and let $C_{\mathsf{FB1}} = \lambda C_{\mathsf{FB}}$ and

$C_{\mathsf{FB2}} = (1 - \lambda)C_{\mathsf{FB}}$, such that $0 \leq \lambda \leq 1$. Under these assumptions, for any value of λ, the outer bounds on sum rate in (31a)–(31c) would be the same, call the minimum of them C^*. Assuming that we can achieve to within 14.8 bits/s/Hz of this outer bound in the extreme cases, then, with the described achievability scheme for general feedback assignment, we can achieve

$$R_{\mathsf{SUM}} = \lambda R_{\mathsf{SUM}}^{C_{\mathsf{FB2}}=0} + (1 - \lambda) R_{\mathsf{SUM}}^{C_{\mathsf{FB1}}=0} \qquad (74)$$
$$\geq \lambda (C^* - 14.8) + (1 - \lambda)(C^* - 14.8) = (C^* - 14.8).$$

We now prove our claim for the extreme cases. By symmetry, we only need to analyze the gap in one case, say $C_{\mathsf{FB1}} = C_{\mathsf{FB}}$ and $C_{\mathsf{FB2}} = 0$. We assume that $\mathsf{INR} \geq 1$, since for the case when $\mathsf{INR} < 1$, by ignoring the feedback and treating interfere as noise, we can achieve a sum rate of

$$2 \log \left(1 + \frac{\mathsf{SNR}}{1 + \mathsf{INR}} \right), \qquad (75)$$

which is at most within 2.6 bits of outer bound (31b) in Corollary 1

$$\log \left(1 + \frac{\mathsf{SNR}}{1 + \mathsf{INR}} \right) + \log \left(1 + \mathsf{SNR} + \mathsf{INR} + 2\sqrt{\mathsf{SNR} \cdot \mathsf{INR}} \right)$$
$$- 2 \log \left(1 + \frac{\mathsf{SNR}}{1 + \mathsf{INR}} \right)$$
$$\overset{(\mathsf{INR} \leq 1)}{\leq} \log \left(1 + \mathsf{SNR} + \mathsf{INR} + 2\sqrt{\mathsf{SNR} \cdot \mathsf{INR}} \right)$$
$$- \log \left(1 + \frac{\mathsf{SNR}}{2} \right)$$
$$\overset{(\mathsf{INR} \leq 1)}{\leq} \log \left(2 + 3\mathsf{SNR} \right) - \log \left(1 + \mathsf{SNR} \right) + 1$$
$$= \log \left(\frac{2 + 3\mathsf{SNR}}{1 + \mathsf{SNR}} \right) + 1$$
$$\leq \log \left(3 \right) + 1 \leq 2.6. \qquad (76)$$

We consider five different subcases.

Case (a) $\log (\mathsf{INR}) \leq \frac{1}{2} \log (\mathsf{SNR})$: For this case, we pick the following set of power levels [7]:

$$P_1^{(1)} = \left(\frac{1}{\mathsf{INR}} - \frac{1}{\mathsf{SNR}} \min\{2^{C_{\mathsf{FB}}}, \mathsf{INR} - 1\} \right)^+$$
$$P_1^{(2)} = \frac{1}{\mathsf{SNR}} \min\{2^{C_{\mathsf{FB}}}, \mathsf{INR} - 1\}$$
$$P_1^{(3)} = \frac{1}{\mathsf{INR}} \min\{2^{C_{\mathsf{FB}}}, \mathsf{INR} - 1\} \qquad (77)$$
$$P_2^{(1)} = \frac{1}{\mathsf{INR}}$$
$$P_2^{(2)} = \frac{1}{2\mathsf{INR}} \min\{2^{C_{\mathsf{FB}}}, \mathsf{INR} - 1\}.$$

Note that the power levels are nonnegative, and at transmitter 1, we have

$$P_1 = P_1^{(1)} + P_1^{(2)} + P_1^{(3)}$$
$$= \left(\frac{1}{\mathsf{INR}} - \frac{1}{\mathsf{SNR}} \min\{2^{C_{\mathsf{FB}}}, \mathsf{INR} - 1\} \right)^+$$
$$+ \frac{1}{\mathsf{SNR}} \min\{2^{C_{\mathsf{FB}}}, \mathsf{INR} - 1\} + \frac{1}{\mathsf{INR}} \min\{2^{C_{\mathsf{FB}}}, \mathsf{INR} - 1\}$$

---

[7]Remember that starting the beginning of Section IV-B2, we have assumed that $\mathsf{INR} \geq 1$. Hence, we are not encountering division by zero in power assignments of (77).

