ECE 361: Lecture 8: Energy-Efficient Communication – Part II

In Lecture Note 7, we considered three different forms of M-ary orthogonal communication schemes. In all three systems, the receiver is confronted by M independent Gaussian random variables $\mathbb{Y}_0, \mathbb{Y}_1, \ldots, \mathbb{Y}_{n-1}$ with common variance σ^2. *One* of these random variables has a large mean $\sqrt{\mathcal{E}_s}$, all others have mean 0, and it is the receiver's job to decide which of the random variables is the one with the large mean. In this Lecture, we show that the receiver can make such a decision with very high reliability, and that M-ary orthogonal communication schemes are *energy efficient*: provided that $\mathcal{E}_b$, the energy transmitted *per bit*, is larger than $2\sigma^2 \ln 2$, the error probability can be made as small as desired by increasing M. Of course, as noted previously, the data rate $(\log_2 M)/M$ bits per channel use decreases rapidly towards 0 as M increases, and so the schemes are *not rate efficient*.

8.1. Error Probabilities in M-ary Orthogonal Signaling

The receiver in an M-ary orthogonal communication schemes has available to it M random variables $\mathbb{Y}_0$, $\mathbb{Y}_1, \ldots, \mathbb{Y}_{M-1}$. Given that signal s_i was transmitted, these are conditionally independent Gaussian random variables with common variance σ^2 and with means $\mathsf{E}[\mathbb{Y}_i] = \sqrt{\mathcal{E}_s}$, $\mathsf{E}[\mathbb{Y}_j] = 0$ for $j \neq i$. It should be intuitively obvious[1] that if the largest of the observations$\mathbb{Y}_0, \mathbb{Y}_1, \ldots, \mathbb{Y}_{M-1}$ is $\mathbb{Y}_j$, then the receiver decides that s_j was the transmitted signal. If $\mathbb{Y}_i$ is the largest observation, then the receiver decision (that s_i was the transmitted signal) is correct. What is the probability of this occurring?

8.1.1. The Probability of a Correct Decision

Let C denote the event that the receiver decision is correct. Since we are assuming that s_i is the transmitted signal, we have that

$$P(C \mid s_i \text{ transmitted}) = P\big\{\mathbb{Y}_i > \max\{\mathbb{Y}_0, \mathbb{Y}_1, \ldots, \mathbb{Y}_{i-1}, \mathbb{Y}_{i+1}, \ldots, \mathbb{Y}_{M-1}\}\big\}$$

To calculate this probability, let us suppose that $\mathbb{Y}_i = \alpha$. Then,

$$\begin{aligned} P(C \mid s_i \text{ transmitted}, \mathbb{Y}_i = \alpha) &= P\{\mathbb{Y}_0 < \alpha, \mathbb{Y}_1 < \alpha, \ldots, \mathbb{Y}_{i-1} < \alpha, \mathbb{Y}_{i+1} < \alpha, \ldots, \mathbb{Y}_{M-1} < \alpha\} \\ &= P\{\mathbb{Y}_0 < \alpha\}P\{\mathbb{Y}_1 < \alpha\} \cdots P\{\mathbb{Y}_{i-1} < \alpha\}P\{\mathbb{Y}_{i+1} < \alpha\} \cdots P\{\mathbb{Y}_{M-1} < \alpha\} \\ &= \big[\Phi(\alpha/\sigma)\big]^{M-1} \end{aligned}$$

where we have used the independence of the $\mathbb{Y}$'s and the fact that they all are $\mathcal{N}(0, \sigma^2)$ random variables. The law of total probability tells us that we can remove the conditioning on the value of $\mathbb{Y}_i$ by multiplying by the pdf $\sigma^{-1}\phi((\alpha - \sqrt{\mathcal{E}_s})/\sigma)$ of $\mathbb{Y}_i$ and integrating with respect to α. In fact, the symmetry of the problem suggests that $P(C \mid s_i \text{ transmitted})$ is the same for all choices of i and so $P(C)$ is the same as $P(C \mid s_i \text{ transmitted})$. Thus we have (after making a change of variable $x = \alpha/\sigma$) that

