Article

Capacity Region of a New Bus Communication Model

Bin Dai [1]*, **A. J. Han Vinck** [3], **Yuan Luo** [2] **and Zhuojun Zhuang** [2]

[1] School of Information Science and Technology, Southwest JiaoTong University, Northbound Section Second Ring Road 111, Chengdu, 610031, China

[2] Computer Science and Engineering Department, Shanghai Jiao Tong University, Dongchuan road 800, Shanghai, 200240, China; E-Mails: luoyuan@cs.sjtu.edu.cn (Y.L.); zhuojunzzj@sjtu.edu.cn (Z.Z.)

[3] Institute for Experimental Mathematics, Duisburg-Essen University, Ellernstr.29, Essen, 45326, Germany; E-Mail: vinck@iem.uni-due.de

* Author to whom correspondence should be addressed; E-Mail: daibin007@sjtu.edu.cn; Tel.: +86 21 3420 5477.

Abstract: In this paper, we study a new bus communication model, where two transmitters wish to send their corresponding private messages and a common message to a destination, while they also wish to send the common message to another receiver connected to the same wire. From an information-theoretical point of view, we first study a general case of this new model (with discrete memoryless channels). The capacity region composed of all achievable (R_0, R_1, R_2) triples is determined for this general model, where R_1 and R_2 are the transmission rates of the private messages and R_0 is the transmission rate of the common message. Then, the result is further explained via the Gaussian example. Finally, we give the capacity region for the new bus communication model with additive Gaussian noises and attenuation factors. This new bus communication model captures various communication scenarios, such as the bus systems in vehicles, and the bus type of communication channel in power line communication (PLC) networks.

Keywords: bus communication system; capacity region; multiple-access channel; degraded broadcast channel; Gaussian broadcast channel; power line communication networks

1. Introduction

The bus communication model has been widely studied for many years. It captures various communication scenarios, such as the bus systems in vehicles, and the bus type of communication channel in power line communication (PLC) networks (see [1–5]).

Let us consider the bus communication model of Figure 1 from an information-theoretical point of view. Figure 1 can be equivalent to the model of the broadcast channel (see Figure 2). Note that Figure 2 implies that $f_1 = f_2 = f$.

Figure 1. The bus communication model (Gains f_1 and f_2, the power of the sender σ_x^2 and the power of the noise σ_n^2).
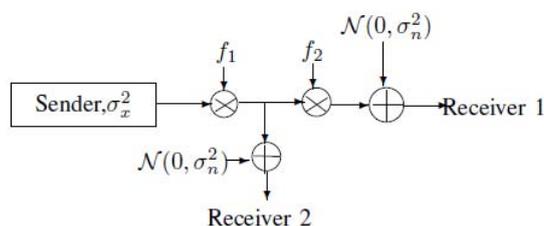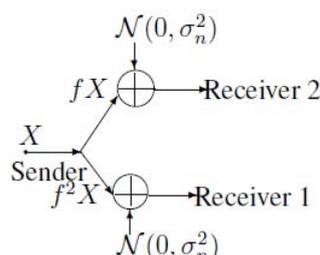

Figure 2. The broadcast presentation of the bus communication model.

The model of the broadcast channel was first investigated by Cover [6], and the capacity region of the general case (two private messages and one common message) is still not known. After the publication of Cover's work, Körner and Marton [7] studied the broadcast channel with a degraded message set (one private and one common message) and found its capacity region. For the degraded broadcast channel, the capacity region is totally determined (see [8–10]). In addition, Gamal and Cover [11] showed that the Gaussian broadcast channel is a kind of degraded broadcast channel, and therefore, the capacity region for the Gaussian case can be directly obtained from the result of the degraded broadcast channel.

The following theorem 1 shows the capacity region of the model of Figure 2, which is a kind of Gaussian broadcast channel.

Theorem 1 *The capacity region of the model of Figure 2 is the set of rate pairs (R_1, R_2), such that*

$$R_2 \leq \frac{1}{2}\ln(1 + \frac{\alpha\sigma_x^2 f^2}{\sigma_n^2})$$

$$R_1 \leq \frac{1}{2}\ln(1 + \frac{(1-\alpha)\sigma_x^2 f^4}{\alpha\sigma_x^2 f^4 + \sigma_n^2})$$

for some $\alpha \in [0,1]$. Note that σ_x^2 is the power constraint of the channel input, X, f and f^2 are channel gains ($0 < f \le 1$) and σ_n^2 is the power of the noise.

Theorem 1 is directly obtained from the capacity region of the Gaussian broadcast channel [11], and therefore, the proof is omitted here.

In this paper, we study a two-sender bus communication model (see Figure 3). Two transmitters wish to send their corresponding private messages and a common message to receiver 1, while they also wish to send the common message to receiver 2.

Figure 3. A new bus communication model with two transmitters.
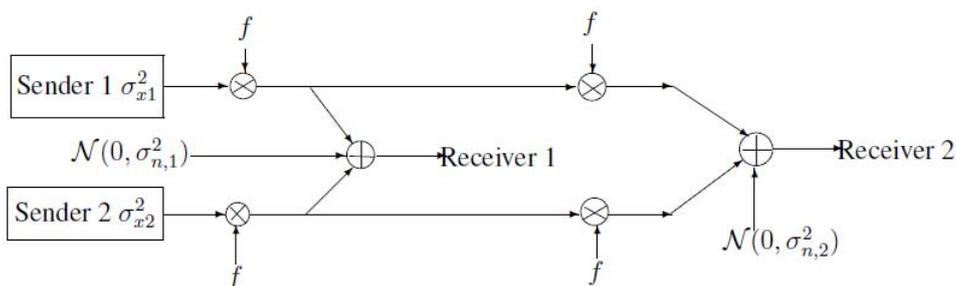

Figure 3 can be equivalent to the following Figure 4. Note that the capacity region of the Gaussian broadcast channel (BC) can be obtained from the capacity region of the discrete memoryless degraded broadcast channel. Therefore, first, we study the discrete memoryless case of the model of Figure 4, where two transmitters wish to send their private messages, W_1 and W_2, and a common message, W_0, to receiver 1, and meanwhile, they also wish to send the common message, W_0, to receiver 2. Receiver 2 can receive a degraded version of the output of the multiple-access channel (MAC) via a discrete memoryless channel (DMC) (see Figure 5). This model can be viewed as a combination of multiple-access channel and degraded broadcast channel. For convenience, we call it MAC-DBC in this paper.

**Figure 4.** An equivalent model for the model of Figure 3.
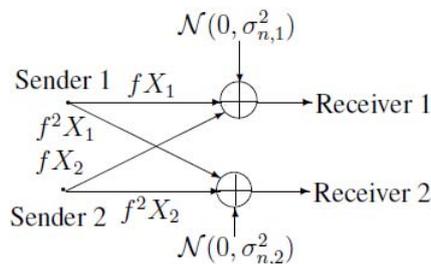
Then, we study the model of Figure 4, which is a Gaussian example of the MAC-DBC in Figure 5. The capacity regions of the MAC-DBC and the model of Figure 4 are totally determined.

The study of MAC-DBC from an information-theoretical point of view is due to the fact that the network information theory has recently become an active research area. Both MAC and BC play an important role in the network information theory, and they have been extensively studied separately. However, the cascade of MAC and BC (MAC-BC) has seldom drawn people's attention. To investigate

the capacity region and the capacity-achieving coding scheme for the MAC-BC is the motivation of this work.

Figure 5. A combination of multiple-access channel and degraded broadcast channel (MAC-DBC).
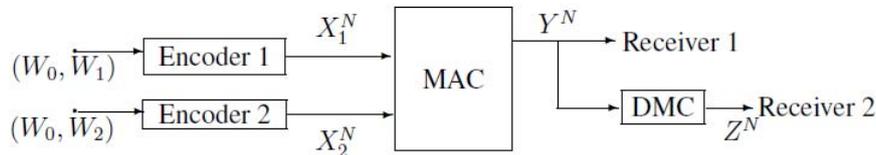
In this paper, random variab1es, sample values and alphabets are denoted by capital letters, lower case letters and calligraphic letters, respectively. A similar convention is applied to the random vectors and their sample values. For example, U^N denotes a random N-vector $(U_1, ..., U_N)$, and $u^N = (u_1, ..., u_N)$ is a specific vector value in $\mathcal{U}^N$ that is the Nth Cartesian power of $\mathcal{U}$. U_i^N denotes a random $N-i+1$-vector $(U_i, ..., U_N)$, and $u_i^N = (u_i, ..., u_N)$ is a specific vector value in $\mathcal{U}_i^N$. Let $p_V(v)$ denote the probability mass function $Pr\{V = v\}$. Throughout the paper, the logarithmic function is to the base 2.

The remainder of this paper is organized as follows. In Section 2, we present the basic definitions and the main result on the capacity region of MAC-DBC. In Section 3, we provide the capacity region of the model of Figure 4. Final conclusions are presented in Section 4. The proofs are provided from Section A to Section E.

2. Notations, Definitions and the Main Results of MAC-DBC

In this section, a description of the MAC-DBC is given by Definition 1 to Definition 3. The capacity region, $\mathcal{R}^{(\mathcal{A})}$, composed of all achievable (R_0, R_1, R_2) triples is given in Theorem 2, where the achievable (R_0, R_1, R_2) triple is defined in Definition 4.

Definition 1 (*Encoders*) *The private messages, W_1 and W_2, take values in $\mathcal{W}_1$ and $\mathcal{W}_2$, respectively. The common message, W_0, takes values in $\mathcal{W}_0$. W_0, W_1 and W_2 are independent and uniformly distributed over their ranges. The channel encoders are two mappings:*

$$f_1^N : \mathcal{W}_0 \times \mathcal{W}_1 \to \mathcal{X}_1^N \tag{1}$$

where $f_1^N(w_0, w_1) = x_1^N \in \mathcal{X}_1^N$, $w_1 \in \mathcal{W}_1$, and $w_0 \in \mathcal{W}_0$.

$$f_2^N : \mathcal{W}_0 \times \mathcal{W}_2 \to \mathcal{X}_2^N \tag{2}$$

where $f_2^N(w_0, w_2) = x_2^N \in \mathcal{X}_2^N$, $w_2 \in \mathcal{W}_2$, and $w_0 \in \mathcal{W}_0$. Note that W_1 and X_2^N are independent, and W_2 is independent of X_1^N.

