Information Theory for Wireless Communications Part-II

Lecture 1: Discrete Memoryless Broadcast Channels (DM-BC)

Taught by Dr. Saif K. Mohammed

Scribed by Reza Moosavi

Autumn 2012

I. Introduction

In this lecture, we consider the discrete memoryless broadcast channel (DM-BC) with two users, depicted in Fig. 1. The channel is memoryless, and thus we have

$$P_{Y_1^n,Y_2^n|X^n}(Y_1^n = y_1^n, Y_2^n = y_2^n | X^n = x^n) = \prod_{i=1}^{n} p_{Y_1,Y_2|X}(y_{1,i}, y_{2,i}|x_i). \tag{1}$$

We assume that the probability mass function (PMF) $p_{Y_1,Y_2|X}(y_1, y_2|x)$ is known at the decoders. We also assume that the messages to user 1 and 2 are independent of one another. Error event for user i, $i = 1, 2$, is when $\left(\hat{M}_i \neq M_i\right)$, and thus we define the average error probability for user i as $P_{e_i}^{(n)} = \Pr\left\{\hat{M}_i \neq M_i\right\}$. The overall average error probability is therefore

$$P_e^{(n)} = \left\{\left(\hat{M}_1 \neq M_1\right) \text{ or } \left(\hat{M}_2 \neq M_2\right)\right\}.$$

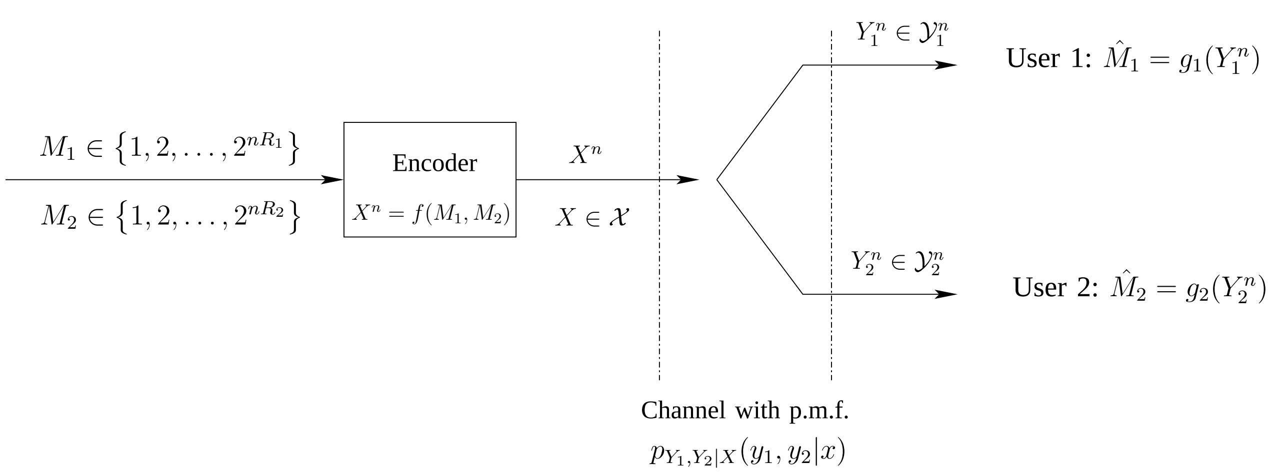

Fig. 1. Discrete memoryless broadcast channel model.

Definition. A rate pair (R_1, R_2) is said to be achievable if there exists a sequence of $\left(2^{nR_1}, 2^{nR_2}\right)$ codes such that,

$$\lim_{n\to\infty} P_e^{(n)} = 0.$$

Definition. For a discrete memoryless broadcast channel with a given PMF $p_{Y_1,Y_2|X}(y_1, y_2|x)$, the *capacity region* is the closure of the set of all achievable rate pairs.

In general, the capacity region of DM-BC is not known in closed-form. However, we have the following useful lemma.

Lemma 1. *The capacity of the DM-BC depends upon the channel PMF* $p_{Y_1,Y_2|X}(y_1, y_2|x)$ *only through the conditional marginal PMF:s* $p_{Y_1|X}(y_1|x)$ *and* $p_{Y_2|X}(y_2|x)$.

