ECE 361: Lecture 7: Energy-Efficient Communication – Part I

In Lecture Note 6, we considered the transmission of data bits over a discrete-time Gaussian channel with a peak power constraint and found that if the received energy per channel use was too small, then binary repetition coding could increase the SNR and reduce the error probability to as small a value as desired, but only at the price of making the data rate very small. Use of 2^B-ary ASK with repetition coding allows us to reduce the probability of error for a B-bit block of data, but as the formula

$$P(E) = 2\left(1 - \frac{1}{2^B}\right) Q\left(\frac{\sqrt{n\mathcal{E}/\sigma^2}}{2^B - 1}\right)$$

shows, n must be on the order of 2^{2B} to make the argument of $Q(\cdot)$ small enough, that is, a huge number of pulses must be transmitted to get reliable communication of B bits. Thus, both the delay and the total energy required with repetition coding increase exponentially with B. In particular, $n\mathcal{E}$ is used to transmit B bits, making $\mathcal{E}_b$, the *energy per bit* have value $n\mathcal{E}/B \approx 2^{2B}\mathcal{E}/B$ which also increases rapidly, and in fact, is infeasible for any value of B that we might reasonably expect as the packet size. In other words, while repetition coding can reduce the error probability to any small value that we desire, this form of coding is *not energy efficient* because the required energy per bit $\mathcal{E}_b$ is large. Repetition coding is *not rate efficient* either because the data rate $R = B/n \approx B/2^{2B}$ bits per channel use is very small. Indeed, $\mathcal{E}_b$ increases without bound, and R decreases towards 0 as the block size B increases.

7.1. Average Power Constraint

We have thus far considered the performance of repetition coding with a peak power constraint, and the reader might well wonder if matters would be any better with the less stringent average power constraint. If the system is required to satisfy only

$$\frac{1}{n}\sum_{i=0}^{n-1} \mathsf{E}[\mathbb{X}_i^2] \le \mathcal{E} \tag{7.1}$$

but no peak power constraint, then of course we could still insist on using the same repetition coding method. With binary repetition coding, the received energy per channel use would be exactly $\mathcal{E}$, the decision would be made on the basis of $\sum \mathbb{Y}_i$ etc, and the bit error probability would be $Q(\sqrt{n\mathcal{E}/\sigma^2})$. The details with 2^B-ary ASK would be slightly different because the average energy is roughly one-third of the maximum energy but the conclusions would still be the same. But, there is an *alternative* form of communication that can be used in the absence of a peak power constraint. This method requires an *enormous* peak power, and is not at all practical. Unfortunately, this method does not improve the performance at all as compared to repetition coding. But we nonetheless study it, albeit briefly, because it allows a good approach to a better form of coding than repetition coding. In particular, this better form of coding, called *orthogonal* signaling is *energy efficient*. Provided that $\mathcal{E}_b/\sigma^2$ is larger than a small threshold value, the error probability can be made as small as we wish. Thus, unlike with repetition coding, a fixed amount of energy per bit suffices to give arbitrarily small error probabilities. Unfortunately, this orthogonal coding scheme is *not rate efficient*: we still need 2^B channel uses to transmit B bits.

Given that the communication system must satisfy (7.1) but that, subject to this constraint being satisfied, an arbitrarily large amount of energy may be transmitted during any channel use, consider a 2-ASK signaling scheme in which *all* the available energy $n\mathcal{E}$ is transmitted using $\mathbb{X}_0$ that takes on values $\pm\sqrt{n\mathcal{E}}$ with equal probability $\frac{1}{2}$ while $\mathbb{X}_1 = \mathbb{X}_2 = \mathbb{X}_{n-1} = 0$ with probability 1. Note that (7.1) is satisfied. The receiver thus observes $\mathbb{Y}_0 \sim \mathcal{N}(\pm\sqrt{n\mathcal{E}}, \sigma^2)$ but ignores $\mathbb{Y}_1, \mathbb{Y}_2, \ldots, \mathbb{Y}_{n-1}$ since they contain no information about the transmitted bit. The decision is made from the observation $\mathbb{Y}_0$ and the error probability is obviously $Q(\sqrt{n\mathcal{E}/\sigma^2})$, the same as with repetition coding with a peak power constraint.

