MULTIPLE USER INFORMATION THEORY FOR THE GAUSSIAN CHANNEL

Thomas M. Cover

Stanford University

ABSTRACT. We consider several multiple user information channels.
Included are Shannon channels, broadcast channels, multiple access
channels, and relay channels. The key ideas needed to achieve
capacity are presented for the Gaussian version of these channels.

1. INTRODUCTION

The general problem in multiple user information theory consists
of many users attempting to communicate with one another in the
presence of arbitrary interference and noise. See van der Meulen
[1] for a review of this field.

Here we attempt to isolate some of the basic proof techniques.
These techniques involve the following ideas: 1) Random coding,
2) Superposition coding, 3) List codes, 4) Slepian-Wolf binning,
5) Block Markov decoding, 6) Joint typicality, 7) Convexifica-
tion through time-sharing.

We restrict attention to Gaussian channels, where joint typi-
cality is equivalent to a Euclidean distance condition. This dis-
cussion is expanded in a paper coauthored with Abbas El Gamal,
which may appear in the Proceedings of the IEEE [2].

2. GAUSSIAN MULTIPLE USER CHANNELS

The basic discrete time additive white Gaussian noise channel with
input power P and noise variance N is modeled by

54

$$Y_i = x_i + Z_i, \quad i = 1,2,\ldots,$$

where Z_i are independent identically distributed Gaussian random variables with mean zero and variance N. The signal $\underline{x} = (x_1, x_2, \ldots, x_n)$ has a power constraint

$$1/n \sum_{i=1}^{n} x_i^2 \leq P.$$

The Shannon capacity C, obtained by maximizing $I(X;Y)$ over all random variables X such that $E X^2 \leq P$ is given by

$$C = 1/2 \log(1 + P/N) \quad \text{bits/transmission.} \tag{2.1}$$

The continuous time Gaussian channel capacity is simply related to the discrete time capacity. If the signal $x(t)$, $0 \leq t \leq T$, has power constraint P and bandwidth constraint W, and the white noise $Z(t)$, $0 \leq t \leq T$, has power spectral density N, then the capacity of the channel $Y(t) = x(t) + Z(t)$, $0 \leq t \leq T$, is given by

$$C = W \log(1 + P/NW) \quad \text{bits/second.} \tag{2.2}$$

The relationship between (2.1) and (2.2) can be seen informally by replacing the continuous time processes by $n = 2TW$ independent samples from the process and calculating the noise variance per sample. The full theory establishing (2.2) can be found in Wyner [3], Gallager [4], and Pollack, Landau, and Slepian [5].

Having said this, we restrict our treatment to time discrete Gaussian channels.

Random Codebook: Shannon observed in 1948 that a randomly selected codebook is good with high probability when the rate R of the codebook is less than the channel capacity $C = \max I(X;Y)$. As mentioned above, for the Gaussian channel the capacity is given by $C = (1/2) \log(1 + P/N)$ bits per transmission.

We now set up a codebook that will be used in all of the multiple user channel models below. The codewords comprising the codebook are vectors of length n and power P. To generate such a random codebook, simply choose 2^{nR} independent identically distributed random n-vectors $\{\underline{x}(1), \underline{x}(2), \ldots, \underline{x}(2^{nR})\}$, each consisting of n independent Gaussian random variables with mean zero and variance P. The rate R will be specified later. Sometimes we will need two or more independently generated codebooks.

In the continuous channel case, one simply lets the white noise generator of power P and bandwidth W run for T seconds.

Every T seconds, a new codeword is generated and we list them until we fill up the codebook.