Proof:

$$\begin{aligned}
P_{e_1}^{(n)} &= \sum_{m_1=1}^{2^{nR_1}} \sum_{m_2=1}^{2^{nR_2}} \Pr\left\{M_1 = m_1, M_2 = m_2\right\} \Pr\left\{\hat{M}_1 \neq m_1, M_2 = m_2 | M_1 = m_1, M_2 = m_2\right\} \\
&= \sum_{m_1=1}^{2^{nR_1}} \sum_{m_2=1}^{2^{nR_2}} \Pr\left\{M_1 = m_1, M_2 = m_2\right\} \sum_{y_1^n | g_1(y_1^n \neq m_1)} \Pr\left\{Y_1^n = y_1^n | X^n = f(m_1, m_2)\right\} \\
&= \sum_{m_1=1}^{2^{nR_1}} \sum_{m_2=1}^{2^{nR_2}} \Pr\left\{M_1 = m_1, M_2 = m_2\right\} \sum_{y_1^n | g_1(y_1^n \neq m_1)} \prod_{i=1}^{n} p_{Y_1|X}(y_{1,i}|x_i),
\end{aligned} \tag{2}$$

and thus $P_{e_1}^{(n)}$ only depends on the marginal PMF $p_{Y_1|X}(y_1|x)$. Similarly, $P_{e_2}^{(n)}$ only depends on the marginal PMF $p_{Y_2|X}(y_2|x)$. Now since we have,

$$P_e^{(n)} \leq P_{e_1}^{(n)} + P_{e_2}^{(n)}, \tag{3}$$

$$P_e^{(n)} \geq \max\left\{P_{e_1}^{(n)}, P_{e_2}^{(n)}\right\}, \tag{4}$$

we conclude that,

$$\lim_{n\to\infty} P_e^{(n)} = 0 \qquad \Longleftrightarrow \qquad \begin{cases} \lim_{n\to\infty} P_{e_1}^{(n)} = 0, \\ \lim_{n\to\infty} P_{e_2}^{(n)} = 0, \end{cases} \tag{5}$$

and the proof is complete. ■

The direct implication of the above lemma is that the capacity region of the DM-BC depends only on the conditional marginal PMF:s. In other words, two DM-BC:s with distinct channel PMF:s but the same marginal PMF:s, have the same capacity region.

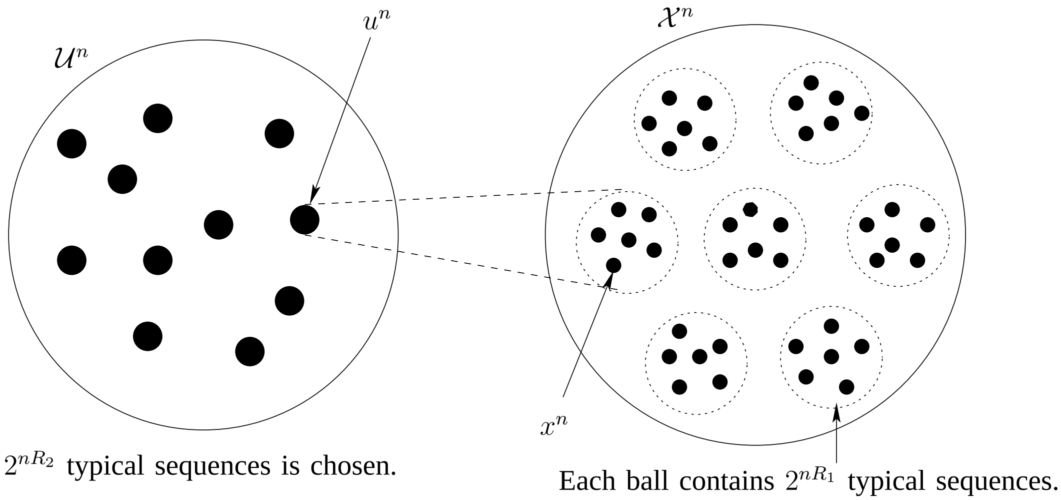

Fig. 2. Superposition Coding.

