Capacity Theorems for the "Z" Channel

Hon-Fah Chong, *Student Member, IEEE*, Mehul Motani, *Member, IEEE*, and Hari Krishna Garg, *Senior Member, IEEE*

***Abstract*—We consider the two-user "Z" channel (ZC), where there are two senders and two receivers. One of the senders transmits information to its intended receiver (without interfering with the unintended receiver), while the other sender transmits information to both receivers. The complete characterization of the discrete memoryless ZC remains unknown to date. For the Gaussian ZC, the capacity has only been established for a crossover link gain of 1. In this work, we study both the discrete memoryless ZC and the Gaussian ZC. We first establish achievable rates for the general discrete memoryless ZC. The coding strategy uses rate-splitting and superposition coding at the sender with information for both receivers. At the receivers, we use joint decoding. We then specialize the rates obtained to two different types of degraded discrete memoryless ZCs and also derive respective outer bounds to their capacity regions. We show that as long as a certain condition is satisfied, the achievable rate region is the capacity region for one type of degraded discrete memoryless ZC. The results are then extended to the two-user Gaussian ZC with different crossover link gains. We determine an outer bound to the capacity region of the Gaussian ZC with strong crossover link gain and establish the capacity region for moderately strong crossover link gain.**

***Index Terms*—Gaussian "Z" channel (ZC), rate-splitting, simultaneous decoding, superposition coding, "Z" channel (ZC).**

I. Introduction

In the past, the study of multiuser information theory has largely been motivated by wireline and cellular systems. Multiuser channel configurations often revolved around the multiple-access channel (cellular uplink), broadcast channel (cellular downlink), and interference channel (IC) (wireline systems). Researchers have also studied the two-way channel, the relay channel, and various other multiuser channel configurations (see [1, Ch. 14] and the references therein). With recent advances in noncentralized networks such as sensor networks and wireless *ad hoc* networks, there has been a growing interest in the study of other multiuser channels. Recently, Vishwanath, Jindal, and Goldsmith [2] introduced the "Z" channel (ZC) shown in Fig. 1. The ZC consists of two senders and two receivers. The transmission of sender TX_1 can reach only receiver RX_1, while that of sender TX_2 can reach both receivers.

The Z interference channel (ZIC) has the same topology as the ZC shown in Fig. 1. In both the ZC and ZIC, there is no cooperation between the two senders or between the two receivers. However, in the ZIC, sender TX_2 has no information to transmit to receiver RX_1, while the ZC allows transmission of information from sender TX_2 to receiver RX_1. Hence, the ZC models a more general multiuser network compared to the ZIC. The capacity region of the ZC includes the capacity region of the broadcast channel (sender TX_2 is transmitting information to both receivers), the capacity region of the multiple access channel (sender TX_1 and TX_2 are both transmitting information to receiver RX_1), and the capacity region of the ZIC (both senders are transmitting information to their own intended receivers).

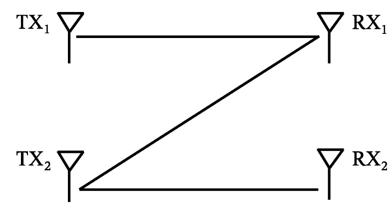

Fig. 1. The configuration of the ZC.

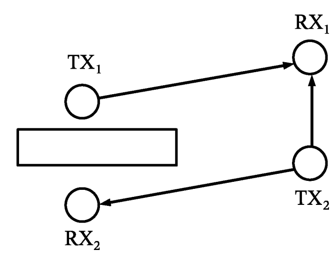

Fig. 2. A ZC: transmission of sender TX_1 is unable to reach receiver RX_2 due to an obstacle.

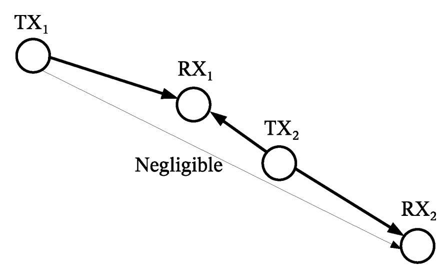

Fig. 3. A ZC: transmission of sender TX_1 is unable to reach receiver RX_2 due to distance.

Such a multiuser configuration may correspond to a local scenario (with two users and two receivers) in a large sensor or wireless *ad hoc* network. As shown in Fig. 2, sender TX_1 is unable to transmit to receiver RX_2 due to an obstacle, while sender TX_2 is able to transmit to both receivers. Another possible scenario is shown in Fig. 3, where sender TX_1 is so far away from receiver RX_2 that its transmission is negligible.

In this paper, we also study the Gaussian ZC shown in Fig. 4. We use the term *weak crossover link gain* to describe the scenario $0 < a^2 < 1$ and the term *strong crossover link gain* to describe the scenario $a^2 \geq 1$. Furthermore, we use the terms *moderately strong crossover link gain* and *very strong crossover link*

Manuscript received December 23, 2005; revised September 21, 2006. This work was supported in part by the National University of Singapore under Grants R-263-000-261-112 and R-263-000-293-112.

The authors are with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117576, Singapore (e-mail: g0305962@nus.edu.sg; motani@nus.edu.sg; eleghk@nus.edu.sg).

Communicated by K. Kobayashi, Associate Editor for Shannon Theory.
Digital Object Identifier 10.1109/TIT.2006.890779

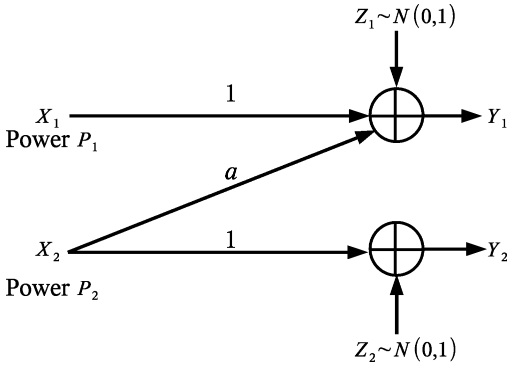

Fig. 4. Standard form Gaussian ZC.

gain to differentiate between the two scenarios $1 \le a^2 \le 1+P_1$ and $a^2 > 1+P_1$, respectively. Vishwanath, Jindal, and Goldsmith [2] established an achievable rate region for the Gaussian ZC with very strong crossover link gain. In [3], Liu and Ulukus determined an inner bound and an outer bound to the capacity region of the Gaussian ZC with weak crossover link gain. To date, the capacity region of the Gaussian ZC is only known when the crossover link gain is 1 [3].

The outline of the paper is as follows.

- We first give a mathematical model for the discrete memoryless ZC in Section II. We then describe three different types of degraded ZCs. We also describe the Gaussian ZC model.
- Next, we review past results on the ZC in Section III. We describe a problem in one of the proofs in [2] for the capacity region of one type of degraded discrete memoryless ZC.
- In Section IV, we establish an achievable rate region for the general discrete memoryless ZC using rate-splitting and joint decoding.
- In Section V, we specialize the result for the general setting to one type of degraded discrete memoryless ZC. We also determine an outer bound to the capacity region. The result is extended directly to the two-user Gaussian ZC with weak crossover link gain.
- In Section VI, we specialize the result for the general setting to another type of degraded discrete memoryless ZC. The result is extended directly to the Gaussian ZC with strong crossover link gain. We also determine respective outer bounds to their capacity regions. We establish the capacity region of the Gaussian ZC with moderately strong crossover link gain. In the discrete case, we show that the achievable rate region is the capacity region if a certain condition is satisfied. Finally, we show that the achievable rate region, determined in [2], for the Gaussian ZC with very strong crossover link gain can be enlarged.

II. Mathematical Preliminaries

A two-user discrete ZC consists of four finite sets $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2$, and a conditional joint distribution $p(y_1, y_2|x_1, x_2)$, with the conditional marginal distributions given by

$$p(y_1|x_1, x_2) = \sum_{y_2 \in \mathcal{Y}_2} p(y_1, y_2|x_1, x_2) \tag{1}$$

$$p(y_2|x_2) = \sum_{y_1 \in \mathcal{Y}_1} p(y_1, y_2|x_1, x_2). \tag{2}$$

The ZC is said to be memoryless if

$$p\left(y_1^N, y_2^N|x_1^N, x_2^N\right) = \prod_{n=1}^{N} p(y_{1n}|x_{1n}, x_{2n})\, p(y_{2n}|x_{2n}, y_{2n}).$$

Throughout the paper, we assume the ZC to be memoryless. From (2), we see that

$$X_1 \to X_2 \to Y_2 \tag{3}$$

form a Markov chain. Since there is no cooperation between the two receivers, the capacity region of the ZC depends on the conditional joint distribution $p(y_1, y_2|x_1, x_2)$ only through the conditional marginal distributions. In addition, we note that X_1 and Y_2 are independent for all input distributions of the form $p(x_1)\,p(x_2)$.

$$\begin{aligned}\sum_{x_2 \in \mathcal{X}_2} \sum_{y_1 \in \mathcal{Y}_1} & p(y_1, y_2|x_1, x_2)\, p(x_1)\, p(x_2) \\ &= p(x_1) \sum_{x_2 \in \mathcal{X}_2} p(y_2|x_2)\, p(x_2) \\ &= p(x_1)\, p(y_2).\end{aligned} \tag{4}$$

Similarly, X_1^N and Y_2^N are independent for all input distributions of the form $p\left(x_1^N\right) p\left(x_2^N\right)$. In the ZC, sender TX_1 produces an integer $W_1 \in \left\{1, \ldots, 2^{NR_1}\right\}$. Sender TX_2 produces an integer pair

$$(W_{21}, W_{22}) \in \left\{1, \ldots, 2^{NR_{21}}\right\} \times \left\{1, \ldots, 2^{NR_{22}}\right\}.$$

W_1 denotes the message sender TX_1 intends to transmit to receiver RX_1, W_{21} denotes the message sender TX_2 intends to transmit to receiver RX_1, and W_{22} denotes the message sender TX_2 intends to transmit to receiver RX_2. A $\left(2^{NR_1}, 2^{NR_{21}}, 2^{NR_{22}}, N\right)$ code for a ZC with independent messages consists of two encoders

$$\begin{aligned} f_1 &: \left\{1, \ldots, 2^{NR_1}\right\} \to \mathcal{X}_1^N \\ f_2 &: \left\{1, \ldots, 2^{NR_{21}}\right\} \times \left\{1, \ldots, 2^{NR_{22}}\right\} \to \mathcal{X}_2^N \end{aligned}$$

and two decoders

$$\begin{aligned} g_1 &: \mathcal{Y}_1^N \to \left\{1, \ldots, 2^{NR_1}\right\} \times \left\{1, \ldots, 2^{NR_{21}}\right\} \\ g_2 &: \mathcal{Y}_2^N \to \left\{1, \ldots, 2^{NR_{22}}\right\}. \end{aligned}$$

The average probability of error is defined as the probability that the decoded messages are not equal to the transmitted messages, i.e.,

$$P_e^{(N)} = \Pr\left(g_1\left(Y_1^N\right) \neq (W_1, W_{21}) \text{ or } g_2\left(Y_2^N\right) \neq W_{22}\right).$$

The distributions of W_1, W_{21}, and W_{22} are assumed to be uniform. A rate triplet (R_1, R_{21}, R_{22}) is said to be achievable for the ZC if there exists a sequence of $\left(2^{NR_1}, 2^{NR_{21}}, 2^{NR_{22}}, N\right)$ codes with $P_e^{(N)} \to 0$.

Willems and Van Der Meulen proved that stochastic encoders and decoders do not increase the capacity region of the discrete memoryless multiple-access channel with cribbing encoders [4]. The same argument can be extended to the ZC.

Proposition 1: Stochastic encoders and decoders do not increase the capacity region of the ZC.

Proof: For stochastic encoders and decoders, we may assume that the encoding and decoding functions are given by

$$x_1^N = f_1\left(W_1, A^{E1}\right) \tag{5}$$

$$x_2^N = f_2\left(W_{21}, W_{22}, A^{E2}\right) \tag{6}$$

$$\left(\hat{W}_1, \hat{W}_{21}\right) = g_1\left(Y_1^N, A^{D1}\right) \tag{7}$$

$$\hat{W}_{22} = g_2\left(Y_2^N, A^{D2}\right) \tag{8}$$

where A^{E1}, A^{E2}, A^{D1}, and A^{D2} are random variables independent of each other and all other random variables. Now, define

$$A \triangleq \left(A^{E1}, A^{E2}, A^{D1}, A^{D2}\right) \tag{9}$$

where A ranges over $\mathcal{A}$ and $p(\cdot)$ is A's density function. If a $\left(2^{NR_1}, 2^{NR_{21}}, 2^{NR_{22}}, N\right)$-code exists for stochastic encoders and decoders, and achieves a probability of error P_e, we then have

$$\begin{aligned} P_e &= \Pr\left\{\left(\hat{W}_1, \hat{W}_{21}, \hat{W}_{22}\right) \neq (W_1, W_{21}, W_{22})\right\} \\ &= \int_{a\in\mathcal{A}} p(A=a) \Pr\Big\{\left(\hat{W}_1, \hat{W}_{21}, \hat{W}_{22}\right) \\ &\quad \neq (W_1, W_{21}, W_{22}) \,|A=a\Big\} da. \end{aligned} \tag{10}$$

It then readily follows that there must exist an $a \in \mathcal{A}$ such that

$$\Pr\left\{\left(\hat{W}_1, \hat{W}_{21}, \hat{W}_{22}\right) \neq (W_1, W_{21}, W_{22})\,|A=a\right\} \leq P_e. \tag{11}$$

Hence, the capacity region of the ZC is unaffected if we assume deterministic encoders and decoders. □

A. Some Useful Properties of Markov Chains

We state some useful properties of Markov chains that we will use throughout the paper (see [5, Sec. 1.1.5]).

- Decomposition: $X \to Z \to YW \Rightarrow X \to Z \to Y$;
- Weak Union: $X \to Z \to YW \Rightarrow X \to ZW \to Y$;
- Contraction: $(X \to Z \to Y)$ and $(X \to ZY \to W) \Rightarrow X \to Z \to YW$.

B. Degraded ZC

We first define three types of physically degraded ZCs. A ZC is said to be stochastically degraded if its conditional marginal distributions are the same as that of a physically degraded ZC. Since $\Pr((\hat{W}_1, \hat{W}_{21}) \neq (W_1, W_{21}))$ and $\Pr(\hat{W}_{22} \neq W_{22})$ depend only on the conditional marginal distributions $p(y_1|x_1, x_2)$ and $p(y_2|x_2)$, the capacity region of the stochastically degraded ZC is the same as that of the corresponding physically degraded ZC. In the rest of the paper, we assume that the ZCs are physically degraded.

Definition 1: We define a ZC to be a *degraded* ZC of *type I* if

$$X_2 \to (X_1, Y_2) \to Y_1 \tag{12}$$

form a Markov chain.