$$\overset{(a)}{=} \frac{1}{\mathsf{INR}} - \frac{1}{\mathsf{SNR}} \min\{2^{C_{\mathsf{FB}}}, \mathsf{INR} - 1\}$$
$$+ \frac{1}{\mathsf{SNR}} \min\{2^{C_{\mathsf{FB}}}, \mathsf{INR} - 1\} + \frac{1}{\mathsf{INR}} \min\{2^{C_{\mathsf{FB}}}, \mathsf{INR} - 1\}$$
$$= \frac{1}{\mathsf{INR}} + \frac{1}{\mathsf{INR}} \min\{2^{C_{\mathsf{FB}}}, \mathsf{INR} - 1\}$$
$$\leq \frac{1}{\mathsf{INR}} + \frac{\mathsf{INR} - 1}{\mathsf{INR}} \leq 1 \qquad (78)$$

where (a) follows from the fact that

$$\frac{1}{\mathsf{SNR}} \min\{2^{C_{\mathsf{FB}}}, \mathsf{INR} - 1\} \leq \frac{\mathsf{INR} - 1}{\mathsf{SNR}} \overset{(\mathsf{INR}^2 \leq \mathsf{SNR})}{\leq} \frac{1}{\mathsf{INR}}. \quad (79)$$

At transmitter 2

$$P_2 = P_2^{(1)} + P_2^{(2)}$$
$$= \frac{1}{\mathsf{INR}} + \frac{1}{2\mathsf{INR}} \min\{2^{C_{\mathsf{FB}}}, \mathsf{INR} - 1\}$$
$$\leq \frac{1}{\mathsf{INR}} + \frac{\mathsf{INR} - 1}{\mathsf{INR}} \leq 1. \qquad (80)$$

By plugging the given values of power levels into our achievable sum rate $R_{\mathsf{SUM}}^{(a)}$ defined in (46), we get

$$R_{\mathsf{SUM}}^{(a)} = \log \left(1 + \frac{\mathsf{SNR}P_1^{(1)}}{1 + \mathsf{INR}P_2^{(1:2)} + \mathsf{SNR}P_1^{(2)}} \right)$$
$$+ \log \left(\frac{\mathsf{SNR}P_1^{(2)}}{2} \right)^+ + \log \left(1 + \frac{\mathsf{SNR}P_2^{(1)}}{2} \right)$$
$$+ \log \left(\frac{\mathsf{INR}P_2^{(2)}}{1 + \mathsf{INR}P_2^{(1)}} \right)^+$$
$$= \log \left(\frac{2 + \mathsf{INR}P_2^{(2)} + \mathsf{SNR}P_1^{(1:2)}}{2 + \mathsf{INR}P_2^{(2)} + \mathsf{SNR}P_1^{(2)}} \right)$$
$$+ \log \left(\frac{\mathsf{SNR}P_1^{(2)}}{2} \right)^+ + \log \left(1 + \frac{\mathsf{SNR}}{2\mathsf{INR}} \right)$$
$$+ \log \left(\frac{\mathsf{INR}P_2^{(2)}}{1 + \mathsf{INR}P_2^{(1)}} \right)^+$$
$$\overset{(a)}{\geq} \log \left(\frac{2 + \mathsf{INR}P_2^{(2)} + \mathsf{SNR}P_1^{(1:2)}}{2 \left(1 + \mathsf{SNR}P_1^{(2)} \right)} \times \frac{1 + \mathsf{SNR}P_1^{(2)}}{4} \right)$$
$$+ \log \left(1 + \frac{\mathsf{SNR}}{2\mathsf{INR}} \right) + \log \left(\frac{\mathsf{INR}P_2^{(2)}}{1 + \mathsf{INR}P_2^{(1)}} \right)^+$$
$$= \log \left(2 + \mathsf{INR}P_2^{(2)} + \mathsf{SNR}P_1^{(1:2)} \right) + \log \left(1 + \frac{\mathsf{SNR}}{2\mathsf{INR}} \right)$$
$$+ \log \left(\frac{\min\{2^{C_{\mathsf{FB}}}, \mathsf{INR} - 1\}}{4} \right)^+ - 3$$
$$\geq \log \left(2 + \mathsf{INR}P_2^{(2)} + \mathsf{SNR}P_1^{(1:2)} \right) + \log \left(1 + \frac{\mathsf{SNR}}{\mathsf{INR}} \right)$$
$$+ \min\{C_{\mathsf{FB}}, \log (\mathsf{INR} - 1)^+\} - 6$$
$$\geq \log \left(2 + \mathsf{INR}P_2^{(2)} + \mathsf{SNR}P_1^{(1:2)} \right) + \log \left(1 + \frac{\mathsf{SNR}}{\mathsf{INR}} \right)$$
$$+ \min\{C_{\mathsf{FB}}, \log (1 + \mathsf{INR})\} - \log (3) - 6 \qquad (81)$$

where (a) follows from the assumption $\mathsf{SNR} \geq \mathsf{INR} \geq 1$, and the last inequality holds since

$$\log\left(\mathsf{INR}-1\right)^+ \geq \log\left(1+\mathsf{INR}\right) - 3 \forall\ \mathsf{INR} \geq 1. \tag{82}$$

