Last Time

- Channel coding for continuous channels
- Additive Gaussian white noise (AGWN)

Capacity of AGWN, continued

Recall that for a AGWN channel with $E[X^2] \leq S$,

$$C(S) = \frac{1}{2} \ln(1 + \frac{S}{\sigma^2})$$

Now it turns out that this channel is the worst among all channels with $E[N_j] = 0$ and $E[N_j^2] = \sigma^2$.

5.6.29 Theorem. If $\gamma$ is an additive memoryless channel with noise $\{N_j\}_{j=1}^\infty$ of zero mean and variance $\sigma^2$ then,

$$\frac{1}{2} \ln(1 + \frac{S}{\sigma^2}) \leq C(S)$$

Proof. Let $\tilde{\gamma}$ be a AWGN channel with $\tilde{N}_j$ having mean zero and variance $\sigma^2$. Also, consider a Gaussian sequence $\{\tilde{X}_j\}_{j=1}^\infty$ with $E[\tilde{X}_j^2] = S$. Now let us compare the mutual information between the Gaussian and the (possibly) non-Gaussian Channel.

$$I(\tilde{X}; \gamma(\tilde{X})) - I(\tilde{X}; \tilde{\gamma}(\tilde{X}))$$

$$= \int \int p_{\tilde{X}}(x)p_{\tilde{N}}(y-x) \ln(\frac{p_{\tilde{N}}(y-x)}{p_{\tilde{Y}}(y)}) dydx - \int \int p_{\tilde{X}}(x)p_{\tilde{N}}(y-x) \ln(\frac{p_{\tilde{N}}(y-x)}{p_{\tilde{Y}}(y)}) dydx$$

By equality of second moment we can replace $\tilde{N}$ with $N$ and get,

$$= \int \int p_{\tilde{X}}(x)p_{N}(y-x) \ln(\frac{p_{N}(y-x)}{p_{Y}(y)}) dydx - \int \int p_{\tilde{X}}(x)p_{N}(y-x) \ln(\frac{p_{N}(y-x)}{p_{Y}(y)}) dydx$$
\[
\int \int p_X(x)p_N(y-x) \ln \left( \frac{p_N(y-x)p_X(y)}{p_Y(y)p_N(y-x)} \right) dy dx \geq \int \int p_X(x)p_N(y-x)(1 - \frac{p_Y(y)p_N(y-x)}{p_N(y-x)p_Y(y)}) dy dx
\]

\[
= 1 - \int \int \frac{p_X(x)p_N(y-x)p_Y(y)}{p_Y(y)} dy dx = 1 - \int p_Y(y) \left( \int p_X(x)p_N(y-x) dx \right) dy
\]

\[
= 1 - \int p_Y(y) \frac{p_N(y-x)}{p_Y(y)} [p_{X+N=Y}(y)] dy = 1 - \int p_Y(y) dy = 0
\]

and equality holds if and only if
\[
p_Y(y) = \frac{p_N(y-x)}{p_Y(y)}
\]

Now if we pick a \( y \) such that this equality holds, then the left hand side is fixed and the right hand side is a constant for almost every \( x \), so \( p_N(y-x) = c \cdot p_N(y-x) \) and normalization forces \( c = 1 \). This shows equality holds if and only if \( \{N_j\}_{j=1}^{\infty} \) is Gaussian. Next lets compare the Gaussian and the (possibly) non-Gaussian Capacity,

\[
\max_{E[X^2] \leq S} I(X; \tilde{\gamma}(\tilde{X})) = I(\tilde{X}; \tilde{\gamma}(\tilde{X})) \leq I(\tilde{X}; \gamma(\tilde{X})) \leq \max_{E[X^2] \leq S} I(X; \gamma(X)) = C(S)
\]

\[\square\]

### 5.7 Partially Noisy Channel

Suppose you have \( Y_j = X_j + N_j \) where

\[
N_j = \begin{cases} 
\tilde{N}_j & \text{with probability 0.1} \\
0 & \text{with probability 0.9}
\end{cases}
\]

where \( \tilde{N}_j \) are Gaussian with mean zero and variance \( \sigma^2 \).

What is the Capacity of this Channel?

We have that

\[
p_N(x) = 0.1 \delta(x) + \frac{0.9}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}x^2}
\]

hence,

\[
h(N) = \lim_{\Delta \to 0} (H(N^\Delta) + \ln(\Delta)) = -\infty
\]

Now if we choose \( X \) so that \( h(Y) \) is finite we get,

\[
C(S) \geq h(Y) - h(N) = \infty
\]

How is this possible?

Pick \( X \) uniformly distributed between \(-\sqrt{S}\) and \(\sqrt{S}\) and transmit rational numbers repeatedly.

If the receiver gets a rational number as an output, the \( \tilde{N}_j = 0 \) almost surely. As a result, using the channel \( n \) times gives decoding error probability of \( (0.9)^n \xrightarrow{n \to \infty} 0 \).
5.8 Capacity for parallel AGWN channels

5.8.30 Theorem. Suppose we have \( k \) Channels with Gaussian white noisy variables having variance \( \sigma_1, \ldots, \sigma_k \) and an overall power constraint \( \sum_{i=1}^{k} E[X_i^2] \leq S \), then

\[
C(S) = \sum_{i=1}^{k} \frac{1}{2} \ln(1 + \frac{S_i}{\sigma_i^2})
\]

where \( S_i = \max\{0, \theta - \sigma_i^2\} \) and \( \theta \) is chosen such that \( \sum_{i=1}^{k} S_i = S \). The process of choosing this \( \theta \) is often known as “water filling algorithm”.

Proof. By definition,

\[
\max_{P_{X^\otimes k}} I(X; Y) \quad \text{subject to} \quad \sum_{i=1}^{k} E[X_i^2] \leq S
\]

Now since noise is independent of \( X \),

\[
I(X, Y) = h(Y) - h(Y|X) = h(Y) - h(X + N|X) = h(Y) - h(N|X)
\]

\[
= h(Y) - h(N) \leq \sum_{i=1}^{k} h(Y_i) - h(N_i) = \sum_{i=1}^{k} I(Y_i; N_i) \leq \sum_{i=1}^{k} \frac{1}{2} \ln(1 + \frac{S_i}{\sigma_i^2})
\]

Now if we maximize the right hand side subject to \( \sum_{i=1}^{k} S_i = S \) we get our desired result.

To this end, we note that the sum of the logarithms is concave in \( \{(S_i + \sigma_i^2)/\sigma_i^2\} \), thus averaging among \( S_i + \sigma_i^2 \) for indices with \( S_i > 0 \) increases the right hand side. For a given \( \theta \), then \( S_i + \sigma_i^2 = \theta \) when \( S_i > 0 \) achieves the maximum. Using the monotonicity of the logarithm, we can choose \( \theta \) so that \( \sum_{i=1}^{k} S_i = S \).