# An Application of a Wringing Lemma to the Multiple Access Channel with Cooperative Encoders

Shirin Saeedi Bidokhti
Lehrstuhl für Nachrichtentechnik
Technische Universität München, Germany
shirin.saeedi@tum.de

Gerhard Kramer
Lehrstuhl für Nachrichtentechnik
Technische Universität München, Germany
gerhard.kramer@tum.de

***Abstract*—The problem of communicating over a multiple access channel with cooperative encoders is studied. A new upper bound is derived on the capacity which is motivated by the regime of operation where the relays start to cooperate. The proof technique is based on a wringing lemma by Dueck and Ahlswede which was used for the multiple description problem with no excess rate. Previous upper bounds are shown to be loose in general, and may be improved.**

## I. INTRODUCTION

Consider a multiple access channel (MAC) with cooperative encoders, where cooperation is facilitated through the two-hop network of Fig 1. Over this network, the source communicates with a sink with the help of two relay nodes that have no information of their own to communicate. The problem of interest is a special case of the diamond network [1], where the broadcast channel is modelled by independent bit-pipes.

The problem was initially studied in [2] where lower and upper bounds were derived on the ultimate rate of communication. The bounds are improved in the recent works of [3], [4]. In particular, the cut-set bound is shown to be loose for a Gaussian MAC and a binary adder MAC for certain regimes of the bit-pipe capacities.

The problem of finding the capacity of this network is unresolved in general. The underlying challenge may be described as follows. In order to fully utilize the MAC to the receiver, we would ideally like full cooperation between the relays. On the other hand, in order to communicate the maximum amount of information and better use the diversity that is offered by the relays, we would like to send independent information to the relays over the broadcast channel.

In this work, we use a wringing lemma [5], [6] to study the regime of operation where independent inputs to the MAC stop being optimal, and cooperation between the relay nodes becomes necessary. Using this technique, we show through an example that previous bounds (the cut set bound in [2] and the bound in [4]) are not generally tight and may be improved.

The paper is organized as follows. We first formally describe the problem in Section II. In Section III, we study a regime of operation where the target communication rate is close to the total capacity of the bit pipes. We find necessary and sufficient conditions for achievability of such rates. As a by-product, we show that the upper bounds in [2], [4] are not always tight. Finally, in Section IV we use a wringing lemma to characterize a new upper bound on the achievable rate.

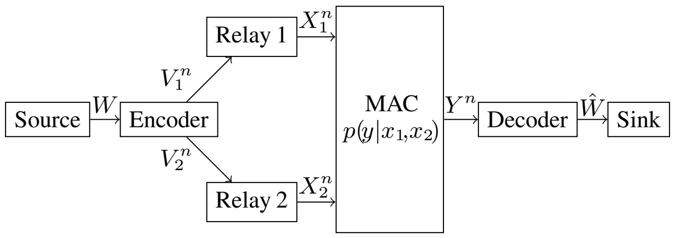


Fig. 1: Problem setup

## II. PROBLEM FORMULATION

### A. Notation

We use standard notation for random variables, e.g., $X$, probabilities, e.g., $p_X(x)$ or $p(x)$, entropies, e.g., $H(X)$ and $H(X|Y)$, and mutual information, e.g., $I(X;Y)$. We denote the sequence $X_1, \ldots, X_n$ by $X^n$. Sets are denoted by script letters.

### B. Model

Consider the diamond network in Fig. 1. A source communicates a message of rate $R$ to a sink. The source is connected to two relays via noiseless bit-pipes of capacities $C_1$ and $C_2$, and the relays communicate with the receiver over a MAC.

The source encodes a message $W$ with $nR$ bits into a sequence $V_1^n$, which is available at encoder 1, and a sequence $V_2^n$, which is available at encoder 2. $V_1^n$ and $V_2^n$ are such that $H(V_1^n) \leq nC_1$ and $H(V_2^n) \leq nC_2$.

Each relay $i$, $i = 1, 2$, maps its $V_i^n$ into a sequence $X_i^n$ which is sent over the MAC. The MAC is characterized by its input alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$, output alphabet $\mathcal{Y}$, and transition probabilities $p(y|x_1, x_2)$, for each $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$. From the received sequence, the sink puts out an estimate $\hat{W}$ of $W$.