The transmission rates of the private messages and the common message are $\frac{\log \|\mathcal{W}_1\|}{N}$, $\frac{\log \|\mathcal{W}_2\|}{N}$ and $\frac{\log \|\mathcal{W}_0\|}{N}$, respectively.

Definition 2 (*Channels*) *The MAC is a DMC with finite input alphabet $\mathcal{X}_1 \times \mathcal{X}_2$, finite output alphabet $\mathcal{Y}$ and transition probability $Q_1(y|x_1, x_2)$, where $x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2, y \in \mathcal{Y}$. $Q_1(y^N|x_1^N, x_2^N) = \prod_{n=1}^N Q_1(y_n|x_{1,n}, x_{2,n})$. The inputs of the MAC are X_1^N and X_2^N, while the output is Y^N.*

Receiver 2 has access to the output of the MAC via a discrete memoryless channel (DMC). The input of this DMC is Y^N, and the output is Z^N. The transition probability satisfies that

$$p_{Z^N|Y^N}(z^N|y^N) = \prod_{i=1}^{N} p_{Z|Y}(z_i|y_i) \tag{3}$$

where $z_i \in \mathcal{Z}$ and $y_i \in \mathcal{Y}$.

Definition 3 (*Decoders*) *The decoder for receiver 1 is a mapping,* $f_{D1} : \mathcal{Y}^N \to \mathcal{W}_1 \times \mathcal{W}_2 \times \mathcal{W}_0$, *with input* Y^N *and outputs* $\widehat{W}_1$, $\widehat{W}_2$ *and* $\widehat{W}_0$. *Let* P_{e1} *be the error probability of the receiver 1 , and it is defined as* $Pr\{(W_0, W_1, W_2) \neq (\widehat{W}_0, \widehat{W}_1, \widehat{W}_2)\}$.

The decoder for receiver 2 is a mapping, $f_{D2} : \mathcal{Z}^N \to \mathcal{W}_0$, *with input* Z^N *and output* $\widehat{W}_0$. *Let* P_{e2} *be the error probability of the receiver 2 , and it is defined as* $Pr\{W_0 \neq \widehat{W}_0\}$.

Definition 4 (*Achievable* (R_0, R_1, R_2) *triple in the model of Figure 5*) *A triple* (R_0, R_1, R_2) *(where* $R_0, R_1, R_2 > 0$*) is called achievable if, for any* $\epsilon > 0$ *(where* ϵ *is an arbitrary small positive real number and* $\epsilon \to 0$*), there exists channel encoders-decoders* $(N, P_{e1} and P_{e2})$, *such that*

$$\frac{\log \|\mathcal{W}_0\|}{N} \geq R_0 - \epsilon, \frac{\log \|\mathcal{W}_1\|}{N} \geq R_1 - \epsilon, \frac{\log \|\mathcal{W}_2\|}{N} \geq R_2 - \epsilon, P_{e1} \leq \epsilon, P_{e2} \leq \epsilon \tag{4}$$

Theorem 2 gives a single-letter characterization of the set $\mathcal{R}^{(\mathcal{A})}$, which is composed of all achievable (R_0, R_1, R_2) triples in the model of Figure 5, and it is proved in Section A and Section B.

Theorem 2 *A single-letter characterization of the region* $\mathcal{R}^{(\mathcal{A})}$ *is as follows,*

$$\mathcal{R}^{(\mathcal{A})} = \left\{ \begin{array}{l} (R_0, R_1, R_2) : \\ 0 \leq R_0 \leq I(U; Z) \\ 0 \leq R_1 \leq I(X_1; Y|X_2, U) \\ 0 \leq R_2 \leq I(X_2; Y|X_1, U) \\ R_0 + R_1 \leq I(X_1, X_2; Y|V_2) \\ R_0 + R_2 \leq I(X_1, X_2; Y|V_1) \\ R_1 + R_2 \leq I(X_1, X_2; Y|U) \\ R_0 + R_1 + R_2 \leq I(X_1, X_2; Y). \end{array} \right\}$$

where $(U, V_1, V_2) \to (X_1, X_2) \to Y \to Z$.

Remark 1 *There are some notes on Theorem 2; see the following:*

- *The region* $\mathcal{R}^{(\mathcal{A})}$ *is convex, and the proof is in Section C.*
- *The ranges of the random variables* U, V_1 *and* V_2 *satisfy*

$$\|\mathcal{U}\| \leq \|\mathcal{X}_1\|\|\mathcal{X}_2\| + 2$$

$$\|\mathcal{V}_1\| \leq \|\mathcal{X}_1\|\|\mathcal{X}_2\|$$

$$\|\mathcal{V}_2\| \leq \|\mathcal{X}_1\|\|\mathcal{X}_2\|$$

The proof is in Section D.

- *The auxiliary random variables U, V_1 and V_2, in fact, are corresponding to W_0, W_1 and W_2, respectively.*
- *If $R_0 = 0$, receiver 2 is useless, and the model of Figure 5 reduces to the multiple-access channel (MAC).*

  Let $R_0 = 0$ and the corresponding $U = const$, the region, $\mathcal{R}^{(\mathcal{A})}$, reduces to

$$
\left\{
\begin{array}{l}
(R_1, R_2): \\
0 \leq R_1 \leq I(X_1; Y|X_2) \\
0 \leq R_2 \leq I(X_2; Y|X_1) \\
R_1 \leq I(X_1, X_2; Y|V_2) \\
R_2 \leq I(X_1, X_2; Y|V_1) \\
R_1 + R_2 \leq I(X_1, X_2; Y).
\end{array}
\right\}
\tag{5}
$$

Note that the Markov chains, $V_1 \rightarrow X_1 \rightarrow Y$ and $V_2 \rightarrow X_2 \rightarrow Y$, then the inequalities $R_1 \leq I(X_1, X_2; Y|V_2)$ and $R_2 \leq I(X_1, X_2; Y|V_1)$ are included in $R_1 \leq I(X_1; Y|X_2)$ and $R_2 \leq I(X_2; Y|X_1)$, respectively. Therefore, the above region (5) is simplified as

$$
\left\{
\begin{array}{l}
(R_1, R_2): \\
0 \leq R_1 \leq I(X_1; Y|X_2) \\
0 \leq R_2 \leq I(X_2; Y|X_1) \\
R_1 + R_2 \leq I(X_1, X_2; Y).
\end{array}
\right\}
\tag{6}
$$

, and this is coincident with the capacity region of the MAC [12,13].

3. A Gaussian Example of MAC-DBC and the Capacity Region of the Model of Figure 4

In this section, we first study a Gaussian example of Figure 5, where the channel input-output relationships at each time instant i ($1 \leq i \leq N$) are given by

$$
Y_i = X_{1,i} + X_{2,i} + Z_{1,i}
\tag{7}
$$

and

$$
Z_i = X_{1,i} + X_{2,i} + Z_{1,i} + Z_{2,i}
\tag{8}
$$

where $Z_{1,i} \sim \mathcal{N}(0, \sigma_{n,1}^2)$ and $Z_{2,i} \sim \mathcal{N}(0, \sigma_{n,2}^2)$. The random vectors, Z^N, Z_1^N and Z_2^N, are independent with i.i.d. components. The channel inputs, X_1^N and X_2^N, are subject to the average power constraints, $\sigma_{x,1}^2$ and $\sigma_{x,2}^2$, respectively, *i.e.*,

$$
\frac{1}{N} \sum_{i=1}^{N} E[X_{1,i}^2] \leq \sigma_{x,1}^2, \qquad \frac{1}{N} \sum_{i=1}^{N} E[X_{2,i}^2] \leq \sigma_{x,2}^2
\tag{9}
$$

Theorem 3 *For the Gaussian case of Figure 5, the capacity-equivocation region,* $\mathcal{R}^{(\mathcal{B})}$, *is given by*

$$
\mathcal{R}^{(\mathcal{B})} = \bigcup_{0 \le \alpha \le 1}
\left\{
\begin{array}{l}
(R_0, R_1, R_2) : \\[4pt]
0 \le R_0 \le \frac{1}{2} \log(1 + \frac{\alpha(\sigma_{x,1}^2 + \sigma_{x,2}^2)}{(1-\alpha)(\sigma_{x,1}^2 + \sigma_{x,2}^2) + \sigma_{n,1}^2 + \sigma_{n,2}^2}) \\[8pt]
0 \le R_1 \le \frac{1}{2} \log(1 + \frac{(1-\alpha)\sigma_{x,1}^2}{\sigma_{n,1}^2}) \\[8pt]
0 \le R_2 \le \frac{1}{2} \log(1 + \frac{(1-\alpha)\sigma_{x,2}^2}{\sigma_{n,1}^2}) \\[8pt]
R_1 + R_2 \le \frac{1}{2} \log(1 + \frac{(1-\alpha)(\sigma_{x,1}^2 + \sigma_{x,2}^2)}{\sigma_{n,1}^2}) \\[8pt]
R_0 + R_1 \le \frac{1}{2} \log(1 + \frac{\sigma_{x,1}^2 + \alpha\sigma_{x,2}^2}{\sigma_{n,1}^2}) \\[8pt]
R_0 + R_2 \le \frac{1}{2} \log(1 + \frac{\sigma_{x,2}^2 + \alpha\sigma_{x,1}^2}{\sigma_{n,1}^2}) \\[8pt]
R_0 + R_1 + R_2 \le \frac{1}{2} \log(1 + \frac{\sigma_{x,1}^2 + \sigma_{x,2}^2}{\sigma_{n,1}^2}).
\end{array}
\right\}
\tag{10}
$$

The proof of Theorem 3 is in Section E.

