Information Theory for Wireless Communications. Lecture 10
Discrete Memoryless Multiple Access Channel (DM-MAC): The Converse Theorem

Instructor: Dr. Saif Khan Mohammed
Scribe: Antonios Pitarokoilis

I. The Discrete Memoryless Multiple Access Channel

We consider a two-user discrete memoryless multiple access channel (DM-MAC) with distribution $p(y|x_1, x_2)$, as shown in Fig. 1. Users 1 and 2 select their messages, X_1 and X_2, according to the probability density functions $p_{x_1}(\cdot)$ and $p_{x_2}(\cdot)$ at rates R_1 and R_2, respectively, while the channel outputs the symbol Y. We define the rate region, $\mathfrak{R}(p_{x_1}(\cdot), p_{x_2}(\cdot))$, as the pair of rates (R_1, R_2) that satisfy

$$\mathfrak{R}(p_{x_1}(\cdot), p_{x_2}(\cdot)) \triangleq \{(R_1, R_2)|R_1 \leq I(X_1; Y|X_2), R_2 \leq I(X_2; Y|X_1), R_1 + R_2 \leq I(X_1, X_2; Y)\}. \quad (1)$$

Further, the capacity region, $\mathfrak{C}_{\text{DM-MAC}}$ is given by the convex hull of the union of all the possible rate regions, $\mathfrak{R}(p_{x_1}(\cdot), p_{x_2}(\cdot))$, over all the possible input distributions $p_{x_1}(\cdot)$ and $p_{x_2}(\cdot)$, namely

$$\mathfrak{C}_{\text{DM-MAC}} \triangleq \text{Co}\left(\bigcup_{p_{x_1}(\cdot), p_{x_2}(\cdot)} \mathfrak{R}(p_{x_1}(\cdot), p_{x_2}(\cdot))\right). \quad (2)$$

We also define the average probablity of error, $\bar{\lambda}_n$ of a code. Let a code of length n and a set of codewords $(W_i, W_j),\ i = \{1, \ldots, 2^{nR_1}\},\ j = \{1, \ldots, 2^{nR_2}\}$. The encoders of the two users transmit the codewords W_i and W_j. The decoder detects the output, Y, of the DM-MAC channel and decides for the pair of codewords $\{\hat{W}_i, \hat{W}_j\}$. Then, the probability of an error even is given by $\mathbb{P}\left\{(\hat{W}_i, \hat{W}_j) \neq (W_i, W_j)|(W_i, W_j) \text{ was sent}\right\}$. The average error probability of this code is given by

$$\bar{\lambda}_n \triangleq \sum_{i,j} \mathbb{P}\{(W_i, W_j)\}\, \mathbb{P}\left\{(\hat{W}_i, \hat{W}_j) \neq (W_i, W_j)|(W_i, W_j) \text{ was sent}\right\}.$$

The channel coding theorem for the DM-MAC can now be stated as follows.

Theorem 1: If there is a sequence of codes $\left(n, \left(2^{nR_1}, 2^{nR_2}\right), \bar{\lambda}_n\right)$ such that $\bar{\lambda}_n \to 0$, then $(R_1, R_2) \in \mathfrak{C}_{\text{DM-MAC}}$.

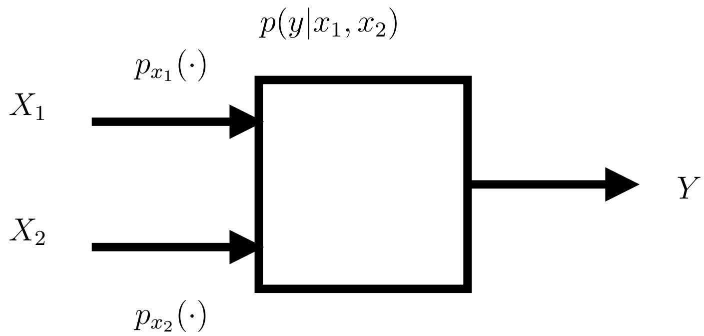

Fig. 1. Two user discrete memoryless multiple access channel

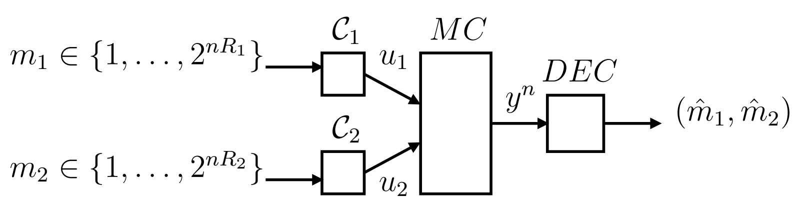

Fig. 2. Coding for the DM-MAC

In the following two important inequalities are given for future reference. The data processing inequality for a system as the one shown in Fig. 2 is given by