A slightly different scheme can be used in channels in which the $\mathbb{X}_i$ are constrained to be nonnegative. Such restrictions arise in systems where signals cannot take on both positive and negative values (e.g. optical channels where we can "transmit light but not dark".) Consider an ON-OFF scheme in which *all* the available energy $n\mathcal{E}$ is transmitted using $\mathbb{X}_0$ that takes on values $\sqrt{2n\mathcal{E}}$ and 0 with equal probability $\frac{1}{2}$ while $\mathbb{X}_1 = \mathbb{X}_2 = \mathbb{X}_{n-1} = 0$ with probability 1. Note that (7.1) is satisfied. The receiver thus observes $\mathbb{Y}_0$ (which is a Gaussian random variable with variance σ^2 and mean $\sqrt{2n\mathcal{E}}$ or 0 depending on the bit is being transmitted) but ignores $\mathbb{Y}_1, \mathbb{Y}_2, \ldots, \mathbb{Y}_{n-1}$ since they contain no information about the transmitted bit. The decision is made from the observation $\mathbb{Y}_0$ and the error probability is obviously $Q((\sqrt{2n\mathcal{E}} - 0)/2\sigma) = Q(\sqrt{n\mathcal{E}/2\sigma^2})$ which is not quite as good as with 2-ASK, but then, 2-ASK is not usable on the channel being considered here.

Thus, weakening the energy constraint does not seem to improve matters very much, but a simple modification to the schemes presented in the above two paragraphs yields energy-efficient coding systems. We discuss these in the next section.

7.2. Orthogonal Signaling

7.2.1. ON-OFF Keying

The ON-OFF scheme discussed in the previous section blasts a huge amount of energy across the channel during the first channel use and is silent thereafter for the remaining $n-1$ channel uses, and merely manages to send one bit of data with all this effort. The receiver makes its decision based on $\mathbb{Y}_0$ and ignores $\mathbb{Y}_1, \mathbb{Y}_2, \ldots, \mathbb{Y}_{n-1}$. But now consider a different communication scheme in which we transmit B bits of data across the channel by setting $n = 2^B$ and transmitting a pulse of energy $\mathcal{E}_s$ during the i-th channel use to let the receiver know that the B bits of data represent the integer i, $0 \le i \le 2^B - 1$. This technique is referred to as *pulse position modulation* (PPM) since the information is being conveyed by the position of the pulse of energy among the 2^B channel uses. Now the receiver has 2^B Gaussian random variables $\mathbb{Y}_0$, $\mathbb{Y}_1$, $\ldots$, $\mathbb{Y}_{2^B-1}$ of variance σ^2, of which *one* random variable has mean $\mathcal{E}_s$ while all the other random variables have zero mean. If the receiver can identify with high reliability which of the $\mathbb{Y}$'s is the one with large mean, then we will have succeeded in transmitting B bits of data across the channel using energy $\mathcal{E}_b = \mathcal{E}_s/B$ per bit. We will discuss the probability of error of this scheme in the next Lecture, and here merely remark that this scheme is *energy efficient* in that B bits can be transmitted with high reliability for a *fixed* amount of energy per bit. The reliability increases as we increase B, and $\mathcal{E}_s$ also increases as we increase B, but $\mathcal{E}_s$ increases as a *linear* function of B, that is, $\mathcal{E}_b = \mathcal{E}_s/B$ is fixed instead of increasing exponentially with B as in repetition coding. However, this communication scheme is nonetheless *not rate efficient* because we still need 2^B channel uses to transmit B bits, so that the data rate is $R = B/2^B$ bits per channel use, larger than the $\approx B/2^{2B}$ rate achieved with repetition coding, but nevertheless decreasing rapidly towards 0 as we increase B.

The scheme described above is one of a class of schemes referred to as *orthogonal signaling schemes* because the 2^B signals being used are

$$\begin{array}{lclllll} \underline{s}_0 & = & [\sqrt{\mathcal{E}_s}, & 0, & \ldots & 0 &] \\ \underline{s}_1 & = & [\ 0, & \sqrt{\mathcal{E}_s}, & \ldots & 0 &] \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \\ \underline{s}_{M-1} & = & [\ 0, & 0, & \ldots & \sqrt{\mathcal{E}_s} &] \end{array} \tag{7.2}$$

and these are *orthogonal vectors*; their inner product is 0. It turns out that the essential properties of the scheme depend on using orthogonal signals so that the decision variables are (conditionally) independent Gaussian random variables only one of which has a nonzero mean $\sqrt{\mathcal{E}_s}$. It does not really matter *which* orthogonal vectors are used. Now, if the signals are constrained to have nonnegative amplitudes, there is no other choice than the one shown above. However, when signals can be both positive and negative, then a very interesting choice of orthogonal signals is possible, and even makes orthogonal signaling schemes useable with peak power constraints.