Then, we will show that the capacity region of the model of Figure 4 can be obtained from the above Theorem 3. The channel input-output relationships of Figure 4 at each time instant i ($1 \le i \le N$) are given by

$$
Y_i = f X_{1,i} + f X_{2,i} + Z_{1,i}
\tag{11}
$$

and

$$
Z_i = f^2 X_{1,i} + f^2 X_{2,i} + Z_{2,i}
\tag{12}
$$

, where $Z_{1,i} \sim \mathcal{N}(0, \sigma_{n,1}^2)$, $Z_{2,i} \sim \mathcal{N}(0, \sigma_{n,2}^2)$ and $\sigma_{n,1}^2 \le \sigma_{n,2}^2$. The channel inputs, X_1^N and X_2^N, are subject to the average power constraints, $\sigma_{x,1}^2$ and $\sigma_{x,2}^2$, respectively.

Note that the additive Gaussian noise, $Z_{2,i}$, can be viewed as a cascade of $Z_{1,i}$ and $Z'_{2,i}$, where $Z'_{2,i} \sim \mathcal{N}(0, \sigma_{n,2}^2 - \sigma_{n,1}^2)$. Moreover, Equations (11) and (12) are equivalent to Equations (13) and (14), respectively, where

$$
Y_i = X_{1,i} + X_{2,i} + \frac{1}{f} Z_{1,i}
\tag{13}
$$

and

$$
Z_i = X_{1,i} + X_{2,i} + \frac{1}{f^2} Z_{2,i}
\tag{14}
$$

Therefore, the model of Figure 4 is analogous to the above Gaussian example of MAC-DBC. The capacity region is as follows.

Theorem 4 *For the model of Figure 4, the capacity region* $\mathcal{R}^{(\mathcal{C})}$ *is given by*

$$
\mathcal{R}^{(\mathcal{C})} = \bigcup_{0 \le \alpha \le 1}
\left\{
\begin{array}{l}
(R_0, R_1, R_2) : \\[4pt]
0 \le R_0 \le \frac{1}{2} \log(1 + \frac{\alpha f^4 (\sigma_{x,1}^2 + \sigma_{x,2}^2)}{(1-\alpha)(\sigma_{x,1}^2 + \sigma_{x,2}^2) f^4 + \sigma_{n,2}^2}) \\[8pt]
0 \le R_1 \le \frac{1}{2} \log(1 + \frac{(1-\alpha)\sigma_{x,1}^2 f^2}{\sigma_{n,1}^2}) \\[8pt]
0 \le R_2 \le \frac{1}{2} \log(1 + \frac{(1-\alpha)\sigma_{x,2}^2 f^2}{\sigma_{n,1}^2}) \\[8pt]
R_1 + R_2 \le \frac{1}{2} \log(1 + \frac{(1-\alpha)(\sigma_{x,1}^2 + \sigma_{x,2}^2) f^2}{\sigma_{n,1}^2}) \\[8pt]
R_0 + R_1 \le \frac{1}{2} \log(1 + \frac{(\sigma_{x,1}^2 + \alpha\sigma_{x,2}^2) f^2}{\sigma_{n,1}^2}) \\[8pt]
R_0 + R_2 \le \frac{1}{2} \log(1 + \frac{(\sigma_{x,2}^2 + \alpha\sigma_{x,1}^2) f^2}{\sigma_{n,1}^2}) \\[8pt]
R_0 + R_1 + R_2 \le \frac{1}{2} \log(1 + \frac{(\sigma_{x,1}^2 + \sigma_{x,2}^2) f^2}{\sigma_{n,1}^2}).
\end{array}
\right\}
\tag{15}
$$

The proof is directly obtained from Theorem 3, and it is omitted here.

Figures 6 to 10 plot the capacity region, $\mathcal{R}^{(\mathcal{C})}$, with different values of f, $\sigma_{x,1}^2$, $\sigma_{x,2}^2$, $\sigma_{n,1}^2$ and $\sigma_{n,2}^2$. It is easy to see that $\mathcal{R}^{(\mathcal{C})}$ reduces to the capacity region of the Gaussian MAC when $R_0 = 0$. From these figures, we can see that $\mathcal{R}^{(\mathcal{C})}$ enlarges as $\sigma_{n,2}^1$ and $\sigma_{n,2}^2$ decrease. Moreover, for fixed f, $\sigma_{n,1}^2$ and $\sigma_{n,2}^2$, $\mathcal{R}^{(\mathcal{C})}$ enlarges as $\sigma_{x,1}^2$ and $\sigma_{x,2}^2$ increase.

Figure 6. The capacity region of Figure 4 with $\sigma_{x,1}^2 = \sigma_{x,2}^2 = 1$, $\sigma_{n,1}^2 = 1$, $\sigma_{n,2}^2 = 2$ and $f = 1$.
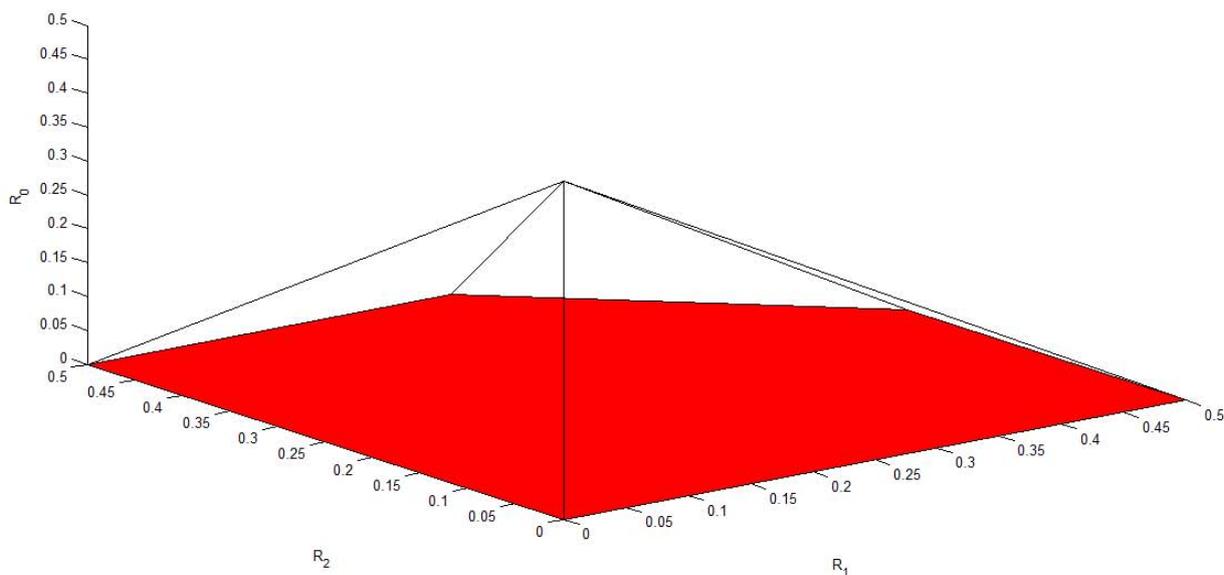

Figure 7. The capacity region of Figure 4 with $\sigma_{x,1}^2 = \sigma_{x,2}^2 = 1$, $\sigma_{n,1}^2 = 1$, $\sigma_{n,2}^2 = 100$ and $f = 1$.

Figure 8. The capacity region of Figure 4 with $\sigma_{x,1}^2 = \sigma_{x,2}^2 = 1$, $\sigma_{n,1}^2 = 1$, $\sigma_{n,2}^2 = 0.01$ and $f = 1$.