Remark 1: The conditional joint distribution $p(y_1, y_2|x_1, x_2)$ can be written as

$$\begin{aligned} p(y_1, y_2|x_1, x_2) &= p(y_2|x_1, x_2)\, p(y_1|x_1, x_2, y_2) \\ &= p(y_2|x_2)\, p(y_1|x_1, y_2). \end{aligned} \tag{13}$$

For the degraded ZC of type I, the following inequality holds:

$$I(U; Y_2) \geq I(U; Y_1|X_1) \tag{14}$$

for all input distributions $p(x_1)\,p(u)\,p(x_2|u)$.

Example 1: Fig. 7 shows a degraded Gaussian ZC of type I. One may easily verify that the two Markov chains given by (3) and (12) are simultaneously satisfied.

Definition 2: We define a ZC to be a *degraded* ZC of *type II* if

$$X_2 \to (X_1, Y_1) \to Y_2 \tag{15}$$

form a Markov chain.

Remark 2: For the degraded ZC of type II, the conditional joint distribution $p(y_1, y_2|x_1, x_2)$ can be written as

$$\begin{aligned} p(y_1, y_2|x_1, x_2) &= p(y_1|x_1, x_2)\, p(y_2|x_1, x_2, y_1) \\ &= p(y_1|x_1, x_2)\, p(y_2|x_1, y_1). \end{aligned} \tag{16}$$

The following inequality holds:

$$I(U; Y_1|X_1) \geq I(U; Y_2) \tag{17}$$

for all input distributions $p(x_1)\,p(u)\,p(x_2|u)$.

Example 2: Fig. 9 shows a degraded Gaussian ZC of type II. One may easily verify that the two Markov chains given by (3) and (15) are simultaneously satisfied.

Definition 3: We define a ZC to be a *degraded* ZC of *type III* if

$$(X_1, X_2) \to Y_1 \to Y_2 \tag{18}$$

form a Markov chain.

Remark 3: The degraded ZC of type III was first defined in [2] and corresponds to the case where the output of receiver RX_2 (Y_2) is a degraded version of the output of receiver RX_1 (Y_1). By applying the weak union property for Markov chains, we see that the Markov chain $X_2 \to (X_1, Y_1) \to Y_2$ holds for the degraded ZC of type III. Hence, a degraded ZC of type III is also a degraded ZC of type II. However, the converse may not necessarily be true.

Example 3: We consider the degraded ZC of type III shown in Fig. 5 where $\mathcal{X}_1 = \{x_{11}, x_{12}\}$, $\mathcal{X}_2 = \{x_{21}, x_{22}\}$, $\mathcal{Y}_1 = \{y_{11}, y_{12}, y_{13}, y_{14}\}$, and $\mathcal{Y}_2 = \{y_{21}, y_{22}\}$. We note that receiver

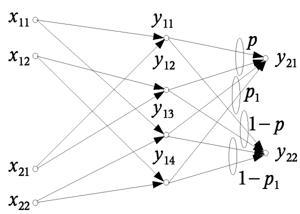

Fig. 5. An example of a degraded ZC of type III.

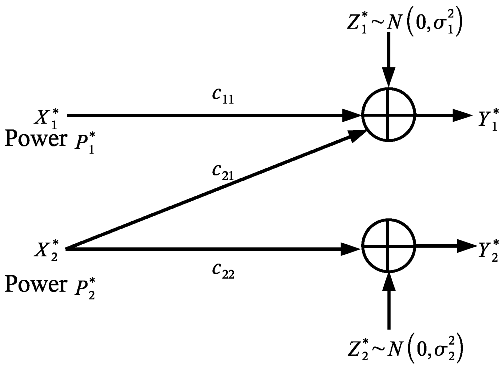

Fig. 6. General Gaussian ZC.

RX_1 is able to decode X_1 and X_2 without error. We also have $p(Y_2|Y_1 = y_{11}) = p(Y_2|Y_1 = y_{12})$ and $p(Y_2|Y_1 = y_{13}) = p(Y_2|Y_1 = y_{14})$. One may easily verify that the two Markov chains given by (3) and (18) are simultaneously satisfied.

C. Gaussian ZC

For a general Gaussian ZC, the inputs and outputs are related by

$$Y_1^* = c_{11}X_1^* + c_{21}X_2^* + Z_1^* \quad (19)$$
$$Y_2^* = c_{22}X_2^* + Z_2^*. \quad (20)$$

as depicted in Fig. 6. The channel outputs and inputs are real valued and have power constraints $\mathbb{E}[|X_1^*|^2] \le P_1^*$ and $\mathbb{E}[|X_2^*|^2] \le P_2^*$. Z_1^* and Z_2^* are zero-mean Gaussian random variables with variance σ_1^2 and σ_2^2, respectively. Similar to the Gaussian IC, one can use a scaling transformation to convert the Gaussian ZC into its standard form as shown in Fig. 4. The inputs and outputs of the standard form Gaussian ZC are related by

$$\begin{aligned} Y_1 &= X_1 + aX_2 + Z_1 \\ Y_2 &= X_2 + Z_2 \end{aligned} \quad (21)$$

where

$$\begin{aligned} X_1 &= \frac{c_{11}}{\sigma_1}X_1^*, & Y_1 &= \frac{Y_1^*}{\sigma_1}, & Z_1 &= \frac{Z_1^*}{\sigma_1} \\ X_2 &= \frac{c_{22}}{\sigma_2}X_2^*, & Y_2 &= \frac{Y_2^*}{\sigma_2}, & Z_2 &= \frac{Z_2^*}{\sigma_2} \end{aligned} \quad (22)$$

and the new power constraints and channel gain are

$$P_1 = \frac{c_{11}^2}{\sigma_1^2}P_1^*, \qquad P_2 = \frac{c_{22}^2}{\sigma_2^2}P_2^*, \qquad a = \frac{c_{12}}{c_{11}}\frac{\sigma_1}{\sigma_2}. \quad (23)$$

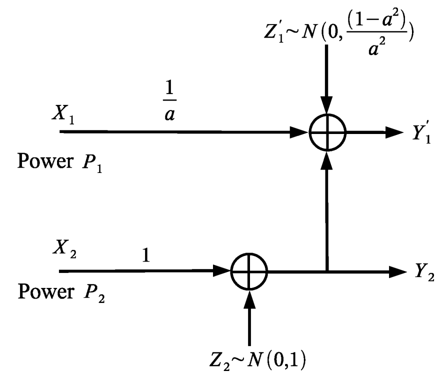

Fig. 7. Degraded Gaussian ZC of type I.

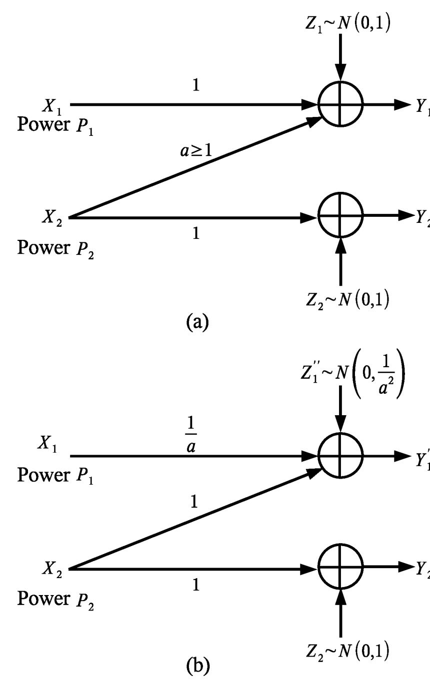

Fig. 8. Transformation of the Gaussian ZC ($a^2 \ge 1$).

1) Equivalent Gaussian ZC With Weak Crossover Link Gain ($0 < a^2 < 1$): In [6], Costa showed that the class of Gaussian ZIC with weak interference ($a^2 \in (0, 1)$) and the class of degraded Gaussian IC are equivalent, i.e., for every Gaussian ZIC with weak interference, there is a degraded Gaussian IC with the same capacity region. Using the same arguments as in [6], we can deduce that the class of Gaussian ZC with weak crossover link gain and the class of degraded Gaussian ZC of type I are equivalent, i.e., for every Gaussian ZC with weak crossover gain, there is a degraded Gaussian ZC of type I with the same capacity region. Hence, the capacity region of the channel shown in Fig. 7 is equivalent to that of the model shown in Fig. 4 when $0 < a^2 < 1$. An achievable rate region for the degraded discrete memoryless ZC of type I can be readily extended to the Gaussian ZC with weak crossover link gain. The assumption $0 < a^2 < 1$ ensures that the term $\frac{1-a^2}{a^2}$ is nonnegative.

2) Equivalent Gaussian ZC With Strong Crossover Link Gain ($a^2 \ge 1$): Consider the two channels shown in Fig. 8. The second channel is equivalent to the first since scaling the output of a channel does not affect its capacity. The channel shown in

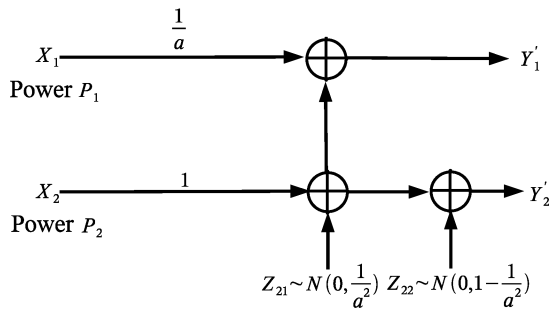

Fig. 9. A degraded Gaussian ZC of type II.

Fig. 9 is equivalent to the channel shown in Fig. 8(b) since they have identical conditional marginal distributions. In Fig. 9, the outputs are related to the inputs by

$$\begin{aligned} Y_1' &= \frac{X_1}{a} + X_2 + Z_{21} \\ Y_2' &= X_2 + Z_{21} + Z_{22} \end{aligned} \tag{24}$$

where $Z_{21} \sim \mathcal{N}\left(0, \frac{1}{a^2}\right)$ and $Z_{22} \sim \mathcal{N}\left(0, 1 - \frac{1}{a^2}\right)$. We will make use of this equivalent channel to determine an outer bound to the capacity region of the Gaussian ZC with strong crossover link gain. Here, we have made the assumption that $a^2 \geq 1$ to ensure that the term $1 - \frac{1}{a^2}$ is nonnegative. Since the class of Gaussian ZC with strong crossover link gain and the class of degraded Gaussian ZC of type II are equivalent, an achievable rate region for the degraded discrete memoryless ZC of type II can be readily extended to the Gaussian ZC with strong crossover link gain.

D. Jointly Typical Sequences

We review some basic results for typical sequences. Let $(X_1, \ldots, X_k)$ denote a finite collection of discrete random variables with some fixed joint distribution, $p(x_1, x_2, \ldots, x_k)$, $(x_1, x_2, \ldots, x_k) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_k$. Let S denote an ordered subset of these random variables and consider N independent copies of S. Thus

$$p\left(s^N\right) = \prod_{n=1}^{n=N} p(s_n), \quad s^N \in \mathcal{S}^N.$$

Definition 4: The set $A_\epsilon^{(N)}(X_1, X_2, \ldots, X_k)$ of ϵ-typical N-sequences $\left(x_1^N, x_2^N, \ldots, \boldsymbol{x}_k^N\right)$ is defined by

$$\begin{aligned} & A_\epsilon^{(N)}(X_1, \ldots, X_k) \\ & = A_\epsilon^{(N)} \\ & = \Big\{ \left(x_1^N, \ldots, \boldsymbol{x}_k^N\right) : \\ & \left| -\frac{1}{N} \log_2 p\left(s^N\right) - H(S) \right| < \epsilon, \forall S \subseteq \{X_1, \ldots, X_k\} \Big\}. \end{aligned}$$

For any $\epsilon > 0$, there exists an integer N such that $A_\epsilon^{(N)}(S)$ satisfies the following.

1) $\Pr\left(A_\epsilon^{(N)}(S)\right) \geq 1 - \epsilon, \forall S \subseteq \{X_1, X_2, \ldots, X_k\}$.
2) $s^N \in A_\epsilon^{(N)}(S) \Rightarrow \left| -\frac{1}{N} \log_2 p\left(s^N\right) - H(S) \right| < \epsilon$.
3) $(1-\epsilon) 2^{N(H(S)-\epsilon)} \leq \|A_\epsilon^{(N)}(S)\| \leq 2^{N(H(S)+\epsilon)}$.
4) Let S_1 and S_2 be two subsets of $\{X_1, X_2, \ldots, X_k\}$. If $\left(s_1^N, s_2^N\right) \in A_\epsilon^{(N)}(S_1, S_2)$, then
$$p\left(s_1^N | s_2^N\right) \doteq 2^{-N(H(S_1|S_2) \pm 2\epsilon)}.$$
5) Let S_1, S_2, and S_3 be three subsets of $\{X_1, X_2, \ldots, X_k\}$. If $\left(s_1^N, s_2^N, s_3^N\right) \in A_\epsilon^{(N)}(S_1, S_2, S_3)$, and if
$$p\left(s_1^N, s_2^N, s_3^N\right) = \prod_{n=1}^{n=N} p(s_{1n}|s_{3n})\, p(s_{2n}|s_{3n})\, p(s_{3n})$$
then
$$p\left\{\left(S_1^N, S_2^N, S_3^N\right) \in A_\epsilon^{(N)}(S_1, S_2, S_3)\right\} \doteq 2^{-N(I(S_1;S_2|S_3) \pm 6\epsilon)}.$$

E. Notation

We denote a discrete random variable with capital letter X and its realization with lower case letter x. A discrete random variable X takes values in a finite discrete set $\mathcal{X}$ and we use $\|\mathcal{X}\|$ to denote the cardinality of $\mathcal{X}$. We denote vectors with superscripts, e.g., X^N denotes a random vector and x^N denotes a realization of the random vector. The nth element of a random vector X^N is denoted by X_n, while the nth element of a vector x^N is denoted by x_n. We also denote by X_j^{j+M} the sequence of random variables $X_j, X_{j+1}, \ldots, X_{j+M}$. $p_X(x)$ denotes the probability distribution function of X on $\mathcal{X}$, while $p_{X^N}\left(\boldsymbol{x}^N\right)$ denotes the probability distribution function of X^N on $\mathcal{X}^N$. For brevity, we may omit the subscript X or X^N when it is clear from the context. We denote the entropy of a discrete random variable X by $H(X)$ and the differential entropy of a continuous random variable Y by $h(Y)$.

III. Review of Past Results

In this section, we review some known results for the ZC.

A. Degraded ZC of Type I

In [3, Larger Achievable Region 2], Liu and Ulukus determined a lower bound to the capacity region of the Gaussian ZC with weak crossover link gain. This corresponds to the degraded ZC of Type I. Liu and Ulukus make use of rate splitting and successive decoding technique similar to Carleial for the Gaussian IC [7]. Let us denote the information sender TX_2 intends to transmit to receiver RX_1 by W_{21} and the information sender TX_2 intends to transmit to receiver RX_2 by W_{22}. W_{21} has rate T_{21}. Sender TX_2 splits W_{22} in $[W_{221}, W_{222}]$, where W_{221} and W_{222} have rates S_{22} and T_{22}, respectively. W_{221} represents the information that only receiver RX_2 can decode, while W_{21} and W_{222} represent the information that both receivers can decode.