Considering the outer bound in (31c), in this case, we can write

$$R_1 + R_2 \leq 2\log\left(1+\mathsf{INR}+\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right)$$
$$+ \min\{C_{\mathsf{FB}}, \log\left(1+\mathsf{INR}\right)\}$$
$$\leq 2\log\left(1+\mathsf{INR}+\frac{\mathsf{SNR}}{\mathsf{INR}}\right) + \min\{C_{\mathsf{FB}}, \log\left(1+\mathsf{INR}\right)\}$$
$$= 2\log\left(1+\frac{\mathsf{INR}^2+\mathsf{SNR}}{\mathsf{INR}}\right) + \min\{C_{\mathsf{FB}}, \log\left(1+\mathsf{INR}\right)\}$$
$$\overset{\left(\mathsf{INR}^2 \leq \mathsf{SNR}\right)}{\leq} 2\log\left(1+\frac{2\mathsf{SNR}}{\mathsf{INR}}\right) + \min\{C_{\mathsf{FB}}, \log\left(1+\mathsf{INR}\right)\}$$
$$\leq 2\log\left(1+\frac{\mathsf{SNR}}{\mathsf{INR}}\right) + \min\{C_{\mathsf{FB}}, \log\left(1+\mathsf{INR}\right)\} + 2. \tag{83}$$

The gap between the achievable sum rate of (81) and the outer bound in (83) is upper bounded by

$$8 + \log\left(3\right) + 2\log\left(1+\frac{\mathsf{SNR}}{\mathsf{INR}}\right) + \min\{C_{\mathsf{FB}}, \log\left(1+\mathsf{INR}\right)\}$$
$$- \log\left(2 + \mathsf{INR}P_2^{(2)} + \mathsf{SNR}P_1^{(1:2)}\right)$$
$$- \log\left(1+\frac{\mathsf{SNR}}{\mathsf{INR}}\right) - \min\{C_{\mathsf{FB}}, \log\left(1+\mathsf{INR}\right)\}$$
$$\leq 8 + \log\left(3\right) + \log\left(1+\frac{\mathsf{SNR}}{\mathsf{INR}}\right) - \log\left(2+\frac{\mathsf{SNR}}{\mathsf{INR}}\right)$$
$$\leq 8 + \log\left(3\right).$$

Hence, the gap between the inner bound and the outer bound is at most $8 + \log\left(3\right) \leq 9.6$ bits/s/Hz.

Case (b) $\frac{1}{2}\log\left(\mathsf{SNR}\right) \leq \log\left(\mathsf{INR}\right) \leq \frac{2}{3}\log\left(\mathsf{SNR}\right)$*:* For this case, we pick the following set of power levels[8]:

$$\begin{cases} P_1^{(1)} = \left(\frac{1}{\mathsf{INR}} - \frac{1}{\mathsf{SNR}}\min\{2^{C_{\mathsf{FB}}}, \frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\}\right)^+ \\ P_1^{(2)} = \frac{1}{\mathsf{SNR}}\min\{2^{C_{\mathsf{FB}}}, \frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\} \\ P_1^{(3)} = \frac{1}{\mathsf{INR}}\min\{2^{C_{\mathsf{FB}}}, \frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\} \\ P_1^{(4)} = \left(1 - P_1^{(1:3)}\right)^+ \\ P_2^{(1)} = \frac{1}{\mathsf{INR}} \\ P_2^{(2)} = \frac{1}{2\mathsf{INR}}\min\{2^{C_{\mathsf{FB}}}, \frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\} \\ P_2^{(3)} = \left(1 - P_2^{(1:2)}\right)^+ \end{cases} \tag{84}$$

All the power levels are nonnegative. Also, we have

$$P_1^{(1)} + P_1^{(2)} + P_1^{(3)}$$
$$= \left(\frac{1}{\mathsf{INR}} - \frac{1}{\mathsf{SNR}}\min\{2^{C_{\mathsf{FB}}}, \frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\}\right)^+$$
$$+ \frac{1}{\mathsf{SNR}}\min\{2^{C_{\mathsf{FB}}}, \frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\}$$

[8]$\mathsf{INR} \geq 1$.