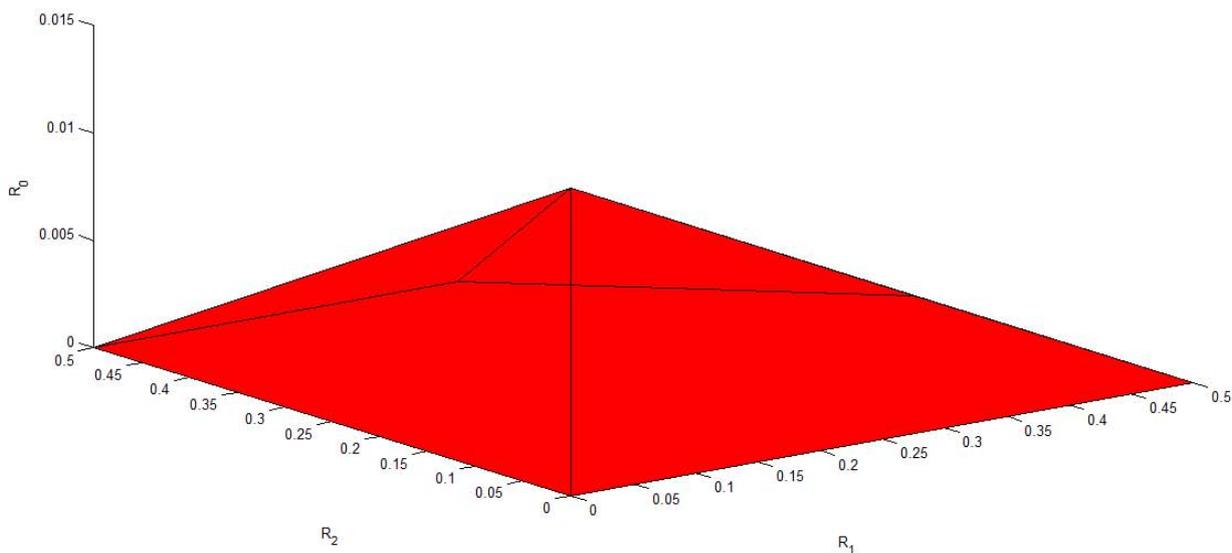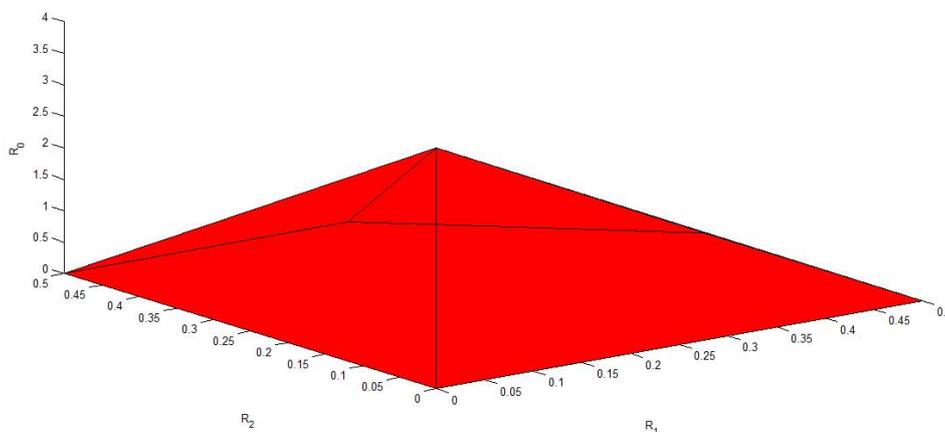**Figure 9.** The capacity region of Figure 4 with $\sigma_{x,1}^2 = \sigma_{x,2}^2 = 2$, $\sigma_{n,1}^2 = 0.1$, $\sigma_{n,2}^2 = 0.01$ and $f = 1$.
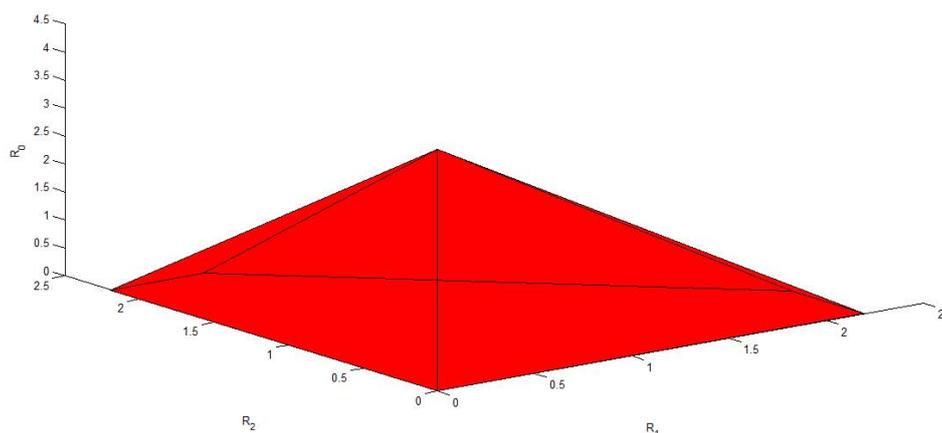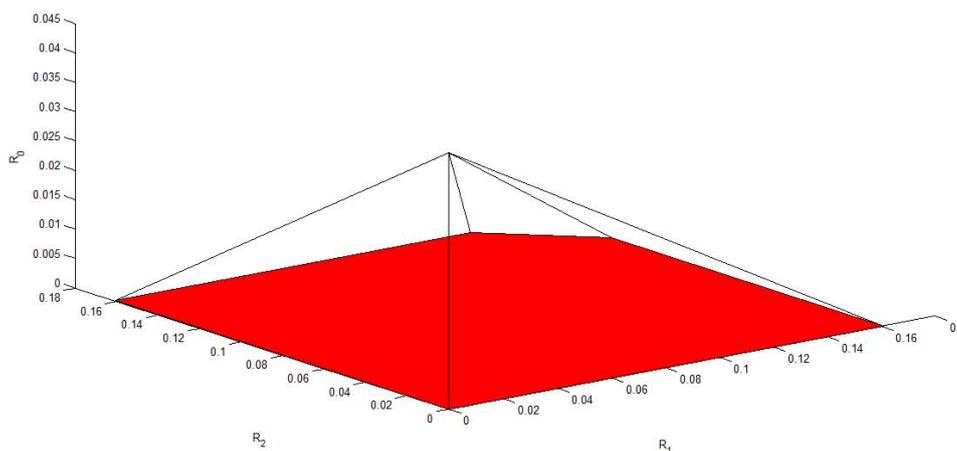

Figure 10. The capacity region of Figure 4 with $\sigma_{x,1}^2 = \sigma_{x,2}^2 = 1$, $\sigma_{n,1}^2 = 1$, $\sigma_{n,2}^2 = 2$ and $f = 0.5$.

4. Conclusions

In this paper, we first study the discrete memoryless MAC-DBC (the model of Figure 5). The capacity region is totally determined for this new model. Then, we study the model of Figure 4 and show that the capacity region of Figure 4 can be directly obtained from the Gaussian example of the MAC-DBC.

Acknowledgment

References

1. Lampe, L.; Schober, R.; Yiu, S. Distributed space-time block coding for multihop transmission in power line communication networks. *IEEE J. Select. Areas Commun.* **2006**, *24*, 1389–1400.
2. Bumiller, G.; Lampe, L.; Hrasnica, H. Power line communications for large-scale control and automation systems. *IEEE Commun. Mag.* **2010**, *48*, 106–113.
3. Lampe, L.; Han Vinck, A.J. On cooperative coding for narrow band PLC networks. *Int. J. Electron. Commun. (AEÜ)* **2011**, *8*, 681–687.
4. Balakirsky, V.; Vinck, A. Potential performance of PLC systems composed of several communication links. *Intl. Symp. Power Line Commun. Appl. Vancouver* **2005**, *10*, 12–16.
5. Johansson, M.; Nielsen, L. Vehicle applications of controller area network. *Tech. Rep.* **2003**, *7*, 27–52.
6. Cover, T.M. Broadcast channels. *IEEE Trans. Inf. Theory* **1972**, *18*, 2–14.
7. Körner, J.; Marton, K. General broadcast channels with degraded message sets. *IEEE Trans. Inf. Theory* **1977**, *23*, 60–64.
8. Gallager, R.G. Capacity and coding for degraded broadcast channels. *Probl. Inf. Transm.* **1974**, *10*, 3–14.
9. Bergmans, P.P. A simple converse for broadcast channels with additive white Gaussian noise. *IEEE Trans. Inf. Theory* **1974**, *20*, 279–280.
10. Bergmans, P.P. Random coding theorem for broadcast channels with degraded components. *IEEE Trans. Inf. Theory* **1973**, *19*, 197–207.
11. El Gamal, A.; Cover, T.M. Multiple user information theory. *Proc. IEEE* **1980**, *68*, 1466–1483.
12. Ahlswede, R. Multiway Communication Channels. In Proceedings of the 2nd International Symposium Information Theory, Tsahkadsor, Armenia, /hlconference date, day and month 1971; pp. 23–52.
13. Liao, H.H.J. Multiple access channels. Ph.D. Thesis, University of Hawaii: Honolulu 1972.
14. Csiszár, I.; Körner, J. *Information Theory. Coding Theorems for Discrete Memoryless Systems*; Academic: London, UK, 1981.

15. El Gamal, A.; Kim, Y. Lecture notes on network information theory. Available on-line: http://arxiv.org/abs/1001.3404 (accessed on 20 January 2010).

16. Yeung, R.W. *Information Theory and Network Coding*; Springer: New York, NY, USA, 1981.

17. Liang, Y.; Poor, H.V. Multiple-access channels with confidential messages. *IEEE Trans. Inf. Theory* **2008**, *54*, 976–1002.

Appendix

A. Proof of the Converse Part of Theorem 2

In this section, we establish the converse part of Theorem 2: all the achievable (R_0, R_1, R_2) triples are contained in the set $\mathcal{R}^{(\mathcal{A})}$, *i.e.*, for any achievable triple, there exists random variables U, V_1, V_2, X_1, X_2, Y and Z, such that the inequalities in Theorem 2 hold, and $(U, V_1, V_2) \to (X_1, X_2) \to Y \to (Y_1, Y_2)$ forms a Markov chain. We will prove the inequalities of Theorem 2 in the remainder of this section.

(Proof of $0 \leq R_0 \leq I(U; Z)$**)** The proof of this inequality is as follows:

$$
\begin{aligned}
\frac{1}{N}H(W_0) &\overset{(a)}{\leq} \frac{1}{N}(I(W_0; Z^N) + \delta(P_{e2})) \\
&= \frac{1}{N}(H(Z^N) - H(Z^N|W_0) + \delta(P_{e2})) \\
&= \frac{1}{N}\left(\sum_{i=1}^{N}(H(Z_i|Z^{i-1}) - H(Z_i|W_0, Z^{i-1}) + \delta(P_{e2})\right) \\
&\leq \frac{1}{N}\left(\sum_{i=1}^{N}(H(Z_i|Z^{i-1}) - H(Z_i|W_0, Z^{i-1}, Y^{i-1}) + \delta(P_{e2})\right) \\
&\overset{(b)}{=} \frac{1}{N}\left(\sum_{i=1}^{N}(H(Z_i|Z^{i-1}) - H(Z_i|W_0, Y^{i-1}) + \delta(P_{e2})\right) \\
&\leq \frac{1}{N}\left(\sum_{i=1}^{N}(H(Z_i) - H(Z_i|W_0, Y^{i-1}) + \delta(P_{e2})\right) \\
&\overset{(c)}{=} \frac{1}{N}\left(\sum_{i=1}^{N}(H(Z_i) - H(Z_i|U_i) + \delta(P_{e2})\right) \\
&\overset{(d)}{=} \frac{1}{N}\left(\sum_{i=1}^{N}(H(Z_i|J = i) - H(Z_i|U_i, J = i) + \delta(P_{e2})\right) \\
&\overset{(e)}{=} H(Z_J|J) - H(Z_J|U_J, J) + \frac{1}{N}\delta(P_{e2}) \\
&\leq H(Z_J) - H(Z_J|U_J, J) + \frac{1}{N}\delta(P_{e2}) \\
&\overset{(f)}{=} H(Z) - H(Z|U) + \frac{1}{N}\delta(P_{e2}) \\
&= I(U; Z) + \frac{\delta(P_{e2})}{N}
\end{aligned}
\tag{A1}
$$

where (a) is from the Fano's inequality, (b) is from $Z^{i-1} \to (Y^{i-1}, W_0) \to Z_i$, (c) is from the definition that $U_i \triangleq (W_0, Y^{i-1})$, (d) is from J as a random variable (uniformly distributed over $\{1, 2, ..., N\}$), and

J is independent of Z_i and U_i, (e) is from J as uniformly distributed over $\{1, 2, ..., N\}$ and (f) is from the definitions that $Z \triangleq Z_J$ and $U \triangleq (U_J, J)$.

