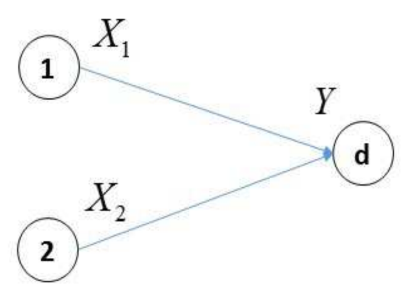
A Proof of the Strong Converse Theorem for Gaussian Multiple Access Channels

SILAS L. FONG AND VINCENT Y. F. TAN

Department of Electrical and Computer Engineering, National University of Singapore, Singapore

Multiple Access Channel (MAC)



- 2 sources transmit messages to a single destination.
- Each source transmits 1 message.
- The destination decodes all the messages.

Discrete Memoryless MAC (DM-MAC)

- Characterized by a transition matrix $q_{Y|X_1,X_2}$
- Capacity region $\mathcal{C}_{\text{AL}}$ was derived by [Ahlswede, 1971] and [Liao, 1972] in the early 1970s, which is the convex closure of

$$\bigcup_{p_{X_1}p_{X_2}} \left\{ \begin{matrix}(R_1, R_2) \\ \in \mathbb{R}_+^2\end{matrix} \middle| \begin{array}{l} R_1 \le I_{p_{X_1}p_{X_2}q_{Y|X_1,X_2}}(X_1; Y|X_2), \\ R_2 \le I_{p_{X_1}p_{X_2}q_{Y|X_1,X_2}}(X_2; Y|X_1), \\ R_1 + R_2 \le I_{p_{X_1}p_{X_2}q_{Y|X_1,X_2}}(X_1, X_2; Y) \end{array} \right\}$$

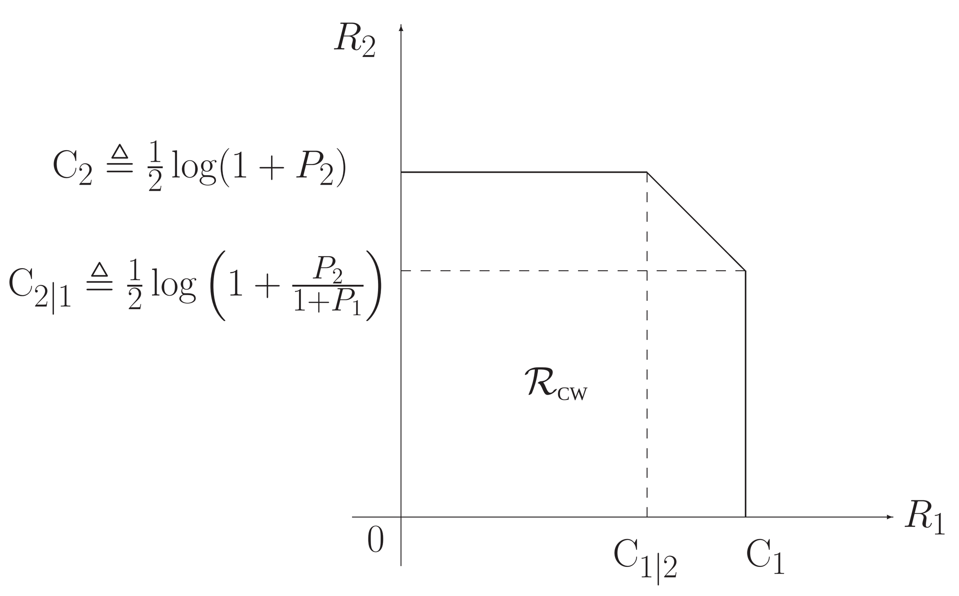
Weak Converse vs Strong Converse

- **Weak Converse**
  - For a fixed rate pair (R_1, R_2), if the decoding error probability *vanishes* as the blocklength increases, then $(R_1, R_2) \in \mathcal{C}_{\text{AL}}$.
  - If the rate pair falls outside $\mathcal{C}_{\text{AL}}$, then the decoding error is *bounded away from 0* as the blocklength increases.
- **Strong Converse**
  - For a fixed rate pair (R_1, R_2), if the decoding error probability is *upper bounded by some* $\varepsilon \in (0, 1)$ as the blocklength increases, then $(R_1, R_2) \in \mathcal{C}_{\text{AL}}$.
  - If the rate pair falls outside $\mathcal{C}_{\text{AL}}$, then the decoding error *tends to 1* as the blocklength increases.
- The original proofs in [Ahlswede, 1971] and [Liao, 1972] are weak converse results.
- Strong converse was proved by [Dueck, 1981] and [Ahlswede, 1982].