$$H(U_1, U_2|Y^n) \leq H(U_1, U_2|(\hat{m}_1, \hat{m}_2)). \tag{3}$$

Fano's inequality is given by

$$H(U_1, U_2|(\hat{m}_1, \hat{m}_2)) \leq n(R_1 + R_2)\bar{\lambda}_n + 1. \tag{4}$$

For equally-likely codewords

$$
\begin{aligned}
H(U_1,U_2) = n(R_1+R_2) &= H(U_1,U_2) - H(U_1,U_2|Y^n) + H(U_1,U_2|Y^n) \\
&\leq I(U_1,U_2;Y^n) + n(R_1+R_2)\bar{\lambda}_n + 1 \qquad (5)\\
&= H(Y^n) - H(Y^n|U_1,U_2) + n(R_1+R_2)\bar{\lambda}_n + 1 \\
&\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|Y^{i-1},U_1,U_2) + n(R_1+R_2)\bar{\lambda}_n + 1 \\
&\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|U_1,U_2) + n(R_1+R_2)\bar{\lambda}_n + 1 \qquad (6)\\
&= \sum_{i=1}^{n} (H(Y_i) - H(Y_i|U_1,U_2)) + n(R_1+R_2)\bar{\lambda}_n + 1 \\
&= \sum_{i=1}^{n} I(U_1,U_2;Y_i) + n(R_1+R_2)\bar{\lambda}_n + 1, \qquad (7)
\end{aligned}
$$

where in (5) we made use of the Data Proceessing inequality, (3), and the Fano's inequality (see (4)). Eq. (6) follows from the the memoryless property of the channel. Finally, from (7) we get

$$
R_1 + R_2 \leq \frac{1}{n}\sum_{i=1}^{n} I(U_1,U_2;Y_i) + (R_1+R_2)\bar{\lambda}_n + \frac{1}{n}. \qquad (8)
$$

Further, the rate of user 1 can be bounded as follows.

$$
\begin{aligned}
nR_1 = H(U_1) = H(U_1|U_2) &= H(U_1|U_2) - H(U_1|Y^n,U_2) + H(U_1|Y^n,U_2) \qquad (9)\\
&= I(U_1;Y^n|U_2) + H(U_1|Y^n,U_2). \qquad (10)
\end{aligned}
$$

In (9) we have used the fact that conditioning on an independent random variable does not affect the entropy. By the chain rule for entropy we get

$$
H(U_1,U_2|Y^n) = H(U_2|Y^n) + H(U_1|U_2,Y^n),
$$

and due to the non-negativity of the entropy

$$
H(U_1|U_2,Y^n) \leq H(U_1,U_2|Y^n) < n(R_1+R_2)\bar{\lambda}_n + 1
$$

$$nR_1 < I(U_1;Y^n|U_2) + n(R_1+R_2)\bar{\lambda}_n + 1 = \sum_{i=1}^{n} I(U_1;Y_i|Y^{i-1},U_2) + n(R_1+R_2)\bar{\lambda}_n + 1 \tag{11}$$

$$= \sum_{i=1}^{n} \left(H(Y_i|Y^{i-1},U_2) - H(Y_i|U_{1i},U_{2i},Y^{i-1})\right) + n(R_1+R_2)\bar{\lambda}_n + 1 \tag{12}$$

$$\le \sum_{i=1}^{n} (H(Y_i|U_2) - H(Y_i|U_{1i},U_{2i})) + n(R_1+R_2)\bar{\lambda}_n + 1 \tag{13}$$

$$= \sum_{i=1}^{n} (H(Y_i|U_{2i}) - H(Y_i|U_{1i},U_{2i})) + n(R_1+R_2)\bar{\lambda}_n + 1 \tag{14}$$

