ECE 361: Digital Communications

Lecture 15: Transmitter-Centric ISI Management: Precoding

Introduction

The ISI channel with L taps,

$$y[m] = \sum_{\ell=0}^{L-1} h_\ell x[m-\ell] + w[m] \tag{1}$$

can alternatively be represented as

$$y[m] = h_0 x[m] + I[m] + w[m], \tag{2}$$

where

$$I[m] = \sum_{\ell=1}^{L-1} h_\ell x[m-\ell] \tag{3}$$

is the ISI. So far, we have looked at how the receiver can work to overcome the impact of ISI; all this while, the transmitter was continuing with sequential communication. In this lecture, we will take a different line of attack towards mitigating ISI: we do this by enlisting the transmitter as well in our efforts. The key point is that the ISI caused at time m, denoted by $I[m]$, is known *noncausally* at the transmitter:

- the channel tap coefficients $h_1, \ldots, h_{L-1}$ are constant and known to the transmitter ahead of time;
- the previously transmitted symbols $x[m-1], \ldots, x[m-L+1]$ are surely known to the transmitter.

Thus, the transmitter is in a position to *adapt* its present transmitted symbol $x[m]$ as a *function* of what ISI is being caused. Such an adaption is known in the literature as *precoding* and in this lecture we will investigate the most basic form of precoding.

A Simple Precoding Strategy

Since the transmitter *knows* what the ISI is going to be, the simplest idea is to just cancel it off: if we wanted to send a voltage $d[m]$ on an AWGN channel (with no ISI) we could transmit

$$x[m] = d[m] - \frac{I[m]}{h_0}. \tag{4}$$

This way the received voltage is simply

$$y[m] = h_0 d[m] + w[m], \tag{5}$$

a regular AWGN channel with no-ISI! This is an easy way to get rid of ISI, particularly when compared to the travails of the receiver-centric schemes we have seen in the last couple of lectures. But there is a catch, and that is we need more energy at the transmitter to cancel ISI than that present in $d[m]$.

For concreteness, let us suppose sequential communication is used to generate the voltages $d[m]$. Denoting the average energy used in the data symbol $d[m]$ by E, the average transmit energy used by the simple precoding scheme (cf. Equation (4)) at time sample m is

$$\mathbb{E}\left[(x[m])^2\right] = E + \frac{\mathbb{E}\left[(I[m])^2\right]}{h_0^2}, \tag{6}$$

where we supposed, as usual, that the mean of the voltage $d[m]$ is zero. Now, the average energy in the ISI at time m is (from Equation (3))

$$\mathbb{E}\left[(I[m])^2\right] = \sum_{\ell=1}^{L-1} h_\ell^2 \mathbb{E}\left[(x[m-\ell])^2\right] \tag{7}$$

where we used the fact, again, that the mean of the voltage $d[m]$ is zero. Substituting from Equation (6), we now get

$$\mathbb{E}\left[(I[m])^2\right] = \sum_{\ell=1}^{L-1} h_\ell^2 \left(E + \frac{\mathbb{E}\left[(I[m-\ell])^2\right]}{h_0^2}\right). \tag{8}$$

To get a quick feel for the solution to this *recursive* definition, suppose that $L = 2$. Then

$$\begin{aligned} \mathbb{E}\left[(I[m])^2\right] &= Eh_1^2 + \frac{h_1^2}{h_0^2}\mathbb{E}\left[(I[m-1])^2\right] && (9) \\ &= \sum_{k=1}^{m-1} \frac{h_1^{2k}}{h_0^{2k-2}} E, && (10) \end{aligned}$$

since the ISI at the very first time sample, $I[1]$ is zero. Substituting this back into Equation (6) we arrive at

$$\mathbb{E}\left[(x[m])^2\right] = E + \sum_{k=1}^{m-1} \frac{h_1^{2k}}{h_0^{2k}} E. \tag{11}$$

Now we see that if $h_1^2 \geq h_0^2$ then the average transmit energy of the simple precoding scheme grows to infinity over time. Otherwise, the average transmit energy grows to

$$E + \frac{E\frac{h_1^2}{h_0^2}}{1 - \frac{h_1^2}{h_0^2}}, \tag{12}$$

strictly larger than the energy E used by the sequential precoding scheme. This calculation is for the special case of $L = 2$; a homework exercise will explore the solution for general number of taps L.

The main conclusion is that the average transmit energy required to execute the simple precoding scheme grows with time; how much it grows to (a finite amount or even an infinite value) depends on the channel encountered and thus cannot be predicted beforehand. This is quite a problem for the communication engineer since transmit power requirements are decided ahead of time while the filter tap coefficients change based on the wireline channel encountered. It would be good if we could eliminate the ISI for *all* wireline channels encountered using a *constant* increase in the transmit power. The key idea to achieve this goal occurred independently to two graduate students (working towards their PhD in two different continents) in the early 1970s: Tomlinson and Harashima. We will see their elegant improvement to the simple precoding scheme next.