Gaussian MAC

- $Y = X_1 + X_2 + Z$ where Z is a standard Gaussian noise.
- The codewords X_1^n and X_2^n should satisfy $\|X_1^n\|^2 \le nP_1$ and $\|X_2^n\|^2 \le nP_2$ respectively.
- Capacity region was derived by [Cover, 1975] and [Wyner, 1974]:

$$\mathcal{R}_{\text{CW}} \triangleq \left\{ \begin{matrix}(R_1, R_2) \\ \in \mathbb{R}_+^2\end{matrix} \middle| \begin{array}{l} R_1 \le \frac{1}{2}\log(1 + P_1), \\ R_2 \le \frac{1}{2}\log(1 + P_2), \\ R_1 + R_2 \le \frac{1}{2}\log(1 + P_1 + P_2) \end{array} \right\}.$$



- The proofs in [Cover, 1975] and [Wyner, 1974] are weak converse results.
- Our contribution is a strong converse proof for the Gaussian MAC.

Network Model

- Sources 1 and 2 transmit information to d in n time slots:
  - Each source i chooses W_i to transmit. Message W_i is uniform on $\{1, 2, \ldots, 2^{nR_i}\}$ where R_i denotes the rate of W_i.
- Each source i transmits $X_{i,k}$ in time slot k and node d receives

$$Y_k = X_{1,k} + X_{2,k} + Z_k\,,$$

  where Z^n are n i.i.d. standard normal random variables.
- Power constraint P_i: X_i^n should satisfy

$$\|X_i^n\|^2 \triangleq \sum_{k=1}^{n} X_{i,k}^2 \le nP_i\,.$$

ε-Capacity Region

- A length-n code operating at rate (R_1, R_2) is called an $(n, (R_1, R_2), \varepsilon_n)$*-code* if the *average* probability of decoding error is less than ε_n .
- (R_1, R_2) is ε-achievable if $\exists$ a sequence of $(n, (R_1, R_2), \varepsilon)$-codes such that $\limsup_{n\to\infty} \varepsilon_n \le \varepsilon$.
- Define ε-capacity region

$$\mathcal{C}_\varepsilon \triangleq \left\{ (R_1, R_2) \in \mathbb{R}_+^2 \,\middle|\, (R_1, R_2) \text{ is } \varepsilon\text{-achievable} \right\}$$

Main Result

Theorem: For each $\varepsilon \in [0, 1)$,

$$\mathcal{C}_\varepsilon = \mathcal{R}_{\text{CW}}. \tag{1}$$

Generalization to N sources: Let $\mathcal{I} \triangleq \{1, 2, \ldots, N\}$ denote the N sources. For each $\varepsilon \in [0, 1)$,

$$\mathcal{C}_\varepsilon = \bigcap_{T \subseteq \mathcal{I}} \left\{ (R_1, \ldots, R_N) \in \mathbb{R}_+^N \,\middle|\, \sum_{i\in T} R_i \le \frac{1}{2}\log\left(1 + \sum_{i\in T} P_i\right) \right\}$$

Major Challenge of Proving the Theorem

- The first step of the proof is to convert the given *average error code* to a *maximal error code* by expurgating appropriate codewords. After the expurgation step, we need a *wringing technique* to "wring out" the dependence between $X_{1,k}$ and $X_{2,k}$.
- After the wringing step, the proof can be completed by using Augustin's converse [Augustin, 1966], a well-known tool for establishing strong converses for the DMC as well as AWGN.
- For the *discrete-alphabet* case, Ahlswede established a wringing technique in [Ahlswede, 1982] to prove the strong converse of the DM-MAC.
- However for the Gaussian case where the alphabet is *continuous*, we cannot directly apply the wringing technique. Instead, we need to quantize the codewords judiciously so that the dependence between $\hat{X}_{1,k}$ and $\hat{X}_{2,k}$ (quantized versions) can be wringed out. This is the main challenge.

Codebook Expurgation [Dueck, 1981]

- Construct a *maximal error code* from the given *average error code* by expurgating a fraction $\frac{2\varepsilon}{1+\varepsilon}$ of codewords, resulting in *maximal probability of decoding error* less than $\frac{1+\varepsilon}{2}$. The messages of the maximal error code are uniformly distributed on some $\mathcal{A} \subset \mathcal{W}_1 \times \mathcal{W}_2$.
- Under the maximal error code,

$$|\mathcal{A}| \ge \left(\frac{1-\varepsilon}{2(1+\varepsilon)}\right) 2^{n(R_1+R_2)}$$

  and for each $(w_1, w_2) \in \mathcal{A}$,

$$p_{W_1,W_2}(w_1, w_2) = \frac{1}{|\mathcal{A}|} \le \frac{1}{2^{n(R_1+R_2)}} \cdot \left(\frac{2(1+\varepsilon)}{1-\varepsilon}\right)$$

Codeword Quantizer

Let $N > 0$ and $\Delta > 0$, and let

$$\mathbb{Z}_{N,\Delta} \triangleq \{-N\Delta, (-N+1)\Delta, \ldots, N\Delta\} \tag{2}$$

be a set of $2N+1$ quantization points where Δ specifies the quantization precision. A scalar quantizer with domain $[-N\Delta, N\Delta]$ and precision Δ is the mapping $\Omega_{N,\Delta} : [-N\Delta, N\Delta] \to \mathbb{Z}_{N,\Delta}$ such that

$$\Omega_{N,\Delta}(x) = \begin{cases} \lfloor x/\Delta \rfloor \Delta & \text{if } x \ge 0, \\ \lceil x/\Delta \rceil \Delta & \text{otherwise.} \end{cases}$$