One strategy is to have receiver RX_1 decode W_{21} followed by W_{222} and finally W_1. Receiver RX_2 decodes W_{21} followed by W_{222} and finally W_{221}. Another strategy is to have receiver RX_1 decode W_{222} followed by W_{21} and finally W_1, while receiver RX_2 decodes W_{222} followed by W_{21} and finally W_{221}. The Larger Achievable Region 2 determined by Liu and Ulukus is the union of the achievable rate regions of these two strategies for the Gaussian ZC with weak crossover link gain. When this strategy is applied to the degraded discrete memoryless ZC of type I, an achievable rate region is given by the set $\mathcal{R}_{\mathrm{LU}}$, which is the closure of the convex hull of all rate triplets (R_1, R_{21}, R_{22}) satisfying

$$R_1 \le S_1 \quad (25)$$
$$R_{21} \le T_{21} \quad (26)$$
$$R_{22} \le S_{22} + T_{22} \quad (27)$$

where S_1, T_{21}, S_{22}, and T_{22} are subject to the constraints

$$T_{21} + T_{22} \le I(U; Y_1) \quad (28)$$
$$S_1 \le I(X_1; Y_1|U) \quad (29)$$
$$S_{22} \le I(X_2; Y_2|U) \quad (30)$$

for all input distributions $p(u, x_1, x_2) = p(x_1)\, p(u, x_2)$. In [3], Liu and Ulukus also determined an outer bound to the capacity region of the Gaussian ZC with weak crossover link gain. By making use of the entropy power inequality, Liu and Ulukus obtained the following theorem.

Theorem 1: [Liu and Ulukus] For the Gaussian ZC with weak crossover link gain ($0 \le a^2 \le 1$), the achievable rate triplets (R_1, R_{21}, R_{22}) have to satisfy

$$R_{21} \le \gamma\left(\frac{a^2\beta P_2}{a^2(1-\beta)P_2 + 1}\right) \quad (31)$$
$$R_{22} \le \gamma((1-\beta)P_2) \quad (32)$$
$$R_1 + R_{21} \le \gamma\left(\frac{a^2\beta P_2 + P_1}{a^2(1-\beta)P_2 + 1}\right) \quad (33)$$

for some $0 \le \beta \le 1$ and where $\gamma(x) \triangleq \frac{1}{2}\log_2(1+x)$.

Proof: The proof can be found in [3, Theorem 2]. □

Remark 4: This outer bound includes the best outer bound to the capacity region of the Gaussian ZIC under weak interference derived by Kramer [8, Theorem 2]. Kramer makes use of a proposition of Sato for a degraded interference channel, while Liu and Ulukus derived this using the entropy power inequality. To see the equivalence between the two, we can ignore the constraint for R_{21} since $R_{21} = 0$ for an interference channel. Hence, for the Gaussian ZIC under weak interference, the achievable rate pair (R_1, R_2) has to satisfy

$$R_1 \le \gamma\left(\frac{a^2\beta P_2 + P_1}{a^2(1-\beta)P_2 + 1}\right) \quad (34)$$
$$R_2 \le \gamma((1-\beta)P_2) \quad (35)$$

for some $0 \le \beta \le 1$. This is in fact the outer bound determined by Kramer for the capacity region of the degraded Gaussian IC, which is equivalent to that of the Gaussian ZIC under weak interference.

B. Degraded ZC of Type III

It was stated in [2] that the capacity region of a degraded discrete memoryless ZC of type III is the closure of the convex hull of all triplets (R_1, R_{21}, R_{22}) subject to

$$R_1 \le I(X_1; Y_1|X_2) \quad (36)$$
$$R_{21} \le I(X_2; Y_1|UX_1) \quad (37)$$
$$R_1 + R_{21} \le I(X_1X_2; Y_1|U) \quad (38)$$
$$R_{22} \le I(U; Y_2) \quad (39)$$

for some input distributions $p(u, x_1, x_2) = p(x_1)\, p(u, x_2)$.

Remark 5: The rates given by (36)–(39) can readily be seen to be achievable. Since the output of receiver RX_2 (Y_2) is a degraded version of receiver RX_1 (Y_1), we can use superposition coding at sender TX_2, where the auxiliary random variable U represents the information to be transmitted from sender TX_2 to receiver RX_2. Unfortunately, this achievable rate may not be the outer bound in general due to the following problem in the converse.

In [2], the authors define $U_i = (W_{22}, Y_{11}, Y_{12}, \ldots, Y_{1i-1})$ and state that $U_i \to X_{2i} \to (Y_{1i}Y_{2i})$ form a Markov chain. However, this is not necessarily the case as U_i may contain some information about Y_{1i} that is not in X_{2i}. We first observe that U_i contains all the past outputs of receiver RX_1 until time $i-1$. Moreover, the current output of receiver RX_1 (Y_{1i}) is dependent on the current input of sender TX_1 and sender TX_2 ($p(y_1|x_1, x_2)$). Hence, the Markov chain should be given by $U_i \to (X_{1i}X_{2i}) \to (Y_{1i}Y_{2i})$. Therefore, in the derivation of the outer bound, the input distribution $p(u, x_1, x_2)$ may not be equal to $p(x_1)\, p(u, x_2)$ as specified in [2].

IV. Achievable Rate Region for the Discrete Memoryless ZC

Similar to Carleial's treatment of the interference channel [7], we make use of rate splitting and superposition coding. Transmitter 2 splits W_{21} into $[W_{211}, W_{212}]$, where W_{211} and W_{212} have rates S_{21} and T_{21}, respectively. Similarly, transmitter 2 splits W_{22} into $[W_{221}, W_{222}]$, where W_{221} and W_{222} have rates S_{22} and T_{22}, respectively. Referring to Fig. 10, W_{211} represents the information that only receiver RX_1 can decode, while W_{221} represents the information that only receiver RX_2 can decode. W_{212} and W_{222} represent the information that both receivers can decode.

Carleial suggested the use of sequential decoding at the receivers for the interference channel. In [9], Han and Kobayashi refined Carleial's method by using a joint decoder superior to sequential decoding for the interference channel. Rather than using the convex-hull operation, they added a time-sharing random variable Q. Following the ideas of Han and Kobayashi, we use a joint decoder at the receivers and also include a time-sharing random variable Q. We first describe the codebook generation, encoding at the transmitters, and decoding at the receivers before describing our main result in Theorem 2.

A. Random Codebook Construction

We first fix the following input probability distribution:

$$p(q, x_1, u, v_1, v_2, x_2) = p(q)\, p(x_1|q)\, p(u|q)\, p(v_1|u, q) \cdot p(v_2|u, q)\, p(x_2|u, v_1, v_2, q). \quad (40)$$

The auxiliary random variable (r.v.) u carries the common information W_{212} and W_{222}, the auxiliary r.v. v_1 carries the infor-

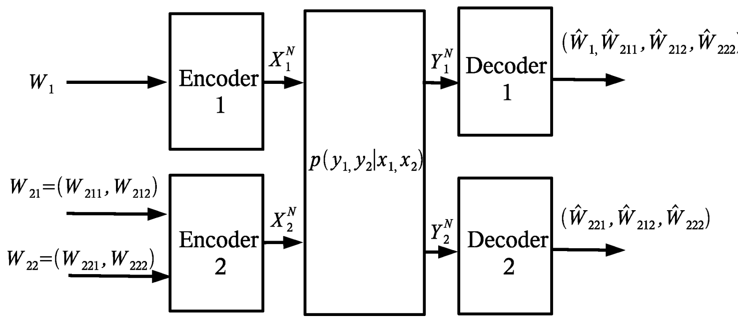

Fig. 10. Encoding and decoding for the ZC.

mation W_{211}, while the auxiliary r.v. v_2 carries the information W_{221}. The codebook is constructed as follows.

1) Generate one N-sequence $q^N=(q_1,q_2,\dots,q_N)$, drawn according to

$$p\left(q^N\right)=\prod_{n=1}^{n=N}p\left(q_n\right).$$

2) Generate 2^{NS_1} conditionally independent N-sequences $x_1^N=(x_{11},x_{12},\dots,x_{1N})$, each drawn according to

$$p\left(x_1^N|q^N\right)=\prod_{n=1}^{n=N}p\left(x_{1n}|q_n\right).$$

Label them $x_1^N\left(w_1\right)$, $w_1\in\left[1,2^{NS_1}\right]$.

3) Then, generate $2^{N(T_{21}+T_{22})}$ conditionally independent N-sequences $u^N=(u_1,u_2,\dots,u_N)$, each drawn according to

$$p\left(u^N|q^N\right)=\prod_{n=1}^{n=N}p\left(u_n|q_n\right).$$

Label them $u^N\left(w_{212},w_{222}\right)$, $w_{212}\in\left[1,2^{NT_{21}}\right]$, $w_{222}\in\left[1,2^{NT_{22}}\right]$.

4) For the codeword q^N and each of the codewords $u^N\left(w_{212},w_{222}\right)$, generate $2^{NS_{21}}$ conditionally independent N-sequences $v_1^N=(v_{11},v_{12},\dots,v_{1N})$, each drawn according to

$$p\left(v_1^N|u^N\left(w_{212},w_{222}\right),q^N\right)=\prod_{n=1}^{n=N}p\left(v_{1n}|u_n\left(w_{212},w_{222}\right),q_n\right).$$

Label them $v_1^N\left(w_{211},w_{212},w_{222}\right)$, $w_{211}\in\left[1,2^{NS_{21}}\right]$.

5) For the codeword q^N and each of the codewords $u^N\left(w_{212},w_{222}\right)$, generate $2^{NS_{22}}$ conditionally independent N-sequences $v_2^N=(v_{21},v_{22},\dots,v_{2N})$, each drawn according to

$$p\left(v_2^N|u^N\left(w_{212},w_{222}\right),q^N\right)=\prod_{n=1}^{n=N}p\left(v_{2n}|u_n\left(w_{212},w_{222}\right),q_n\right).$$

Label them $v_2^N\left(w_{221},w_{212},w_{222}\right)$, $w_{221}\in\left[1,2^{NS_{22}}\right]$.

6) Finally, for the codeword q^N and each of the following codewords $u^N\left(w_{212},w_{222}\right)$, $v_1^N\left(w_{211},w_{212},w_{222}\right)$, and $v_2^N\left(w_{221},w_{212},w_{222}\right)$, generate an N-sequence $x_2^N=(x_{21},x_{22},\dots,x_{2N})$, drawn according to

$$\begin{aligned}p\big(x_2^N|v_1^N&\left(w_{211},w_{212},w_{222}\right),v_2^N\left(w_{221},w_{212},w_{222}\right),\\&u^N\left(w_{212},w_{222}\right),q^N\big)\\&=\prod_{n=1}^{n=N}p\big(x_{2n}|v_{1n}\left(w_{211},w_{212},w_{222}\right),\\&\quad v_{2n}\left(w_{221},w_{212},w_{222}\right),u_n\left(w_{212},w_{222}\right),q_n\big).\end{aligned}$$

Label them $x_2^N\left(w_{211},w_{212},w_{221},w_{222}\right)$

B. Encoding and Decoding

To send the index w_1, sender TX_1 sends the codeword $x_1^N\left(w_1\right)$. To send the pair $\left(w_{211},w_{212}\right)$ to receiver RX_1 and the pair $\left(w_{221},w_{222}\right)$ to receiver RX_2, sender TX_2 sends the codeword $x_2^N\left(w_{211},w_{212},w_{221},w_{222}\right)$. For decoding, receiver RX_1 determines the unique $\left(\hat{w}_1,\hat{w}_{211},\hat{w}_{212},\hat{w}_{222}\right)$ such that

$$\begin{aligned}\big(q^N,u^N\left(\hat{w}_{212},\hat{w}_{222}\right),v_1^N\left(\hat{w}_{211},\hat{w}_{212},\hat{w}_{222}\right),\\x_1^N\left(\hat{w}_1\right),y_1^N\big)\in A_\epsilon^{(N)}\left(Q,U,V_1,X_1,Y_1\right).\end{aligned}\tag{41}$$

For the other decoder, receiver RX_2 determines the unique $\left(\hat{w}_{221},\hat{w}_{212},\hat{w}_{222}\right)$ such that

$$\begin{aligned}\big(q^N,u^N\left(\hat{w}_{212},\hat{w}_{222}\right),v_2^N\left(\hat{w}_{221},\hat{w}_{212},\hat{w}_{222}\right),y_2^N\big)\\\in A_\epsilon^{(N)}\left(Q,U,V_2,Y_2\right).\end{aligned}\tag{42}$$

C. Main Result

We may then state the main result.

Theorem 2: An achievable rate region for sending information over the discrete memoryless ZC is given by the set $\mathcal{R}_{\mathrm{G}}$, which is the closure of all rate triplets (R_1,R_{21},R_{22}) satisfying

$$R_1\le S_1\tag{43}$$
$$R_{21}\le S_{21}+T_{21}\tag{44}$$
$$R_{22}\le S_{22}+T_{22}\tag{45}$$

where S_1, S_{21}, S_{22}, T_{21}, and T_{22} are subject to the following constraints:

$$S_1+S_{21}+T_{21}+T_{22}\le I\left(X_1UV_1;Y_1|Q\right)\tag{46}$$
$$S_{21}+T_{21}+T_{22}\le I\left(UV_1;Y_1|X_1Q\right)\tag{47}$$
$$S_1+S_{21}\le I\left(X_1V_1;Y_1|UQ\right)\tag{48}$$

$$S_1 \le I(X_1;Y_1|UV_1Q) \quad (49)$$
$$S_{21} \le I(V_1;Y_1|X_1UQ) \quad (50)$$
$$S_{22}+T_{21}+T_{22} \le I(UV_2;Y_2|Q) \quad (51)$$
$$S_{22} \le I(V_2;Y_2|UQ) \quad (52)$$

for all input distributions of the form (40).

Proof: Refer to Appendix I. □

It is easy to see that $\mathcal{R}_{\mathrm{G}}$ is convex. In addition, we note that Theorem 2 is not limited to the ZC. It also applies to the general two-sender two-receiver channel (without the constraint in (3)) where one sender has information to transmit to both receivers, while the other sender has information to transmit to only one receiver. Next, we show that $\mathcal{R}_{\mathrm{G}}$ includes the capacity regions of the multiple-access channel and the degraded broadcast channel. It also includes the best known achievable rate region for the ZIC.

Remark 6: We obtain the multiple-access channel when $R_{22}=0$. By setting $S_{22}=T_{21}=T_{22}=0$, $R_1=S_1$, $R_{21}=S_{21}$, $Q=U=V_2=\phi$ and $V_1=X_2$, we obtain the capacity of the multiple-access channel, which is the closure of the convex hull of all rate pairs (R_1,R_{21}) satisfying

$$R_1 \le I(X_1;Y_1|X_2) \quad (53)$$
$$R_{21} \le I(X_2;Y_1|X_1) \quad (54)$$
$$R_1+R_{21} \le I(X_1X_2;Y_2) \quad (55)$$

for some input distributions $p(x_1,x_2)=p(x_1)p(x_2)$.