II. Superposition Coding

In this section we first give a coding scheme which is known as the *superposition coding*. We then give a theorem that describes the achievable rate region of this coding scheme. In the next section, we show that this coding scheme is actually capacity achieving for the so-called *degraded DM-BC*.

Definition. Superposition encoding is done through the following steps:

(i) Choose a random variable U with an arbitrary joint distribution

$$p_{U,X}(U=u,X=x)=p_U(U=u)p_{X|U}(X=x|U=u).$$

(ii) For each $m_2 \in \{1,2,\ldots,2^{nR_2}\}$, generate randomly and independently 2^{nR_2} *cloud centers* $u^n(m_2) \in \mathcal{U}^n$ using PMF $p_U(U=u)$.

(iii) For each cloud center $u^n(m_2)$, generate 2^{nR_1} *satellite* codewords $x(m_1,m_2) \in \mathcal{X}^n$, using $p_{X|U}(X=x|U=u)$.

The above procedure is illustrated in Fig. 2. Note that user 2, decodes for the cloud center and user 1 decodes for the satellite codeword.

Theorem 1. *With superposition coding, a rate pair* (R_1,R_2) *is achievable for DM-BC with* $p_{Y_1,Y_2|X}(y_1,y_2|x)$

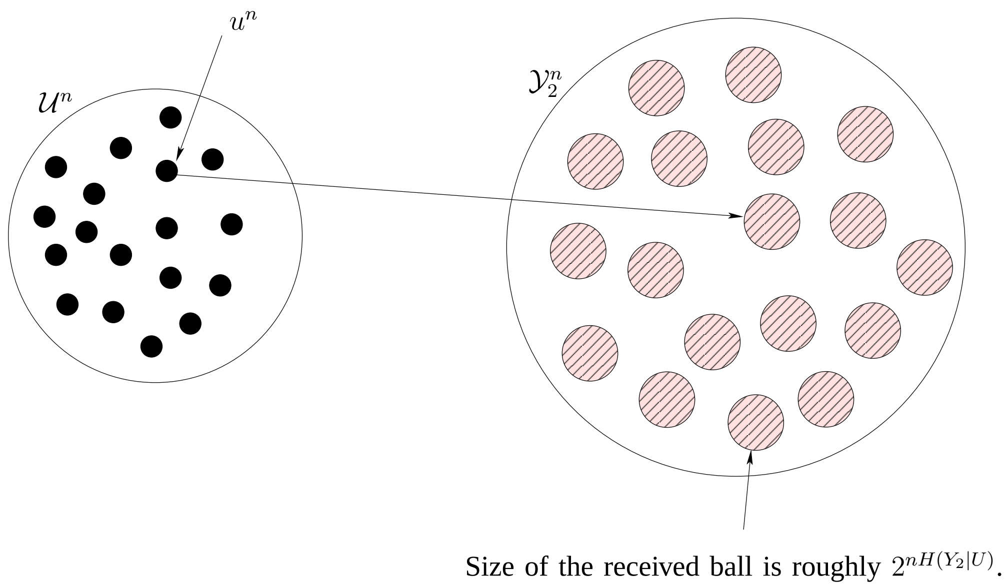

Fig. 3. Decoding for the second user.

if,

$$R_2 \le I(U;Y_2), \tag{6}$$

$$R_1 \le I(X;Y_1|U), \tag{7}$$

$$R_1 + R_2 \le I(X;Y_1), \tag{8}$$

for some PMF $P_{U,X}(u,x) = p_U(u)p_{X|U}(x|u)$.