Therefore,

$$R_1 \le \frac{1}{n}\sum_{i=1}^{n} I(U_{1i};Y_i|U_{2i}) + (R_1+R_2)\bar{\lambda}_n + \frac{1}{n}. \tag{15}$$

Similarly, we can prove that

$$R_2 \le \frac{1}{n}\sum_{i=1}^{n} I(U_{2i};Y_i|U_{1i}) + (R_1+R_2)\bar{\lambda}_n + \frac{1}{n}. \tag{16}$$

Definition 1: For given $p_{X_1}(\cdot),\ p_{X_2}(\cdot)$, define the vector

$$\begin{aligned} I(p_{X_1}(\cdot),\ p_{X_2}(\cdot)) &\triangleq (I(X_1;Y|X_2), I(X_2;Y|X_1), I(X_1,X_2|Y)) \in \mathbb{R}^3 \\ &= (I(1), I(2), I(3)) \end{aligned}$$

Based on definition 1 the set $\mathtt{C}_I$ of rate pairs (R_1,R_2) can be defined as follows

$$\mathtt{C}_I \triangleq \{(R_1,R_2)|R_1 \ge 0, R_2 \ge 0, R_1 \le I(1), R_2 \le I(2), R_1+R_2 \le I(3)\} \tag{17}$$

Further, it can be shown that $I(1)+I(2) \ge I(3) \ge \max\{I(1), I(2)\}$.

Lemma 1: Let $I_1, I_2 \in \mathbb{R}^3$ with $\mathtt{C}_{I_1},\ \mathtt{C}_{I_2}$ as in (17). Consider $0 \le \lambda \le 1$ and define $I^\lambda \triangleq \lambda I_1 + (1-\lambda)I_2$. Then $\mathtt{C}_{I^\lambda} = \lambda \mathtt{C}_{I_1} + (1-\lambda)\mathtt{C}_{I_2}$.

Theorem 2: The convex hull of the union of the rate regions defined by the individual I vectors, which is $\mathfrak{C}_{\text{DMAC}}$, is equal to the rate region defined by the convex hull of the I vectors $\mathtt{C}_{\text{Co}(I)}$, i.e.

$$\mathfrak{C}_{\text{DMAC}} = \mathtt{C}_{\text{Co}(I)}$$

where

$$\mathtt{C}_{\text{Co}(I)} \triangleq \bigcup_{i \in \text{Co}(I)} \mathtt{C}_i.$$

II. The Gaussian MAC Channel

We specify the Gaussian MAC Channel as

$$Y = X_1 + X_2 + Z \tag{18}$$

where the power constraints of user 1 and user 2 are $\mathbb{E}X_1^2 \leq P_1$ and $\mathbb{E}X_2^2 \leq P_2$, respectively. Further, we have for the additive noise, Z, that $Z \sim \mathcal{N}(0, \sigma^2)$. For fixed input distributions, $f_{X_1}(\cdot)$, $f_{X_2}(\cdot)$ the rate region is given by

$$\mathfrak{R}(f_{X_1}(\cdot), f_{X_2}(\cdot)) = \{(R_1, R_2)|R_1 \leq I(X_1; Y|X_2), R_2 \leq I(X_2; Y|X_1), R_1 + R_2 \leq I(X_1, X_2; Y)\}$$

Then, the capacity region of the Gaussian MAC is given by the convex hull of the union of the rate regions, $\mathfrak{R}(f_{X_1}(\cdot), f_{X_2}(\cdot))$, for every possible choice of the input distributions $f_{X_1}(\cdot)$, $f_{X_2}(\cdot)$ that satisfy the power constraints, i.e.

$$\mathfrak{C}_{\text{GMAC}} \triangleq \text{Co}\left\{ \bigcup_{\substack{f_{X_1}(\cdot): \int f_{X_1}(x_1) x_1^2 \mathrm{d}x_1 \leq P_1 \\ f_{X_2}(\cdot): \int f_{X_2}(x_2) x_2^2 \mathrm{d}x_2 \leq P_2}} \mathfrak{R}(f_{X_1}(\cdot), f_{X_2}(\cdot)) \right\}. \tag{19}$$