Remark 7: We obtain the broadcast channel if Y_1 is independent of the input X_1. If Y_2 is a degraded version of Y_1, we obtain the degraded broadcast channel. By setting $S_{22}=T_{21}=S_1=R_1=0$, $R_{21}=S_{21}$, $R_{22}=T_{22}$, $V_2=Q=\phi$, and $V_1=X_2$, we obtain the capacity region of the degraded broadcast channel, which is the closure of the convex hull of all rate pairs (R_{21},R_{22}) satisfying

$$R_{21} \le I(X_2;Y_1|U) \quad (56)$$
$$R_{22} \le I(U;Y_2) \quad (57)$$

for some input distributions $p(u,x_2)=p(u)p(x_2|u)$.

Remark 8: We obtain the ZIC when $R_{21}=0$. By setting $S_{21}=T_{21}=0$, $V_1=\phi$, and $V_2=X_2$, we obtain the Han–Kobayashi rate region (the best rate region to date) for the ZIC which is the closure of all rate pairs (R_1,R_{22}) satisfying

$$R_1 \le S_1 \quad (58)$$
$$R_{22} \le S_{22}+T_{22} \quad (59)$$

where S_1, S_{22}, and T_{22} are subject to the following constraints:

$$S_1+T_{22} \le I(X_1U;Y_1|Q) \quad (60)$$
$$T_{22} \le I(U;Y_1|X_1Q) \quad (61)$$
$$S_1 \le I(X_1;Y_1|UQ) \quad (62)$$
$$S_{22}+T_{22} \le I(X_2;Y_2|Q) \quad (63)$$
$$S_{22} \le I(X_2;Y_2|UQ) \quad (64)$$

for some input probability distributions of the following form: $p(q,u,x_1,x_2)=p(q)p(x_1|q)p(u|q)p(x_2|u,q)$. By using Fourier–Motzkin elimination, we can reduce this to a set of bounds containing only R_1 and R_{22}. (Refer to [10], [11].)

V. Rate Regions for the Degraded Discrete Memoryless ZC of Type I

As we have mentioned in Section II, the capacity region of a Gaussian ZC with weak crossover link gain is equivalent to that of a degraded Gaussian ZC of type I. We shall first determine an achievable rate region for the degraded discrete memoryless ZC of type I. We note that receiver RX_2 is able to decode all the information meant for receiver RX_1. Hence, we may set $S_{21}=0$. We are then able to establish the following lemma.

Lemma 1: An achievable rate region for sending information over the degraded discrete memoryless ZC of type I $(\mathcal{X}_1\times\mathcal{X}_2, p(y_1,y_2|x_1x_2), \mathcal{Y}_1\times\mathcal{Y}_2)$ is given by the set $\mathcal{R}_1$, which is the closure of all rate triplets (R_1,R_{21},R_{22}) satisfying

$$R_1 \le S_1 \quad (65)$$
$$R_{21} \le T_{21} \quad (66)$$
$$R_{22} \le S_{22}+T_{22} \quad (67)$$

where S_1, T_{21}, S_{22}, and T_{22} are subject to the following constraints:

$$S_1+T_{21}+T_{22} \le I(UX_1;Y_1|Q) \quad (68)$$
$$T_{21}+T_{22} \le I(U;Y_1|X_1Q) \quad (69)$$
$$S_1 \le I(X_1;Y_1|UQ) \quad (70)$$
$$S_{22} \le I(X_2;Y_2|UQ) \quad (71)$$

for all input probability distributions of the following form: $p(q,u,x_1,x_2)=p(q)p(x_1|q)p(u|q)p(x_2|u,q)$. Furthermore, the region is unchanged if we impose the following constraints on the cardinalities of the auxiliary sets:

$$\|\mathcal{U}\| \le \|\mathcal{X}_2\|+2 \text{ and } \|\mathcal{Q}\| \le 4. \quad (72)$$

Proof: Set $S_{21}=0$, $V_1=U$, and $V_2=X_2$ in Theorem 2. We note that for a degraded discrete memoryless ZC of type I, $I(U;Y_1|X_1) \le I(U;Y_2)$ for all input distributions $p(x_1)p(u)p(x_2|u)$. This implies that

$$I(U;Y_1|X_1Q)+I(X_2;Y_2|UQ) \le I(U;Y_2|Q)+I(X_2;Y_2|UQ)=I(X_2;Y_2|Q). \quad (73)$$

Hence, the following constraint:

$$S_{22}+T_{21}+T_{22} \le I(X_2;Y_2|Q) \quad (74)$$

is redundant for a degraded discrete memoryless ZC of type I. The assertions about the cardinalities of $\mathcal{U}$ and $\mathcal{Q}$ follow directly from the application of Caratheodory's theorem to the expressions (68)–(71). □

Remark 9: By observing that $I(U;Y_1) \le I(U;Y_1|X_1)$, we readily see that the achievable rate region of Lemma 1 will always include the achievable rate region determined by Liu and Ulukus, i.e., $\mathcal{R}_{\mathrm{LU}} \subseteq \mathcal{R}_1$.

A. Outer Bound to the Capacity Region of the Degraded Discrete Memoryless ZC of Type I

The following is an outer bound to the capacity region of the degraded discrete memoryless ZC of type I.

Theorem 3: The set of rate triplets (R_1, R_{21}, R_{22}) satisfying

$$R_{21} \le I(U; Y_1|X_1Q) \tag{75}$$
$$R_{22} \le I(X_2; Y_2|UQ) \tag{76}$$
$$R_1 + R_{21} \le I(UX_1; Y_1|Q) \tag{77}$$

for some input probability distributions of the following form $p(q,u,x_1,x_2) = p(q)\,p(x_1|q)\,p(u|q)\,p(x_2|u,q)$ constitutes an outer bound to the capacity region of the degraded discrete memoryless ZC of type I. Furthermore, the region is unchanged if we impose the following constraints on the cardinalities of the auxiliary sets:

$$\|\mathcal{U}\| \le \|\mathcal{X}_2\| + 1 \text{ and } \|\mathcal{Q}\| \le 3. \tag{78}$$

Proof: Refer to Appendix II. □

B. Achievable Rate Region for the Gaussian ZC With Weak Crossover Link Gain ($0 < a^2 < 1$)

We have already established an achievable rate region for the degraded discrete memoryless ZC of type I. Lemma 1 can then be readily extended to a Gaussian ZC with weak crossover link gain.

Corollary 1: For $0 < a^2 < 1$, an achievable rate region for the Gaussian ZC is given by the set $\mathcal{R}_2$, which is the closure of the convex hull of all rate triplets (R_1, R_{21}, R_{22}) satisfying

$$R_1 \le S_1 \tag{79}$$
$$R_{21} \le T_{21} \tag{80}$$
$$R_{22} \le S_{22} + T_{22} \tag{81}$$

where S_1, T_{21}, S_{22}, and T_{22} are subject to the constraints

$$S_1 + T_{21} + T_{22} \le \gamma\left(\frac{a^2\beta P_2 + P_1}{a^2(1-\beta)P_2 + 1}\right) \tag{82}$$
$$T_{21} + T_{22} \le \gamma\left(\frac{a^2\beta P_2}{a^2(1-\beta)P_2 + 1}\right) \tag{83}$$
$$S_1 \le \gamma\left(\frac{P_1}{a^2(1-\beta)P_2 + 1}\right) \tag{84}$$
$$S_{22} \le \gamma((1-\beta)P_2) \tag{85}$$

for any $0 \le \beta \le 1$.

Proof: The proof follows directly from Lemma 1 with $\|\mathcal{Q}\| = 1$, $X_2 = U + W$ where U, W, and X_1 are independent Gaussian random variables, and $\beta = \frac{\mathbb{E}(U^2)}{\mathbb{E}(X_2^2)}$. □

VI. Rate Regions for the Degraded Discrete Memoryless ZC of Type II

As we have mentioned in Section II, the capacity region of a Gaussian ZC with strong crossover link gain is equivalent to that of a degraded Gaussian ZC of type II. Hence, we shall first determine an achievable rate region for the degraded discrete memoryless ZC of type II. In addition, the achievable rate region in Lemma 2 is also applicable to the degraded discrete memoryless ZC of type III.

Lemma 2: An achievable rate region for sending information over the degraded discrete memoryless ZC of type II and type III $(\mathcal{X}_1 \times \mathcal{X}_2, p(y_1, y_2|x_1x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$ is given by the set $\mathcal{R}_3$, which is the closure of all triplets (R_1, R_{21}, R_{22}) satisfying

$$R_{21} \le I(X_2; Y_1|UX_1Q) \tag{86}$$
$$R_{22} \le I(U; Y_2|Q) \tag{87}$$
$$R_1 \le I(X_1; Y_1|X_2Q) \tag{88}$$
$$R_1 + R_{21} \le I(X_1X_2; Y_1|UQ) \tag{89}$$
$$R_1 + R_{21} + R_{22} \le I(X_1X_2; Y_1|Q) \tag{90}$$

for all input probability distributions of the following form: $p(q,u,x_1,x_2) = p(q)\,p(x_1|q)\,p(u|q)\,p(x_2|u,q)$. Furthermore, the region is unchanged if we impose the following constraints on the cardinalities of the auxiliary sets:

$$\|\mathcal{U}\| \le \|\mathcal{X}_2\| + 2 \text{ and } \|\mathcal{Q}\| \le 5. \tag{91}$$

Proof: Set $T_{21} = S_{22} = 0$, $R_1 = S_1$, $R_{21} = S_{21}$, $R_{22} = T_{22}$, $V_2 = U$, and $V_1 = X_2$ in Theorem 2. Since

$$I(U; Y_2|Q) + I(X_2; Y_1|UX_1Q) \le I(X_2; Y_1|X_1Q) \tag{92}$$

for a degraded discrete memoryless ZC of type II and type III, the constraint

$$R_{21} + R_{22} \le I(X_2; Y_1|X_1Q) \tag{93}$$

is redundant. The assertions about the cardinalities of $\mathcal{U}$ and $\mathcal{Q}$ follow directly from the application of Caratheodory's theorem to the expressions (86)–(90). □

A. Outer Bound to the Capacity Region of the Degraded Discrete Memoryless ZC of Type II and Type III

The following is an outer bound to the capacity region of the degraded discrete memoryless ZC of type II and type III.

Theorem 4: The set of rate triplets (R_1, R_{21}, R_{22}) satisfying

$$R_{21} \le I(X_2; Y_1|UX_1Q) \tag{94}$$
$$R_{22} \le I(U; Y_2|Q) \tag{95}$$
$$R_1 \le I(X_1; Y_1|X_2Q) \tag{96}$$
$$R_1 + R_{21} + R_{22} \le I(X_1X_2; Y_1|Q) \tag{97}$$

for some input probability distributions of the following form: $p(q,u,x_1,x_2) = p(q)\,p(x_1|q)\,p(u|q)\,p(x_2|u,q)$ constitutes an outer bound to the capacity region of the degraded discrete memoryless ZC of type II and type III. Furthermore, the region is unchanged if we impose the following constraints on the cardinalities of the auxiliary sets:

$$\|\mathcal{U}\| \le \|\mathcal{X}_2\| + 1 \text{ and } \|\mathcal{Q}\| \le 4. \tag{98}$$

Proof: Refer to Appendix III. □

We note that the outer bound of Theorem 4 has one less constraint than the achievable rate region of Lemma 2. A natural question is under what conditions do the inner bound and outer bound meet. This is given in the following theorem.

Theorem 5: The capacity region of the class of discrete memoryless ZC of type II, with the condition that $I(U;Y_1) \leq I(U;Y_2)$ for all input distributions of the form $p(u,x_1,x_2) = p(x_1)p(u,x_2)$, is the set $\mathcal{R}_3'$, which is the closure of the set of rate triplets satisfying (94)–(97) for some input probability distributions of the following form: $p(q,u,x_1,x_2) = p(q)p(x_1|q)p(u|q)p(x_2|u,q)$. Furthermore, the region is unchanged if we impose the same constraints on the cardinalities of the auxiliary sets as (98).

Proof: Let us first assume that a certain rate triplet (R_1', R_{21}', R_{22}') satisfies (94)–(97) for a fixed input distribution

$$p_1(q,u,x_1,x_2) = p_1(q)p_1(x_1|q)p_1(u|q)p_1(x_2|u,q). \tag{99}$$

Let the joint distribution of the set of random variables $\left(Q^1U^1X_1^1X_2^1\right)$ be given by (99). Let the joint distribution of the set of random variables $\left(Q^2U^2X_1^2X_2^2\right)$, where $U^2 = \phi$, be given by

$$\begin{aligned} p_2(q,x_1,x_2) &= \sum_{u\in\mathcal{U}} p_1(q,u,x_1,x_2) \\ &= p_1(q)p_1(x_1|q)p_1(x_2|q). \end{aligned} \tag{100}$$

Now, let the random variable I range over $\{1, 2\}$, where $0 \leq \Pr(I=1) = \alpha \leq 1$ and $\Pr(I=2) = 1-\alpha$. Furthermore, we define $X_1 \triangleq X_1^I$, $X_2 \triangleq X_2^I$, $U \triangleq U^I$, and $Q \triangleq \left(Q^I, I\right)$. Next, we need to set an appropriate value for α. If $I\left(U^1;Y_2^1|Q^1\right) = 0$, set $\alpha = 0$. Otherwise, we set α as follows:

$$\alpha = \frac{R_{22}'}{I(U^1;Y_2^1|Q^1)}. \tag{101}$$

We note that R_{21}' satisfies

$$\begin{aligned} R_{21}' &\leq I\left(X_2^1;Y_1^1|U^1X_1^1Q^1\right) \\ &= \alpha I\left(X_2^1;Y_1^1|U^1X_1^1Q^1\right) + (1-\alpha)I\left(X_2^1;Y_1^1|U^1X_1^1Q^1\right) \\ &\leq \alpha I\left(X_2^1;Y_1^1|U^1X_1^1Q^1\right) + (1-\alpha)I\left(X_2^1;Y_1^1|X_1^1Q^1\right) \\ &= \alpha I\left(X_2^1;Y_1^1|U^1X_1^1Q^1\right) + (1-\alpha)I\left(X_2^2;Y_1^2|X_1^2Q^2\right) \\ &= I(X_2;Y_1|UX_1Q). \end{aligned} \tag{102}$$

We also note that $R_1' + R_{21}'$ satisfies

$$\begin{aligned} R_1' + R_{21}' &\leq I\left(X_1^1X_2^1;Y_1^1|Q^1\right) - R_{22}' \\ &= I(X_1X_2;Y_1|Q) - R_{22}' \\ &= I(X_1X_2;Y_1|Q) - \alpha I\left(U^1;Y_2^1|Q^1\right) \\ &= I(X_1X_2;Y_1|Q) - \alpha I\left(U^1;Y_2^1|Q^1\right) \\ &\quad - (1-\alpha)I\left(U^2;Y_2^2|Q^2\right) \\ &= I(X_1X_2;Y_1|Q) - I(U;Y_2|Q) \\ &\leq I(X_1X_2;Y_1|Q) - I(U;Y_1|Q) \\ &= I(X_1X_2;Y_1|UQ) \end{aligned} \tag{103}$$

if $I(U;Y_1) \leq I(U;Y_2)$ for all input probability distributions $p(x_1)p(u)p(x_2|u)$. We see that the same rate triplet (R_1', R_{21}', R_{22}') satisfies

$$R_{21}' \leq I(X_2;Y_1|UX_1Q) \tag{104}$$
$$R_{22}' \leq I(U;Y_2|Q) \tag{105}$$
$$R_1' \leq I(X_1;Y_1|X_2Q) \tag{106}$$
$$R_1' + R_{21}' \leq I(X_1X_2;Y_1|UQ) \tag{107}$$
$$R_1' + R_{21}' + R_{22}' \leq I(X_1X_2;Y_1|Q). \tag{108}$$

Hence, all rate triplets in the set $\mathcal{R}_3'$ are achievable. □

The region $\mathcal{R}_3'$ is in fact also the capacity region of a certain class of degraded discrete memoryless ZC of type I.