$$+ \frac{1}{\mathsf{INR}}\min\{2^{C_{\mathsf{FB}}}, \frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\}$$
$$\overset{(a)}{=} \frac{1}{\mathsf{INR}} - \frac{1}{\mathsf{SNR}}\min\{2^{C_{\mathsf{FB}}}, \frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\}$$
$$+ \frac{1}{\mathsf{SNR}}\min\{2^{C_{\mathsf{FB}}}, \frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\}$$
$$+ \frac{1}{\mathsf{INR}}\min\{2^{C_{\mathsf{FB}}}, \frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\}$$
$$= \frac{1}{\mathsf{INR}} + \frac{1}{\mathsf{INR}}\min\{2^{C_{\mathsf{FB}}}, \frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\}$$
$$\leq \frac{\mathsf{SNR}^2}{\mathsf{INR}^4} \overset{(\sqrt{\mathsf{SNR}} \leq \mathsf{INR})}{\leq} \frac{\sqrt{\mathsf{SNR}}}{\mathsf{INR}} \leq 1 \tag{85}$$

where (a) follows from the fact that

$$\frac{1}{\mathsf{SNR}}\min\{2^{C_{\mathsf{FB}}}, \frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\} \leq \frac{1}{\sqrt{\mathsf{SNR}}} \leq \frac{1}{\mathsf{INR}}. \tag{86}$$

Since $P_1^{(4)} = \left(1 - P_1^{(1:3)}\right)^+$, we conclude that $P_1 \leq 1$. Similarly, we can show that $P_2 \leq 1$.

By plugging the given values of power levels into our achievable sum rate $R_{\mathsf{SUM}}^{(b)}$ defined in (55), we have

$$R_{\mathsf{SUM}}^{(b)} = \log\left(\frac{1 + \mathsf{INR}P_2^{(1)} + \mathsf{SNR}P_1^{(1:2)}}{1 + \mathsf{INR}P_2^{(1)}}\right)$$
$$+ \log\left(\frac{2\mathsf{INR} + \mathsf{SNR} + \mathsf{INR}^2 - 2\mathsf{INR}}{2\mathsf{INR} + \mathsf{SNR}}\right)$$
$$+ \log\left(1 + \frac{\mathsf{SNR}}{2\mathsf{INR}}\right)$$
$$+ \log\left(\frac{2\mathsf{INR} + \mathsf{SNR} + \mathsf{INR}^2 - 3/2\mathsf{INR}}{2\mathsf{INR} + \mathsf{SNR}}\right)$$
$$+ \min\{C_{\mathsf{FB}}, \log\left(\frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\right)^+\} - 3$$
$$\geq \log\left(\frac{2 + \mathsf{SNR}P_1^{(1:2)}}{2}\right)$$
$$+ \log\left(1 + \frac{\mathsf{SNR}}{2\mathsf{INR}}\right) + 2\log\left(\frac{\mathsf{INR}^2}{3\mathsf{SNR}}\right)$$
$$+ \min\{C_{\mathsf{FB}}, \log\left(\frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\right)^+\} - 3$$
$$\geq 2\log\left(1 + \frac{\mathsf{SNR}}{2\mathsf{INR}}\right) + 2\log\left(\frac{\mathsf{INR}^2}{3\mathsf{SNR}}\right)$$
$$+ \min\{C_{\mathsf{FB}}, \log\left(\frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\right)^+\} - 3$$
$$\geq 2\log\left(\frac{1 + \mathsf{SNR}}{2\mathsf{INR}}\right) + 2\log\left(\mathsf{INR}^2\right) - 2\log\left(3\mathsf{SNR}\right)$$
$$+ \min\{C_{\mathsf{FB}}, \log\left(\frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\right)^+\} - 3$$
$$= 2\log\left(1 + \mathsf{SNR}\right) - 2\log\left(\mathsf{INR}\right)$$
$$+ 4\log\left(\mathsf{INR}\right) - 2\log\left(3\mathsf{SNR}\right)$$
$$+ \min\{C_{\mathsf{FB}}, \log\left(\frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\right)^+\} - 3$$

$$\geq 2\log\left(1 + \mathsf{INR}\right)$$
$$+ \min\{C_{\mathsf{FB}}, \log\left(\left(\frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\right)^+\right\} - 5 - 2\log(3). \quad (87)$$

We first simplify the outer bounds in (31b) and (31c). Considering the outer bound in (31c), for this case, we can write

$$R_1 + R_2 \leq 2\log\left(1 + \mathsf{INR} + \frac{\mathsf{SNR}}{1 + \mathsf{INR}}\right)$$
$$+ \min\{C_{\mathsf{FB}}, \log\left(1 + \mathsf{INR}\right)\}$$
$$\leq 2\log\left(1 + \mathsf{INR} + \frac{\mathsf{SNR}}{2\mathsf{INR}}\right) + \min\{C_{\mathsf{FB}}, \log\left(1 + \mathsf{INR}\right)\}$$
$$\overset{(\mathsf{SNR}\leq\mathsf{INR}^2)}{\leq} 2\log\left(1 + \mathsf{INR} + \frac{\mathsf{INR}}{2}\right)$$
$$+ \min\{C_{\mathsf{FB}}, \log\left(1 + \mathsf{INR}\right)\}$$
$$= 2\log\left(1 + \frac{3\mathsf{INR}}{2}\right) + \min\{C_{\mathsf{FB}}, \log\left(1 + \mathsf{INR}\right)\}$$
$$\leq 2\log\left(1 + \mathsf{INR}\right) + \min\{C_{\mathsf{FB}}, \log\left(1 + \mathsf{INR}\right)\} + 2\log(1.5) \quad (88)$$