We can compute

$$\begin{aligned} I(X_1; Y|X_2) &= h(Y|X_2) - h(Y|X_1, X_2) = h(X_1 + Z) - \frac{1}{2}\log_2(2\pi e \sigma^2) \\ &\leq \frac{1}{2}\log_2(2\pi e(P_1 + \sigma^2)) - \frac{1}{2}\log_2(2\pi e \sigma^2) = \frac{1}{2}\log_2\left(1 + \frac{P_1}{\sigma^2}\right), \end{aligned}$$

with equality when $X_1 \sim \mathcal{N}(0, P_1)$. Similarly, we can show that $I(X_2; Y|X_1) \leq \frac{1}{2}\log_2\left(1 + \frac{P_2}{\sigma^2}\right)$ with equality for $X_2 \sim \mathcal{N}(0, P_2)$. For the sum rate bound it holds

$$I(X_1, X_2; Y) = h(Y) - h(Y|X_1, X_2) = h(Y) - \frac{1}{2}\log_2 2\pi e \sigma^2 \leq \frac{1}{2}\log_2\left(1 + \frac{P_1 + P_2}{\sigma^2}\right), \tag{20}$$

with equality when $X_1 \sim \mathcal{N}(0, P_1)$, $X_2 \sim \mathcal{N}(0, P_2)$. In (20) the differential entropy of Y, $h(Y)$, is bounded by

$$h(Y) \leq \frac{1}{2}\log_2 2\pi e \text{Var}(Y) \leq \frac{1}{2}\log_2 2\pi e \mathbb{E}Y^2 = \frac{1}{2}\log_2 2\pi e\left(P_1 + P_2 + \sigma^2\right).$$

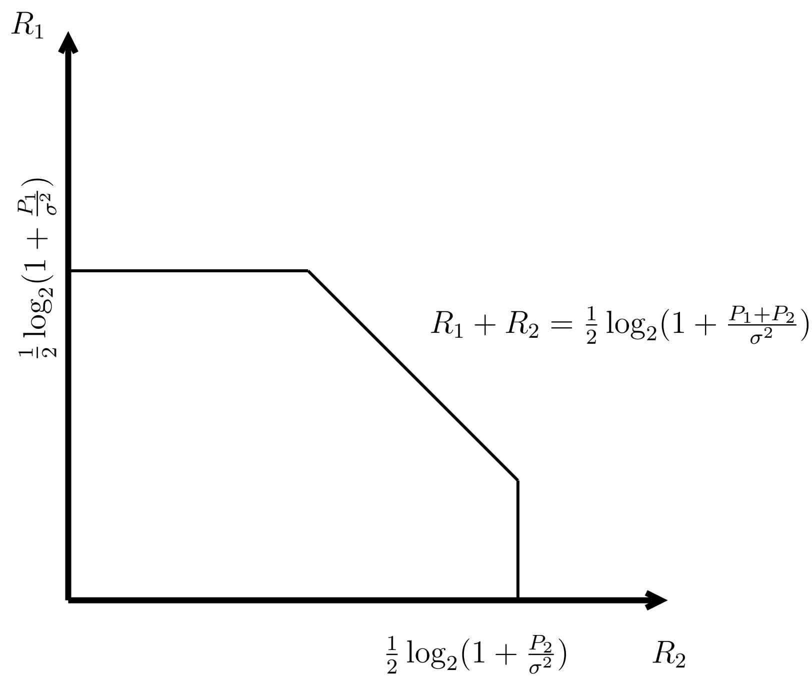

Fig. 3. Capacity Region of a two user Gaussian MAC channel.

Finally, the capacity region of the 2-user Gaussian MAC is given by

$$\mathfrak{C}_{\text{GMAC}} \triangleq \left\{(R_1,R_2)|R_1 \leq \frac{1}{2}\log_2\left(1+\frac{P_1}{\sigma^2}\right),\ R_2 \leq \frac{1}{2}\log_2\left(1+\frac{P_2}{\sigma^2}\right),\ R_1+R_2 \leq \frac{1}{2}\log_2\left(1+\frac{P_1+P_2}{\sigma^2}\right)\right\}. \tag{21}$$

