ECE 361: Digital Communication

Lecture 12: Intersymbol Interference Management: Low SNR Regime

Introduction

So far, we have seen that the wireline channel is modeled as an FIR (finite impulse response) filter acting on the transmit voltage sequence. The main difference between the AWGN channel and wireline channel is the presence of *inter-symbol interference* (ISI), i.e., transmit voltages of the previous symbols also mix along with the additive Gaussian noise with the voltage of the current symbol. The main challenge in wireline channel is to handle noise and ISI simultaneously in the quest to achieve rate efficient reliable communication.

Our approach to wireline channel will be to "simply" process the transmit and receive voltage sequence (at the transmitter and receiver, respectively) to mitigate ISI and harness the block codes developed for rate efficient reliable communication on the AWGN channel. In the next few lectures we study the *receiver centric* methods to deal with ISI. The transmitter will be more or less the same as that in the case of AWGN channel. In this lecture, we focus on the low SNR regime: in this scenario, the noise power dominates the total signal power. Thus noise dominates the ISI.

A First Approach

The received voltage is

$$\begin{aligned} y[m] &= \sum_{l=0}^{L-1} h_l x[m-l] + w[m] && (1) \\ &= h_0 x[m] + \sum_{l=1}^{L-1} h_l x[m-l] + w[m] && (2) \\ &= S[m] + I[m] + N[m] && (3) \end{aligned}$$

where

$$S[m] \stackrel{\text{def}}{=} h_0 x[m] \quad (4)$$

is the "signal";

$$I[m] \stackrel{\text{def}}{=} \sum_{l=1}^{L-1} h_l x[m-l] \quad (5)$$

is interference and

$$N[m] \stackrel{\text{def}}{=} w[m] \quad (6)$$

is noise. The implication of the low SNR regime is that

$$\mathbb{E}\left[(I[m])^2\right] \ll \mathbb{E}\left[(N[m])^2\right]. \quad (7)$$

In this regime, the natural approach is to ignore the interference completely and treat it as noise. In other words, the receiver could just do the nearest neighbor detection (even though it may not be optimal since the discrete interference definitely does not have Gaussian statistics). A natural question is:

> What reliable rate of communication is feasible with this approach?

In an AWGN channel, SNR is the only parameter of interest: the capacity is a direct function of the SNR. If we are to use that intuition here, we could look at the SINR (signal to interference plus noise ratio) defined as:

$$\begin{aligned}\mathsf{SINR} &\stackrel{\text{def}}{=} \frac{\mathbb{E}\left[(S[m])^2\right]}{\mathbb{E}\left[(I[m])^2\right] + \mathbb{E}\left[(N[m])^2\right]} && (8)\\ &= \frac{h_0^2 \mathbb{E}\left[x[m]^2\right]}{\mathbb{E}\left[\left(\sum_{l=1}^{L-1} h_l x[m-l]\right)^2\right] + \sigma^2}. && (9)\end{aligned}$$

Here σ^2 is equal to $\mathbb{E}(w[m]^2)$, the variance of the discrete-time noise. In deriving this expression, we implicitly used the statistical independence between transmitted voltages and the noise. Denoting

$$\mathbb{E}(x[m]^2) = E \tag{10}$$

to be the signal power (the student is encouraged to think of binary modulation for concreteness), the SINR can be written as

$$\begin{aligned}\mathsf{SINR} &= \frac{h_0^2 E}{\left(\sum_{l=1}^{L-1} h_l^2\right) E + \sigma^2} && (11)\\ &= \frac{h_0^2 \mathsf{SNR}}{1 + \left(\sum_{l=1}^{L-1} h_l^2\right) \mathsf{SNR}}. && (12)\end{aligned}$$

If we continue the analogy with the AWGN channel, we expect the largest rate of communication when we treat the interference as noise to be

$$\frac{1}{2}\log_2(1 + \mathsf{SINR}) \quad \text{bits/channel use.} \tag{13}$$

How do we use our previous knowledge to conclude this result? Our previous result on the capacity (cf. Lecture 8a) depended on two basic assumptions:

- noise is independent of the signal.
- statistics of the noise is Gaussian.