$$P(C) = \int_{-\infty}^{\infty} \left[\Phi\left(\frac{\alpha}{\sigma}\right)\right]^{M-1} \frac{1}{\sigma}\phi\left(\frac{\alpha - \sqrt{\mathcal{E}_s}}{\sigma}\right) d\alpha = \int_{-\infty}^{\infty} [\Phi(x)]^{M-1} \phi(x - \mu)\, dx \tag{8.1}$$

where $\mu = \sqrt{\mathcal{E}_s/\sigma^2}$ is a measure of the signal to noise ratio SNR. Except when $M = 2$, this integral must be evaluated numerically. Fortunately, since $\phi(\cdot)$ is a rapidly decreasing function of its argument, it is not too hard to evaluate $P(C)$ quite accurately (e.g. to four or five significant figures, that is, to within $\pm 0.01\%$ or better accuracy) via numerical integration.

[1] ...and if it is not, work out the details and prove it for yourself!

8.1.2. The Probability of Error

From (8.1), it follows readily that the error probability of an M-ary orthogonal communication scheme is

$$P(E) = 1 - P(C) = 1 - \int_{-\infty}^{\infty} [\Phi(x)]^{M-1} \phi(x-\mu)\, dx. \tag{8.2}$$

This equation is much beloved of textbook writers, but is an absolutely terrible way of computing $P(E)$! We are interested in *small* values of $P(E)$ and taking the difference of two nearly equal numbers to find $P(E)$ is not a good idea. The "answer" is likely to be mostly round-off errors encountered in the numerical integration routines used in evaluating (8.1). It is far better to develop a different integral that will allow us to find $P(E)$ directly. Integrating the right side of (8.2) by parts, we have

$$\begin{aligned} P(E) &= 1 - \int_{-\infty}^{\infty} [\Phi(x)]^{M-1} \phi(x-\mu)\, dx \\ &= 1 - \Big[[\Phi(x)]^{M-1} \Phi(x-\mu) \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} (M-1) [\Phi(x)]^{M-2} \phi(x)\Phi(x-\mu)\, dx \end{aligned}$$

and since $\Phi(x)$ and $\Phi(x-\mu)$ converge to 1 and 0 respectively as $x \to \infty$ and $x \to -\infty$, we have that

$$P(E) = (M-1) \int_{-\infty}^{\infty} [\Phi(x)]^{M-2} \Phi(x-\mu)\phi(x)\, dx \text{ where } \mu = \sqrt{\mathcal{E}_s/\sigma^2}. \tag{8.3}$$

The integral in (8.3) also requires numerical evaluation but yields $P(E)$ directly instead of as the difference of two nearly equal numbers. Next, note that symmetry suggests that when an error occurs, *any* of the other $M-1$ signals is equally likely to be chosen by the receiver, and thus the probability that the receiver decides that s_j was transmitted when s_i was in fact the transmitted signal is the integral shown in (8.3). This can also be obtained more directly. The receiver decides that s_j was the transmitted signal exactly when $\mathbb{Y}_j$ is the largest observation. Conditioned on $\mathbb{Y}_j = \alpha$, the probability that all the other observations are smaller is just $[\Phi(\alpha/\sigma)]^{M-2} \Phi((\alpha - \sqrt{\mathcal{E}_s})/\sigma)$. The unconditional probability is obtained by multiplying by the pdf $\sigma^{-1}\phi(\alpha/\sigma)$ of $\mathbb{Y}_j$ and integrating, and with the change of variable $x = \alpha/\sigma$ (just as before!), we get

$$P\{\text{receiver decides } s_j \text{ transmitted} \mid s_i \text{ transmitted}\} = \int_{-\infty}^{\infty} [\Phi(x)]^{M-2} \Phi(x-\mu)\phi(x)\, dx = \frac{1}{M-1} P(E). \tag{8.4}$$

Thus, when an error occurs, the receiver decision is equally likely to be any of the other $M-1$ signals.