7.2.2. Hadamard matrices and Hadamard coding

A Hadamard matrix $H = [h_{i,j}]$ is a $n \times n$ matrix in which $h_{i,j} \in \{+1, -1\}$ for all i and j, and the rows are mutually orthogonal vectors. Note that the orthogonality of the rows is equivalent to the assertion that $HH^T = nI$ where the superscript T denotes the transpose and I denotes the identity matrix. The 1×1 matrix $[1]$ is trivially a Hadamard matrix since the requirement that the rows be mutually orthogonal is satisfied vacuously. Obviously, if a $n \times n$ Hadamard matrix exists for $n > 1$, then n is necessarily an even integer. More strongly, if $n > 2$, then n must be a multiple of 4. The existence of a $n \times n$ Hadamard matrix has not been demonstrated for all integers n that are multiples of 4, but fortunately for us, it is easy to construct a $2^k \times 2^k$ Hadamard matrix (we call it H_k) for all positive integers k. Let H_1 denote the 2×2 matrix with rows $[1, 1]$ and $[1, -1]$ which is obviously a Hadamard matrix. The method of construction of a $2^k \times 2^k$ Hadamard matrix that is shown below is called Sylvester's method.

$$H_1 = \begin{bmatrix} +1, +1 \\ +1, -1 \end{bmatrix}, \quad H_k = \begin{bmatrix} H_{k-1} & H_{k-1} \\ H_{k-1} & -H_{k-1} \end{bmatrix} = H_1 \otimes H_{k-1}$$

where $\otimes$ denotes the *Kronecker product* of the matrices. We shall number the rows and columns of H_k from 0 to $2^k - 1$ instead of the more conventional 1 to 2^k. Also, note that H_k is a *symmetric* matrix: $h_{i,j} = h_{j,1}$, and more particularly that the columns of H_k are the same as the rows of H_k and thus are also mutually orthogonal vectors.

Let us now consider a communication system operating over a discrete-time Gaussian channel in which B bits are transmitted in 2^B channel uses via the signals

$$\underline{s}_{i,H} = \sqrt{\mathcal{E}_s/2^B}\left[h_{i,0}, h_{i,1}, \ldots, h_{i,2^B-1}\right], \quad 0 \le i \le 2^B - 1. \tag{7.3}$$

where $[h_{i,j}] = H_B$ is the $2^B \times 2^B$ Hadamard matrix resulting from Sylvester's construction. This is called *Hadamard coding* for transmitting B bits of data using 2^B channel uses, or more succinctly, the use of a $(2^B, B)$ Hadamard code. Clearly, the energy transmitted during each channel use is $\mathcal{E}_s/2^B$ for a total energy of $\mathcal{E}_s$ which is the same as in (7.2) but now a peak power constraint clearly applies even if $\mathcal{E}_s$ were to increase proportionally to 2^B. As before, the receiver observes $\mathbb{Y}_0$, $\mathbb{Y}_1$, $\ldots$, $\mathbb{Y}_{2^B-1}$, and if $\underline{s}_{i,H}$ has been transmitted, these are (conditionally) independent Gaussian random variables with $\mathbb{Y}_j \sim \mathcal{N}(\sqrt{\mathcal{E}_s/2^B}h_{i,j}, \sigma^2)$, $0 \le j \le 2^B - 1$. This is quite different from what we had previously with one decision variable having a large mean and the rest having mean 0. We can remedy this by taking linear combinations of the $\mathbb{Y}$'s and using these linear combinations as decision variables. In particular, suppose that

$$[\mathbb{Z}_0, \mathbb{Z}_1, \ldots, \mathbb{Z}_{2^B-1}] = 2^{-B/2}[\mathbb{Y}_0, \mathbb{Y}_1, \ldots, \mathbb{Y}_{2^B-1}]H_B. \tag{7.4}$$

In particular, we have that

$$\mathbb{Z}_k = 2^{-B/2}\sum_{j=0}^{2^B-1} h_{k,j}\mathbb{Y}_j, \quad 0 \le k \le 2^B - 1.$$