Both these issues are in contention in this scenario when we have considered interference as part of the overall noise:

1. the overall noise (interference plus noise) is *dependent* on the signal; indeed, the interference is just previously transmitted signals!
2. the statistics of the overall noise is not Gaussian.

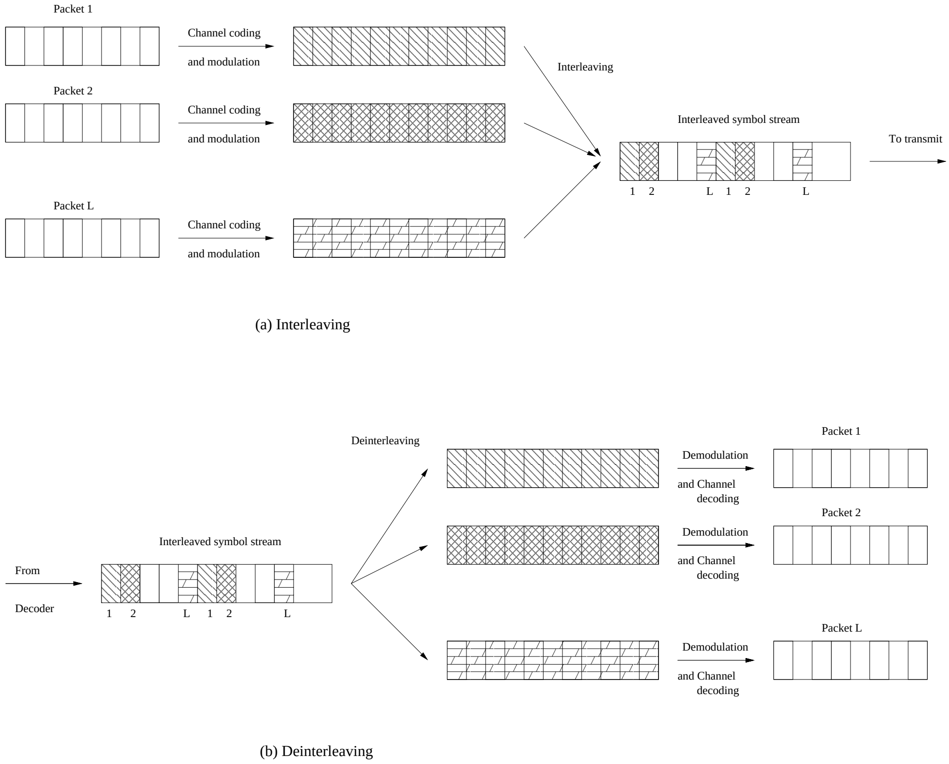

Figure 1: Interleaving and deinterleaving.

We can overcome the first issue by the following simple but *important* strategy. The overall information packet is broken up into sub-packets (that are all statistically independent of each other). The sub-packets are coded and modulated. The transmit voltages corresponding to each sub-packet are then *interleaved* at the transmitter. The receiver *deinterleaves* the received voltages. This operation is illustrated in Figure 1 where the number of sub-packets is exactly chosen to be the length of the ISI.

Now the consecutive transmitted symbols over a block of length L are statistically independent since they are generated based on statistically independent sub-packets. So, from the perspective of each of the transmit symbols, the interference is statistically independent of it.

However, interference is still *not* Gaussian. But it turns out that the Gaussian statistics is the "worst case" scenario for noise, i.e., by treating interference as noise with Gaussian statistics, the error probability of the nearest neighbor receiver (though not the optimal ML receiver anymore) for good AWGN channel codes remains the same. Look at it another way, and we can conclude that the nearest neighbour detector is quite robust. As a rule, we derive every statement we make from first principles in these lectures. However, this fact is a bit too deep for us to delve into right now (and will lead us astray from our main purpose: reliable communication over wireline channels). So, we provide the appropriate reference to this result which the interested student can read and mull over.

A. Lapidoth, "Nearest neighbor decoding for additive non-Gaussian noise channels," *IEEE Transactions on Information Theory*, Vol. 42(5), pp. 1520-1529, 1996.