Letting $\epsilon \to 0$ and noting that $P_{e2} \le \epsilon$, $\frac{H(W_0)}{N} \ge R_0 - \epsilon$, it is easy to see that $0 \le R_0 \le I(U; Z)$.

(Proof of $0 \le R_1 \le I(X_1; Y | X_2, U)$**)** The proof of this inequality is as follows:

$$
\begin{aligned}
\frac{1}{N} H(W_1) \quad &\overset{(1)}{\le}\quad \frac{1}{N}(H(W_1|X_2^N, W_0) + \delta(P_{e1}) - H(W_1|Y^N, X_2^N, W_0)) \\
&=\quad \frac{1}{N}(I(W_1; Y^N | X_2^N, W_0) + \delta(P_{e1})) \\
&=\quad \frac{1}{N}(H(Y^N|X_2^N, W_0) - H(Y^N|X_2^N, W_0, W_1) + \delta(P_{e1})) \\
&\le\quad \frac{1}{N}(H(Y^N|X_2^N, W_0) - H(Y^N|X_2^N, W_0, W_1, X_1^N) + \delta(P_{e1})) \\
&\overset{(2)}{=}\quad \frac{1}{N}(H(Y^N|X_2^N, W_0) - H(Y^N|X_2^N, X_1^N) + \delta(P_{e1})) \\
&\overset{(3)}{=}\quad \frac{1}{N}(\sum_{i=1}^{N}(H(Y_i|X_2^N, W_0, Y^{i-1}) - H(Y_i|X_{1,i}, X_{2,i})) + \delta(P_{e1})) \\
&\le\quad \frac{1}{N}(\sum_{i=1}^{N}(H(Y_i|X_{2,i}, W_0, Y^{i-1}) - H(Y_i|X_{1,i}, X_{2,i})) + \delta(P_{e1})) \\
&\overset{(4)}{=}\quad \frac{1}{N}(\sum_{i=1}^{N}(H(Y_i|X_{2,i}, W_0, Y^{i-1}) - H(Y_i|X_{1,i}, X_{2,i}, W_0, Y^{i-1})) + \delta(P_{e1})) \\
&\overset{(5)}{=}\quad H(Y|X_2, U) - H(Y|X_1, X_2, U) + \frac{1}{N}\delta(P_{e1}) \\
&=\quad I(X_1; Y|X_2, U) + \frac{1}{N}\delta(P_{e1}) \qquad\qquad\qquad\qquad (A2)
\end{aligned}
$$

where (1) is from the Fano's inequality and the fact that W_1 is independent of X_2^N and W_0; (2) is from $(W_0, W_1) \to (X_2^N, X_1^N) \to Y^N$; (3) is from the discrete memoryless property of the channel; (4) is from $(W_0, Y^{i-1}) \to (X_{1,i}, X_{2,i}) \to Y_i$; and (5) is from the definitions that $X_1 \triangleq X_{1,J}$, $X_2 \triangleq X_{2,J}$, $Y \triangleq Y_J$, $U \triangleq (W_0, Y^{J-1}, J)$, where J is a random variable (uniformly distributed over $\{1, 2, ..., N\}$) and is independent of $X_{1,i}$, $X_{2,i}$, W_0, Y^{i-1} and Y_i.

Letting $\epsilon \to 0$ and noting that $P_{e1} \le \epsilon$, $\frac{H(W_1)}{N} \ge R_1 - \epsilon$, it is easy to see that $0 \le R_1 \le I(X_1; Y|X_2, U)$.

(Proof of $0 \le R_2 \le I(X_2; Y|X_1, U)$**)** The proof is similar to the proof of $0 \le R_1 \le I(X_1; Y|X_2, U)$, and it is omitted here.

(Proof of $0 \leq R_1 + R_2 \leq I(X_1, X_2; Y|U)$**)**

$$
\begin{aligned}
\frac{1}{N}H(W_1, W_2) &\overset{(a)}{\leq} \frac{1}{N}(I(W_1, W_2; Y^N|W_0) + \delta(P_{e1})) \\
&= \frac{1}{N}(H(Y^N|W_0) - H(Y^N|W_0, W_1, W_2) + \delta(P_{e1})) \\
&\leq \frac{1}{N}(H(Y^N|W_0) - H(Y^N|W_0, W_1, W_2, X_1^N, X_2^N) + \delta(P_{e1})) \\
&\overset{(b)}{=} \frac{1}{N}(H(Y^N|W_0) - H(Y^N|X_1^N, X_2^N) + \delta(P_{e1})) \\
&\overset{(c)}{=} \frac{1}{N}(\sum_{i=1}^{N}(H(Y_i|W_0, Y^{i-1}) - H(Y_i|X_{1,i}, X_{2,i})) + \delta(P_{e1})) \\
&\overset{(d)}{=} \frac{1}{N}(\sum_{i=1}^{N}(H(Y_i|W_0, Y^{i-1}) - H(Y_i|X_{1,i}, X_{2,i}, W_0, Y^{i-1})) + \delta(P_{e1})) \\
&\overset{(e)}{=} H(Y|U) - H(Y|X_1, X_2, U) + \frac{1}{N}\delta(P_{e1})) \\
&= I(X_1, X_2; Y|U) + \frac{\delta(P_{e1})}{N}
\end{aligned}
\tag{A3}
$$

where (a) is from the Fano's inequality and the fact that W_0 is independent of W_1 and W_2; (b) is from $(W_0, W_1, W_2) \rightarrow (X_1^N, X_2^N) \rightarrow Y^N$; (c) is from the discrete memoryless property of the channel; (d) is from $(W_0, Y^{i-1}) \rightarrow (X_{1,i}, X_{2,i}) \rightarrow Y_i$; and (e) is from the definitions that $X_1 \triangleq (X_{1,J}, J)$, $X_2 \triangleq (X_{2,J}, J)$, $Y \triangleq Y_J$ and $U \triangleq (W_0, Y^{J-1}, J)$.

Letting $\epsilon \rightarrow 0$ and noting that $P_{e1} \leq \epsilon$, $\frac{H(W_1)}{N} \geq R_1 - \epsilon$, $\frac{H(W_2)}{N} \geq R_2 - \epsilon$ and W_1 is independent of W_2, it is easy to see that $0 \leq R_1 + R_2 \leq I(X_1, X_2; Y|U)$.

(Proof of $R_0 + R_1 \leq I(X_1, X_2; Y|V_2)$**)** The proof is obtained by the following Equation (A4).

$$
\begin{aligned}
\frac{1}{N}H(W_0, W_1) &\overset{(a)}{\leq} \frac{1}{N}(I(W_0, W_1; Y^N|W_2) + \delta(P_{e1})) \\
&= \frac{1}{N}(H(Y^N|W_2) - H(Y^N|W_0, W_1, W_2) + \delta(P_{e1})) \\
&\leq \frac{1}{N}(H(Y^N|W_2) - H(Y^N|W_0, W_1, W_2, X_1^N, X_2^N) + \delta(P_{e1})) \\
&\overset{(b)}{=} \frac{1}{N}(H(Y^N|W_2) - H(Y^N|X_1^N, X_2^N) + \delta(P_{e1})) \\
&\overset{(c)}{=} \frac{1}{N}(\sum_{i=1}^{N}(H(Y_i|W_2, Y^{i-1}) - H(Y_i|X_{1,i}, X_{2,i})) + \delta(P_{e1})) \\
&\overset{(d)}{=} \frac{1}{N}(\sum_{i=1}^{N}(H(Y_i|W_2, Y^{i-1}) - H(Y_i|X_{1,i}, X_{2,i}, W_2, Y^{i-1})) + \delta(P_{e1})) \\
&\overset{(e)}{=} H(Y|V_2) - H(Y|X_1, X_2, V_2) + \frac{1}{N}\delta(P_{e1})) \\
&= I(X_1, X_2; Y|V_2) + \frac{\delta(P_{e1})}{N}
\end{aligned}
\tag{A4}
$$

where (a) is from the Fano's inequality and the fact that W_2 is independent of W_0 and W_1; (b) is from $(W_0, W_1, W_2) \rightarrow (X_1^N, X_2^N) \rightarrow Y^N$; (c) is from the discrete memoryless property of the channel;

(d) is from $(W_2, Y^{i-1}) \rightarrow (X_{1,i}, X_{2,i}) \rightarrow Y_i$; and (e) is from the definitions that $X_1 \triangleq (X_{1,J}, J)$, $X_2 \triangleq (X_{2,J}, J)$, $Y \triangleq Y_J$ and $V_2 \triangleq (W_2, Y^{J-1}, J)$.

Letting $\epsilon \rightarrow 0$ and noting that $P_{e1} \leq \epsilon$, $\frac{H(W_1)}{N} \geq R_1 - \epsilon$, $\frac{H(W_0)}{N} \geq R_0 - \epsilon$ and W_0 is independent of W_1, it is easy to see that $R_0 + R_1 \leq I(X_1, X_2; Y|V_2)$.

(Proof of $R_0 + R_2 \leq I(X_1, X_2; Y|V_1)$) The proof is analogous to the proof of $R_0 + R_1 \leq I(X_1, X_2; Y|V_2)$, and it is omitted here.