8.1.3. The Union Bound on Error Probability

Instead of doing a numerical integration, we can massage (8.3) to get an easily computable upper bound on $P(E)$ that is generally called the *union bound* on error probability. First, let us replace $[\Phi(x)]^{M-2} < 1$ in (8.3) by 1 to get that

$$P(E) < (M-1) \int_{-\infty}^{\infty} \phi(x)\Phi(x-\mu)\, dx. \tag{8.5}$$

Now, if $\mathbb{X} \sim \mathcal{N}(0,1)$ and $\mathbb{W} \sim \mathcal{N}(\mu,1)$ are independent random variables, then $P\{\mathbb{W} < \mathbb{X} \mid \mathbb{X} = x\} = \Phi(x-\mu)$ and thus $P\{\mathbb{W} < \mathbb{X}\}$ can be obtained by multipliying $P\{\mathbb{W} < \mathbb{X} \mid \mathbb{X} = x\} = \Phi(x-\mu)$ by the pdf of $\mathbb{X}$ and integrating. Thus, the integral in (8.5) can be recognized as $P\{\mathbb{W} < \mathbb{X}\}$. But, $\mathbb{W} - \mathbb{X} \sim \mathcal{N}(\mu, 2)$ and so $P\{\mathbb{W} < \mathbb{X}\} = P\{\mathbb{W} - \mathbb{X} < 0\} = Q(\mu/\sqrt{2})$. Substituting into (8.5), we get

$$P(E) < (M-1)Q(\mu/\sqrt{2}) = (M-1)Q(\sqrt{\mathcal{E}_s/2\sigma^2}). \tag{8.6}$$

The reason that (8.6) is called the union bound is that if we define $M-1$ events $A_j = \{\mathbb{Y}_j > \mathbb{Y}_i\}$, $j \neq i$, then we have that $P(A_j) = P\{\mathbb{Y}_j > \mathbb{Y}_i\} = P\{\mathbb{Y}_j - \mathbb{Y}_i > 0\} = Q(\sqrt{\mathcal{E}_s/2\sigma^2})$ since $\mathbb{Y}_j - \mathbb{Y}_i \sim \mathcal{N}(-\sqrt{\mathcal{E}_s}, 2\sigma^2)$. But, an error occurs if and only if *at least one* of the events A_j occurs, that is, $E = \bigcup_j A_j$. But *the probability* *of a union is no larger than the sum of the probabilities:* $P(E) \le \sum_j P(A_j)$ with equality exactly when $P(A_j \cap A_k) = 0$ for all $j \ne k$, which does not hold in this instance. Therefore,

$$P(E) < \sum_j P(A_j) = (M-1)Q(\sqrt{\mathcal{E}_s/2\sigma^2})$$

as claimed in (8.6). In fact, the union bound is typically derived using the second argument presented above. Numerically, the union bound on $P(E)$ is *very tight* for large SNR. But at low SNR, the bound can be much larger than the actual value of $P(E)$.

A tighter result than the union bound, which we present without a proof, is

$$P(E) < 1 - \left[1 - Q(\sqrt{\mathcal{E}_s/2\sigma^2})\right]^{M-1}. \tag{8.7}$$

Of course, this seems to be taking the difference of two nearly equal numbers, a practice that we have excoriated above, but expanding out the right side of (8.7) via the binomial theorem gives

$$P(E) < (M-1)Q(\sqrt{\mathcal{E}_s/2\sigma^2}) - \binom{M-1}{2}\left[Q(\sqrt{\mathcal{E}_s/2\sigma^2})\right]^2 + \binom{M-1}{3}\left[Q(\sqrt{\mathcal{E}_s/2\sigma^2})\right]^3 - \cdots$$

where the first term is the union bound, the first two terms give a lower bound on the right side of (8.7), the first three give an upper bound on the right side of (8.7), and so on. Note also that if we write $\Phi(x) = 1 - Q(x)$ in (8.2) and expand via the binomial theorem, we get the *exact* value of $P(E)$ as

$$P(E) = (M-1)\int_{-\infty}^{\infty} \phi(x-\mu)Q(x)\,dx - \binom{M-1}{2}\int_{-\infty}^{\infty} \phi(x-\mu)[Q(x)]^2\,dx + \cdots$$

where the first integral can be shown to have value $Q(\mu/\sqrt{2})$ via the same argument that we used in getting to (8.6) from (8.5), that is, the first term above is the union bound.