We are interested in finding the highest rate $R$ that permits arbitrarily small positive error probability $\Pr(\hat{W} \neq W)$.

## III. THE CUT-SET BOUND IS NOT TIGHT

### A. The regime of operation where $R \approx C_1 + C_2$

We find necessary and sufficient conditions for the rate $R = C_1 + C_2$ to be approachable. Our approach is similar to and motivated by [5] that treats the multiple description problem with no excess rate.

**Theorem 1.** *The rate $R = C_1 + C_2$ is achievable if and only if there exists a pmf $p(u)p(x_1|u)p(x_2|u)p(y|x_1, x_2)$ so that the following inequalities hold.*

$$C_1 \leq I(X_1; Y|X_2U) \tag{1}$$
$$C_2 \leq I(X_2; Y|X_1U) \tag{2}$$
$$C_1 + C_2 \leq I(X_1X_2; Y|U) \tag{3}$$

*In the above characterization, we have $U \in \mathcal{U}$ where $|\mathcal{U}| \leq 2$.*

The sufficiency part of Theorem 1 follows from [2, Theorem 1]. We prove the necessity next; i.e., we prove that the conditions in Theorem 1 are necessary for $R = C_1 + C_2$. Although these conditions resemble those of the standard cut-set bound, we remark that the optimization is over the product distribution $p(u)p(x_1|u)p(x_2|u)$ rather than the joint distribution $p(u, x_1, x_2)$.

*Proof of the necessity of the conditions in* (1)-(3)*:* First, we have the following chain of inequalities

$$\begin{aligned} nR &\leq H(V_1^n V_2^n) \\ &= H(V_1^n) + H(V_2^n) - I(V_1^n; V_2^n) \\ &\leq nC_1 + nC_2 - I(X_1^n; X_2^n) \end{aligned} \tag{4}$$

where the last step follows because $I(V_1^n; V_2^n) \geq I(X_1^n; X_2^n)$ by the data processing inequality.

So in order to approach $R = C_1 + C_2 - \gamma$ for any $\gamma > 0$, we need to have

$$I(X_1^n; X_2^n) \leq n\gamma. \tag{5}$$

We now use a wringing lemma [5] to construct a random variable $U$ and make $X_1^n$ and $X_2^n$ conditionally almost independent on a letter-by-letter basis. More precisely, we use the following lemma.

**Lemma 1** (Dueck)**.** *Suppose $I(X_1^n; X_2^n) \leq \sigma$. For any $\delta > 0$, there exist $k$ indices $t_1, \ldots, t_k \in \{1, \ldots, n\}$ such that*

$$I(X_{1,i}; X_{2,i}|U) \leq \delta \quad \forall i \in \{1, \ldots, n\}, \tag{6}$$

*where $U = (X_{1,t_1}, \ldots, X_{1,t_k}, X_{2,t_1}, \ldots, X_{2,t_k})$, and $k < \frac{\sigma}{\delta}$.*

Setting $\sigma = n\gamma$ and $\delta = \sqrt{\gamma}$, Lemma 1 provides us with a random variable $U$ such that (6) is satisfied and

$$\begin{aligned} H(U) &= H(X_{1,t_1}, \ldots, X_{1,t_k}, X_{2,t_1}, \ldots, X_{2,t_k}) \\ &\leq k \log |\mathcal{X}_1||\mathcal{X}_2| \\ &< \frac{n\gamma}{\delta} \log |\mathcal{X}_1||\mathcal{X}_2| \\ &\leq n\sqrt{\gamma} \log |\mathcal{X}_1||\mathcal{X}_2|. \end{aligned} \tag{7}$$