Theorem 6: $\mathcal{R}_3'$ is the capacity region of the class of degraded discrete memoryless ZC of type I with Y_2 being a deterministic function of X_1 and Y_1, i.e., $Y_2 = f(X_1, Y_1)$.

Proof: Since $Y_2 = f(X_1,Y_1)$, we note that $X_2 \to (X_1,Y_1) \to Y_2$ form a Markov chain. In fact, this special class of ZC is a degraded discrete memoryless ZC of both type I and type II. It is easy to verify that for the degraded discrete memoryless ZC of type I, $I(U;Y_1) \leq I(U;Y_2)$ for all input distributions $p(x_1)p(u)p(x_2|u)$. □

B. Achievable Rate Region for the Gaussian ZC With Strong Crossover Link Gain ($a^2 \geq 1$)

So far, we have established an achievable rate region for the degraded discrete memoryless ZC of type II and type III. Since the capacity region of the Gaussian ZC with strong crossover link gain corresponds to that of a degraded Gaussian ZC of type II, we see that Lemma 2 is readily applicable with obvious modifications.

Corollary 2: For $a^2 \geq 1$, an achievable rate region for the Gaussian ZC is given by the set $\mathcal{R}_4$, which is the closure of the convex hull of all rate triplets (R_1, R_{21}, R_{22}) satisfying

$$R_{21} \leq \gamma\left(a^2\beta P_2\right) \tag{109}$$
$$R_{22} \leq \gamma\left(\frac{(1-\beta)P_2}{1+\beta P_2}\right) \tag{110}$$
$$R_1 \leq \gamma(P_1) \tag{111}$$
$$R_1 + R_{21} \leq \gamma\left(a^2\beta P_2 + P_1\right) \tag{112}$$
$$R_1 + R_{21} + R_{22} \leq \gamma\left(a^2 P_2 + P_1\right) \tag{113}$$

for any $0 \leq \beta \leq 1$.

Proof: The proof follows directly from Lemma 2 with $\|\mathcal{Q}\| = 1$. We also assume that $X_2 = U + W$ where U, W, and X_1 are independent, zero-mean, Gaussian random variables and where $\beta = \frac{\mathbb{E}(W^2)}{\mathbb{E}(X_2^2)}$. □

Remark 10: Corollary 2 was derived in [2] for the Gaussian ZC with very strong crossover link gain. We note that the last constraint (113) is redundant for the Gaussian ZC with very strong crossover link gain.

C. Outer Bound to the Capacity Region of the Gaussian ZC With Strong Crossover Link Gain ($a^2 \geq 1$)

In the previous section, we derived an achievable rate region for the Gaussian ZC with strong crossover link gain. Next, we proceed to establish an outer bound to the capacity region of the Gaussian ZC with strong crossover link gain. We make use of the equivalent channel shown in Fig. 9 and Shannon's entropy power inequality to derive an outer bound.

Theorem 7: For a Gaussian ZC with power constraints P_1 and P_2, and $a^2 \geq 1$, any achievable rate triplet (R_1, R_{21}, R_{22}) has to satisfy

$$R_{21} \leq \gamma\left(a^2\beta P_2\right) \tag{114}$$

$$R_{22} \leq \gamma\left(\frac{(1-\beta)P_2}{1+\beta P_2}\right) \tag{115}$$

$$R_1 \leq \gamma(P_1) \tag{116}$$

$$R_1 + R_{21} + R_{22} \leq \gamma\left(a^2 P_2 + P_1\right). \tag{117}$$

for some $0 \leq \beta \leq 1$.

Proof: Refer to Appendix IV. □

D. Capacity of the Gaussian ZC With Moderately Strong Crossover Link Gain ($1 \leq a^2 \leq P_1 + 1$)

We have derived an achievable rate region and an outer bound for the Gaussian ZC when $a^2 \geq 1$. In this subsection, we show that the achievable rate region coincides with the outer bound when the crossover link gain is moderately strong, i.e, when $1 \leq a^2 \leq 1 + P_1$.

Theorem 8: The capacity region of the Gaussian ZC with moderately strong crossover link gain is given by the closure of all rate triplets (R_1, R_{21}, R_{22}) satisfying (114)–(117) for some $0 \leq \beta \leq 1$.

Proof: Let us first assume a particular rate triplet (R'_1, R'_{21}, R'_{22}) satisfies (114)–(117) for $\beta = \beta_0$. Next, let us set β_1 as follows:

$$\beta_1 = \frac{\frac{1+P_2}{2^{2R'_{22}}} - 1}{P_2}$$

$$\Longrightarrow R'_{22} = \gamma\left(\frac{(1-\beta_1)P_2}{1+\beta_1 P_2}\right). \tag{118}$$

We note that $\beta_0 \leq \beta_1 \leq 1$ since

$$R'_{22} \leq \gamma\left(\frac{(1-\beta_0)P_2}{1+\beta_0 P_2}\right).$$

Let us consider the last constraint given by (117). We obtain

$$\begin{aligned}
&R'_1 + R'_{21}\\
&\leq \frac{1}{2}\log_2\left(1 + a^2P_2 + P_1\right) - \frac{1}{2}\log_2\left(\frac{1+P_2}{1+\beta_1P_2}\right)\\
&= \frac{1}{2}\log_2\left(\frac{1 + a^2P_2 + P_1 + \beta_1P_2 + a^2\beta_1{P_2}^2 + \beta_1P_1P_2}{1+P_2}\right)\\
&= \gamma\left(a^2\beta_1P_2 + P_1 + \frac{(1-\beta_1)P_2\left(a^2-1-P_1\right)}{1+P_2}\right)\\
&\leq \gamma\left(a^2\beta_1P_2 + P_1\right), \ \ a^2 \leq 1 + P_1.
\end{aligned} \tag{119}$$

We see that the same rate triplet (R'_1, R'_{21}, R'_{22}) also satisfies (109)–(113) for $\beta = \beta_1$.

$$R'_{21} \leq \gamma\left(a^2\beta_0P_2\right) \leq \gamma\left(a^2\beta_1P_2\right), \ \ \beta_1 \geq \beta_0 \tag{120}$$

$$R'_{22} = \gamma\left(\frac{(1-\beta_1)P_2}{1+\beta_1P_2}\right) \tag{121}$$

$$R'_1 \leq \gamma(P_1) \tag{122}$$

$$R'_1 + R'_{21} \leq \gamma\left(a^2\beta_1P_2 + P_1\right) \tag{123}$$

$$R'_1 + R'_{21} + R'_{22} \leq \gamma\left(a^2P_2 + P_1\right). \tag{124}$$

Hence, any rate triplet in the outer bound is achievable. □

E. Achievable Rates and Numerical Computation For the Gaussian ZC With Very Strong Crossover Link Gain ($a^2 \geq P_1 + 1$)

In [2], Vishwanath, Jindal, and Goldsmith determined an achievable rate region for very strong crossover link gain using superposition coding at sender TX_2 and successive decoding at receiver RX_1. In fact, the achievable rate region of Vishwanath, Jindal, and Goldsmith corresponds to that of Corollary 2 with very strong crossover link gain. However, their technique does not apply to the case of moderately strong crossover link gain. This is because their successive decoding method would require receiver RX_1 to be able to decode all the information intended for receiver RX_2. This is possible only with very strong crossover link gain.

We have already determined the capacity of the Gaussian ZC with moderately strong crossover link gain. A very natural question that comes to mind is whether Corollary 2 also gives us the capacity region of the Gaussian ZC with very strong crossover link gain. Our experience with the Gaussian ZIC under very strong interference may influence one to think that Corollary 2 would also give us the capacity region of the Gaussian ZC with very strong crossover link gain. However, in this subsection, we show that this is not the case in general. In fact, this is suggested by the time-sharing random variable Q in the converse proof in [2]. We can enlarge the achievable rate region of Corollary 2 for the Gaussian ZC with very strong crossover link gain by allowing $\|\mathcal{Q}\| > 1$. We could theoretically compute an achievable rate region for larger values of $\|\mathcal{Q}\|$ but for computational reasons, we restrict our attention to $\|\mathcal{Q}\| = 2$.

Corollary 3: For $a^2 \geq 1 + P_1$, an achievable rate region for the Gaussian ZC is given by the set $\mathcal{R}_5$, which is the closure of the convex hull of all (R_1, R_{21}, R_{22}) triplets satisfying

$$R_{21} \leq \lambda.\gamma\left(\frac{a^2\beta\rho P_2}{\lambda}\right) + \overline{\lambda}.\gamma\left(\frac{a^2\sigma\overline{\rho}P_2}{\overline{\lambda}}\right) \tag{125}$$

$$R_{22} \leq \lambda.\gamma\left(\frac{\overline{\beta}\rho P_2}{\lambda + \beta\rho P_2}\right) + \overline{\lambda}.\gamma\left(\frac{\overline{\sigma}\overline{\rho}P_2}{\overline{\lambda} + \sigma\overline{\rho}P_2}\right) \tag{126}$$

$$R_1 \leq \lambda.\gamma\left(\frac{\alpha P_1}{\lambda}\right) + \overline{\lambda}.\gamma\left(\frac{\overline{\alpha}P_1}{\overline{\lambda}}\right) \tag{127}$$

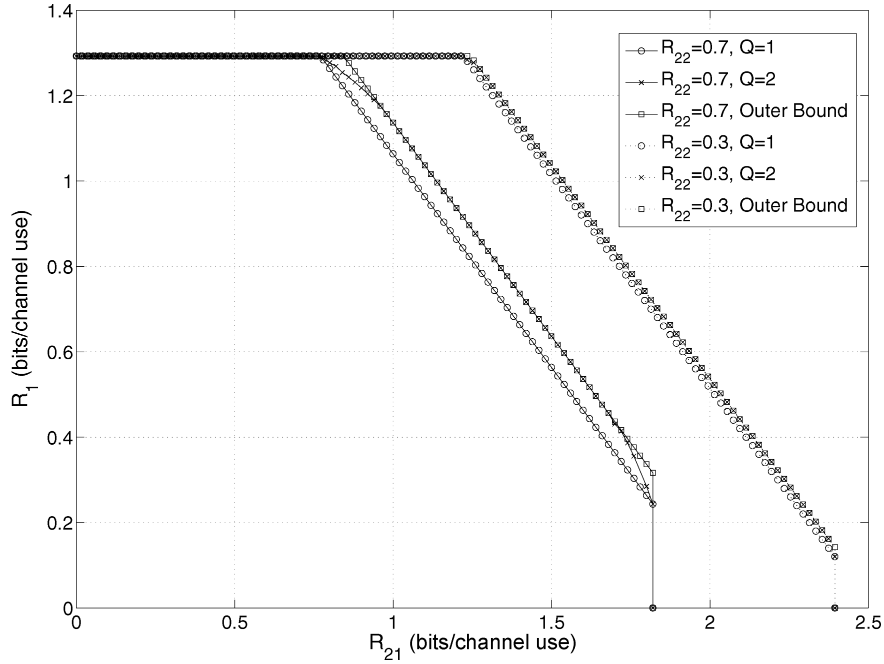

Fig. 11. Numerical computations ($P_1 = 5$, $P_2 = 5$, $a^2 = 9$, $R_{22} = 0.3/0.7$).

$$R_1 + R_{21} \le \lambda.\gamma\left(\frac{a^2\beta\rho P_2 + \alpha P_1}{\lambda}\right) + \overline{\lambda}.\gamma\left(\frac{a^2\sigma\overline{\rho}P_2 + \overline{\alpha}P_1}{\overline{\lambda}}\right) \tag{128}$$

$$R_1 + R_{21} + R_{22} \le \lambda.\gamma\left(\frac{a^2\rho P_2 + \alpha P_1}{\lambda}\right) + \overline{\lambda}.\gamma\left(\frac{a^2\overline{\rho}P_2 + \overline{\alpha}P_1}{\overline{\lambda}}\right). \tag{129}$$

for any $0 \le \lambda, \alpha, \beta, \rho, \sigma \le 1$.

Proof: The result follows directly from Lemma 2 with $\|\mathcal{Q}\| = 2$. We assume that $X_2^N = U^N + W^N$ where U^N, W^N, and X_1^N are independent. During a fraction λ of the time, the symbols of X_1^N, U^N, and W^N are Gaussian distributed with zero mean, and variances $\frac{\alpha P_1}{\lambda}$, $\frac{\overline{\beta}\rho P_2}{\lambda}$, and $\frac{\beta\rho P_2}{\lambda}$, respectively

$$X_{1n} \sim \mathcal{N}\left(0, \frac{\alpha P_1}{\lambda}\right), \; U_n \sim \mathcal{N}\left(0, \frac{\overline{\beta}\rho P_2}{\lambda}\right), \quad W_n \sim \mathcal{N}\left(0, \frac{\beta\rho P_2}{\lambda}\right), \quad 0 \le \alpha, \beta, \rho \le 1, \; n = 1, 2, \ldots, N\lambda \tag{130}$$

and during the remaining fraction $\overline{\lambda} \triangleq 1 - \lambda$ of the time

$$X_{1n} \sim \mathcal{N}\left(0, \frac{\overline{\alpha}P_1}{\overline{\lambda}}\right), \; U_n \sim \mathcal{N}\left(0, \frac{\overline{\sigma}\overline{\rho}P_2}{\overline{\lambda}}\right), \quad W_n \sim \mathcal{N}\left(0, \frac{\sigma\overline{\rho}P_2}{\overline{\lambda}}\right), \quad 0 \le \sigma \le 1, \; n = N\lambda + 1, \ldots, N \tag{131}$$

which ensures that the power constraints are satisfied. □

Remark 11: Fig. 11 shows numerical computations of the achievable rates for the Gaussian ZC with $P_1 = 5, P_2 = 5, a^2 = 9\,(a^2 > 1 + P_1)$. Instead of plotting rate triplets (R_1, R_{21}, R_{22}), we fix $R_{22} = \{0.3, 0.7\}$ and plot the rate pair (R_1, R_{21}). From Fig. 11, we see that when R_{22} is fixed, Corollary 2 gives rate pairs (R_1, R_{21}) that correspond to a Gaussian multiple-access channel. However, we see that when we increase $\|\mathcal{Q}\|$ from 1 to 2, Corollary 3 gives an achievable rate region that is even larger than that of Corollary 2 for the Gaussian ZC with very strong crossover link gain. Moreover, we note that for the parameters chosen, setting $\|\mathcal{Q}\| = 2$ suffices to achieve the capacity for most rate triplets. In general, Corollary 2 is not the capacity region of the Gaussian ZC with very strong crossover link gain.