We do not give the proof of the theorem in this lecture notes. However, we sketch an outline of the proof in the followings. For the formal proof of the theorem, we refer the readers to [1, pp. 109]. Note that the second user decodes for the cloud center. Since the size of the typical set at $\mathcal{Y}_2^n$ is roughly $2^{nH(Y_2)}$ and since the received signal corresponding to each of the 2^{nR_2} messages lies in the regions of volume roughly $2^{nH(Y_2|U)}$, as in Fig. 3, we need to have

$$2^{nR_2} \times 2^{nH(Y_2|U)} \le 2^{nH(Y_2)}, \tag{9}$$

in order to have low probability or error. From (9), we can write

$$R_2 \le H(Y_2) - H(Y_2|U) = I(Y_2;U). \tag{10}$$

In order for user 1 to have low error probability, two constraints need to be satisfied, as can be seen from Fig. 4:

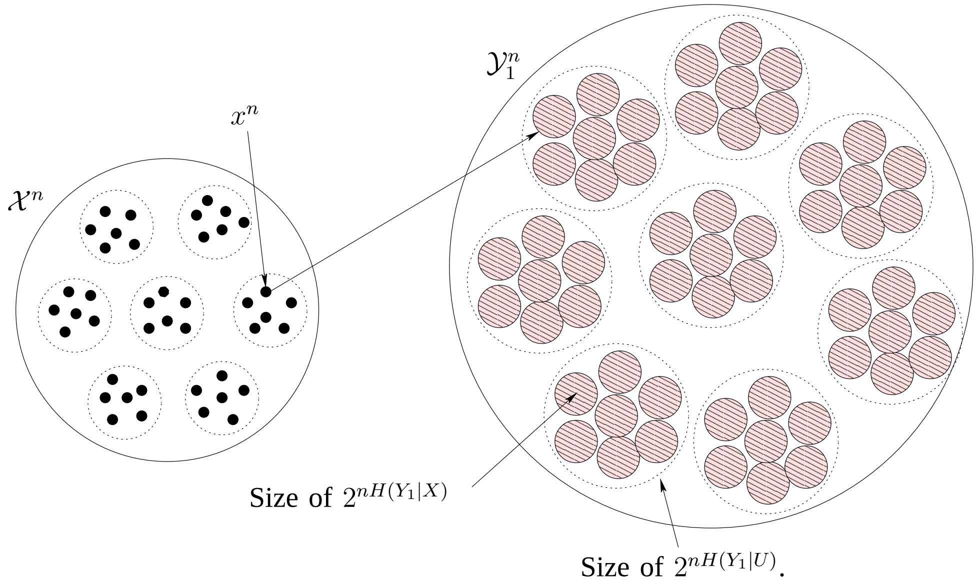

Fig. 4. Received signal at the first user.

1. 2^{nR_1} balls of sizes roughly $2^{nH(Y_1|X)}$ should be packed in a region of size $2^{nH(Y_1|U)}$. Thus,

$$2^{nR_1} \times 2^{nH(Y_1|X)} \leq 2^{nH(Y_1|U)}, \tag{11}$$

and therefore,

$$R_1 \leq H(Y_1|U) - H(Y_1|X) = H(Y_1|U) - H(Y_1|X,U) = I(X;Y_1|U), \tag{12}$$

where we use the fact that X contains all the information of U due to the code construction and thus $H(Y_1|X) = H(Y_1|X,U)$.

2. $2^{nR_1} \times 2^{nR_2}$ balls of sizes roughly $2^{nH(Y_1|X)}$ should be packed in a region of size $2^{n(H(Y_1)}$, and thus

$$2^{nR_1} \times 2^{nR_2} \times 2^{nH(Y_1|X)} \leq 2^{n(H(Y_1)}, \tag{13}$$

or equivalently,

$$R_1 + R_2 \leq H(Y_1) - H(Y_1|X) = I(X;Y_1). \tag{14}$$

III. Degraded Broadcast Channel

There are two definitions for degraded discrete memoryless broadcast channel.

- **Physical degraded DM-BC:** A DM-BC is physically degraded if

$$X \to Y_1 \to Y_2$$

forms a Markov chain, or equivalently if,

$$p_{Y_1,Y_2|X}\left(y_1,y_2|x\right)=p_{Y_1|X}\left(y_1|x\right)p_{Y_2|Y_1}\left(y_2|y_1\right). \tag{15}$$

- **Stochastically degraded DM-BC:** A DM-BC is called stochastically degraded, if there exists $\tilde{Y}_2 \in \mathcal{Y}_2$ such that,

$$1. \qquad p_{\tilde{Y}_2|X}(y_2|x)=p_{Y_2|X}(y_2|x), \tag{16}$$

$$2. \qquad X \to Y_1 \to \tilde{Y}_2. \tag{17}$$

The capacity region of degraded DM-BC is known and given by the following theorem.