With the two roadblocks resolved, we have now justified the expression in Equation (13) as a rate of reliable communication that is possible to achieve over the wireline channel.[1] Since we are operating at low SNR, we can approximate the rate of reliable communication by this first order approach (cf. Equation (12)) by

$$C \approx \frac{1}{2} (\log_2 e)\, h_0^2 \mathsf{SNR}. \tag{14}$$

Matched Filter : A Refined Approach

Can we reliably communicate at better rates even while continuing to treat interference as noise? The key to answer this question is to recall that we are power limited at low SNRs (cf. Lecture 8a). So any boost in the received SNR will translate directly and linearly into a boost in the rate of reliable communication. How do we boost the received SNR of any symbol? A careful look at the received voltage (cf. Equation 1) shows that we are receiving L different copies of *each* transmitted symbol. These copies are scaled differently and shifted in time (example: symbol $x[1]$ will appear in $y[1], \cdots, y[L]$). The scenario is similar to repetition coding except that now each copy of the transmitted symbol is scaled differently and sees different interference. However, since we are in low SNR regime, noise dominates the interference and we are justified in ignoring the interference. Consider the following "approximate" channel where we have just removed the interference altogether.

$$y[m+l] \quad \approx \quad h_l x[m] + w[m+l] \quad l = 0, 1, \cdots, L-1. \tag{15}$$

A natural desire is to combine these received voltages so as to maximize the SINR for $x[m]$. Since the noise dominates in the low SNR regime, maximizing SINR is same as maximizing SNR. To be concrete and explicit, let us restrict our strategies to the *linear combinations* of the received voltages:

$$\hat{x}[m] \stackrel{\text{def}}{=} \sum_{l=0}^{L-1} c_l y[m+l] \tag{16}$$

$$= \sum_{l=0}^{L-1} c_l h_l x[m] + \sum_{l=0}^{L-1} c_l w[m+l]. \tag{17}$$

The idea is to choose the vector $\mathbf{c}$, i.e., the coefficients $c_0, \ldots, c_{L-1}$ that maximizes the resulting SNR of the estimate:

$$\mathsf{SNR}(\mathbf{c}) \stackrel{\text{def}}{=} \frac{\left(\sum_{l=0}^{L-1} c_l h_l\right)^2 \mathbb{E}\left[x[m]\right]^2}{\left(\sum_{l=0}^{L-1} c_l^2\right) \sigma^2} \tag{18}$$

$$= \frac{\left(\mathbf{c}^T \mathbf{h}\right)^2}{\mathbf{c}^T \mathbf{c}} \mathsf{SNR}. \tag{19}$$

[1]Strictly speaking, the first L symbols of each of the L sub-packets do not see the same interference as rest of the symbols. However, the fraction of these L symbols to the total data is negligible.

Here we have written $\mathbf{h}$ to be the vector $(h_0, \ldots, h_{L-1})^T$ of channel tap coefficients.

Key Optimization Problem

What is the solution to the optimization problem

$$\max_{\mathbf{c}} \quad \frac{\left(\mathbf{c}^T\mathbf{h}\right)^2}{\mathbf{c}^T\mathbf{c}}? \tag{20}$$

To get a feel for this problem, first observe that the *magnitude* of the vector of combining coefficients $\mathbf{c}$ has no role to play: the function being maximized in the problem in Equation (20) is the *same* for all arguments $a\mathbf{c}$, as long as the scaling coefficient $a \neq 0$. So, without loss of generality let us suppose that

$$\mathbf{c}^T\mathbf{c} = 1. \tag{21}$$

Now we have a slightly refined picture of the optimization problem in Equation (20). We aim to maximize the magnitude of the inner product between a fixed vector $\mathbf{h}$ and a unit magnitude vector $\mathbf{c}$. Figure 2 depicts the situation pictorially. The solution is fairly obvious there: we should align the unit magnitude vector $\mathbf{c}$ in the *direction* of $\mathbf{h}$. In other words, we should choose

$$\mathbf{c} = \frac{\mathbf{h}}{||\mathbf{h}||}. \tag{22}$$

Here we have denoted the Euclidean magnitude $\sqrt{\mathbf{h}^T\mathbf{h}}$ of the vector $\mathbf{h}$ by $||\mathbf{h}||$. To use more words to say the same thing: in two dimensions, $\mathbf{c}$ lies on the unit circle. If $\mathbf{c}$ makes an angle θ with $\mathbf{h}$, then

$$|\mathbf{h}^T\mathbf{c}| = ||\mathbf{h}|| |cos\theta|. \tag{23}$$

This is maximized when $\theta = 0$.