Remark 12: However, Corollary 2 gives us the capacity region of the Gaussian ZIC under strong interference. We can ignore the constraint for R_{21} since $R_{21} = 0$ for an interference channel. By setting $\beta = 0$, we obtain the capacity region of the Gaussian ZIC [9] under strong interference.

VII. Concluding Remarks

In this paper, we derive an achievable rate region for the general discrete memoryless ZC. We make use of rate-splitting and superposition coding at sender TX_2 and joint decoding at receiver RX_1. We specialize this general result to obtain achievable rate regions for degraded discrete memoryless ZCs of type I, type II, and type III. We also obtain outer bounds to the capacity regions of these three types of degraded discrete memoryless ZCs. We show that as long as a certain condition is satisfied, the achievable rate region is the capacity region of the degraded discrete memoryless ZC of type II. The results for the degraded discrete memoryless ZCs of type I and II are then extended to the Gaussian ZC with different crossover link gains. We determine an outer bound to the capacity region of the Gaussian ZC with strong crossover link gain. For moderately strong crossover link gain, we show that the inner and outer bounds for the capacity region meet.

APPENDIX I
PROOF OF THEOREM 2

By the symmetry of the random code generation, the conditional probability of error does not depend on which indices are sent. Therefore, we may assume that the message

$$(w_1,(w_{211},w_{212}),(w_{221},w_{222}))=(1,(1,1),(1,1))$$

is sent. Let $P(\cdot)$ denote the conditional probability of the event that $(1,(1,1),(1,1))$ is sent. For receiver RX_1, we define the following events:

$$E_{ijkm}=\left\{\left(q^N,x_1^N(i),u^N(k,m),v_1^N(j,k,m),y_1^N\right)\in A_\epsilon^{(N)}(Q,X_1,U,V_1,Y_1)\right\}. \quad (132)$$

Then we can bound the probability of error as follows:

$$\begin{aligned}P_e^{(N)}(1)&=P\left(E_{1111}^c\bigcup\cup_{(i,j,k,m)\neq(1,1,1,1)}E_{ijkm}\right)\\&\le P(E_{1111}^c)+\sum_{\substack{i\neq1,j\neq1,\\(k,m)\neq(1,1)}}P(E_{ijkm})\\&\quad+\sum_{\substack{i=1,j\neq1,\\(k,m)\neq(1,1)}}P(E_{1jkm})+\sum_{\substack{i\neq1,j=1,\\(k,m)\neq(1,1)}}P(E_{i1km})\\&\quad+\sum_{\substack{i\neq1,j\neq1,\\(k,m)=(1,1)}}P(E_{ij11})+\sum_{\substack{i\neq1,j=1,\\(k,m)=(1,1)}}P(E_{i111})\\&\quad+\sum_{\substack{i=1,j\neq1,\\(k,m)=(1,1)}}P(E_{1j11})+\sum_{\substack{i=1,j=1,\\(k,m)\neq(1,1)}}P(E_{11km}).\end{aligned} \quad (133)$$

For $(i,j,(k,m))\neq(1,1,(1,1))$, we have

$$\begin{aligned}&P(E_{ijkm})\\&=P\left(\left(q^N,x_1^N(i),u^N(k,m),v_1^N(j,k,m),y_1^N\right)\in A_\epsilon^{(N)}\right)\\&=\sum_{(q^N,x_1^N,u^N,v_1^N,y_1^N)\in A_\epsilon^{(N)}}p(q^N)\,p(x_1^N,u^N,v_1^N|q^N)\\&\quad\cdot p(y_1^N|q^N)\\&\le\left\|A_\epsilon^{(N)}\right\|2^{-N(H(Q)-\epsilon+H(X_1UV_1|Q)-2\epsilon+H(Y_1|Q)-2\epsilon)}\\&\le2^{-N(H(Q)+H(X_1UV_1|Q)+H(Y_1|Q)-H(QX_1UV_1Y_1)-6\epsilon)}\\&=2^{-N(I(X_1UV_1;Y_1|Q)-6\epsilon)}.\end{aligned} \quad (134)$$

For $(j,(k,m))\neq(1,(1,1))$, we have

$$\begin{aligned}&P(E_{1jkm})\\&=P\left(\left(q^N,x_1^N(1),u^N(k,m),v_1^N(j,k,m),y_1^N\right)\in A_\epsilon^{(N)}\right)\\&=\sum_{(q^N,x_1^N,u^N,v_1^N,y_1^N)\in A_\epsilon^{(N)}}p(q^N,x_1^N)\,p(u^N,v_1^N|q^N)\\&\quad\cdot p(y_1^N|x_1^N,q^N)\\&\le\left\|A_\epsilon^{(N)}\right\|2^{-N(H(QX_1)-\epsilon+H(UV_1|Q)-2\epsilon+H(Y_1|X_1Q)-2\epsilon)}\\&\le2^{-N(H(QX_1)+H(UV_1|Q)+H(Y_1|X_1Q)-H(QX_1UV_1Y_1)-6\epsilon)}\\&=2^{-N(I(UV_1;Y_1|X_1Q)-6\epsilon)}.\end{aligned} \quad (135)$$

For $(i,(k,m))\neq(1,(1,1))$, we have

$$\begin{aligned}&P(E_{i1km})\\&=P\left(\left(q^N,x_1^N(i),u^N(k,m),v_1^N(1,k,m),y_1^N\right)\in A_\epsilon^{(N)}\right)\\&=\sum_{(q^N,x_1^N,u^N,v_1^N,y_1^N)\in A_\epsilon^{(N)}}p(q^N)\,p(x_1^N,u^N,v_1^N|q^N)\\&\quad\cdot p(y_1^N|q^N)\\&\le\left\|A_\epsilon^{(N)}\right\|2^{-N(H(Q)-\epsilon+H(X_1UV_1|Q)-2\epsilon+H(Y_1|Q)-2\epsilon)}\\&\le2^{-N(H(Q)+H(X_1UV_1|Q)+H(Y_1|Q)-H(QX_1UV_1Y_1)-6\epsilon)}\\&=2^{-N(I(X_1UV_1;Y_1|Q)-6\epsilon)}.\end{aligned} \quad (136)$$

For $(i,j)\neq(1,1)$, we have

$$\begin{aligned}&P(E_{ij11})\\&=P\left(\left(q^N,x_1^N(i),u^N(1,1),v_1^N(j,1,1),y_1^N\right)\in A_\epsilon^{(N)}\right)\\&=\sum_{(q^N,x_1^N,u^N,v_1^N,y_1^N)\in A_\epsilon^{(N)}}p(q^N,u^N)\,p(x_1^N,v_1^N|u^N,q^N)\\&\quad\cdot p(y_1^N|u^N,q^N)\\&\le\left\|A_\epsilon^{(N)}\right\|2^{-N(H(QU)-\epsilon+H(X_1V_1|UQ)-2\epsilon+H(Y_1|UQ)-2\epsilon)}\\&\le2^{-N(H(QU)+H(X_1V_1|UQ)+H(Y_1|UQ)-H(QX_1UV_1Y_1)-6\epsilon)}\\&=2^{-N(I(X_1V_1;Y_1|UQ)-6\epsilon)}.\end{aligned} \quad (137)$$

For $i\neq1$, we have

$$\begin{aligned}&P(E_{i111})\\&=P\left(\left(q^N,x_1^N(i),u^N(1,1),v_1^N(1,1,1),y_1^N\right)\in A_\epsilon^{(N)}\right)\\&=\sum_{(q^N,x_1^N,u^N,v_1^N,y_1^N)\in A_\epsilon^{(N)}}p(q^N,u^N,v_1^N)\\&\quad\cdot p(x_1^N|q^N)\,p(y_1^N|u^N,v_1^N,q^N)\\&\le\left\|A_\epsilon^{(N)}\right\|2^{-N(H(QUV_1)-\epsilon+H(X_1|Q)-2\epsilon+H(Y_1|UV_1Q)-2\epsilon)}\\&\le2^{-N(H(QUV_1)+H(X_1|Q)+H(Y_1|UV_1Q)-H(QX_1UV_1Y_1)-6\epsilon)}\\&=2^{-N(I(X_1;Y_1|UV_1Q)-6\epsilon)}.\end{aligned} \quad (138)$$

For $j\neq1$, we have

$$\begin{aligned}&P(E_{1j11})\\&=P\left(\left(q^N,x_1^N(1),u^N(1,1),v_1^N(j,1,1),y_1^N\right)\in A_\epsilon^{(N)}\right)\\&=\sum_{(q^N,x_1^N,u^N,v_1^N,y_1^N)\in A_\epsilon^{(N)}}p(q^N,u^N,x_1^N)\,p(v_1^N|u^N,q^N)\\&\quad\cdot p(y_1^N|x_1^N,u^N,q^N)\\&\le\left\|A_\epsilon^{(N)}\right\|2^{-N(H(QUX_1)-\epsilon+H(V_1|UQ)-2\epsilon+H(Y_1|X_1UQ)-2\epsilon)}\\&\le2^{-N(H(QUX_1)+H(V_1|UQ)+H(Y_1|X_1UQ)-H(QX_1UV_1Y_1)-6\epsilon)}\\&=2^{-N(I(V_1;Y_1|X_1UQ)-6\epsilon)}.\end{aligned} \quad (139)$$

For $(k,m) \neq (1,1)$, we have

$$\begin{aligned}
&P(E_{11km}) \\
&= P\left(\left(q^N, x_1^N(1), u^N(k,m), v_1^N(1,k,m), y_1^N\right) \in A_\epsilon^{(N)}\right) \\
&= \sum_{\left(q^N, x_1^N, u^N, v_1^N, y_1^N\right) \in A_\epsilon^{(N)}} p\left(q^N, x_1^N\right) p\left(u^N, v_1^N | q^N\right) \\
&\quad \cdot p\left(y_1^N | x_1^N, q^N\right) \\
&\leq \left\| A_\epsilon^{(N)} \right\| 2^{-N(H(QX_1)-\epsilon+H(UV_1|Q)-2\epsilon+H(Y_1|X_1Q)-2\epsilon)} \\
&\leq 2^{-N(H(QX_1)+H(UV_1|Q)+H(Y_1|X_1Q)-H(QX_1UV_1Y_1)-6\epsilon)} \\
&= 2^{-N(I(UV_1;Y_1|X_1Q)-6\epsilon)}. \qquad (140)
\end{aligned}$$

We may then bound the probability of error at receiver RX_1 as follows:

$$\begin{aligned}
&P_e^{(N)}(1) \\
&\quad \leq P\left(E_{1111}^c\right) + 2^{N(S_1+S_{21}+T_{21}+T_{22})} 2^{-N(I(X_1UV_1;Y_1|Q)-6\epsilon)} \\
&\qquad + 2^{N(S_{21}+T_{21}+T_{22})} 2^{-N(I(UV_1;Y_1|X_1Q)-6\epsilon)} \\
&\qquad + 2^{N(S_1+T_{21}+T_{22})} 2^{-N(I(X_1UV_1;Y_1|Q)-6\epsilon)} \\
&\qquad + 2^{N(S_1+S_{21})} 2^{-N(I(X_1V_1;Y_1|UQ)-6\epsilon)} \\
&\qquad + 2^{NS_1} 2^{-N(I(X_1;Y_1|UV_1Q)-6\epsilon)} \\
&\qquad + 2^{NS_{21}} 2^{-N(I(V_1;Y_1|X_1UQ)-6\epsilon)} \\
&\qquad + 2^{N(T_{21}+T_{22})} 2^{-N(I(UV_1;Y_1|X_1Q)-6\epsilon)}. \qquad (141)
\end{aligned}$$

For receiver RX_2, we define the following events:

$$E_{lkm} = \left\{\left(q^N, u^N(k,m), v_2^N(l,k,m), y_2^N\right) \in A_\epsilon^{(N)}(Q,U,V_2,Y_2)\right\}. \qquad (142)$$

Then we can bound the probability of error as follows:

$$\begin{aligned}
P_e^{(N)}(2) &= P\left(E_{111}^c \bigcup \cup_{(l,k,m)\neq(1,1,1)} E_{lkm}\right) \\
&\leq P\left(E_{111}^c\right) + \sum_{l\neq 1,(k,m)\neq(1,1)} P\left(E_{lkm}\right) \\
&\quad + \sum_{\substack{l=1,\\ (k,m)\neq(1,1)}} P\left(E_{1km}\right) + \sum_{\substack{l\neq 1,\\ (k,m)=(1,1)}} P\left(E_{l11}\right). \qquad (143)
\end{aligned}$$

For $(l,(k,m)) \neq (1,(1,1))$, we have

$$\begin{aligned}
&P(E_{lkm}) \\
&= P\left(\left(q^N, u^N(k,m), v_2^N(l,k,m), y_2^N\right) \in A_\epsilon^{(N)}\right) \\
&= \sum_{\left(q^N, u^N, v_2^N, y_2^N\right) \in A_\epsilon^{(N)}} p\left(q^N\right) p\left(u^N, v_2^N | q^N\right) p\left(y_2^N | q^N\right) \\
&\leq \left\| A_\epsilon^{(N)} \right\| 2^{-N(H(Q)-\epsilon)} 2^{-N(H(UV_2|Q)-2\epsilon)} 2^{-N(H(Y_2|Q)-2\epsilon)} \\
&\leq 2^{-N(H(Q)+H(UV_2|Q)+H(Y_2|Q)-H(QUV_2Y_2)-6\epsilon)} \\
&= 2^{-N(I(UV_2;Y_2|Q)-6\epsilon)}. \qquad (144)
\end{aligned}$$

For $(k,m) \neq (1,1)$, we have

$$\begin{aligned}
&P(E_{1km}) \\
&= P\left(\left(q^N, u^N(k,m), v_2^N(1,k,m), y_2^N\right) \in A_\epsilon^{(N)}\right) \\
&= \sum_{\left(q^N, u^N, v_2^N, y_2^N\right) \in A_\epsilon^{(N)}} p\left(q^N\right) p\left(u^N, v_2^N | q^N\right) p\left(y_2^N | q^N\right) \\
&\leq \left\| A_\epsilon^{(N)} \right\| 2^{-N(H(Q)-\epsilon)} 2^{-N(H(UV_2|Q)-2\epsilon)} 2^{-N(H(Y_2|Q)-2\epsilon)} \\
&\leq 2^{-N(H(Q)+H(UV_2|Q)+H(Y_2|Q)-H(QUV_2Y_2)-6\epsilon)} \\
&= 2^{-N(I(UV_2;Y_2|Q)-6\epsilon)}. \qquad (145)
\end{aligned}$$

For $l \neq 1$, we have

$$\begin{aligned}
&P(E_{l11}) \\
&= P\left(\left(q^N, u^N(1,1), v_2^N(l,1,1), y_2^N\right) \in A_\epsilon^{(N)}\right) \\
&= \sum_{\left(q^N, u^N, v_2^N, y_2^N\right) \in A_\epsilon^{(N)}} p\left(q^N, u^N\right) p\left(v_2^N | u^N, q^N\right) \qquad (146) \\
&\quad \cdot p\left(y_2^N | u^N, q^N\right) \\
&\leq \left\| A_\epsilon^{(N)} \right\| 2^{-N(H(QU)-\epsilon+H(V_2|UQ)-2\epsilon+H(Y_2|UQ)-2\epsilon)} \\
&\leq 2^{-N(H(QU)+H(V_2|UQ)+H(Y_2|UQ)-H(QUV_2Y_2)-6\epsilon)} \\
&= 2^{-N(I(V_2;Y_2|UQ)-6\epsilon)}. \qquad (147)
\end{aligned}$$

We may then bound the probability of error at receiver RX_2 as follows:

$$\begin{aligned}
P_e^{(N)}(2) &\leq P\left(E_{111}^c\right) + 2^{N(S_{22}+T_{21}+T_{22})} 2^{-N(I(UV_2;Y_2|Q)-6\epsilon)} \\
&\quad + 2^{N(T_{21}+T_{22})} 2^{-N(I(UV_2;Y_2|Q)-6\epsilon)} \\
&\quad + 2^{NS_{22}} 2^{-N(I(V_2;Y_2|UQ)-6\epsilon)}. \qquad (148)
\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, the conditions of Theorem 2 ensure that each of the terms in (141) and (148) tends to 0 as $N \to \infty$.