(Proof of $R_0 + R_1 + R_2 \leq I(X_1, X_2; Y)$)

$$
\begin{aligned}
\frac{1}{N} H(W_0, W_1, W_2) &\overset{(a)}{\leq} \frac{1}{N}(I(W_0, W_1, W_2; Y^N) + \delta(P_{e1})) \\
&= \frac{1}{N}(H(Y^N) - H(Y^N|W_0, W_1, W_2) + \delta(P_{e1})) \\
&\leq \frac{1}{N}(H(Y^N) - H(Y^N|W_0, W_1, W_2, X_1^N, X_2^N) + \delta(P_{e1})) \\
&\overset{(b)}{=} \frac{1}{N}(H(Y^N) - H(Y^N|X_1^N, X_2^N) + \delta(P_{e1})) \\
&\overset{(c)}{=} \frac{1}{N}\left(\sum_{i=1}^N (H(Y_i|Y^{i-1}) - H(Y_i|X_{1,i}, X_{2,i})) + \delta(P_{e1})\right) \\
&\leq \frac{1}{N}\left(\sum_{i=1}^N (H(Y_i) - H(Y_i|X_{1,i}, X_{2,i})) + \delta(P_{e1})\right) \\
&\overset{(d)}{\leq} H(Y) - H(Y|X_1, X_2, U) + \frac{1}{N}\delta(P_{e1})) \\
&= I(X_1, X_2; Y) + \frac{\delta(P_{e1})}{N} \qquad (A5)
\end{aligned}
$$

where (a) is from the Fano's inequality; (b) is from $(W_0, W_1, W_2) \rightarrow (X_1^N, X_2^N) \rightarrow Y^N$; (c) is from the discrete memoryless property of the channel; and (d) is from the definitions that $X_1 \triangleq (X_{1,J}, J)$, $X_2 \triangleq (X_{2,J}, J)$, $Y \triangleq Y_J$.

Letting $\epsilon \rightarrow 0$ and noting that $P_{e1} \leq \epsilon$, $\frac{H(W_0)}{N} \geq R_0 - \epsilon$, $\frac{H(W_1)}{N} \geq R_1 - \epsilon$, $\frac{H(W_2)}{N} \geq R_2 - \epsilon$, it is easy to see that $R_0 + R_1 + R_2 \leq I(X_1, X_2; Y)$.

The Markov chain, $(U, V_1, V_2) \rightarrow (X_1, X_2) \rightarrow Y \rightarrow Z$, is directly obtained from the definitions $U \triangleq (W_0, Y^{J-1}, J)$, $V_1 \triangleq (W_1, Y^{J-1}, J)$, $V_2 \triangleq (W_2, Y^{J-1}, J)$ $X_1 \triangleq (X_{1,J}, J)$, $X_2 \triangleq (X_{2,J}, J)$, $Y \triangleq Y_J$ and $Z \triangleq Z_J$.

The proof of the converse part of Theorem 2 is completed.

B. Proof of the Direct Part of Theorem 2

In this section, we establish the direct part of Theorem 2 (about existence). Suppose $(R_0, R_1, R_2) \in \mathcal{R}^\mathcal{A}$, we will show that (R_0, R_1, R_2) is achievable.

The coding scheme for Theorem 2 is in the following Figure A1. Now, the remainder of this section is organized as follows. Some preliminaries about typical sequences are introduced in Subsection B.1. The construction of the code is introduced in Subsection B.2. For any given $\epsilon > 0$, the proofs of $\frac{\log\|\mathcal{W}_1\|}{N} \geq R_1 - \epsilon$, $\frac{\log\|\mathcal{W}_2\|}{N} \geq R_2 - \epsilon$, $\frac{\log\|\mathcal{W}_0\|}{N} \geq R_0 - \epsilon$, $P_{e1} \leq \epsilon$ and $P_{e2} \leq \epsilon$ are given in Subsection B.3.

Figure A1. Coding scheme for MAC-DBC.
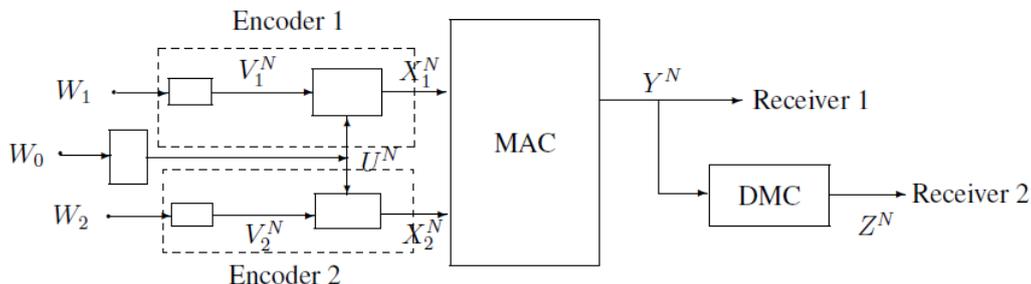
B.1. Preliminaries

- Given a probability mass function, $p_V(v)$, for any $\eta > 0$, let $T_V^N(\eta)$ be the strong typical set of all v^N, such that $|p_V(v) - \frac{c_{v^N}(v)}{N}| < \eta$ for all $v \in \mathcal{V}$, where $c_{v^N}(v)$ is the number of occurences of the letter v in the v^N. We say that the sequences, $v^N \in T_V^N(\eta)$, are V-**typical**.

- Analogously, given a joint probability mass function, $p_{VW}(v, w)$, for any $\eta > 0$, let $T_{VW}^N(\eta)$ be the joint strong typical set of all pairs (v^N, w^N), such that $|p_{VW}(v, w) - \frac{c_{v^N, w^N}(v, w)}{N}| < \eta$ for all $v \in \mathcal{V}$ and $w \in \mathcal{W}$, where $c_{v^N, w^N}(v, w)$ is the number of occurences of (v, w) in the pair of sequences (v^N, w^N). We say that the pairs of sequences, $(v^N, w^N) \in T_{VW}^N(\eta)$, are VW-**typical**.

- Moreover, w^N is called $W|V$-**generated** by v^N if v^N is V- typical and $(v^N, w^N) \in T_{VW}^N(\eta)$. For any given $v^N \in T_V^N(\eta)$, define $T_{W|V}^N(\eta) = \{w^N : w^N \text{ is } W|V\text{-generated by } v^N\}$.

- **Lemma 1** *For any $v^N \in T_V^N(\eta)$,*

$$2^{-N(H(V)+\eta^*)} \le p_{V^N}(v^N) \le 2^{-N(H(V)-\eta^*)}$$

where $\eta^* \to 0$ as $\eta \to 0$.

B.2. Coding Construction

Given a triple (R_0, R_1, R_2), choose a joint probability mass function, $p_{U, V_1, V_2, X_1, X_2, Y, Z}(u, v_1, v_2, x_1, x_2, y, z)$, such that

$$0 \le R_1 \le I(X_1; Y|X_2, U), \ \ 0 \le R_2 \le I(X_2; Y|X_1, U), \ \ R_1 + R_2 \le I(X_1, X_2; Y|U),$$

$$R_0 \le I(U; Z), \ \ R_0 + R_1 \le I(X_1, X_2; Y|V_2), \ \ R_0 + R_2 \le I(X_1, X_2; Y|V_1),$$

$$R_0 + R_1 + R_2 \le I(X_1, X_2; Y)$$

The message sets, $\mathcal{W}_0$, $\mathcal{W}_1$ and $\mathcal{W}_2$, satisfy the following conditions:

$$\frac{1}{N} \log \| \mathcal{W}_0 \| = R_0, \ \ \frac{1}{N} \log \| \mathcal{W}_1 \| = R_1 \ \ \frac{1}{N} \log \| \mathcal{W}_2 \| = R_2 \tag{A6}$$

Code-book generation:

- For a given $w_0 \in \mathcal{W}_0$, generate a corresponding $u^N(w_0)$ i.i.d., according to the probability mass function $p_U(u)$.
- For a given $w_1 \in \mathcal{W}_1$, generate a corresponding $v_1^N(w_1)$ i.i.d., according to the probability mass function $p_{V_1}(v_1)$.
- For a given $w_2 \in \mathcal{W}_2$, generate a corresponding $v_2^N(w_2)$ i.i.d., according to the probability mass function $p_{V_2}(v_2)$.
- $x_1^N(w_0, w_1)$ is generated according to a new discrete memoryless channel (DMC), with inputs $v_1^N(w_1)$ and $u^N(w_0)$ and output $x_1^N(w_0, w_1)$. The transition probability of this new DMC is $p_{X_1|V_1,U}(x_1|v_1, u)$.

   Similarly, $x_2^N(w_0, w_2)$ is generated according to a new discrete memoryless channel (DMC), with inputs $v_2^N(w_2)$ and $u^N(w_0)$ and output $x_2^N(w_0, w_2)$. The transition probability of this new DMC is $p_{X_2|V_2,U}(x_2|v_2, u)$.

Decoding scheme:

- (Receiver 1) Receiver 1 declares that messages, $\hat{w}_0$, $\hat{w}_1$ and $\hat{w}_2$, are sent if they are the unique messages, such that $(u^N(\hat{w}_0), v_1^N(\hat{w}_1), v_2^N(\hat{w}_2), x_1^N(\hat{w}_0, \hat{w}_1), x_2^N(\hat{w}_0, \hat{w}_2), y^N) \in T_{UV_1V_2X_1X_2Y}^N(\epsilon)$; otherwise, it declares an error.
- (Receiver 2) Receiver 2 declares that a message $\hat{w}_0$ is sent if it is the unique message, such that $(u^N(\hat{w}_0), z^N) \in T_{UZ}^N(\epsilon)$; otherwise it declares an error.

B.3. Achievability Proof

By using the above Equation (A6), it is easy to verify that $\frac{\log\|\mathcal{W}_0\|}{N} \geq R_0 - \epsilon$, $\frac{\log\|\mathcal{W}_1\|}{N} \geq R_1 - \epsilon$ and $\frac{\log\|\mathcal{W}_2\|}{N} \geq R_2 - \epsilon$. It remains to show that $P_{e1} \leq \epsilon$ and $P_{e2} \leq \epsilon$; see the following.