This argument holds true for $n > 2$ as well, since we can project each of the vectors $\mathbf{h}$ and $\mathbf{c}$ to the same two dimensional plane without altering the inner product between them. A more verbose, but mathematically complete, proof of this claim is derived in the appendix. To conclude, the SNR in Equation (19) is optimized by choosing the coefficient as in Equation (22). Since the combining coefficients are *matched* to the channel coefficients, this operation is also called a *matched filter*.

Now let us revert back to the original wireline channel of Equation 1. This time we will account for the interference as well. Denoting $\hat{y}_{MF}[m]$ as the output of the matched filter at

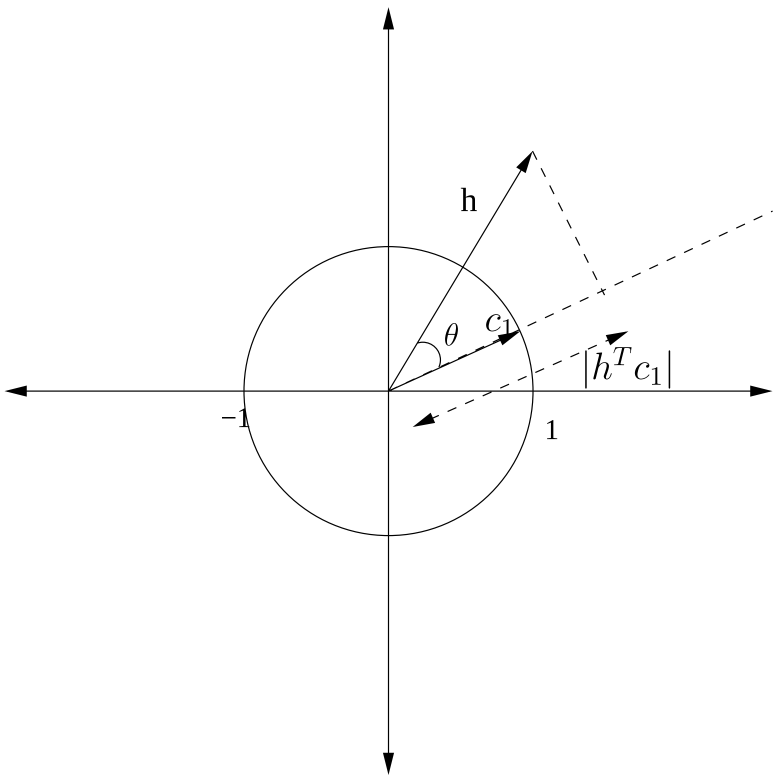

Figure 2: Pictorial explanation of Cauchy Shwartz inequality.

time index m, we have

$$\begin{aligned}
\hat{y}_{\mathrm{MF}}[m] &\stackrel{\text{def}}{=} \sum_{l=0}^{L-1} h_l y[m+l] && (24)\\
&= \sum_{l=0}^{L-1} h_l \left(\sum_{k=0}^{L-1} h_k x[m+l-k] + w[m+l] \right) && (25)\\
&= \sum_{l=0}^{L-1} \sum_{k=0}^{L-1} h_l h_k x[m+l-k] + \sum_{l=0}^{L-1} h_l w[m+l] && (26)\\
&= \left(\sum_{l=0}^{L-1} h_l^2 \right) x[m] + \sum_{l=0}^{L-1} \sum_{k=0;k\neq l}^{L-1} h_l h_k x[m+l-k] + \sum_{l=0}^{L-1} h_l w[m+l]. && (27)
\end{aligned}$$

The first term in the last equation is the desired signal, second term is the interference and the third term is the noise. The corresponding SINR is

$$\mathsf{SINR}_{\mathsf{MF}} = \frac{\left(\sum_{l=0}^{L-1} h_l^2 \right)^2 \mathbb{E}(x[m]^2)}{\sum_{l=0}^{L-1} \sum_{k=0;k\neq l}^{L-1} h_l^2 h_k^2 \mathbb{E}(x[m+l-k]^2) + \left(\sum_{l=0}^{L-1} h_l^2 \right) \sigma^2}. \tag{28}$$

Note that here we have used the facts that noise samples $w[m+l]$ are i.i.d. with variance σ^2 and that the interference terms are statistically independent and are also independent of noise.