Now, for every $\epsilon > 0$, we have

$$\begin{aligned} nR &\leq H(V_1^n V_2^n) \\ &= I(V_1^n V_2^n; Y^n) - H(V_1^n V_2^n|Y^2) \\ &\overset{(a)}{\leq} I(V_1^n V_2^n; Y^n) - n\epsilon \\ &= I(V_2^n; Y^n) + I(V_1^n; Y^n|V_2^n) - n\epsilon \\ &\overset{(b)}{\leq} I(V_2^n; Y^n) + I(V_1^n X_1^n; Y^n|V_2^n X_2^n) - n\epsilon \\ &\leq nC_2 + I(X_1^n; Y^n|X_2^n) - n\epsilon \\ &= nC_2 + I(UX_1^n; Y^n|X_2^n) - n\epsilon \\ &= nC_2 + I(X_1^n; Y^n|X_2^n U) + I(U; Y^n) - n\epsilon \\ &\leq nC_2 + I(X_1^n; Y^n|X_2^n U) + H(U) - n\epsilon \\ &\overset{(c)}{<} nC_2 + I(X_1^n; Y^n|X_2^n U) \\ &\quad + n\sqrt{\gamma} \log |\mathcal{X}_1||\mathcal{X}_2| - n\epsilon \\ &\leq nC_2 + \sum_{i=1}^{n} I(X_{1i}; Y_i|X_{2i}U) \\ &\quad + n\sqrt{\gamma} \log |\mathcal{X}_1||\mathcal{X}_2| - n\epsilon. \end{aligned}$$

In the above chain of inequalities $(a)$ follows from Fano's inequality, $(b)$ follows because each $X_i^n$, $i = 1, 2$, is a function of $V_i^n$, and $(c)$ follows from (7).

Similarly, we have

$$nR < nC_1 + \sum_{i=1}^{n} I(X_{2i}; Y_i|X_{1i}U) + n\sqrt{\gamma} \log |\mathcal{X}_1||\mathcal{X}_2| - n\epsilon,$$

and

$$\begin{aligned} nR &< I(X_1^n, X_2^n; Y^n) \\ &\leq \sum_{i=1}^{n} I(X_{1i}, X_{2i}; Y_i|U) + n\sqrt{\gamma} \log |\mathcal{X}_1||\mathcal{X}_2| - n\epsilon. \end{aligned}$$

Finally, a standard time sharing argument and small $\gamma$ and $\epsilon$ give the converse part of Theorem 1. ∎

*B. Example: the binary adder MAC*

Consider a binary adder MAC, where $X_1, X_2$ are the binary inputs and $Y = X_1 + X_2$ is its ternary output. Let $C_1 = C_2 = C$. We show that the bounds in [2], [4] are not tight for some ranges of $C$.

For this symmetric diamond network, [4] gives an upper bound which does not match the lower bound for $0.75 \leq C \leq 0.7928$. The regime where

$$C \leq h_2\left(\frac{1}{1 + \sqrt{2}}\right) - \frac{1}{2 + 2\sqrt{2}} \approx 0.7716$$

is particularly interesting in that the upper bound is characterized by $R \leq 2C$ and matches the cut-set bound. We wonder if the bound is tight in this regime.

We apply Theorem 1 and characterize all $C$ for which $R = 2C$ is approachable. This is given by

$$C \leq \max_{p(u)p(x_1|u)p(x_2|u)} \min \left\{ H(X_1|U), H(X_2|U), \frac{1}{2} H(Y|U) \right\}.$$

It is easy to calculate the right hand side (RHS) and see its equality to 0.75. Hence, the bounds in [2], [4] are not tight for

$$0.75 < C \leq .7716.$$

## IV. An upper bound

We generalize the proof technique of Section III to find a new upper bound. For simplicity, consider $|\mathcal{X}_1| = |\mathcal{X}_2|$ and relabel so that $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{X}$ for some $\mathcal{X}$.