Appendix II
Proof of Theorem 3

By Fano's inequality, we have

$$H\left(W_{21}|Y_1^N\right) \leq N\epsilon_{1N} \qquad (149)$$
$$H\left(W_{22}|Y_2^N\right) \leq N\epsilon_{2N} \qquad (150)$$
$$H\left(W_1|Y_1^N\right) \leq N\epsilon_{3N} \qquad (151)$$

where $\epsilon_{1N}, \epsilon_{2N}, \epsilon_{3N} \to 0$ as $N \to \infty$. We first bound R_{21} as follows:

$$\begin{aligned}
NR_{21} &= I\left(W_{21};Y_1^N\right) + H\left(W_{21}|Y_1^N\right) \\
&\leq I\left(W_{21};Y_1^N\right) + N\epsilon_{1N} \\
&= H\left(W_{21}\right) - H\left(W_{21}|Y_1^N\right) + N\epsilon_{1N} \\
&\overset{(a)}{=} H\left(W_{21}|X_1^N(W_1)\right) - H\left(W_{21}|Y_1^N\right) + N\epsilon_{1N} \\
&\leq H\left(W_{21}|X_1^N\right) - H\left(W_{21}|X_1^N Y_1^N\right) + N\epsilon_{1N} \\
&= I\left(W_{21};Y_1^N|X_1^N\right) + N\epsilon_{1N} \\
&= \sum_{n=1}^{n=N} I\left(W_{21};Y_{1n}|X_1^N Y_1^{n-1}\right) + N\epsilon_{1N} \\
&= \sum_{n=1}^{n=N} H\left(Y_{1n}|X_1^N Y_1^{n-1}\right) - H\left(Y_{1n}|X_1^N Y_1^{n-1} W_{21}\right) \\
&\quad + N\epsilon_{1N} \\
&\leq \sum_{n=1}^{n=N} H\left(Y_{1n}|X_{1n}\right) - H\left(Y_{1n}|X_1^N Y_1^{n-1} W_{21} Y_2^{n-1}\right) \\
&\quad + N\epsilon_{1N} \\
&\overset{(b)}{=} \sum_{n=1}^{n=N} H\left(Y_{1n}|X_{1n}\right) - H\left(Y_{1n}|X_1^N W_{21} Y_2^{n-1}\right) \\
&\quad + N\epsilon_{1N}
\end{aligned}$$

$$\overset{(c)}{=} \sum_{n=1}^{n=N} H(Y_{1n}|X_{1n}) - H\left(Y_{1n}|X_{1n}W_{21}Y_2^{n-1}\right) + N\epsilon_{1N}$$
$$= \sum_{n=1}^{n=N} H(Y_{1n}|X_{1n}) - H(Y_{1n}|X_{1n}U_n) + N\epsilon_{1N}$$
$$= \sum_{n=1}^{n=N} I(U_n;Y_{1n}|X_{1n}) + N\epsilon_{1N}. \tag{152}$$

where we define the random variable $U_n = \left(W_{21}, Y_2^{n-1}\right)$ for all n, (a) follows from the fact that since W_{21} and W_1 are independent, so are W_{21} and $X_1^N(W_1)$, and (b) follows from the fact that $\left(W_{21}X_{1n}^N Y_{1n}\right) \to \left(X_1^{n-1}Y_2^{n-1}\right) \to Y_1^{n-1}$ form a Markov chain. This is due to the memoryless property of the channel and the fact that for any i, Y_{1i} depends only on Y_{2i} and X_{1i} (refer to (13)). Finally, (c) follows from the fact that $\left(X_1^{n-1}X_{1n+1}^N\right) \to \left(W_{21}Y_2^{n-1}X_{1n}\right) \to Y_{1n}$ form a Markov chain. We can prove this using the functional dependence graph technique introduced in [12]. Alternatively, we first note the following Markov chain

$$\left(X_1^{n-1}X_{1n+1}^N U_n\right) \to (X_{1n}Y_{2n}) \to Y_{1n} \tag{153}$$

which follows from the fact that Y_{1n} depends only on Y_{2n} and X_{1n}. Using the weak union property, we obtain the following Markov chain:

$$\left(X_1^{n-1}X_{1n+1}^N\right) \to (X_{1n}U_nY_{2n}) \to Y_{1n}. \tag{154}$$

Next, we note that X_1^N and Y_2^N are independent. Hence, (U_n, Y_{2n}) is independent of X_1^N. Coupled with the contraction property [5], we obtain the following Markov chain:

$$\left(X_1^{n-1}X_{1n+1}^N\right) \to X_{1n} \to (U_nY_{2n}Y_{1n}). \tag{155}$$

Finally, using the weak union property and the decomposition property [5], we obtain $\left(X_1^{n-1}X_{1n+1}^N\right) \to (U_nX_{1n}) \to Y_{1n}$ as desired. Next, we bound R_{22} as follows:

$$NR_{22} = I\left(W_{22};Y_2^N|W_{21}\right) + H\left(W_{22}|Y_2^N W_{21}\right)$$
$$= I\left(W_{21}W_{22};Y_2^N|W_{21}\right) + H\left(W_{22}|Y_2^N W_{21}\right)$$
$$\le I\left(X_2^N;Y_2^N|W_{21}\right) + N\epsilon_{2N}$$
$$= \sum_{n=1}^{n=N} I\left(X_2^N;Y_{2n}|W_{21}Y_2^{n-1}\right) + N\epsilon_{2N}$$
$$= \sum_{n=1}^{n=N} H\left(Y_{2n}|W_{21}Y_2^{n-1}\right) - H\left(Y_{2n}|W_{21}Y_2^{n-1}X_2^N\right) + N\epsilon_{2N}$$
$$\overset{(a)}{=} \sum_{n=1}^{n=N} H\left(Y_{2n}|W_{21}Y_2^{n-1}\right) - H\left(Y_{2n}|W_{21}Y_2^{n-1}X_{2n}\right) + N\epsilon_{2N}$$
$$= \sum_{n=1}^{n=N} H(Y_{2n}|U_n) - H(Y_{2n}|U_nX_{2n}) + N\epsilon_{2N}$$
$$= \sum_{n=1}^{n=N} I(X_{2n};Y_{2n}|U_n) + N\epsilon_{2N} \tag{156}$$

where (a) follows immediately from the Markov chain given by $\left(X_2^{n-1}X_{2n+1}^N\right) \to (U_nX_{2n}) \to Y_{2n}$. We first note the following Markov chain:

$$\left(X_2^{n-1}X_{2n+1}^N U_n\right) \to (X_{1n}X_{2n}) \to Y_{1n}Y_{2n}. \tag{157}$$

Using the weak union property, we obtain

$$\left(X_2^{n-1}X_{2n+1}^N\right) \to (U_nX_{1n}X_{2n}) \to Y_{1n}Y_{2n}. \tag{158}$$

Using the fact that $U_nX_2^N$ and X_1^N are independent, and applying the contraction property, we obtain

$$\left(X_2^{n-1}X_{2n+1}^N\right) \to (U_nX_{2n}) \to (X_{1n}Y_{1n}Y_{2n}). \tag{159}$$

Applying the decomposition property, we obtain the desired Markov chain $\left(X_2^{n-1}X_{2n+1}^N\right) \to (U_nX_{2n}) \to Y_{2n}$. Finally, we bound $R_{21}+R_1$ as follows:

$$N(R_{21}+R_1)$$
$$= I\left(W_1W_{21};Y_1^N\right) + H\left(W_{21}|Y_1^N\right) + H\left(W_1|Y_1^N W_{21}\right)$$
$$\le I\left(W_{21}X_1^N;Y_1^N\right) + N\epsilon_{1N} + N\epsilon_{3N}$$
$$= \sum_{n=1}^{n=N} I\left(W_{21}X_1^N;Y_{1n}|Y_1^{n-1}\right) + N\epsilon_{1N} + N\epsilon_{3N}$$
$$= \sum_{n=1}^{n=N} H\left(Y_{1n}|Y_1^{n-1}\right) - H\left(Y_{1n}|X_1^N Y_1^{n-1}W_{21}\right) + N\epsilon_{1N} + N\epsilon_{3N}$$
$$\le \sum_{n=1}^{n=N} H(Y_{1n}) - H\left(Y_{1n}|X_1^N Y_1^{n-1}W_{21}Y_2^{n-1}\right) + N\epsilon_{1N} + N\epsilon_{3N}$$
$$= \sum_{n=1}^{n=N} H(Y_{1n}) - H\left(Y_{1n}|X_1^N W_{21}Y_2^{n-1}\right) + N\epsilon_{1N} + N\epsilon_{3N}$$
$$= \sum_{n=1}^{n=N} H(Y_{1n}) - H\left(Y_{1n}|X_{1n}W_{21}Y_2^{n-1}\right) + N\epsilon_{1N} + N\epsilon_{3N}$$
$$= \sum_{n=1}^{n=N} H(Y_{1n}) - H(Y_{1n}|X_{1n}U_n) + N\epsilon_{1N} + N\epsilon_{3N}$$
$$= \sum_{n=1}^{n=N} I(U_nX_{1n};Y_{1n}) + N\epsilon_{1N} + N\epsilon_{3N}. \tag{160}$$

By the Markovity of $U_n \to (X_{1n}X_{2n}) \to (Y_{1n}Y_{2n})$ and the independence of (U_n, X_{2n}) and X_{1n}, we observe that

$$p(u_n, x_{1n}, x_{2n}, y_{1n}, y_{2n}) = p(u_n, x_{2n})\, p(x_{1n})\, p(y_{1n}, y_{2n}|x_{1n}, x_{2n}).$$

By introducing a time-sharing random variable Q similar to the proof for the converse of the capacity region of the multiple-access channel [1, p. 402], we obtain Theorem 3. The assertions about the cardinalities of $\mathcal{U}$ and $\mathcal{Q}$ follow directly from the application of Caratheodory's theorem to the expressions (75)–(77).

APPENDIX III
PROOF OF THEOREM 4

By Fano's inequality, we again have

$$H\left(W_{21}|Y_1^N\right) \le N\epsilon_{1N} \tag{161}$$
$$H\left(W_{22}|Y_2^N\right) \le N\epsilon_{2N} \tag{162}$$
$$H\left(W_1|Y_1^N\right) \le N\epsilon_{3N} \tag{163}$$

where $\epsilon_{1N}, \epsilon_{2N}, \epsilon_{3N} \to 0$ as $N \to \infty$. We first bound R_{21} as follows:

$$\begin{aligned} &NR_{21} \\ &= I\left(W_{21};Y_1^N\right) + H\left(W_{21}|Y_1^N\right) \\ &\le I\left(W_{21};Y_1^N\right) + N\epsilon_{1N} \\ &\overset{(a)}{=} H\left(W_{21}|W_{22}X_1^N\left(W_1\right)\right) - H\left(W_{21}|Y_1^N\right) + N\epsilon_{1N} \\ &\le H\left(W_{21}|W_{22}X_1^N\right) - H\left(W_{21}|W_{22}X_1^N Y_1^N\right) + N\epsilon_{1N} \\ &= I\left(W_{21};Y_1^N|W_{22}X_1^N\right) + N\epsilon_{1N} \\ &\le H\left(Y_1^N|W_{22}X_1^N\right) - H\left(Y_1^N|W_{21}W_{22}X_1^N X_2^N\right) + N\epsilon_{1N} \\ &= H\left(Y_1^N|W_{22}X_1^N\right) - H\left(Y_1^N|W_{22}X_1^N X_2^N\right) + N\epsilon_{1N} \\ &= \sum_{n=1}^{n=N} H\left(Y_{1n}|X_1^N W_{22} Y_1^{n-1}\right) - H\left(Y_{1n}|X_{1n}X_{2n}\right) \\ &\quad + N\epsilon_{1N} \\ &\overset{(b)}{=} \sum_{n=1}^{n=N} H\left(Y_{1n}|X_1^N W_{22} Y_1^{n-1} Y_2^{n-1}\right) - H\left(Y_{1n}|X_{1n}X_{2n}\right) \\ &\quad + N\epsilon_{1N} \\ &\le \sum_{n=1}^{n=N} H\left(Y_{1n}|X_{1n}W_{22}Y_2^{n-1}\right) - H\left(Y_{1n}|X_{1n}X_{2n}\right) + N\epsilon_{1N} \\ &= \sum_{n=1}^{n=N} H\left(Y_{1n}|X_{1n}U_n\right) - H\left(Y_{1n}|X_{1n}X_{2n}U_n\right) \\ &\quad + N\epsilon_{1N} \\ &\le \sum_{n=1}^{n=N} I\left(X_{2n};Y_{1n}|U_n X_{1n}\right) + N\epsilon_{1N} \end{aligned} \tag{164}$$

where we define the random variable $U_n = \left(W_{22}, Y_2^{n-1}\right)$ for all n, (a) follows from the fact that since W_{21}, W_{22}, and W_1 are independent, so are W_{21}, W_{22}, and $X_1^N\left(W_1\right)$, and (b) follows from the fact that $Y_2^{n-1} \to \left(X_1^{n-1}Y_1^{n-1}\right) \to \left(W_{22}X_{1n}^N Y_{1n}\right)$ form a Markov chain. This follows from the discrete memoryless property of the channel and the fact that for any i, Y_{2i} depends only on X_{1i} and Y_{1i} (refer to (16)). Next, we bound R_{22} as follows:

$$\begin{aligned} NR_{22} &= I\left(W_{22};Y_2^N\right) + H\left(W_{22}|Y_2^N\right) \\ &\le I\left(W_{22};Y_2^N\right) + N\epsilon_{2N} \\ &= \sum_{n=1}^{N} I\left(W_{22};Y_{2n}|Y_2^{n-1}\right) + N\epsilon_{2N} \\ &\le \sum_{n=1}^{N} I\left(W_{22}Y_2^{n-1};Y_{2n}\right) + N\epsilon_{2N} \\ &= \sum_{n=1}^{N} I\left(U_n;Y_{2n}\right) + N\epsilon_{2N}. \end{aligned} \tag{165}$$