and for the outer bound in (31b), we have

$$R_1 + R_2 \leq \log\left(1 + \frac{\mathsf{SNR}}{1 + \mathsf{INR}}\right)$$
$$+ \log\left(1 + \mathsf{SNR} + \mathsf{INR} + 2\sqrt{\mathsf{SNR}\cdot\mathsf{INR}}\right)$$
$$\overset{(\mathsf{INR}\leq\mathsf{SNR})}{\leq} \log\left(1 + \mathsf{INR} + \mathsf{SNR}\right) - \log\left(1 + \mathsf{INR}\right)$$
$$+ \log\left(1 + 4\mathsf{SNR}\right)$$
$$\overset{(\mathsf{INR}\leq\mathsf{SNR})}{\leq} 2\log\left(1 + \mathsf{SNR}\right) - \log\left(1 + \mathsf{INR}\right) + 3. \quad (89)$$

We consider two possible scenarios based on C_{FB}:

1) $C_{\mathsf{FB}} \leq \log\left(\left(\frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\right)^+\right)$:
   With this assumption, we have

$$R_{\mathsf{SUM}}^{(b)} \geq 2\log\left(1 + \mathsf{INR}\right) + C_{\mathsf{FB}} - 5 - 2\log(3). \quad (90)$$

Note that, in this case

$$C_{\mathsf{FB}} \leq \log\left(\left(\frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\right)^+\right) \leq \log\left(1 + \frac{\mathsf{SNR}^2}{\mathsf{INR}^3}\right)$$
$$\overset{(\sqrt{\mathsf{SNR}}\leq\mathsf{INR})}{\leq} \log\left(1 + \mathsf{INR}\right). \quad (91)$$

Therefore, from (88) we get

$$R_1 + R_2 \leq 2\log\left(1 + \mathsf{INR}\right) + C_{\mathsf{FB}} + 2\log(1.5). \quad (92)$$

Hence, the gap between the achievable sum rate of (90) and the outer bound in (92) is at most $7 + 4\log(1.5) \leq 9.4$ bits/s/Hz.

2) $C_{\mathsf{FB}} \geq \log\left(\left(\frac{\mathsf{SNR}^2}{\mathsf{INR}^3} - 1\right)^+\right)$:
   We have

$$R_{\mathsf{SUM}}^{(b)} \geq 2\log\left(1 + \mathsf{INR}\right) + \log\left(1 + \left(\frac{\mathsf{SNR}^2}{\mathsf{INR}^3}\right)^+\right) - 5 - 3\log(3)$$
$$\geq 2\log\left(1 + \mathsf{INR}\right) + 2\log\left(\mathsf{SNR}\right)$$
$$- 3\log\left(\mathsf{INR}\right) - 5 - 3\log(3) \quad (93)$$

$$\geq 2\log\left(1 + \mathsf{INR}\right) + 2\log\left(1 + \mathsf{SNR}\right)$$
$$- 3\log\left(1 + \mathsf{INR}\right) - 7 - \log(3)$$
$$= 2\log\left(1 + \mathsf{SNR}\right) - \log\left(1 + \mathsf{INR}\right)$$
$$- 7 - 3\log(3).$$

Hence, the gap between the achievable sum rate of (93) and the outer bound in (89) is at most $10 + 3\log(3) \leq 14.8$ bits/s/Hz. As a result, the gap between the inner bound and the minimum of the outer bounds in (31b) and (31c) is at most 14.8 bits/s/Hz.

Case (c) $2\log\left(\mathsf{SNR}\right) \leq \log\left(\mathsf{INR}\right)$:

1) If $\mathsf{SNR} \leq 1$, pick

$$P_2^{(2)} = P_1^{(3)} = \frac{1}{\mathsf{INR}}\min\{2^{C_{\mathsf{FB}}}, \mathsf{INR}\}$$

and set all other power levels equal to zero. By plugging the given values of power levels into our achievable sum rate $R_{\mathsf{SUM}}^{(c)}$ defined in (59), we get

$$R_{\mathsf{SUM}}^{(c)} = \log\left(\frac{\min\{2^{C_{\mathsf{FB}}}, \mathsf{INR}\}}{2}\right)$$
$$= \min\{C_{\mathsf{FB}}, \log\left(\mathsf{INR}\right)\} - 1$$
$$\geq \min\{C_{\mathsf{FB}}, \log\left(1 + \mathsf{INR}\right)\} - 2. \quad (94)$$