**Theorem 2.** *The rate $R$ is achievable only if for every $\gamma$, $0 \leq \gamma \leq 1$, there exists a pmf $p(u, x_1, x_2, y) = p(u, x_1, x_2)p(y|x_1, x_2)$ for which the following conditions hold:*

$$I(X_1; X_2|U) \leq \gamma \log |\mathcal{X}| \tag{8}$$

$$R \leq C_1 + C_2 \tag{9}$$

$$R \leq C_2 + I(X_1; Y|X_2U) + \frac{1-\gamma}{\gamma}(C_1 + C_2 - R) \tag{10}$$

$$R \leq C_1 + I(X_2; Y|X_1U) + \frac{1-\gamma}{\gamma}(C_1 + C_2 - R) \tag{11}$$

$$2R \leq C_1 + C_2 + I(X_1X_2; Y|U) + 2\frac{1-\gamma}{\gamma}(C_1 + C_2 - R) \tag{12}$$

$$R \leq I(X_1, X_2; Y) \tag{13}$$

*In the above bound, $U$ takes its values in $\mathcal{U}$ that can be chosen to satisfy $|\mathcal{U}| \leq \min\{|\mathcal{X}_1||\mathcal{X}_2| + 2, |\mathcal{Y}| + 4\}$.*

Before proving the theorem, we state the following reformulation of Lemma 1. The proof is the same as that of Lemma 1 and is given in the Appendix.

**Proposition 1.** *Suppose $I(X_1^n; X_2^n) \leq \sigma$. For any $\delta > 0$, there exist $k$ indices $t_1, \dots, t_k$ and random variables $U_1 = X_{1,t_1}, \dots, X_{1,t_k}$, $U_2 = X_{2,t_1}, \dots, X_{2,t_k}$, and $U = (U_1, U_2)$, such that*

$$I(X_{1,i}; X_{2,i}|U) \leq \delta, \quad \forall i = 1, \dots, n \tag{14}$$

$$I(U_1; U_2) > k\delta, \tag{15}$$

$$k < \frac{\sigma}{\delta}. \tag{16}$$

*Proof of Theorem 2:* The idea is to constrain the input distribution. From inequality (4) we have

$$I(X_1^n; X_2^n) \leq n(C_1 + C_2 - R).$$

We replace the $\sigma$ in Proposition 1 by $n(C_1+C_2-R)$ and obtain that for every $\delta > 0$, there exist $k$ indices $t_1, \dots, t_k$ and random variables $U_1 = X_{1,t_1}, \dots, X_{1,t_k}$, $U_2 = X_{2,t_1}, \dots, X_{2,t_k}$, and $U = (U_1, U_2)$, such that

$$I(X_{1,i}; X_{2,i}|U) \leq \delta \quad \forall i = 1, \dots, n \tag{17}$$

$$I(U_1; U_2) > k\delta, \tag{18}$$

$$k < \frac{n(C_1 + C_2 - R)}{\delta}. \tag{19}$$

Choose $\delta$ as follows:

$$\delta = \gamma \log |\mathcal{X}|. \tag{20}$$

We now prove inequalities (10)-(13).

To derive (10) we write:

$$\begin{aligned} nR &\leq nC_2 + I(X_1^n; Y^n|X_2^n) - n\epsilon \\ &= nC_2 + I(UX_1^n; Y^n|X_2^n) - n\epsilon \\ &\leq nC_2 + I(X_1^n; Y^n|X_2^nU) + H(U_1|X_2^n) - n\epsilon \\ &\overset{(a)}{<} nC_2 + I(X_1^n; Y^n|X_2^nU) \\ &\quad + k\,(\log |\mathcal{X}| - \delta)^+ - n\epsilon \\ &\overset{(b)}{<} nC_2 + I(X_1^n; Y^n|X_2^nU) \\ &\quad + \frac{n(C_1 + C_2 - R)}{\delta}\,(\log |\mathcal{X}| - \delta)^+ - n\epsilon \\ &\overset{(c)}{\leq} nC_2 + I(X_1^n; Y^n|X_2^nU) \\ &\quad + n\frac{(1-\gamma)^+}{\gamma}(C_1 + C_2 - R) - n\epsilon \\ &\leq nC_2 + n\sum_{i=1}^{n}\frac{1}{n}I(X_{1i}; Y_i|X_{2i}U) \\ &\quad + n\frac{(1-\gamma)^+}{\gamma}(C_1 + C_2 - R) - n\epsilon. \end{aligned} \tag{21}$$

In the above chain of inequalities, step $(a)$ holds by Proposition 1 as follows.

$$\begin{aligned} H(U_1|X_2^n) &= H(U_1) - I(U_1; X_2^n) \\ &\leq H(U_1) - I(U_1; U_2) \\ &< k\,(\log |\mathcal{X}| - \delta)^+ \end{aligned}$$

Also, step $(b)$ follows from (19) and step $(c)$ from (20).