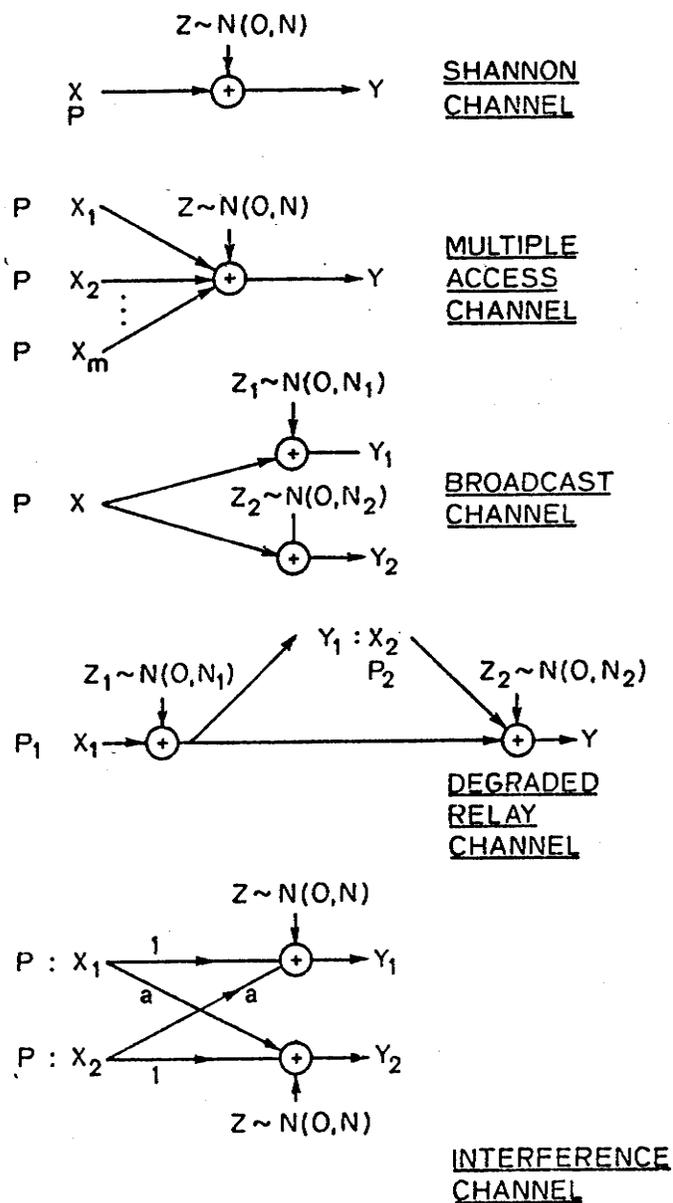Now we analyze the Gaussian channels shown in Figure 2.1.

Fig. 2.1. Gaussian Multiple User Channels

2.1 The Gaussian channel

Here $Y = x + Z$. Choose an $R < C = 1/2 \log(1 + P/N)$. Choose any index i in the set 2^{nR} . Send the ith vector $\underline{x}(i)$ from the codebook generated above. The receiver observes $\underline{Y} = \underline{x}(i) + \underline{Z}$, then finds the index $\hat{i}$ of the closest codeword to $\underline{Y}$. If n is sufficiently large, the probability of error $P(i \neq \hat{i})$ will be arbitrarily small. This minimum distance decoding scheme for the Gaussian channel is essentially equivalent to finding the codeword in the codebook which is jointly typical with the received vector $\underline{Y}$.

2.2 The multiple access channel

We consider m transmitters, each of power P. Let $Y = \sum_{1}^{m} x_i + Z$.

The capacity region for the Gaussian channel takes on the simple form given in the following equations:

$$R_i < C(P/N)$$

$$R_i + R_j < C(2P/N)$$

$$R_i + R_j + R_k < C(3P/N)$$

$$\vdots$$

$$\sum_{1}^{m} R_i \quad < C(mP/N), \tag{2.3}$$

where

$$C(x) = 1/2 \log (1 + x) \tag{2.4}$$

denotes the capacity of the Gaussian channel with signal to noise ratio x. When all the rates are the same, the last inequality dominates the others.

Here we need m codebooks, each with 2^{nR_i} codewords of power P. Transmission is simple. Each of the independent transmitters chooses whatever codeword he wishes from his own codebook. The users simultaneously send these vectors. The receiver sees the codewords added together with the Gaussian noise $\underline{Z}$.

Optimal decoding consists of looking for the m codewords, one from each codebook, such that the vector sum is closest to $\underline{Y}$ in Euclidean distance. The set of m codewords achieving the minimum distance to $\underline{Y}$ corresponds to the hypothesized collection of messages sent.

If $(R_1, R_2, \ldots, R_m)$ is in the capacity region given above, then the probability of error goes to zero as n tends to infinity.

Remarks: It is interesting to observe that the sum of the rates $C(mP/N)$ of the users goes to infinity with m. Thus in a cocktail party with m people each of power P in the presence of ambient noise N, although the interference grows as the number of speakers increases, the intended listener receives an unbounded amount of information as the number of people goes to infinity. A similar conclusion holds of course for ground communications to a satellite.

It is also interesting to note that the optimal transmission scheme here does not involve time division multiplexing. In fact, each of the transmitters utilizes the entire time to send his message.

A practical consideration for ground transmission to a satellite involves the possible inability of the ground communicators to synchronize their transmissions. Nonetheless, it can be shown that the capacity is unchanged when there is a lack of synchronization [6].

2.3 The broadcast channel

Here we assume that we have a sender of power P and two distant receivers, one with noise spectral density N_1 and the other with noise spectral density N_2. Without loss of generality, assume $N_1 < N_2$. Thus in some sense receiver Y_1 is better than receiver Y_2. The model for the channel is $Y_1 = x + Z_1$ and $Y_2 = x + Z_2$, where Z_1 and Z_2 are arbitrarily correlated Gaussian random variables with variances N_1 and N_2 respectively. The sender wishes to send independent messages at rates R_1 and R_2 to receivers Y_1 and Y_2 respectively.