Consider the outer bounds in (31a) and (31b), under the assumptions of case (c) and $\mathsf{SNR} \leq 1$, we have

$$R_1 + R_2 \leq \min\{2\log\left(1 + \mathsf{SNR}\right) + C_{\mathsf{FB}}$$
$$\log\left(1 + \frac{\mathsf{SNR}}{1 + \mathsf{INR}}\right)$$
$$+ \log\left(1 + \mathsf{SNR} + \mathsf{INR} + 2\sqrt{\mathsf{SNR}\cdot\mathsf{INR}}\right)\}$$
$$\leq \min\{2\log(2) + C_{\mathsf{FB}}, \log(2) + \log(2 + 3\mathsf{INR})\}$$
$$\leq \min\{2 + C_{\mathsf{FB}}, 1 + \log(3) + \log\left(1 + \mathsf{INR}\right)\}$$
$$\leq \min\{C_{\mathsf{FB}}, \log\left(1 + \mathsf{INR}\right)\} + 2.6. \quad (95)$$

Therefore, with the given choice of power levels, the achievable sum rate of (94) is within 2.6 bits/s/Hz of the minimum of the outer bounds in (31a) and (31b).

2) If $\mathsf{SNR} \geq 1$,
   pick the following set of power levels:

$$\begin{cases} P_1^{(1)} = 0 \\ P_1^{(2)} = 0 \\ P_1^{(3)} = \frac{\mathsf{SNR}}{\mathsf{INR}}\min\{2^{C_{\mathsf{FB}}}, \frac{\mathsf{INR}}{\mathsf{SNR}^2}\} \\ P_1^{(4)} = 1 - P_1^{(3)} \\ P_2^{(1)} = 0 \\ P_2^{(2)} = \frac{1}{\mathsf{INR}}\min\{2^{C_{\mathsf{FB}}}, \frac{\mathsf{INR}}{\mathsf{SNR}^2}\} \\ P_2^{(3)} = 1 - P_2^{(2)}. \end{cases} \quad (96)$$

It is straightforward to check that the power levels are nonnegative and they satisfy the power constraint at the transmitters. By plugging the given values of power levels into our achievable sum rate $R_{\mathsf{SUM}}^{(c)}$ defined in (59), we get

$$R_{\mathsf{SUM}}^{(c)} = \log\left(1 + \frac{\mathsf{SNR}(1 - \frac{1}{\mathsf{SNR}})}{2}\right)$$

$$+ \log \left(1 + \frac{\mathsf{SNR}(1 - \frac{1}{\mathsf{SNR}})}{2} \right) + \log \left(\frac{\min\{2^{C_{\mathsf{FB}}}, \frac{\mathsf{INR}}{\mathsf{SNR}^2}\}}{2} \right)$$

$$\geq 2 \log (1 + \mathsf{SNR}) + \min\{C_{\mathsf{FB}}, \log \left(\frac{\mathsf{INR}}{\mathsf{SNR}^2} \right)\} - 3$$

$$= \min\{2 \log (1 + \mathsf{SNR}) + C_{\mathsf{FB}}$$

$$2 \log (1 + \mathsf{SNR}) + \log \left(\frac{\mathsf{INR}}{\mathsf{SNR}^2} \right)\} - 3$$

$$\geq \min\{2 \log (1 + \mathsf{SNR}) + C_{\mathsf{FB}}, \log (1 + \mathsf{INR})\} - 4. \quad (97)$$

Consider the outer bounds in (31a) and (31b), under the assumptions of case (c), we have

$$R_1 + R_2 \leq \min\{2 \log (1 + \mathsf{SNR}) + C_{\mathsf{FB}}$$

$$\log \left(1 + \frac{\mathsf{SNR}}{1 + \mathsf{INR}} \right)$$

$$+ \log \left(1 + \mathsf{SNR} + \mathsf{INR} + 2\sqrt{\mathsf{SNR} \cdot \mathsf{INR}} \right)\}$$

$$\leq \min\{2 \log (1 + \mathsf{SNR}) + C_{\mathsf{FB}}$$

$$\log (2) + \log (1 + 4\mathsf{INR})\}$$

$$\leq \min\{2 \log (1 + \mathsf{SNR}) + C_{\mathsf{FB}}, 3 + \log (1 + \mathsf{INR})\}$$

$$\leq \min\{2 \log (1 + \mathsf{SNR}) + C_{\mathsf{FB}}, \log (1 + \mathsf{INR})\} + 3. \quad (98)$$

Therefore, with the given choice of power levels, the achievable sum rate of (59) is within 7 bits/s/Hz of the minimum of the outer bounds in (31a) and (31b).