A similar bound may be written for (11).

The bound in (12) is derived in two steps. First, we refine inequality (9) as follows:

$$\begin{aligned} nR &\leq nC_1 + nC_2 - I(X_1^n; X_2^n) \\ &\leq nC_1 + nC_2 - I(U_1; U_2). \end{aligned} \tag{22}$$

We further have

$$\begin{aligned} nR &\leq I(X_1^nX_2^n; Y^n) - n\epsilon \\ &= I(X_1^nX_2^nU; Y^n) - n\epsilon \\ &\leq I(X_1^nX_2^n; Y^n|U) + H(U) - n\epsilon. \end{aligned} \tag{23}$$

Then, we combine (22) and (23) to obtain

$$\begin{aligned} 2nR &\leq nC_1 + nC_2 + I(X_1^nX_2^n; Y^n|U) \\ &\quad + H(U) - I(U_1; U_2) - n\epsilon \\ &< nC_1 + nC_2 + I(X_1^nX_2^n; Y^n|U) \\ &\quad + 2k(\log |\mathcal{X}| - \delta)^+ - n\epsilon \\ &\leq nC_1 + nC_2 + n\sum_{i=1}^{n}\frac{1}{n}I(X_{1,i}, X_{2,i}; Y_i|U) \\ &\quad + 2k(\log |\mathcal{X}_1| - \delta)^+ - n\epsilon \\ &< nC_1 + nC_2 + n\sum_{i=1}^{n}\frac{1}{n}I(X_{1,i}X_{2,i}; Y_i|U) \\ &\quad + 2n\frac{(1-\gamma)^+}{\gamma}\,(C_1 + C_2 - R) - n\epsilon. \end{aligned} \tag{24}$$

To prove (13), we write

$$nR \leq n \sum_{i=1}^{n} \frac{1}{n} I(X_{1i} X_{2i}; Y_i). \tag{25}$$

It is easy to see that there is no loss of generality in assuming $\gamma \leq 1$ in (21) and (24).

To conclude the proof, we use a standard time sharing argument.

$$I(X_1; X_2|UQ) \leq \gamma \log |\mathcal{X}| \tag{26}$$
$$R \leq C_1 + C_2 \tag{27}$$
$$R \leq C_2 + I(X_1; Y|X_2 UQ) + \frac{1-\gamma}{\gamma}(C_1 + C_2 - R) \tag{28}$$
$$R \leq C_1 + I(X_2; Y|X_1 UQ) + \frac{1-\gamma}{\gamma}(C_1 + C_2 - R) \tag{29}$$
$$2R \leq C_1 + C_2 + I(X_1 X_2; Y|UQ) + 2\frac{1-\gamma}{\gamma}(C_1 + C_2 - R) \tag{30}$$
$$R \leq I(X_1 X_2; Y|Q) \tag{31}$$

Note that $I(X_1X_2; Y|Q) \leq I(X_1X_2; Y)$ and thus renaming $UQ$ to $U$ concludes the proof.

The cardinality bound on the auxiliary random variable is derived using the standard techniques via Carathéodory's theorem. ■

### A. Example

Let us revisit the example in Section III-B. We use Theorem 2 to obtain an upper bound on the communication rate in Table I (for three different values of $C$). The optimizing $\gamma^\star$ which leads to the corresponding bound is also shown in this table. We note that Theorem 2 gives tighter bounds compared to the cut set bound, and improves the bounds of [4] in the regime where $.75 < C \leq .7716$. Although the improvement is small, the underlying technique allows us to study the regime where cooperation between the encoders start to be effective.