The $\mathbb{Z}$'s are linear combinations of Gaussian random variables and hence are Gaussian random variables. Furthermore,

$$\mathsf{E}[\mathbb{Z}_k] = 2^{-B/2}\sum_{j=0}^{2^B-1} h_{k,j}\mathsf{E}[\mathbb{Y}_j] = 2^{-B/2}\sum_{j=0}^{2^B-1} h_{k,j}h_{i,j}\sqrt{\mathcal{E}_s/2^B} = 2^{-B}\sqrt{\mathcal{E}_s}\sum_{j=0}^{2^B-1} h_{k,j}h_{i,j} = \begin{cases} \sqrt{\mathcal{E}_s}, & \text{if } k = i, \\ 0, & \text{if } k \neq i, \end{cases}$$

where we have used the fact that the k-th and i-th columns of H_B are orthogonal to each other if $k \neq i$. Similarly, since the $\mathbb{Y}$'s are (conditionally) independent random variables, we have that

$$\mathsf{var}[\mathbb{Z}_k] = \sum_{j=0}^{2^B-1}\left(2^{-B/2}h_{k,j}\right)^2 \mathsf{var}[\mathbb{Y}_j] = 2^{-B} \times 2^B\sigma^2 = \sigma^2.$$

Finally, $\mathsf{cov}(\mathbb{Z}_k, \mathbb{Z}_\ell) = 2^{-B}\mathsf{cov}\left(\sum_{j=0}^{2^B-1} h_{k,j}\mathbb{Y}_j, \sum_{m=0}^{2^B-1} h_{\ell,m}\mathbb{Y}_m\right) = 2^{-B}\sum_{j=0}^{2^B-1} h_{k,j}h_{\ell,j}\mathsf{cov}(\mathbb{Y}_j, \mathbb{Y}_j) = 0$ if $k \neq \ell$. Thus, the $\mathbb{Z}$ are uncorrelated, and being Gaussian are thus (conditionally) independent random variables. In summary, the $\mathbb{Z}$'s, which are linear combinations of the observed variables $\mathbb{Y}$,s, are (conditionally) independent Gaussian random variables with common variance σ^2, with $\mathsf{E}[\mathbb{Z}_i] = \sqrt{\mathcal{E}_s}$ and all others having mean 0. This is exactly what we had before.

If $\underline{u}$ is a vector of dimension n and H is a $n \times n$ Hadamard matrix, then $\underline{v} = \underline{u}\left(\frac{1}{\sqrt{n}}H\right)$ is called the *Hadamard transform* of $\underline{u}$. The *inverse Hadamard transform* of $\underline{v}$ is $\underline{v}\left(\frac{1}{\sqrt{n}}H^T\right) = \underline{u}\left(\frac{1}{\sqrt{n}}H\right)\left(\frac{1}{\sqrt{n}}H^T\right) = \underline{u}\left(\frac{1}{n}\right)HH^T = \underline{u}I = \underline{u}$. If the Hadamard matrix is symmetric (as the ones obtained by Sylvester's construction are), then the forward and inverse Hadamard transforms are the same. With this concept in mind, notice that the signal $\underline{s}_{i,H}$ in (7.3) is just the Hadamard transform of the signal $\underline{s}_i$ in (7.2) while the *random vector* of $\mathbb{Z}$'s in (7.4) is just the (inverse) Hadamard transform of the random vector of $\mathbb{Y}$'s. In other words, Hadamard coding can be thought of as applying a Hadamard transform to the ON-OFF signals before transmission, and then undoing the change by applying the inverse Hadamard transform to the observations before further processing.

7.2.3. Orthogonal Frequency Shift Keying

One signaling scheme does not quite fit our model of 2^B channel uses resulting in 2^B observations but nonetheless gives rise to 2^B observations only one of which has a large mean. Consider 2^B sinusoidal signals of the form $s_i(t) = \sqrt{2\mathcal{E}_s/T}\sin(2\pi(f_0 + i/T + \theta_i)$, $0 \leq t < T, 0 \leq i \leq 2^B - 1$ differing in frequency by multiples of T^{-1} Hz. These signals are orthogonal over the interval $[0, T)$. One of these signals is transmitted and conveys energy $\mathcal{E}_s$ over the T second interval. The receiver consists of 2^B filters matched to the 2^B possible signals, and one of these has output a Gaussian random variable with mean $\sqrt{\mathcal{E}_s}$ while all other filter outputs are just noise (because the transmitted signal does not produce any response in the other filters). Thus, the problem to be solved is identical to that in the previous two subsections: given 2^B independent Gaussian random variables, exactly one of which has a large mean and all others have mean 0, determine with high reliability which is the random variable with large mean. The next Lecture will discuss the answer to this problem.