In addition, define the scalar quantizer for a real-valued tuple as $\Omega_{N,\Delta}^{(n)} : [-N\Delta, N\Delta]^n \to \mathbb{Z}_{N,\Delta}^n$ such that

$$\Omega_{N,\Delta}^{(n)}(x^n) \triangleq (\Omega_{N,\Delta}(x_1), \Omega_{N,\Delta}(x_2), \ldots, \Omega_{N,\Delta}(x_n))$$

Wringing Technique

- Suppose we are given the $\left(n, \mathcal{A}, \frac{1+\varepsilon}{2}\right)_{\text{max}}$-code constructed after the expurgation step. Then, there exists an $\left(n, \mathcal{A}', \frac{1+\varepsilon}{2}\right)_{\text{max}}$-code with

$$|\mathcal{A}'| \ge n^{\frac{-8(1+3\varepsilon)}{(1-\varepsilon)}\sqrt{\frac{n}{\log n}}} \left(\frac{1-\varepsilon}{2(1+\varepsilon)}\right) 2^{n(R_1+R_2)}$$

  such that the following holds: Let $p_{W_1,W_2,X_1^n,X_2^n,Y^n}$ denote the distribution induced by the $\left(n, \mathcal{A}', \frac{1+\varepsilon}{2}\right)_{\text{max}}$-code. Let

$$\hat{X}_i^n = \Omega_{\lceil n\sqrt{nP_i}\rceil, \frac{1}{n}}(X_i^n),$$

  define the alphabet

$$\hat{\mathcal{X}}_i \triangleq \mathbb{Z}_{\lceil n\sqrt{nP_i}\rceil, \frac{1}{n}}$$

  for each $i \in \{1, 2\}$ and define

$$\begin{aligned} & p_{X_1^n,X_2^n,\hat{X}_1^n,\hat{X}_2^n}(x_1^n, x_2^n, \hat{x}_1^n, \hat{x}_2^n) \\ & \triangleq p_{X_1^n,X_2^n}(x_1^n, x_2^n) \prod_{i\in\{1,2\}} \mathbf{1}\left\{\hat{x}_i^n = \Omega_{\lceil n\sqrt{nP_i}\rceil, \frac{1}{n}}(x_i^n)\right\} \end{aligned}$$

  for all $(x_1^n, x_2^n, \hat{x}_1^n, \hat{x}_2^n) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \hat{\mathcal{X}}_1^n \times \hat{\mathcal{X}}_2^n$. Then there exists a distribution $u_{\hat{X}_1^n,\hat{X}_2^n}$ such that for all $k \in \{1, 2, \ldots, n\}$ and for all $\hat{x}_{1,k}, \hat{x}_{2,k} \in \hat{\mathcal{X}}_1 \times \hat{\mathcal{X}}_2$, we have

$$\begin{aligned} & p_{\hat{X}_{1,k},X_{2,k}}(\hat{x}_{1,k}, \hat{x}_{2,k}) \\ & \le \max\left\{\left(1 + \sqrt{\frac{\log n}{n}}\right) \prod_{i\in\{1,2\}} u_{\hat{X}_{i,k}}(\hat{x}_{i,k}), \frac{1}{n^8}\right\}. \end{aligned}$$

- Therefore, we can approximate $p_{\hat{X}_{1,k},\hat{X}_{2,k}}$ by a product distribution $\prod_{i\in\{1,2\}} u_{\hat{X}_{i,k}}$ through an inequality, which wrings out the independence between $X_{1,k}$ and $X_{2,k}$.

References

[Ahlswede, 1971] Ahlswede, R. (1971). Multi-way communication channels. In *Proc. of IEEE International Symposium on Information Theory*, pages 23–52, Thakadsor, Armenian SSR.

[Ahlswede, 1982] Ahlswede, R. (1982). An elementary proof of the strong converse theorem for the multiple access channel. *Journal of Combinatorics, Information & System Sciences*, 7(3):216–230.

[Augustin, 1966] Augustin, U. (1966). Gedachtnisfreie kannale for diskrete zeit. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, pages 10–61.

[Cover, 1975] Cover, T. M. (1975). Some advances in broadcast channels. In Viterbi, A. J., editor, *Advances in Communication Systems*, volume 4, pages 229–260. Academic Press.

[Dueck, 1981] Dueck, G. (1981). The strong converse coding theorem for the multiple-access channel. *Journal of Combinatorics, Information & System Sciences*, 6(3):187–196.

[Liao, 1972] Liao, H. H. J. (1972). *Multiple access channels*. PhD thesis, University of Hawaii, Honolulu.

[Wyner, 1974] Wyner, A. D. (1974). Recent results in the Shannon theory. *IEEE Transactions on Information Theory*, 20:2–10.