TABLE I: Upper bound of Theorem 2

| C | Cut Set bound | Bound of [4] | Theorem 2 | $\gamma^\star$ |
|---|---|---|---|---|
| .75 | 1.5 | 1.5 | 1.5 | 0 |
| .7716 | 1.5432 | 1.5432 | 1.5431 | .003 |
| .7925 | 1.5850 | 1.5641 | 1.5832 | .028 |

**Remark 1.** *Every choice of* $0 \leq \gamma \leq 1$ *gives an upper bound on* $R$.

**Remark 2.** *The choice* $\gamma = 1$ *gives the cut-set bound.*

**Remark 3.** *For the case* $R \approx C_1 + C_2$*, we obtain Theorem 1 by choosing* $\gamma$ *arbitrarily small.*

## Appendix

The proof to Proposition 1 is straightforward and along the same lines of the proof of Lemma 1. E.g., see [6], [7] . $U$ is constructed from $X_1^n$ and $X_2^n$ algorithmically as follows.

1) Set $U = \{\}$ and $j = 1$.
2) If $I(X_{1i}; X_{2i}|U) \leq \delta$ for all $i = 1, \dots, n$, then we are done.
3) Otherwise, there exists an index $t$ such that $I(X_{1t}; X_{2t}|U) > \delta$. Set $U_1 = U_1 \cup X_{1t}$ and $U_2 = U_2 \cup X_{2t}$. Set $t_j$ to $t$.
4) increase $j$ and go to step 2.

We now show that the three properties stated in Proposition 1 hold.

Recall that $U_1 = X_{1t_1}, \dots, X_{1t_k}$ and $U_2 = X_{2t_1}, \dots, X_{2t_k}$. The first property holds by construction and using the chain rule of mutual information. The second property also holds by construction. The third property is shown as follows.

$$\sigma \geq I(X_1^n; X_2^n) \tag{32}$$
$$\geq I(X_{1t_1}; X_{2t_1}) + I(X_1^n; X_2^n|X_{1t_1}X_{2t_1}) \tag{33}$$
$$\geq I(X_{1t_1}; X_{2t_1}) + I(X_{1t_2}; X_{2t_2}|X_{1t_1}X_{2t_1}) \tag{34}$$
$$+ I(X_1^n; X_2^n|X_{1t_1}X_{2t_1}X_{1t_2}X_{2t_2}) \tag{35}$$
$$\geq \dots \tag{36}$$
$$\geq I(X_{1t_1}; X_{2t_1}) \tag{37}$$
$$+ \sum_{j=2}^{k} I(X_{1t_j}; X_{2t_j}|X_{1t_1}X_{2t_1} \dots X_{1t_{j-1}}X_{2t_{j-1}}) \tag{38}$$
$$+ I(X_1^n; X_2^n|U_1U_2) \tag{39}$$
$$\stackrel{(a)}{>} k\delta. \tag{40}$$

In the above chain of inequalities, $(a)$ holds because $I(X_{1t_1}; X_{2t_1}) > \delta$ and $I(X_{1t_j}; X_{2t_j}|X_{1t_1}X_{2t_1} \dots X_{1t_{j-1}}X_{2t_{j-1}}) > \delta$ by the way $U_1$ and $U_2$ are formed in each step.

## References


[1] B. E. Schein, *Distributed coordination in network information theory*. PhD Dissertation, MIT, 2001.
[2] D. Traskov and G. Kramer, “Reliable communication in networks with multi-access interference,” in *Proc. Inf. Theory Workshop*, 2007.
[3] W. Kang and N. Liu, “The Gaussian multiple access diamond channel,” in *Proc. Int. Symp. Inf. Theory*, 2011.
[4] S. Saeedi Bidokhti and G. Kramer, “Capacity bounds for a class of diamond networks,” in *Proc. Int. Symp. Inf. Theory*, 2014.
[5] R. Ahlswede, “The rate-distortion region for multiple descriptions without excess rate,” *IEEE Trans. Inf. Theory*, vol. 31, no. 6, pp. 721–726, 1985.
[6] ——, “An elementary proof of the strong converse theorem for the multiple access channel,” *Journal of Combinatorics, Information & system Sciences*, vol. 7, no. 3, pp. 216–230, 1982.
[7] E. Telatar, “Wringing lemmas,” unpublished notes.