Fortunately, all Gaussian broadcast channels belong to the class known as degraded broadcast channels. The capacity region for the Gaussian broadcast channel is given by

$$R_1 < C(\alpha P/N_1)$$
$$R_2 < C(\overline{\alpha} P/(\alpha P + N_2)), \tag{2.5}$$

where $0 \leq \alpha \leq 1$, $\overline{\alpha} = 1 - \alpha$. The parameter α may be arbitrarily chosen to trade off rate R_1 for rate R_2 as the transmitter wishes.

To encode the messages, the receiver generates two codebooks, one with power αP at rate R_1, and another codebook with power

$\bar{\alpha} P$ and rate R_2. He has chosen R_1 and R_2 to satisfy the equation above. Then, to send an index $i \in \left\{1,2,\ldots,2^{nR_1}\right\}$ and $j \in \left\{1,2,\ldots,2^{nR_2}\right\}$ to Y_1 and Y_2 respectively, he takes codeword $\underline{x}(i)$ from the first codebook and codeword $\underline{x}(j)$ from the second codebook and computes the sum. He then sends the sum over the channel.

Two receivers must now do the decoding. First consider the bad receiver Y_2. He merely looks through the second codebook for the closest codeword to his received vector $\underline{Y}_2$. His effective signal-to-noise ratio is $\bar{\alpha}P/\alpha P + N_2$, since Y_1's message acts as noise to Y_2. The good receiver Y_1 first decodes Y_2's codewor which he can do because of his lower noise N_1. He subtracts this codeword $\hat{\underline{x}}_2$ from $\underline{Y}_1$. This leaves him with a channel of power αP and noise N_1. He then looks for the closest codeword in the first codebook to $\underline{Y}_1 - \underline{x}_2$. The resulting probability of error can be made as low as wished.

A nice dividend of optimal encoding for degraded broadcast channels is that the better receiver Y_1 always knows the message intended for receiver Y_2 in addition to the extra information intended for himself.

2.4 The relay channel

For the relay channel, we have a sender X_1 and an ultimate intended receiver Y. Also present, however, is the relay channel intended solely to help the sender. The channel is given by

$$Y_1 = x_1 + Z_1$$
$$Y_2 = x_1 + Z_1 + x_2 + Z_2 , \qquad (2.6)$$

where Z_1, Z_2 are independent zero mean Gaussian random variables with variance N_1, N_2 respectively. The allowed encoding by the relay is the causal sequence

$$x_{2i} = f_i(y_{11}, y_{12}, \ldots, y_{1i-1}) . \qquad (2.7)$$

The sender X_1 has power P_1 and the relay X_2 has power P_2.

The capacity is given by

$$C = \max_{0 < \alpha < 1} \min \left\{ C\left(\frac{P_1 + P_2 + 2\sqrt{\alpha P_1 P_2}}{N_1 + N_2}\right), C\left(\frac{\alpha P_1}{N_1}\right) \right\} ,$$

where $\bar{\alpha} = 1 - \alpha$.

Note that if

$$P_2/N_2 \geq P_1/N_1 \, ,$$

it can be seen that $C^* = C(P_1/N_1)$. (This is achieved by $\alpha = 1$.)
The channel appears to be noise free after the relay, and the
capacity $C(P_1/N_1)$ from x_1 to the relay can be achieved. Thus
the rate without the relay $C(P_1/(N_1 + N_2))$ is increased by the
relay to $C(P_1/N_1)$. For large N_2, and for $P_2/N_2 \geq P_1/N_1$, we
see that the increment in rate is from $C(P_1/(N_1 + N_2)) \approx 0$ to
$C(P_1/N_1)$.

<u>Encoding of information</u>: Two codebooks are needed. The
first codebook has 2^{nR_1} words of power αP_1. The second has 2^{nR_0}
codewords of power $\bar{\alpha} P_1$. We shall use words from these codebooks
successively in order to create the opportunity for cooperation
by the relay. We start by sending a codeword from the first code-
book. The relay now knows the index of this codeword since
$R_1 < C(\alpha P_1/N_1)$, but the intended receiver does not. However, the
intended receiver has a list of possible codewords of size
$2^{n(R_1 - C(\alpha P_1/N_1 + N_2))}$. The last calculation involves a result on
list codes.