Theorem 2. *The capacity region of the degraded DM-BC with channel PMF $p_{Y_1,Y_2|X}(y_1,y_2|x)$ is the convex-hull of the closure of all rate pairs (R_1,R_2) satisfying,*

$$R_1 \le I(X;Y_1|U),$$
$$R_2 \le I(U;Y_2),$$

for some PMF $p_{U,X}(u,x)$, where the cardinality of the auxiliary random variable U satisfies

$$|\mathcal{U}| \le \min\{|\mathcal{X}|,|\mathcal{Y}_1|,|\mathcal{Y}_2|\}+1.$$

Note that the capacity region can be found by using superposition coding in this case. Also note that the third constraint that we had for the achievable rate region of the superposition coding (i.e. equation (8) of Theorem 1) is not needed in this case, since,

$$U \to X \to Y_1 \to Y_2,$$

and thus,

$$I(U;Y_2)+I(X;Y_1|U)=H(Y_1)-H(Y_1|U)+H(Y_1|U)-H(Y_1|X,U)=H(Y_1)-H(Y_1|X)=I(X;Y_1).$$

In other words, if $R_1 \le I(X;Y_1|U)$ and $R_2 \le I(U;Y_2)$, we will have $R_1+R_2 \le I(X;Y_1)$ in this special case and hence, the third constraint always holds for the degraded DM-BC.

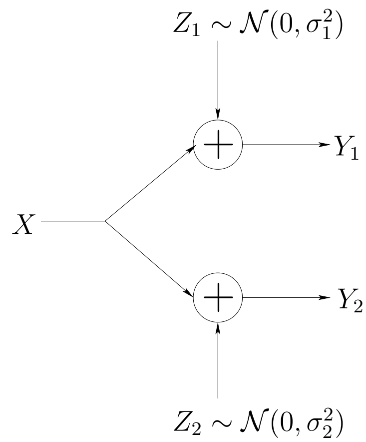

Fig. 5. Gaussian 2-User discrete memoryless broadcast channel.

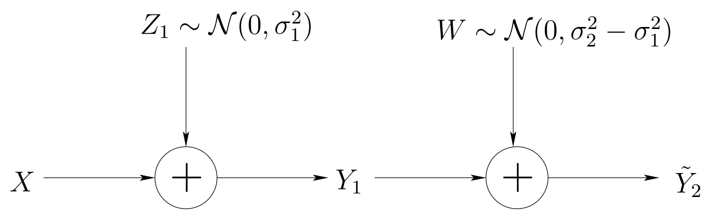

Fig. 6. Corresponding stochastically degraded Gaussian broadcast channel.

IV. Gaussian 2-User Discrete Memoryless Broadcast Channel

Consider the Gaussian DM-BC depicted in Fig. 5, where Z_1 and Z_2 are normal distributed with mean zero and variances σ_1^2 and σ_2^2, respectively. Furthermore, assume that $\sigma_2^2 > \sigma_1^2$ and that the input to the channel, i.e. X, has maximum power P, that is $\mathbb{E}\{X^2\} \leq P$. This channel is stochastically degraded, as can be seen from Fig. 6. Thus, the capacity region can be found through superposition coding, as given by the following theorem.

Theorem 3. *The capacity region of the Gaussian broadcast channel is the set of all rate pairs* (R_1, R_2) *such that*

$$R_1 \leq \frac{1}{2}\log_2\left(1 + \frac{\alpha P}{\sigma_1^2}\right), \tag{18}$$

$$R_2 \leq \frac{1}{2}\log_2\left(1 + \frac{(1-\alpha)P}{\alpha P + \sigma_2^2}\right), \tag{19}$$

for some $\alpha \in [0, 1]$.