Without loss of generality, assume that $w_0 = 1$, $w_1 = 1$ and $w_2 = 1$ are sent.

2.3.1. $P_{e2} \leq \epsilon$

For receiver 2, define the events:

$$B_1 = \{(u^N(1), z^N) \notin T_{UZ}^N(\eta)\}$$

$$B_2 = \{(u^N(w_0), z^N) \in T_{UZ}^N(\eta)\} \text{ for some } w_0 \neq 1$$

The probability of error for receiver 2 is then upper bounded by:

$$P_{e2} = Pr\{B_1 \cup B_2\} \leq Pr\{B_1\} + Pr\{B_2\} \tag{A7}$$

By using LLN, the first term, $Pr\{B_1\} \to 0$ as $N \to \infty$. On the other hand, by using the packing lemma [15, p. 53-54], $Pr\{B_2\} \to 0$ as $N \to \infty$ if $R_0 \leq I(U; Z)$.

Therefore, by choosing sufficiently large N, we have $P_{e2} \leq \epsilon$.

2.3.2. $P_{e1} \leq \epsilon$

The proof of $P_{e1} \leq \epsilon$ is as follows.

Define the sets:

$$A_1 = \{(u^N(1), v_1^N(1), v_2^N(1), x_1^N(1,1), x_2^N(1,1), y^N) \notin T_{UV_1V_2X_1X_2Y}^N(\epsilon)\}$$

$$A_2 = \{(u^N(1), v_1^N(1), v_2^N(w_2), x_1^N(1,1), x_2^N(1,w_2), y^N) \in T_{UV_1V_2X_1X_2Y}^N(\epsilon)\} \text{ for some } w_2 \neq 1$$

$$A_3 = \{(u^N(1), v_1^N(w_1), v_2^N(1), x_1^N(1,w_1), x_2^N(1,1), y^N) \in T_{UV_1V_2X_1X_2Y}^N(\epsilon)\} \text{ for some } w_1 \neq 1$$

$$A_4 = \{(u^N(1), v_1^N(w_1), v_2^N(w_2), x_1^N(1,w_1), x_2^N(1,w_2), y^N) \in T_{UV_1V_2X_1X_2Y}^N(\epsilon)\} \text{ for some } w_1 \neq 1, w_2 \neq 1$$

$$A_5 = \{(u^N(w_0), v_1^N(1), v_2^N(1), x_1^N(w_0,1), x_2^N(w_0,1), y^N) \in T_{UV_1V_2X_1X_2Y}^N(\epsilon)\} \text{ for some } w_0 \neq 1$$

$$A_6 = \{(u^N(w_0), v_1^N(1), v_2^N(w_2), x_1^N(w_0,1), x_2^N(w_0,w_2), y^N) \in T_{UV_1V_2X_1X_2Y}^N(\epsilon)\} \text{ for some } w_0 \neq 1, w_2 \neq 1$$

$$A_7 = \{(u^N(w_0), v_1^N(w_1), v_2^N(1), x_1^N(w_0,w_1), x_2^N(w_0,1), y^N) \in T_{UV_1V_2X_1X_2Y}^N(\epsilon)\} \text{ for some } w_0 \neq 1, w_1 \neq 1$$

$$A_8 = \{(u^N(w_0), v_1^N(w_1), v_2^N(w_2), x_1^N(w_0,w_1), x_2^N(w_0,w_2), y^N) \in T_{UV_1V_2X_1X_2Y}^N(\epsilon)\} \text{ for some } (w_0, w_1, w_2) \neq 1$$

The probability of error for receiver 1 is then upper bounded by:

$$
\begin{aligned}
P_{e2} &= Pr\{A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5 \cup A_6 \cup A_7 \cup A_8\} \\
&\leq \sum_{i=1}^{8} Pr\{A_i\}
\end{aligned}
\tag{A8}
$$

By using LLN, the first term, $Pr\{A_1\} \to 0$ as $N \to \infty$.

For the second term, by using the packing lemma ([15][p. 53–54]), $Pr\{A_2\} \to 0$ as $N \to \infty$ if

$$
\begin{aligned}
R_2 &\leq I(V_2, X_2; Y|U, V_1, X_1) \\
&\overset{(1)}{=} H(Y|U, X_1) - H(Y|X_1, X_2, U) = I(X_2; Y|U, X_1)
\end{aligned}
\tag{A9}
$$

where (1) is from $V_1 \to (U, X_1) \to Y$ and $(V_1, V_2) \to (X_1, X_2, U) \to Y$.

Analogously, for the third term, by using the packing lemma, $Pr\{A_3\} \to 0$ as $N \to \infty$ if $R_1 \leq I(X_1; Y|U, X_2)$.

For the fourth term, by using the packing lemma, $Pr\{A_4\} \to 0$ as $N \to \infty$ if $R_1 + R_2 \leq I(X_1; Y|U, X_2)$.

$$
\begin{aligned}
R_1 + R_2 &\leq I(V_1, V_2, X_1, X_2; Y|U) \\
&\overset{(2)}{=} H(Y|U) - H(Y|X_1, X_2, U) = I(X_1, X_2; Y|U)
\end{aligned}
\tag{A10}
$$

where (2) is from $(V_1, V_2) \to (X_1, X_2, U) \to Y$.

For the fifth term, $Pr\{A_5\} \to 0$ as $N \to \infty$ if $R_0 \leq I(U; Y)$.

For the sixth term, by using the packing lemma, $Pr\{A_6\} \to 0$ as $N \to \infty$ if

$$
\begin{aligned}
R_0 + R_2 &\leq I(U, V_2, X_1, X_2; Y|V_1) \\
&\overset{(3)}{=} H(Y|V_1) - H(Y|X_1, X_2, V_1) = I(X_1, X_2; Y|V_1)
\end{aligned}
\tag{A11}
$$

where (3) is from $(U, V_2) \to (X_1, X_2, V_1) \to Y$.

Analogously, the seventh term, $Pr\{A_7\} \to 0$ as $N \to \infty$ if $R_0 + R_1 \le I(X_1, X_2; Y | V_2)$.

For the eighth term, by using the packing lemma, $Pr\{A_8\} \to 0$ as $N \to \infty$ if $R_0 + R_1 + R_2 \le I(X_1, X_2; Y)$.

Therefore, by choosing sufficiently large N, we have $P_{e1} \le \epsilon$.

The proof of the direct part of Theorem 2 is completed.

C. Proof of the Convexity of $\mathcal{R}^{\mathcal{A}}$

Let $(R_0', R_1', R_2') \in \mathcal{R}^{(\mathcal{A})}$, *i.e.*, (R_0', R_1', R_2'), satisfy the following conditions:

$$R_1' \le I(X_1^{(1)}; Y^{(1)} | X_2^{(1)}, U^{(1)}), R_2' \le I(X_2^{(1)}; Y^{(1)} | X_1^{(1)}, U^{(1)}), R_1' + R_2' \le I(X_1^{(1)}, X_2^{(1)}; Y^{(1)} | U^{(1)}),$$

$$0 \le R_0' \le I(U^{(1)}; Z^{(1)}), R_0' + R_1' \le I(X_1^{(1)}, X_2^{(1)}; Y^{(1)} | V_2^{(1)}), R_0' + R_2' \le I(X_1^{(1)}, X_2^{(1)}; Y^{(1)} | V_1^{(1)}),$$

$$R_0' + R_1' + R_2' \le I(X_1^{(1)}, X_2^{(1)}; Y^{(1)})$$

Let $(R_0'', R_1'', R_2'') \in \mathcal{R}^{(\mathcal{A})}$, *i.e.*, (R_0'', R_1'', R_2''), satisfy the following conditions:

$$R_1'' \le I(X_1^{(2)}; Y^{(2)} | X_2^{(2)}, U^{(2)}), R_2'' \le I(X_2^{(2)}; Y^{(2)} | X_1^{(2)}, U^{(2)}), R_1'' + R_2'' \le I(X_1^{(2)}, X_2^{(2)}; Y^{(2)} | U^{(2)})$$

$$0 \le R_0'' \le I(U^{(2)}; Z^{(2)}), R_0'' + R_1'' \le I(X_1^{(2)}, X_2^{(2)}; Y^{(2)} | V_2^{(2)}), R_0'' + R_2'' \le I(X_1^{(2)}, X_2^{(2)}; Y^{(2)} | V_1^{(2)})$$

$$R_0'' + R_1'' + R_2'' \le I(X_1^{(2)}, X_2^{(2)}; Y^{(2)})$$

Let Q be a switch function, such that $Pr\{Q = 1\} = \theta$ and $Pr\{Q = 2\} = 1 - \theta$, where $0 \le \theta \le 1$. Q is independent of all the random variables.