We can further simplify the SINR expression.

$$\begin{aligned}
\mathsf{SINR}_{\mathsf{MF}} &= \frac{\left(\sum_{l=0}^{L-1} h_l^2 \right)^2 E}{\left(\sum_{l=0}^{L-1} \sum_{k=0;k\neq l}^{L-1} h_l^2 h_k^2 E \right) + \left(\sum_{l=0}^{L-1} h_l^2 \right) \sigma^2} && (29)\\
&= \frac{\left(\sum_{l=0}^{L-1} h_l^2 \right) \mathsf{SNR}}{\frac{\sum_{l=0}^{L-1} \sum_{k=0;k\neq l}^{L-1} h_l^2 h_k^2}{\sum_{l=0}^{L-1} h_l^2} \mathsf{SNR} + 1}. && (30)
\end{aligned}$$

The rate of reliable communication with the matched filter is now

$$\frac{1}{2} \log_2 \left(1 + \mathsf{SINR}_{\mathsf{MF}} \right) \quad \text{bits/channel use.} \tag{31}$$

At low SNRs we can approximate this rate by

$$\frac{1}{2} \left(\log_2 e \right) \mathsf{SINR}_{\mathsf{MF}} \tag{32}$$

that is further approximated by

$$\frac{1}{2} \left(\log_2 e \right) \left(\sum_{\ell=0}^{L-1} h_\ell^2 \right) \mathsf{SNR}. \tag{33}$$

Comparing this with the naive approach first espoused (cf. Equation (14)), we see that the rate has strictly improved: the larger the channel coefficients at later taps, the more the improvement.

Looking ahead

In the low SNR regime, interference is not much of an issue. The focus is hence on collecting as much of the signal energy as possible at the receiver; the matched filter is a natural outcome of this line of thought. In the next lecture, we look at an alternative regime: high SNR. This is the dominant mode of operation over wireline channels. Here treating interference as noise is disastrous. We will need different strategies.

Appendix: Cauchy Shwartz Inequality

We show that the solution to the optimization problem in Equation (20) is the one in Equation (22). We do this by deriving a fundamental inequality known as Cauchy-Shwartz inequality. We will now prove it formally. Fix a and define

$$\mathbf{x} \stackrel{\text{def}}{=} \mathbf{h} - a\mathbf{c}. \tag{34}$$

Now we have

$$\begin{aligned} 0 &\leq \mathbf{x}^T\mathbf{x} && (35) \\ &= (\mathbf{h} - a\mathbf{c})^T(\mathbf{h} - a\mathbf{c}) && (36) \\ &= \mathbf{h}^T\mathbf{h} - 2a\mathbf{h}^T\mathbf{c} + a^2\mathbf{c}^T\mathbf{c}. && (37) \end{aligned}$$

Since this inequality holds for all a, let us choose $a = \frac{\mathbf{h}^T\mathbf{c}}{\mathbf{c}^T\mathbf{c}}$. Substituting this choice, we arrive at

$$0 \leq \mathbf{h}^T\mathbf{h} - 2\frac{|\mathbf{h}^T\mathbf{c}|^2}{\mathbf{c}^T\mathbf{c}} + \frac{|\mathbf{h}^T\mathbf{c}|^2}{\mathbf{c}^T\mathbf{c}}. \tag{38}$$

(39)

Hence

$$|\mathbf{h}^T\mathbf{c}|^2 \leq \mathbf{h}^T\mathbf{h}\mathbf{c}^T\mathbf{c} \tag{40}$$

and equality holds when $\mathbf{c} = a\mathbf{h}$.