Proof:

Achievability: To prove the achievability, we propose the following superposition coding,

$$X = V + U,$$

where V denotes the message to user 1 and U denotes the message to user 2. We take V and U independent of each other with distributions as $V \sim \mathcal{N}(0, \alpha P)$ and $U \sim \mathcal{N}(0, (1-\alpha)P)$. We this encoding scheme, we have,

$$Y_1 = X + Z_1 = V + U + Z_1,$$
$$\tilde{Y}_2 = Y_1 + W = V + U + Z_1 + W.$$

Note that all the involved random variables are Gaussian distributed by the construction. Now using the results of Theorem 2, we can write

$$\begin{aligned} R_1 \leq I(X; Y_1|U) &= h(Y_1|U) - h(Y_1|U, X) \\ &= \frac{1}{2}\log_2\left(2\pi e(\alpha P + \sigma_1^2)\right) - \frac{1}{2}\log_2\left(2\pi e \sigma_1^2\right) \\ &= \frac{1}{2}\log_2\left(1 + \frac{\alpha P}{\sigma_1^2}\right), \end{aligned} \tag{20}$$

and,

$$\begin{aligned} R_2 \leq I(U; \tilde{Y}_2) &= h(\tilde{Y}_2) - h(\tilde{Y}_2|U) \\ &= \frac{1}{2}\log_2\left(2\pi e(P + \sigma_2^2)\right) - \frac{1}{2}\log_2\left(2\pi e(\alpha P + \sigma_2^2)\right) \\ &= \frac{1}{2}\log_2\left(1 + \frac{(1-\alpha)P}{\alpha P + \sigma_2^2}\right). \end{aligned} \tag{21}$$

Converse: Suppose that there exists a $\left\{2^{nR_1}, 2^{nR_2}, n\right\}$ code with $P_e^{(n)}$ such that

$$\lim_{n\to\infty} P_e^{(n)} = 0.$$

We need to show that, the rate pair (R_1, R_2) belongs to the rate region defined by (18) and (19), for some $\alpha \in [0, 1]$. First note that for the message $M_i \in \left\{1, 2, \ldots, 2^{nR_i}\right\}$, ($i = 1, 2$), using the Fano's inequality we can write

$$H(M_i|Y_i^n) \leq (nR_i)P_{e_i}^{(n)} + 1 \leq (nR_i)P_e^{(n)} + 1, \tag{22}$$

and thus assuming equally probable messages, we have

$$I(M_i; Y_i^n) = H(M_i) - H(M_i|Y_i^n) \geq nR_i - \left(1 + nR_i P_e^{(n)}\right), \tag{23}$$

which instantly results in

$$R_1 \leq \frac{1}{n} I(M_1; Y_1^n) + \frac{1}{n} + R_1 P_e^{(n)}, \tag{24}$$

$$R_2 \leq \frac{1}{n} I(M_2; Y_2^n) + \frac{1}{n} + R_2 P_e^{(n)}. \tag{25}$$

It is therefore enough to show that there exists $0 \leq \alpha \leq 1$, such that

$$I(M_1; Y_1^n) \leq n \frac{1}{2} \log_2 \left(1 + \frac{\alpha P}{\sigma_1^2} \right), \tag{26}$$

$$I(M_2; Y_2^n) \leq n \frac{1}{2} \log_2 \left(1 + \frac{(1-\alpha) P}{\alpha P + \sigma_2^2} \right), \tag{27}$$

since in that case as $n \to \infty$, we get (18) and (19).

Note that, since $Y_2^n = X^n + Z_2^n$, we have

$$h(Y_2^n) \leq \frac{n}{2} \log_2 \left(2\pi e (P + \sigma_2^2) \right).$$

Also, we know that

$$h(Y_2^n | M_2) \leq h(Y_2^n) \leq \frac{n}{2} \log_2 \left(2\pi e (P + \sigma_2^2) \right),$$

and

$$h(Y_2^n | M_2) \geq h(Y_2^n | M_2, M_1) = \frac{n}{2} \log_2 (2\pi e \sigma_2^2).$$