8.1.4. Bit error probabilities

M-ary orthogonal signaling communicates $\log_2 M$ bits in M channel uses. Let $m = \lceil \log_2 M \rceil$ so that the messages can be represented as M of the 2^m binary m-bit vectors. Note that $M > 2^{m-1}$. Now suppose that a_0 of these M vectors have a 0 in the k-th position and $a_1 = M - a_0$ have a 1 in the k-th position. Typically, $a_0 \approx a_1 \approx M/2$. (Why?) What is the probability of a bit error in the k-th position? Well, if a message with a 0 in the k-th position is transmitted, then a bit error occurs in the k-th position if the receiver chooses any of the a_1 messages that have a 1 in the k-th position. According to (8.4), this has probability $a_1 P(E)/(M-1)$ of occurring. Similarly, if a message with a 1 in the k-th position is transmitted, the probability of a bit error in the k-th position is $a_0 P(E)/(M-1)$. Now, messages with a 0 in the k-th position are transmitted with probability a_0/M while messages with a 1 in the k-th position are transmitted with probability $a_1/M = 1 - a_0/M$. Thus, the probability of a bit error in the k-th position is

$$P_{b,k} = \frac{a_1 P(E)}{M-1} \times \frac{a_0}{M} + \frac{a_0 P(E)}{M-1} \times \frac{a_1}{M} = \frac{P(E)}{M(M-1)} \times 2a_0(M-a_0) \le \frac{M/2}{M-1}P(E)$$

since $2a_0(M-a_0)$ has maximum value $M^2/2$ when $a_0 = M - a_0 = M/2$. Thus, in each position, the bit error probability is, at worst, *slightly more* than one-half of the symbol error probability $P(E)$. If $M = 2^m$ where m is an integer, then $P_{b,k} = (2^{m-1}/(2^m-1))P(E)$ for all m bit positions.

8.2. Energy Efficiency of M-ary Orthogonal Communications

Throughout this section, we assume that $M = 2^m$ where m is an integer. With 2^m-ary orthogonal communication, we can transmit m bits over 2^m channel uses with a total energy $\mathcal{E}_s$ with probability of error $P(E)$

given by (8.3). The energy *per bit* is $\mathcal{E}_b = \mathcal{E}_s/m$. Now, the union bound (8.6) together with the standard result $Q(x) \le \frac{1}{2}\exp(-x^2/2)$ for $x \ge 0$ gives us that

$$P(E) < (M-1)Q(\sqrt{\mathcal{E}_s/2\sigma^2}) = \frac{2^m-1}{2}\exp(-m\mathcal{E}_b/4\sigma^2) < 2^m \exp(-m\mathcal{E}_b/4\sigma^2).$$

It follows that $P(E) < \exp(-m(\mathcal{E}_b/4\sigma^2 - \ln 2))$ where the argument of the exponential is negative as long as $\mathcal{E}_b > 4\sigma^2 \ln 2$. Thus, if the energy per bit is larger than $4\sigma^2 \ln 2$, we can make $P(E)$ as small as we like by increasing the value of m. In other words, M-ary orthogonal communication schemes are *energy efficient*. Provided that $\mathcal{E}_b$, the received energy per bit, is larger than the threshold value $4\sigma^2 \ln 2$, we can make $P(E)$ as small as we like. This is in contrast to repetition coding where small $P(E)$ is achieved at the expense of (very rapidly) increasing $\mathcal{E}_b$.

It turns out that the minimum energy requirement stated in the previous paragraph is excessive, and can be attributed to the use of the union bound on $P(E)$ and the use of the weak bound $Q(x) \le \frac{1}{2}\exp(-x^2/2)$, which, as we have noted before, is not tight for large x. In fact, a careful analysis of the behavior of $P(C)$ as given in (8.1) shows that bit energy larger that only $2\sigma^2 \ln 2$ suffices to drive $P(E)$ to as small a value as we like. The outline of this argument is as follows.