In the next block the relay and the transmitter would like to
cooperate to resolve the receiver's uncertainty about the previously
sent codeword on the receiver's list. Unfortunately, they cannot
quite be sure what this list is. They do not know the received
signal Y. Thus they randomly partition the first codebook into
2^{nR_0} cells with an equal number of codewords in each cell. The
relay, the receiver, and the transmitter agree on what this parti-
tion is. The relay and the transmitter find the cell of the par-
tition in which the codeword from the first codebook lies and
cooperatively send the codeword from the second codebook with that
index. That is, both X_1 and X_2 send the same designated code-
word. The relay, of course, must scale this codeword so that it
meets his power constraint P_2. They now simultaneously transmit
their codewords. An important point is that the cooperative infor-
mation sent by the relay and the transmitter is sent coherently.
So the power of the sum as seen by the receiver Y is

$$\left(\sqrt{\bar{\alpha} P_1} + \sqrt{P_2} \right)^2 .$$

However, this does not exhaust what the transmitter does in
the second block. He also chooses a fresh codeword from his first

codebook, adds it "on paper" to the cooperative codeword from his second codebook, and sends this sum over the channel.

The reception by the ultimate receiver Y in the second block involves first finding the cooperative index from the second codebook by looking for the closest codeword in the second codebook. He subtracts it off, then calculates a list of indices of size 2^{nR_0} corresponding to all transmitted words from the first book which might have been sent in that second block.

Now it is time for the intended receiver Y to finish computing the codeword from the first codebook sent in the first block. He takes his list of possible codewords that might have been sent in the first block and intersects it with the cell of the partition that he has learned from the cooperative relay transmission in the second block. Since the rates and powers have been chosen judiciously, it is highly probable that there will be only one codeword in this intersection. This is codeword Y's guess about the information sent in the first block.

We are now in steady state. In each new block, the transmitter and the relay cooperate to resolve the list uncertainty from the previous block. In addition, the transmitter adds some fresh information from his first codebook to his transmission from the second codebook and transmit the sum.

The receiver is always one block behind, but for sufficiently many blocks, this does not affect his overall rate of reception.

2.5 The interference channel

In the interference channel, we have two senders and two receivers. Sender 1 wishes to send information to receiver 1. He does not care what receiver 2 receives or understands. Similarly, with sender 2 and receiver 2. As can be seen, this channel involves interference of each user with the other. It is not quite a broadcast channel because there is only one intended receiver for each sender, nor is it quite a multiple access channel because each receiver is only interested in what is being sent by the corresponding transmitter.

This channel has not been solved in general, even in the Gaussian case. But remarkably, in the case of high interference, Carleial [7] has shown that the solution to this channel is the same as if there were no interference whatsoever. To achieve this, generate two codebooks, each with power P and rate C(P/N). Each sender independently chooses a word from his book and sends it. Now, if the interference is sufficiently high, the first receiver

can understand perfectly the index of the second transmitter. He finds it by the usual technique of looking for the closest codeword to his received signal. Once he finds this signal, he subtracts it from his received waveform. Now there is a clean channel between the receiver and his sender. He then searches his sender's codebook to find the closest codeword and declares that codeword to be the one sent.

3. CONCLUSIONS

Random codes were used for all the results stated in the last section. Convexification through time sharing is generally needed for rate regions involving more than one rate. Thus time sharing appears not to be needed for the Shannon channel and the relay channel. Superposition coding achieves nothing new in the multiple access channel (unless there is feedback), but appears to be essential for the broadcast and relay channels. Finally, list codes, Slepian-Wolf binning, and block Markov encoding arise in the relay channel.

It should be mentioned that there are more proof techniques that would have to be mentioned if we were to discuss all of the multiple user channels that have been studied to date.

REFERENCES

1. E.C. van der Meulen, "A Survey of Multi-way Channels in Information Theory: 1961-1976," IEEE Trans. on Information Theory, Vol. IT-23, No. 2, January 1977.
2. A. El Gamal and T. Cover, "Multiple User Information Theory," submitted as an Invited Paper to Proc. IEEE, 1980.
3. A.D. Wyner, "The Capacity of the Band-limited Gaussian Channel," Bell System Tech. J., Vol. 45, March 1965, pp. 359-371.
4. R.G. Gallager, Information Theory and Reliable Communication, Wiley, New York, 1968.
5. D. Slepian and H.O. Pollak, "Prolate Spheroidal Wave Functions, Fourier Analysis, and Uncertainty-I," Bell System Tech. J., Vol. 40, pp. 43-64. (Also see Landau and Pollak for Parts II and III).
6. T. Cover, R.J. McEliece, and E. Posner, "Asynchronous Multiple Access Channel Capacity," Technical Report No. 35, Dept. of Statistics, Stanford University, Stanford, CA, December 1978, to appear IEEE Trans. on Information Theory.
7. A.B. Carleial, "A Case where Interference does not Reduce Capacity," IEEE Trans. on Information Theory, Vol. IT-21, September 1975, pp. 569-570.