We note that the capacity region of the 2-user Gaussian MAC is a pentagon and every point in the region can be achieved when transmitters 1 and 2 use the Gaussian codebooks $X_1 \sim \mathcal{N}(0,P_1)$, $X_2 \sim \mathcal{N}(0,P_2)$, respectively. The optimal decoding strategy is called successive interference cancelation (SIC) and achieves every point in the capacity region. Therefore, SIC achieves also the corner points of the capacity region. The corner point $(R_1,R_2) = \left(\frac{1}{2}\log_2\left(1+\frac{P_1}{P_2+\sigma^2}\right), \frac{1}{2}\log_2\left(1+\frac{P_2}{\sigma^2}\right)\right)$ is achievable when user 1 is decoded first. In this case, the signal of user 2 is treated as interference and the rate $R_1 = \frac{1}{2}\log_2\left(1+\frac{P_1}{P_2+\sigma^2}\right)$ is achievable. Subsequently, since the message of user 1 is decoded correctly, the receiver can subtract it from the received signal and user 2 can be decoded interference free, achieving the rate $R_2 = \frac{1}{2}\log_2\left(1+\frac{P_2}{\sigma^2}\right)$. The other corner point $(R_1,R_2) = \left(\frac{1}{2}\log_2\left(1+\frac{P_1}{\sigma^2}\right), \frac{1}{2}\log_2\left(1+\frac{P_2}{P_1+\sigma^2}\right)\right)$ is achieved at the receiver by SIC with the reverse decoding order.

A. TDMA case

In the following, we compare the capacity region of the Gaussian MAC with the achievable rate region, when time division multiple access (TDMA) is employed. For the TDMA, we split each channel use of

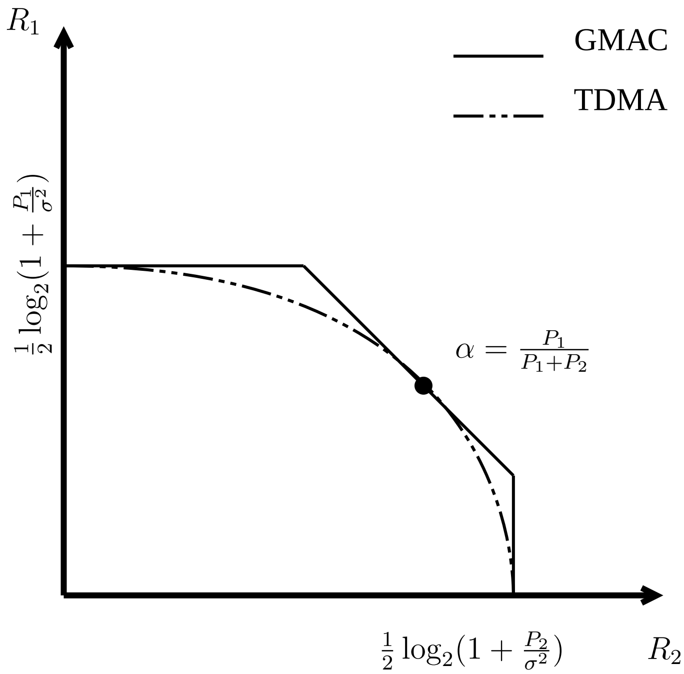

Fig. 4. Capacity Region of a two user Gaussian MAC channel vs. the Rate Region of TDMA.

T_u s into two parts of length αT_u s and $(1-\alpha)T_u$ s, respectively, where $0 \leq \alpha \leq 1$. During the first time slot, only the first user is scheduled to transmit with power of P_1/α and during the second time slot only the second user transmits with power $P_2/(1-\alpha)$. This way, the users transmit in orthogonal resources and the individual average power constraints are satisfied. Therefore, the rate pair $(R_1, R_2) = (\alpha\frac{1}{2}\log_2(1+\frac{P_1}{\alpha\sigma^2}), (1-\alpha)\frac{1}{2}\log_2(1+\frac{P_2}{(1-\alpha)\sigma^2}))$ is achievable. In Fig. 4 it is shown that TDMA is strictly suboptimal compared to the GMAC-capacity-achieving strategy of successive interference cancellation. Note that TDMA achieves the sum capacity of GMAC for the choice $\alpha = \frac{P_1}{P_1+P_2}$.