Therefore, since

$$\frac{n}{2} \log_2 (2\pi e \sigma_2^2) \leq h(Y_2^n | M_2) \leq \frac{n}{2} \log_2 \left(2\pi e (P + \sigma_2^2) \right),$$

there exists some $\alpha \in [0, 1]$ such that,

$$h(Y_2^n | M_2) = \frac{n}{2} \log_2 \left(2\pi e (\alpha P + \sigma_2^2) \right). \tag{28}$$

Thus, we have

$$\begin{aligned} I(M_2; Y_2^n) &= h(Y_2^n) - h(Y_2^n | M_2) \\ &\leq \frac{n}{2} \log_2 \left(2\pi e (P + \sigma_2^2) \right) - \frac{n}{2} \log_2 \left(2\pi e (\alpha P + \sigma_2^2) \right) \\ &= \frac{n}{2} \log_2 \left(1 + \frac{(1-\alpha) P}{\alpha P + \sigma_2^2} \right), \end{aligned} \tag{29}$$

which proves (27). To prove (26), first note that since M_1 and M_2 are independent, $H(M_1|M_2) = H(M_1)$ and thus,

$$I(M_1; Y_1^n) = H(M_1) - H(M_1|Y_1^n) \leq H(M_1) - H(M_1|Y_1^n, M_2)$$
$$= H(M_1|M_2) - H(M_1|Y_1^n, M_2) = I(M_1; Y_1^n|M_2). \quad (30)$$

Since $\tilde{Y}_2^n = Y_1^n + W^n$, by using the entropy power inequality (See the appendix), we can write

$$2^{2h(\tilde{Y}_2^n|M_2)/n} \geq 2^{2h(Y_1^n|M_2)/n} + 2^{2h(W^n|M_2)/n} = 2^{2h(Y_1^n|M_2)/n} + 2\pi e(\sigma_2^2 - \sigma_1^2). \quad (31)$$

Now using (28), we have

$$h(Y_1^n|M_2) \leq \frac{n}{2}\log_2\left(2\pi e(\alpha P + \sigma_1^2)\right). \quad (32)$$

Therefore, we have

$$I(M_1; Y_1^n) \overset{(30)}{\leq} I(M_1; Y_1^n|M_2) = h(Y_1^n|M_2) - h(Y_1^n|M_1, M_2)$$
$$\leq \frac{n}{2}\log_2\left(2\pi e(\alpha P + \sigma_1^2)\right) - \frac{n}{2}\log_2\left(2\pi e\sigma_1^2\right) = \frac{n}{2}\log_2\left(1 + \frac{\alpha P}{\sigma_1^2}\right), \quad (33)$$

and the proof is complete. ■

The implication of Theorem 3 is that since user 1 experiences a better channel ($\sigma_2^2 \geq \sigma_1^2$), he can in principle decode the message U intended to user 2. He can then subtract it from the received signal Y_1 to decode for his own message. Therefore, the received signal after subtracting U is $V + Z_1$ and thus the achievable rate is $R_1 = \frac{1}{2}\log_2\left(1 + \frac{\alpha P}{\sigma_1^2}\right)$. The second user, on the other hand, treats the superposed signal V as an additive noise and hence his achievable rate is $R_2 = \frac{1}{2}\log_2\left(1 + \frac{(1-\alpha)P}{\alpha P + \sigma_2^2}\right)$.

Note that if we use orthogonal time sharing with split factor β with powers P_1 and P_2 such that the average power constraint is satisfied, i.e.,

$$\beta P_1 + (1-\beta)P_2 = P,$$

the achievable rates will be

$$R_1 = \frac{\beta}{2}\log_2\left(1 + \frac{P_1}{\sigma_1^2}\right), \quad (34)$$
$$R_2 = \frac{1-\beta}{2}\log_2\left(1 + \frac{P_2}{\sigma_2^2}\right). \quad (35)$$

It is easy to show that the time sharing scheme described above is strictly suboptimal. (Exercise)