Case (d) $\frac{2}{3} \log (\mathsf{SNR}) \leq \log (\mathsf{INR}) \leq \log (\mathsf{SNR})$: In this case, feedback is not needed. The achievability scheme of [35] for Gaussian IC without feedback results in a sum rate to within 1 bit/s/Hz of

$$\log (1 + \mathsf{SNR}) + \log \left(1 + \frac{\mathsf{SNR}}{1 + \mathsf{INR}} \right). \quad (99)$$

For the outer bound in (31b), in this case, we have

$$R_1 + R_2 \leq \log \left(1 + \frac{\mathsf{SNR}}{1 + \mathsf{INR}} \right)$$

$$+ \log \left(1 + \mathsf{SNR} + \mathsf{INR} + 2\sqrt{\mathsf{SNR} \cdot \mathsf{INR}} \right)$$

$$\leq \log \left(1 + \frac{\mathsf{SNR}}{1 + \mathsf{INR}} \right) + \log (1 + 4\mathsf{SNR})$$

$$\leq \log \left(1 + \frac{\mathsf{SNR}}{1 + \mathsf{INR}} \right) + \log (1 + \mathsf{SNR}) + 2. \quad (100)$$

Therefore, the achievable sum rate of [35] is within 3 bits/s/Hz of the outer bound in (31b).

Case (e) $\log (\mathsf{SNR}) \leq \log (\mathsf{INR}) \leq 2 \log (\mathsf{SNR})$: In this case, feedback is not needed. The achievability scheme of [35] for Gaussian IC without feedback results in a sum-rate to within 1 bit/s/Hz of

$$\log (1 + \mathsf{SNR} + \mathsf{INR}). \quad (101)$$

For the outer bound in (31b), in this case, we have

$$R_1 + R_2 \leq \log \left(1 + \frac{\mathsf{SNR}}{1 + \mathsf{INR}} \right)$$

$$+ \log \left(1 + \mathsf{SNR} + \mathsf{INR} + 2\sqrt{\mathsf{SNR} \cdot \mathsf{INR}} \right)$$

$$\overset{(\mathsf{SNR} \leq \mathsf{INR})}{\leq} \log (2) + \log (1 + 4\mathsf{INR})$$

$$\leq \log (1 + \mathsf{INR}) + 3. \quad (102)$$

Therefore, the achievable sum rate of [35] is within 4 bits/s/Hz of the outer bound in (31b).

REFERENCES

[1] C. E. Shannon, "The zero error capacity of a noisy channel," *IRE Trans. Inf. Theory*, Sep. 1956.

[2] N. T. Gaarder and J. K. Wolf, "The capacity region of a multiple-access discrete memoryless channel can increase with feedback," *IEEE Trans. Inf. Theory*, vol. 21, no. 1, pp. 100–102, Jan. 1975.

[3] L. H. Ozarow, "The capacity of the white Gaussian multiple access channel with feedback," *IEEE Trans. Inf. Theory*, vol. 30, no. 4, pp. 623–629, Jul. 1984.

[4] G. Kramer, "Feedback strategies for white Gaussian interference networks," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1423–1438, Jun. 2002.

[5] G. Kramer, "Correction to feedback strategies for white Gaussian interference networks, and a capacity theorem for Gaussian interference channels with feedback," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1373–1374, Jun. 2004.

[6] M. Gastpar and G. Kramer, "On noisy feedback for interference channels," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Oct. 2006, pp. 216–220.

[7] R. Tandon and S. Ulukus, "Dependence balance based outer bounds for Gaussian networks with cooperation and feedback," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4063–4086, Jul. 2011.

[8] J. Jiang, Y. Xin, and H. K. Garg, "Discrete memoryless interference channels with feedback," in *Proc. 41st Annu. Conf. Inf. Sci. Syst.*, Mar. 2007, pp. 581–584.

[9] C. Suh and D. Tse, "Feedback capacity of the Gaussian interference channel to within 2 bits," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2667–2685, May 2011.

[10] A. Sahai, V. Aggarwal, M. Yuksel, and A. Sabharwal, "On channel output feedback in deterministic interference channels," in *Proc. 41st Annu. Conf. Inf. Sci. Syst.*, Oct. 2009, pp. 298–302.

[11] A. El-Gamal and M. H. Costa, "The capacity region of a class of deterministic interference channels," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 343–346, Mar. 1982.

[12] A. S. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011.

[13] G. Bresler and D. Tse, "The two-user Gaussian interference channel: A deterministic view," *Eur. Trans. Telecommun.*, vol. 19, pp. 1–34, 2008.

[14] G. Bresler, A. Parekh, and D. Tse, "The approximate capacity of the many-to-one and one-to-many Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4566–4592, Sep. 2010.

[15] A. Sezgin, S. Avestimehr, M. Khajehnejad, and B. Hassibi, "Divide-and-conquer: Approaching the capacity of the two-pair bidirectional Gaussian relay network," *IEEE Trans. Inf. Theory*, to be published.