Let us write $P(C) = \int_{-\infty}^{\infty} [\Phi(x)]^{M-1}\,\phi(x-\mu)\,dx = \int_{-\infty}^{\infty} [\Phi(x+\mu)]^{M-1}\,\phi(x)\,dx$ where $\mu = \sqrt{m\mathcal{E}_b/\sigma^2}$. Writing $\beta = \mathcal{E}_b/(2\sigma^2 \ln 2)$ so that $\mu = \sqrt{2m\beta \ln 2} = \sqrt{2\beta}\sqrt{\ln M}$, we have

$$P(C) = \int_{-\infty}^{\infty} \left[\Phi(x+\sqrt{2\beta}\sqrt{\ln M})\right]^{M-1} \phi(x)\,dx$$

It turns out that $\lim_{M\to\infty} \left[\Phi(x+\sqrt{2\beta}\sqrt{\ln M})\right]^{M-1} = \begin{cases} 1, & \text{if } \beta > 1, \\ 0, & \text{if } \beta < 1, \end{cases}$ and consequently we have that

$$\lim_{M\to\infty} P(C) = \begin{cases} 1, & \text{if } \mathcal{E}_b > 2\sigma^2 \ln 2, \\ 0, & \text{if } \mathcal{E}_b < 2\sigma^2 \ln 2. \end{cases} \quad \text{Equivalently,} \quad \lim_{M\to\infty} P(E) = \begin{cases} 0, & \text{if } \mathcal{E}_b > 2\sigma^2 \ln 2, \\ 1, & \text{if } \mathcal{E}_b < 2\sigma^2 \ln 2. \end{cases}$$

This result not only shows that the threshold on $\mathcal{E}_b$ is one-half of what we got from the union bound argument, but also shows the negative result that if $\mathcal{E}_b < 2\sigma^2 \ln 2$, then matters only get worse as M (or equivalently, m) is increased since $P(E)$ increases to 1 rather than decreasing to 0. Since the bit error probability is slightly more than $\frac{1}{2}P(E)$, the bit error probability converges to $\frac{1}{2}$ when $\mathcal{E}_b < 2\sigma^2 \ln 2$. Thus, $2\sigma^2 \ln 2$ is the dividing line between good and bad performance of M-ary orthogonal communication systems.

The rest of this document is not required reading for this course.

The key to deriving all these interesting results is the limit of $\left[\Phi(x+\sqrt{2\beta}\sqrt{\ln M})\right]^{M-1}$. Let us look at

$$\lim_{M\to\infty} \ln \left[\Phi(x+\sqrt{2\beta}\sqrt{\ln M})\right]^{M-1} = \lim_{M\to\infty} \frac{\ln \Phi(x+\sqrt{2\beta}\sqrt{\ln M})}{(M-1)^{-1}}$$

which is of the indeterminate form 0/0 (note that the *argument* of $\Phi(\cdot)$ goes to ∞). Treating M as a real number, applying L'Hôpital's rule, and ignoring factors that we know are converging to 1, it can be shown that

$$\lim_{M\to\infty} \frac{\ln \Phi(x+\sqrt{2\beta}\sqrt{\ln M})}{(M-1)^{-1}} = \lim_{M\to\infty} \frac{(M-1)^2}{M^{1+\beta}}(\ln M)^{-1/2}\frac{1}{\exp(x\sqrt{2\beta}\sqrt{\ln M})}.$$

(Hey, nobody said this was easy stuff!) Further careful analysis of the cases $\beta > 1$ and $\beta < 1$ shows that the limit of $\ln\left[\Phi(x+\sqrt{2\beta}\sqrt{\ln M})\right]^{M-1}$ is 0 or $-\infty$ according as $\beta > 1$ or $\beta < 1$, and so $\left[\Phi(x+\sqrt{2\beta}\sqrt{\ln M})\right]^{M-1}$ converges to 1 or 0 according as $\beta > 1$ or $\beta < 1$.