Next, we bound R_1 as follows:

$$\begin{aligned} NR_1 &= I\left(W_1;Y_1^N\right) + H\left(W_1|Y_1^N\right) \\ &\le I\left(X_1^N;Y_1^N\right) + N\epsilon_{3N} \\ &= H\left(X_1^N\right) - H\left(X_1^N|Y_1^N\right) + N\epsilon_{3N} \\ &= H\left(X_1^N|X_2^N\right) - H\left(X_1^N|Y_1^N\right) + N\epsilon_{3N} \\ &\le H\left(X_1^N|X_2^N\right) - H\left(X_1^N|X_2^N Y_1^N\right) + N\epsilon_{3N} \\ &= I\left(X_1^N;Y_1^N|X_2^N\right) + N\epsilon_{3N} \\ &= \sum_{n=1}^{n=N} H\left(Y_{1n}|X_2^N Y_1^{n-1}\right) - H\left(Y_{1n}|X_2^N Y_1^{n-1} X_1^N\right) \\ &\quad + N\epsilon_{3N} \\ &\le \sum_{n=1}^{n=N} H\left(Y_{1n}|X_{2n}\right) - H\left(Y_{1n}|X_{1n}X_{2n}\right) + N\epsilon_{3N} \\ &= \sum_{n=1}^{n=N} I\left(X_{1n};Y_{1n}|X_{2n}\right) + N\epsilon_{3N}. \end{aligned} \tag{166}$$

For the degraded discrete memoryless ZC of type II, we have

$$I\left(W_{22};Y_1^N|X_1^N\right) \ge I\left(W_{22};Y_2^N\right) \tag{167}$$

from the data processing inequality and the fact that $W_{22} \to X_2^N \to \left(X_1^N Y_1^N\right) \to Y_2^N$ form a Markov chain. The above inequality similarly holds for the discrete memoryless ZC of type III. To bound $R_1 + R_{21} + R_{22}$, we have

$$\begin{aligned} &N(R_1 + R_{21} + R_{22}) \\ &= I\left(W_{22};Y_2^N\right) + H\left(W_{22}|Y_2^N\right) \\ &\quad + I\left(W_1;Y_1^N\right) + H\left(W_1|Y_1^N\right) \\ &\quad + I\left(W_{21};Y_1^N\right) + H\left(W_{21}|Y_1^N\right) \\ &\le I\left(W_{22};Y_1^N|X_1^N\right) + I\left(X_1^N;Y_1^N\right) \\ &\quad + I\left(W_{21};Y_1^N|W_{22}X_1^N\right) + N\epsilon_{1N} + N\epsilon_{2N} + N\epsilon_{3N} \\ &\le I\left(W_{21}W_{22}X_1^N;Y_1^N\right) + N\epsilon_{1N} + N\epsilon_{2N} + N\epsilon_{3N} \\ &\le I\left(X_1^N X_2^N;Y_1^N\right) + N\epsilon_{1N} + N\epsilon_{2N} + N\epsilon_{3N} \\ &\le \sum_{n=1}^{n=N} H\left(Y_{1n}|Y_1^{n-1}\right) - H\left(Y_{1n}|X_{1n}X_{2n}\right) \\ &\quad + N\epsilon_{1N} + N\epsilon_{2N} + N\epsilon_{3N} \\ &\le \sum_{n=1}^{n=N} I\left(X_{1n}X_{2n};Y_{1n}\right) + N\epsilon_{1N} + N\epsilon_{2N} + N\epsilon_{3N}. \end{aligned} \tag{168}$$

By the Markovity of $U_n \to (X_{1n}X_{2n}) \to (Y_{1n}Y_{2n})$ and the independence of (U_n, X_{2n}) and X_{1n}, we observe again that

$$\begin{aligned} &p\left(u_n, x_{1n}, x_{2n}, y_{1n}, y_{2n}\right) \\ &\qquad = p\left(u_n, x_{2n}\right) p\left(x_{1n}\right) p\left(y_{1n}, y_{2n}|x_{1n}, x_{2n}\right). \end{aligned}$$

Finally, we obtain Theorem 4, by introducing a time-sharing random variable Q. The assertions about the cardinalities of $\mathcal{U}$ and $\mathcal{Q}$ follow directly from the application of Caratheodory's theorem to the expressions (94)–(97).

APPENDIX IV
PROOF OF THEOREM 7

We determine an outer bound to the capacity region of the equivalent Gaussian ZC with strong crossover link gain as shown in Fig. 9. By Fano's inequality, we have

$$H\left(W_{21}|Y_1'^N\right) \le N\epsilon_{1N} \quad (169)$$
$$H\left(W_{22}|Y_2'^N\right) \le N\epsilon_{2N} \quad (170)$$
$$H\left(W_1|Y_1'^N\right) \le N\epsilon_{3N} \quad (171)$$

where $\epsilon_{1N}, \epsilon_{2N}, \epsilon_{3N} \to 0$ as $N \to \infty$. We first bound the term $H\left(W_{22}|Y_1'^N W_1\right) = H\left(W_{22}|X_2^N + Z_{21}^N\right)$. From the following Markov chain:

$$(W_{21}, W_{22}) \to X_2^N \to X_2^N + Z_{21}^N \to X_2^N + Z_{21}^N + Z_{22}^N \quad (172)$$

we have by the data processing inequality and Fano's inequality

$$\begin{aligned} I\left(W_{22}; X_2^N + Z_{21}^N\right) &\ge I\left(W_{22}; X_2^N + Z_{21}^N + Z_{22}^N\right) \\ \Longrightarrow H\left(W_{22}|X_2^N + Z_{21}^N\right) &\le H\left(W_{22}|X_2^N + Z_{21}^n + Z_{22}^N\right) \\ &= H\left(W_{22}|Y_2'^N\right) \\ &\le N\epsilon_{2N}. \end{aligned} \quad (173)$$

Next, we bound the following term $h\left(Y_1'^N|W_1W_{22}\right)$. Consider the following inequalities:

$$\begin{aligned} &\frac{N}{2}\log_2\left(\frac{2\pi e}{a^2}\right) \\ &= h\left(Z_{21}^N|W_1W_{21}W_{22}\right) \\ &= h\left(\frac{1}{a}X_1^N(W_1) + X_2^N(W_{21}, W_{22}) + Z_{21}^N|W_1W_{21}W_{22}\right) \\ &= h\left(Y_1'^N|W_1W_{21}W_{22}\right) \\ &\le h\left(Y_1'^N|W_1W_{22}\right) \\ &\le h\left(Y_1'^N|W_1\right) \\ &\le \frac{N}{2}\log_2\left((2\pi e)\left(\frac{a^2P_2+1}{a^2}\right)\right). \end{aligned} \quad (174)$$

Thus, there exists a $\beta \in [0,1]$, such that

$$\begin{aligned} h\left(Y_1'^N|W_1W_{22}\right) &= h\left(X_2^N(W_{21}, W_{22}) + Z_{21}^N|W_{22}\right) \\ &= \frac{N}{2}\log_2\left((2\pi e)\left(\frac{a^2\beta P_2+1}{a^2}\right)\right). \end{aligned} \quad (175)$$

We next obtain a lower bound for $h\left(X_2^n + Z_{21}^n + Z_{22}^n|W_{22}\right)$ by making use of the entropy power inequality

$$\begin{aligned} 2^{\frac{2}{N}h\left(X_2^N+Z_{21}^N+Z_{22}^N|W_{22}\right)} &\ge 2^{\frac{2}{N}h\left(X_2^N+Z_{21}^N|W_{22}\right)} \\ &\quad + 2^{\frac{2}{N}h\left(Z_{22}^N\right)} \\ &= (2\pi e)(\beta P_2 + 1) \\ \Rightarrow h\left(X_2^N + Z_{21}^N + Z_{22}^N|W_{22}\right) &\ge \frac{N}{2}\log_2\left((2\pi e)(\beta P_2+1)\right). \end{aligned} \quad (176)$$

We can now bound R_{21} as follows:

$$\begin{aligned} NR_{21} &= H(W_{21}) \\ &= I\left(W_{21}; Y_1'^N|W_1W_{22}\right) + H\left(W_{21}|Y_1'^N W_1W_{22}\right) \\ &\le I\left(W_{21}; Y_1'^N|W_1W_{22}\right) + H\left(W_{21}|Y_1'^N\right) \\ &\le I\left(W_{21}; Y_1'^N|W_1W_{22}\right) + N\epsilon_{1N} \\ &= h\left(Y_1'^N|W_1W_{22}\right) - h\left(Y_1'^N|W_1W_{21}W_{22}\right) + N\epsilon_{1N} \\ &= h\left(Y_1'^N|W_1W_{22}\right) - h\left(Z_{21}^N\right) + N\epsilon_{1N} \\ &= \frac{N}{2}\log_2\left((2\pi e)\left(\frac{a^2\beta P_2+1}{a^2}\right)\right) - \frac{N}{2}\log_2\left(\frac{2\pi e}{a^2}\right) \\ &\quad + N\epsilon_{1N} \\ &= \frac{N}{2}\log_2\left(a^2\beta P_2+1\right) + N\epsilon_{1N}. \end{aligned} \quad (177)$$

We bound R_{22} as follows:

$$\begin{aligned} NR_{22} &= H(W_{22}) \\ &= I\left(W_{22}; Y_2'^N\right) + H\left(W_{22}|Y_2'^N\right) \\ &\le I\left(W_{22}; Y_2'^N\right) + N\epsilon_{2N} \\ &= h\left(Y_2'^N\right) - h\left(Y_2'^N|W_{22}\right) + N\epsilon_{2N} \\ &= h\left(X_2^N + Z_{21}^N + Z_{22}^N\right) - h\left(X_2^N + Z_{21}^N + Z_{22}^N|W_{22}\right) \\ &\quad + N\epsilon_{2N} \\ &\le \frac{N}{2}\log_2\left(1 + \frac{(1-\beta)P_2}{1+\beta P_2}\right) + N\epsilon_{2N}. \end{aligned} \quad (178)$$

We then bound R_1 as follows:

$$\begin{aligned} NR_1 &= H(W_1) \\ &= I\left(W_1; Y_1'^N|W_{21}W_{22}\right) + H\left(W_1|Y_1'^N W_{21}W_{22}\right) \\ &\le I\left(W_1; Y_1'^N|W_{21}W_{22}\right) + H\left(W_1|Y_1'^N\right) \\ &\le I\left(W_1; Y_1'^N|W_{21}W_{22}\right) + N\epsilon_{3N} \\ &\le h\left(Y_1'^N|W_{21}W_{22}\right) - h\left(Y_1'^N|W_1W_{21}W_{22}\right) + N\epsilon_{3N} \\ &= h\left(\frac{X_1^N}{a} + Z_{21}^N\right) - h\left(Z_{21}^N\right) + N\epsilon_{3N} \\ &\le \frac{N}{2}\log_2(1+P_1) + N\epsilon_{3N}. \end{aligned} \quad (179)$$

Finally, we bound the term $R_1 + R_{21} + R_{22}$ as follows:

$$\begin{aligned} &N(R_1 + R_{21} + R_{22}) \\ &= H(W_1W_{21}W_{22}) \\ &= I\left(W_1W_{21}W_{22}; Y_1'^N\right) + H\left(W_1|Y_1'^N\right) \\ &\quad + H\left(W_{21}|Y_1'^N W_1\right) + H\left(W_{22}|Y_1'^N W_1W_{21}\right) \\ &\le I\left(W_1W_{21}W_{22}; Y_1'^N\right) + H\left(W_1|Y_1'^N\right) \\ &\quad + H\left(W_{21}|Y_1'^N\right) + H\left(W_{22}|Y_1'^N W_1\right) \\ &\le I\left(W_1W_{21}W_{22}; Y_1'^N\right) + N\epsilon_{1N} + N\epsilon_{2N} + N\epsilon_{3N} \\ &= h\left(\frac{X_1^N}{a} + X_2^N + Z_{21}^N\right) - h\left(Z_{21}^N\right) \\ &\quad + N\epsilon_{1N} + N\epsilon_{2N} + N\epsilon_{3N} \\ &\le \frac{N}{2}\log_2\left(1 + a^2P_2 + P_1\right) + N\epsilon_{1N} + N\epsilon_{2N} + N\epsilon_{3N}. \end{aligned} \quad (180)$$

ACKNOWLEDGMENT

The authors are grateful to the anonymous reviewers for constructive comments and suggestions.

REFERENCES

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 1st ed. New York: Wiley-Interscience, 1991, Wiley Series In Telecommunications, ch. 14, pp. 374–458.

[2] S. Vishwanath, N. Jindal, and A. Goldsmith, "The "Z" channel," in *Proc. IEEE Global Telecommunications Conf.*, San Francisco, CA, Dec. 2003, vol. 3, pp. 1726–1730.

[3] N. Liu and S. Ulukus, "On the capacity region of the Gaussian Z-channel," in *Proc. IEEE Global Telecommunications Conf.*, Dallas, TX, Nov. 2004, vol. 1, pp. 415–419.

[4] F. M. J. Willems and E. C. V. D. Meulen, "The discrete memoryless multiple-access channel with cribbing encoders," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 3, pp. 313–327, May 1985.

[5] J. Pearl, *Causality: Models, Reasoning, and Inference*, 1st ed. Cambridge, U.K.: Cambridge Univ. Press, 2000, ch. 1.

[6] M. H. M. Costa, "On the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 5, pp. 607–615, Sep. 1985.

[7] A. B. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 1, pp. 60–70, Jan. 1978.

[8] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 3, pp. 581–586, Mar. 2004.

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

[10] G. Kramer, "Review of rate regions for interference channels," in *Proc. Int. Zurich Seminar*, Zurich, Switzerland, Feb. 2006, pp. 162–165.

[11] H. F. Chong, M. Motani, H. K. Garg, and H. El Gamal, "On a simplification of the Han–Kobayashi rate region for the interference channel," *IEEE Trans. Inf. Theory*, submitted for publication.

[12] G. Kramer, "Capacity results for the discrete memoryless network," *IEEE Trans. Inf. Theory*, vol. 49, no. 1, pp. 4–21, Jan. 2003.