[16] A. P. Hekstra and F. M. J. Willems, "Dependence balance bounds for single-output two-way channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 44–53, Jan. 1989.

[17] T. M. Cover and A. El-Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.

[18] D. Tuninetti, "On interference channel with generalized feedback (IFC-GF)," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2007, pp. 2861–2865.

[19] S. Yang and D. Tuninetti, "Interference channel with generalized feedback (aka with source cooperation): Part I: Achievable region," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2686–2710, May 2011.

[20] Y. Cao and B. Chen, "An achievable rate region for interference channel with conferencing," in *Proc. Int. Symp. Inf. Theory*, Jun. 2007, pp. 1251–1255.

[21] V. Prabhakaran and P. Viswanath, "Interference channels with source cooperation," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 156–186, May 2011.

[22] I.-H. Wang and D. Tse, "Interference mitigation through limited transmitter cooperation," in *Proc. IEEE Trans. Inf. Theory*, Jun. 2010, pp. 380–384, arXiv:1004.5421.

[23] H. Bagheri, A. S. Motahari, and A. K. Khandani, "On the symmetric Gaussian interference channel with partial unidirectional cooperation," 2009, arXiv:0909.2777.

[24] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[25] S. Gel'Fand and M. Pinsker, "Coding for channel with random parameters," *Probl. Contr. Inf. Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[26] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.

[27] F. M. J. Willems, "The feedback capacity region of a class of discrete memoryless multiple access channels," *IEEE Trans. Inf. Theory*, vol. 28, no. 1, pp. 93–95, Jan. 1982.

[28] A. P. Hekstra and F. M. J. Willems, "Dependence balance bounds for single-output two-way channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 44–53, Jan. 1989.

[29] A. E. Gamal and Y.-H. Kim, "Lecture Notes on Network Information Theory" 2010, arXiv:1001.3404.

[30] C. Suh and D. Tse, "The feedback capacity region of El Gamal–Costa deterministic interference channel," in *Proc. Allerton Conf. Commun., Control, Comput.*, Sep. 2009.

[31] A. Vahid and A. S. Avestimehr, "The two-user deterministic interference channel with rate-limited feedback," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2010, pp. 460–464.

[32] H. A. Loeliger, "Averaging bounds for lattices and linear codes," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1767–1773, Nov. 1997.

[33] G. Poltyrev, "On coding without restrictions for the AWGN channel," *IEEE Trans. Inf. Theory*, vol. 40, no. 2, pp. 409–417, Mar. 1994.

[34] U. Erez, S. Litsyn, and R. Zamir, "Lattices which are good for (almost) everything," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3401–3416, Oct. 2005.

[35] R. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.

Alireza Vahid received his B.Sc. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2009. He is currently a Ph.D. student at the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY. His research interests include information theory and wireless communications.

He has received the Director's Ph.D. Teaching Assistant Award in 2010, and Jacobs Scholar Fellowship in 2009.

Changho Suh is a postdoctoral associate at MIT. He received the Ph.D. degree in Electrical Engineering and Computer Sciences from UC-Berkeley in 2011. He received the B.S. and M.S. degrees in Electrical Engineering from Korea Advanced Institute of Science and Technology in 2000 and 2002 respectively. Before joining Berkeley, he had been with the Telecommunication R&D Center, Samsung Electronics.

Dr. Suh received the David J. Sakrison Memorial Prize for outstanding doctoral research from the UC-Berkeley EECS Department in 2011, the Best Student Paper Award of the IEEE International Symposium on Information Theory in 2009 and the Outstanding Graduate Student Instructor Award in 2010. He was awarded several fellowships, including the Vodafone U.S. Foundation Fellowship in 2006 and 2007; the Kwanjeong Educational Foundation Fellowship in 2009; and the Korea Government Fellowship from 1996 to 2002.

A. Salman Avestimehr received the B.S. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2003 and the M.S. degree and Ph.D. degree in electrical engineering and computer science, both from the University of California, Berkeley, in 2005 and 2008, respectively.

He is currently an Assistant Professor at the School of Electrical and Computer Engineering at Cornell University, Ithaca, NY. He was also a postdoctoral scholar at the Center for the Mathematics of Informa-tion (CMI) at the California Institute of Technology, Pasadena, in 2008. His research interests include information theory, communications, and networking.

Dr. Avestimehr has received a number of awards, including the Presidential Early Career Award for Scientists and Engineers (PECASE) in 2011, the Young Investigator Program (YIP) award from the U. S. Air Force Office of Scientific Research (2011), the National Science Foundation CAREER award (2010), the David J. Sakrison Memorial Prize from the U.C. Berkeley EECS Department (2008), and the Vodafone U.S. Foundation Fellows Initiative Research Merit Award (2005). He has been a Guest Associate Editor for the IEEE Transactions on Information Theory Special Issue on Interference Networks.