Tomlinson-Harashima (TH) Precoding

To simplify the exposition, consider sequential communication of one bit at a time used to generate the voltages $d[m]$. Specifically, the voltage $d[m]$ is $\pm\sqrt{E}$ depending on the m^{th} information bit. The basic constellation diagram associated with this communication is illustrated in Figure 1. The TH precoding idea is based on *extending* this constellation diagram, on both sides of the voltage axis; this is illustrated in Figure 2. The idea is that each of the crosses represents the same information bit (say, '1') and each of the triangles represents the same information bit (say, '0').

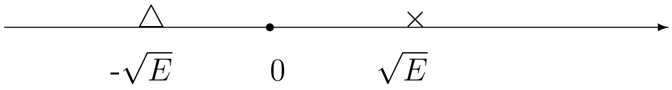

Figure 1: The constellation diagram associated with sequential communication of one bit at a time.

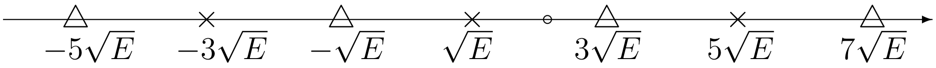

Figure 2: Extended constellation diagram.

To describe the TH precoding scheme, it will help the exposition to suppose that h_0, the first tap coefficient, is 1. There is no loss of generality in this supposition since the receiver can normalize its received voltage $y[m]$ by h_0. Now suppose we want to communicate the information bit '1' at time m. In the simple precoding scheme we would transmit $x[m]$ such that when added to $I[m]$, the resulting sum is the cross situated at $\sqrt{E}$. Since there is only a single cross, this would entail a large transmit energy whenever the ISI $I[m]$ is far away from $\sqrt{E}$. This operation is illustrated in Figure 3.

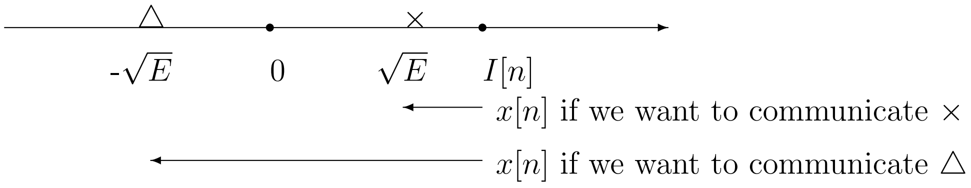

Figure 3: Simple precoding.

On the other hand, with the extended constellation diagram we could transmit $x[m]$ such that when added to the ISI $I[m]$, the resulting sum is *any one* of the many crosses available; specifically we could always pick that cross which is closest to $I[m]$ resulting in bounded transmit energy. This operation is illustrated in Figure 4; we see that the transmit energy is no more than $4E$ for all values of $I[m]$.

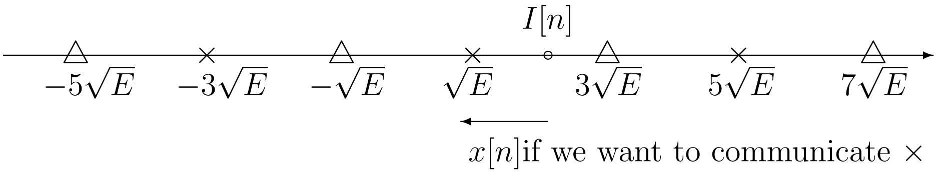

Figure 4: Tomlinson-Harashima precoding

The received voltage at time m is

$$y[m] = d[m] + w[m] \tag{13}$$

where $d[m]$ is one of the crosses in Figure 2. Since the additive noise $w[m]$ is Gaussian, it makes sense to follow the nearest neighbor rule. We decode the information bit transmitted depending on whether a cross or a triangle is closer to the received voltage level. The probability error is

$$\mathbb{P}[\mathcal{E}] \leq 2Q\left(\frac{\sqrt{E}}{\sigma}\right). \tag{14}$$

This is not an exact calculation since there is some chance that the additive noise $w[m]$ is larger than the spacing between nearest points in the extended constellation diagram of Figure 2, but the received voltage is still closer to a cross (if one of the crosses was transmitted).

We have focused on eliminating ISI associated with a sequential communication scheme. However, we have seen in earlier lectures that block communication schemes provide much better performances over the AWGN channel. The extension of the TH precoding scheme is natural; it takes place in higher dimensional spaces since the constellation diagrams themselves are described now in higher dimensional spaces.

Looking Ahead

The view espoused by the precoding technique here is the same one we had in our discussion of the zero forcing equalizer: interference is a nuisance and we work to get rid of it. But we developed a more nuanced and balanced view when we discussed the MMSE equalizer. A natural question is to ask for the transmitter-centric version of that balanced approach. We will see (in the next lecture) an important transmitter-centric technique designed from this perspective: it is called OFDM (orthogonal frequency division modulation) and is the basic technology behind many communication standards around us (examples: DSL – digital subscriber line – and Wi-Fi).