V. K-USER DEGRADED GAUSSIAN BROADCAST CHANNEL

Consider the K user degraded Gaussian broadcast channel depicted in Fig. 7, with the maximum expected power P, i.e. $\mathbb{E}\{X^2\} \leq P$ and assume that,

$$\sigma_1^2 < \sigma_2^2 < \ldots < \sigma_K^2.$$

Using the same argument as that for the 2-user broadcast channel, we conclude that the superposition coding of the form $X = U_1 + U_2 + \ldots U_K$ where the messages intended to the users, i.e. U_i for $i = 1, \ldots, K$ are chosen independently of each other according to $U_i \sim \mathcal{N}(0, P_i)$ with

$$\sum_{i=1}^{P} P_i = P,$$

is capacity achieving, and we have

$$R_K = \frac{1}{2}\log_2\left(1 + \frac{P_K}{\sigma_K^2 + \sum_{i=1}^{K-1} P_i}\right) = \frac{1}{2}\log_2\left(1 + \frac{P_K}{P - P_K + \sigma_K^2}\right), \tag{36}$$

$$\vdots$$

$$R_k = \frac{1}{2}\log_2\left(1 + \frac{P_k}{\sigma_k^2 + \sum_{i=1}^{k-1} P_i}\right) = \frac{1}{2}\log_2\left(1 + \frac{P_k}{P + \sigma_k^2 - \sum_{i=k}^{K} P_i}\right), \tag{37}$$

$$\vdots$$

$$R_1 = \frac{1}{2}\log_2\left(1 + \frac{P_1}{\sigma_1^2}\right). \tag{38}$$

APPENDIX

ENTROPY POWER INEQUALITY (EPI)

- *Scalar EPI:* Let X and Y be two independent random variables. Then for random variable $Z \triangleq X + Y$,

$$2^{2h(Z)} \geq 2^{2h(X)} + 2^{2h(Y)}$$

  with equality if both X and Y are Gaussian.

- *Vector EPI:* Let X^n and Y^n be independent random vectors and $Z^n = X^n + Y^n$. Then

$$2^{2h(Z^n)/n} \geq 2^{2h(X^n)/n} + 2^{2h(Y^n)/n}$$

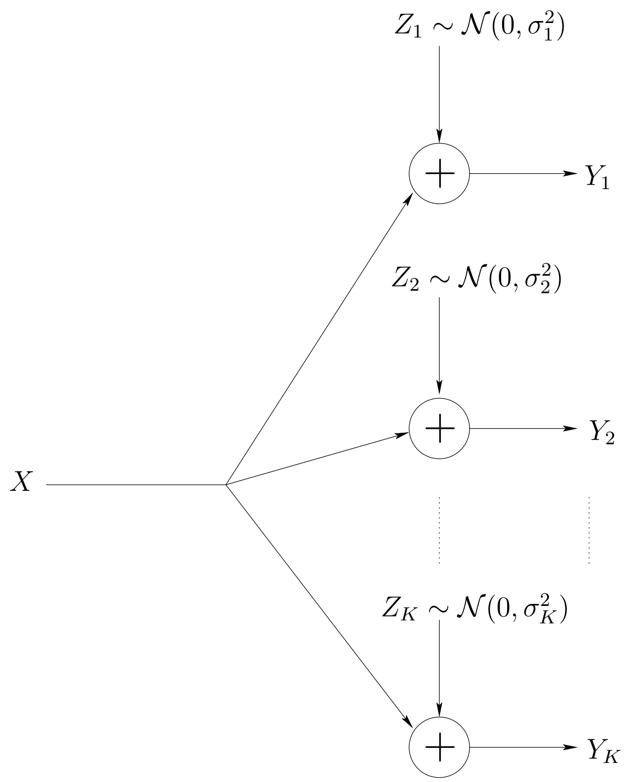

Fig. 7. Gaussian K-User discrete memoryless broadcast channel.

with equality if X^n and Y^n are Gaussian with covariance matrices $K_X = aK_Y$ for some scalar $a > 0$.

- *Conditional EPI:* Let X^n and Y^n be conditionally independent given an arbitrary random variable U, and $Z^n = X^n + Y^n$. Then

$$2^{2h(Z^n|U)/n} \geq 2^{2h(X^n|U)/n} + 2^{2h(Y^n|U)/n}.$$

References

[1] A. El Gamal and Y. H. Kim, *Network Information Theory,*, 1st edition Cambridge University Press, 2011.