Define $V_1 = V_1^{(Q)} Q, V_2 = V_2^{(Q)} Q, U = U^{(Q)} Q, X_1 = X_1^{(Q)} Q, X_2 = X_2^{(Q)} Q, Y = Y^{(Q)}, Z = Z^{(Q)}$. Then we have:

$$
\begin{aligned}
I(X_1; Y | X_2, U) &= I(X_1^{(Q)}; Y^{(Q)} | X_2^{(Q)}, U^{(Q)}, Q) \\
&= \theta I(X_1^{(1)}; Y^{(1)} | X_2^{(1)}, U^{(1)}, Q = 1) + (1 - \theta) I(X_1^{(2)}; Y^{(2)} | X_2^{(2)}, U^{(2)}, Q = 2) \\
&= \theta I(X_1^{(1)}; Y^{(1)} | X_2^{(1)}, U^{(1)}) + (1 - \theta) I(X_1^{(2)}; Y^{(2)} | X_2^{(2)}, U^{(2)}) \quad (A12)
\end{aligned}
$$

$$
\begin{aligned}
I(X_2; Y | X_1, U) &= I(X_2^{(Q)}; Y^{(Q)} | X_1^{(Q)}, U^{(Q)}, Q) \\
&= \theta I(X_2^{(1)}; Y^{(1)} | X_1^{(1)}, U^{(1)}, Q = 1) + (1 - \theta) I(X_2^{(2)}; Y^{(2)} | X_1^{(2)}, U^{(2)}, Q = 2) \\
&= \theta I(X_2^{(1)}; Y^{(1)} | X_1^{(1)}, U^{(1)}) + (1 - \theta) I(X_2^{(2)}; Y^{(2)} | X_1^{(2)}, U^{(2)}) \quad (A13)
\end{aligned}
$$

$$
\begin{aligned}
I(X_1, X_2; Y) &\ge I(X_1^{(Q)}, X_2^{(Q)}; Y^{(Q)} | Q) \\
&= \theta I(X_1^{(1)}, X_2^{(1)}; Y^{(1)} | Q = 1) + (1 - \theta) I(X_1^{(2)}, X_2^{(2)}; Y^{(2)} | Q = 2) \\
&= \theta I(X_1^{(1)}, X_2^{(1)}; Y^{(1)}) + (1 - \theta) I(X_1^{(2)}, X_2^{(2)}; Y^{(2)}) \quad (A14)
\end{aligned}
$$

$$
\begin{aligned}
I(X_1, X_2; Y | U) &= I(X_1^{(Q)}, X_2^{(Q)}; Y^{(Q)} | U^{(Q)}, Q) \\
&= \theta I(X_1^{(1)}, X_2^{(1)}; Y^{(1)} | U^{(1)}, Q = 1) + (1 - \theta) I(X_1^{(2)}, X_2^{(2)}; Y^{(2)} | U^{(2)}, Q = 2) \\
&= \theta I(X_1^{(1)}, X_2^{(1)}; Y^{(1)} | U^{(1)}) + (1 - \theta) I(X_1^{(2)}, X_2^{(2)}; Y^{(2)} | U^{(2)}) \quad (A15)
\end{aligned}
$$

$$
\begin{aligned}
I(X_1, X_2; Y|V_1) &= I(X_1^{(Q)}, X_2^{(Q)}; Y^{(Q)}|V_1^{(Q)}, Q) \\
&= \theta I(X_1^{(1)}, X_2^{(1)}; Y^{(1)}|V_1^{(1)}, Q=1) + (1-\theta)I(X_1^{(2)}, X_2^{(2)}; Y^{(2)}|V_1^{(2)}, Q=2) \\
&= \theta I(X_1^{(1)}, X_2^{(1)}; Y^{(1)}|V_1^{(1)}) + (1-\theta)I(X_1^{(2)}, X_2^{(2)}; Y^{(2)}|V_1^{(2)}) \qquad (A16)
\end{aligned}
$$

$$
\begin{aligned}
I(X_1, X_2; Y|V_2) &= I(X_1^{(Q)}, X_2^{(Q)}; Y^{(Q)}|V_2^{(Q)}, Q) \\
&= \theta I(X_1^{(1)}, X_2^{(1)}; Y^{(1)}|V_2^{(1)}, Q=1) + (1-\theta)I(X_1^{(2)}, X_2^{(2)}; Y^{(2)}|V_2^{(2)}, Q=2) \\
&= \theta I(X_1^{(1)}, X_2^{(1)}; Y^{(1)}|V_2^{(1)}) + (1-\theta)I(X_1^{(2)}, X_2^{(2)}; Y^{(2)}|V_2^{(2)}) \qquad (A17)
\end{aligned}
$$

$$
\begin{aligned}
I(U^{(Q)}; Z^{(Q)}) &\geq I(U^{(Q)}; Z^{(Q)}|Q) \\
&= \theta I(U^{(1)}; Z^{(1)}|Q=1) + (1-\theta)I(U^{(2)}; Z^{(2)}|Q=2) \\
&= \theta I(U^{(1)}; Z^{(1)}) + (1-\theta)I(U^{(2)}; Z^{(2)}) \qquad (A18)
\end{aligned}
$$

From Equations (A12)–(A18), it is easy to see that $(\theta R_1' + (1-\theta)R_1'', \theta R_2' + (1-\theta)R_2'', \theta R_0' + (1-\theta)R_0'') \in \mathcal{R}^{(\mathcal{A})}$, and therefore, $\mathcal{R}^{(\mathcal{A})}$ is convex.

D. Size Constraints of the Auxiliary Random Variables in Theorem 2

By using the support lemma (see [14], p.310), it suffices to show that the random variables U, A and K can be replaced by new ones, preserving the Markovity $(U, V_1, V_2) \rightarrow (X_1, X_2) \rightarrow Y \rightarrow Z$ and the characters $I(U; Z)$, $I(X_2; Y|X_1, U)$, $I(X_1; Y|X_2, U)$, $I(X_1, X_2; Y|U)$, $I(X_1, X_2; Y|V_2)$, $I(X_1, X_2; Y|V_1)$, and furthermore, the range of the new U, A and K satisfies:

$$
\|\mathcal{U}\| \leq \|\mathcal{X}_1\|\|\mathcal{X}_2\| + 2
$$

$$
\|\mathcal{V}_1\| \leq \|\mathcal{X}_1\|\|\mathcal{X}_2\|
$$

$$
\|\mathcal{V}_2\| \leq \|\mathcal{X}_1\|\|\mathcal{X}_2\|
$$

The proof of which is in the reminder of this section.

Let

$$
\bar{p} = p_{X_1X_2}(x_1, x_2) \qquad (A19)
$$

Define the following continuous scalar functions of $\bar{p}$:

$$
f_{X_1X_2}(\bar{p}) = p_{X_1X_2}(x_1, x_2), \quad f_Y(\bar{p}) = H(Y), \quad f_{Y|X_1}(\bar{p}) = H(Y|X_1), \quad f_{Y|X_2}(\bar{p}) = H(Y|X_2)
$$

Since there are $\|\mathcal{X}_1\|\|\mathcal{X}_2\| - 1$ functions of $f_{X_1X_2}(\bar{p})$, the total number of the continuous scalar functions of $\bar{p}$ is $\|\mathcal{X}_1\|\|\mathcal{X}_2\|+2$.

Let $\bar{p}_{X_1X_2|U} = Pr\{X_1 = x_1, X_2 = x_2|U = u\}$. With these distributions $\bar{p}_{X_1X_2|U}$, we have

$$
p_{X_1X_2}(x_1, x_2) = \sum_{u \in \mathcal{U}} p(U = u)f_{X_1X_2}(\bar{p}_{X_1X_2|U}) \qquad (A20)
$$

$$
H(Y|U) = \sum_{u \in \mathcal{U}} p(U = u)f_Y(\bar{p}_{X_1X_2|U}) \qquad (A21)
$$

$$H(Y|X_1, U) = \sum_{u \in \mathcal{U}} p(U = u) f_{Y|X_1}(\bar{p}_{X_1 X_2 | U}) \tag{A22}$$

$$H(Y|X_2, U) = \sum_{u \in \mathcal{U}} p(U = u) f_{Y|X_2}(\bar{p}_{X_1 X_2 | U}) \tag{A23}$$

According to the support lemma ([14], p.310), the random variable, U, can be replaced by new ones, such that the new U takes at most $\|\mathcal{X}_1\|\|\mathcal{X}_2\| + 2$ different values and the expressions (A20)–(A23) are preserved.

Similarly, we can prove that $\|\mathcal{V}_1\| \le \|\mathcal{X}_1\|\|\mathcal{X}_2\|$ and $\|\mathcal{V}_2\| \le \|\mathcal{X}_1\|\|\mathcal{X}_2\|$. The proof is omitted here.

E. Proof of Theorem 3

E.1. Proof of the Achievability The achievability proof follows by computing the mutual information terms in Theorem 2 with the following joint distributions:

$$U \sim \mathcal{N}(0, \alpha(P_1 + P_2)), \; V_1 \sim \mathcal{N}(0, (1 - \alpha)P_1) \text{ and } V_2 \sim \mathcal{N}(0, (1 - \alpha)P_2)$$

$$X_1 = \sqrt{\frac{P_1}{P_1 + P_2}} U + V_1 \text{ and } X_2 = \sqrt{\frac{P_2}{P_1 + P_2}} U + V_2$$

U is independent of V_1 and V_2.

E.2. Proof of the Converse

The proof of $R_0 \le \frac{1}{2} \log(1 + \frac{\alpha(\sigma_{x,1}^2 + \sigma_{x,2}^2)}{(1-\alpha)(\sigma_{x,1}^2 + \sigma_{x,2}^2) + \sigma_{n,1}^2 + \sigma_{n,2}^2})$ and $R_1 + R_2 \le \frac{1}{2} \log(1 + \frac{(1-\alpha)(\sigma_{x,1}^2 + \sigma_{x,2}^2)}{\sigma_{n,1}^2})$ are from the proof of the Gaussian broadcast channel [15], and it is omitted here.

The proof of $R_1 \le \frac{1}{2} \log(1 + \frac{(1-\alpha)\sigma_{x,1}^2}{\sigma_{n,1}^2})$, $R_2 \le \frac{1}{2} \log(1 + \frac{(1-\alpha)\sigma_{x,2}^2}{\sigma_{n,1}^2})$ and $R_0 + R_1 + R_2 \le \frac{1}{2} \log(1 + \frac{\sigma_{x,1}^2 + \sigma_{x,2}^2}{\sigma_{n,1}^2})$ are from the proof of the Gaussian multiple-access channel [16], and it is omitted here.

Then, it remains to show that $R_0 + R_1 \le \frac{1}{2} \log(1 + \frac{\sigma_{x,1}^2 + \alpha\sigma_{x,2}^2}{\sigma_{n,1}^2})$ and $R_0 + R_2 \le \frac{1}{2} \log(1 + \frac{\sigma_{x,2}^2 + \alpha\sigma_{x,1}^2}{\sigma_{n,1}^2})$. The proof of these two inequalities are analogous to ([17] [p. 1000–1001]), and therefore, we omit the proof here.

The proof of Theorem 3 is completed.