Electrical Engineering 229A Lecture 21 Notes

Daniel Raban

November 4, 2021

1 Shannon Capacity of the Parallel Gaussian Channel Model and Power-Constrained Waveform Channels with Colored Noise

1.1 Shannon capacity of the parallel Gaussian channel model

Last time, we began discussing the parallel Gaussian channel model. We are doing communication at times $i = 1, \dots, n$. At time i, we can send inputs $X_i^{(1)}, \dots, X_i^{(K)}$, and the the receiver receives $Y_i^{(1)}, \dots, Y_i^{(K)}$¡ where $Y_i^{(k)} = X_i^{(k)} + Z_i^{(k)}$, and the $Z_i^{(k)} \sim \text{iid}\, \mathcal{N}(0, \sigma_k^2)$. The power constraint is that for each message $m \in [M_n]$, the codeword

$$\begin{bmatrix} x_1^{(1)}(m) & \cdots & x_n^{(1)}(m) \\ \vdots & & \vdots \\ x_1^{(K)}(m) & \cdots & x_n^{(K)}(m). \end{bmatrix}$$

must satisfy

$$\sum_{i=1}^{n} \sum_{k=1}^{K} (x_i^{(k)})^2 \le nP.$$

Theorem 1.1. *The Shannon capacity is*

$$\sup_{\sum_{k=1}^K \mathbb{E}[(X^{(k)})^2] \le nP} I(X^{(1)}, \dots, X^{(K)}; Y^{(1)}, \dots, Y^{(K)}).$$

Proof. We can prove via the usual method of a random coding argument for achievability and Fano's inequality for the converse. □

Choosing the inputs to be independent Gaussians is best (to maximize the mutual information), say $X^{(k)} \sim \mathcal{N}(0, P_k)$ (we must have $\sum_{k=1}^K P_k \le P$). This leads to the problem

$$\max_{\sum_{k=1}^K P_k = P} \sum_{k=1}^{K} \frac{1}{2} \log\left(1 + \frac{P_k}{\sigma_k^2}\right).$$

Use the Lagrange multiplier technique: The Lagrangian is

$$\mathcal{L}(P_1, \ldots, P_k, \lambda) = \sum_{k=1}^{K} \frac{1}{2} \log \left(1 + \frac{P_k}{\sigma_K^2}\right) + \lambda \left(\sum_{k=1}^{K} P_k - P\right).$$

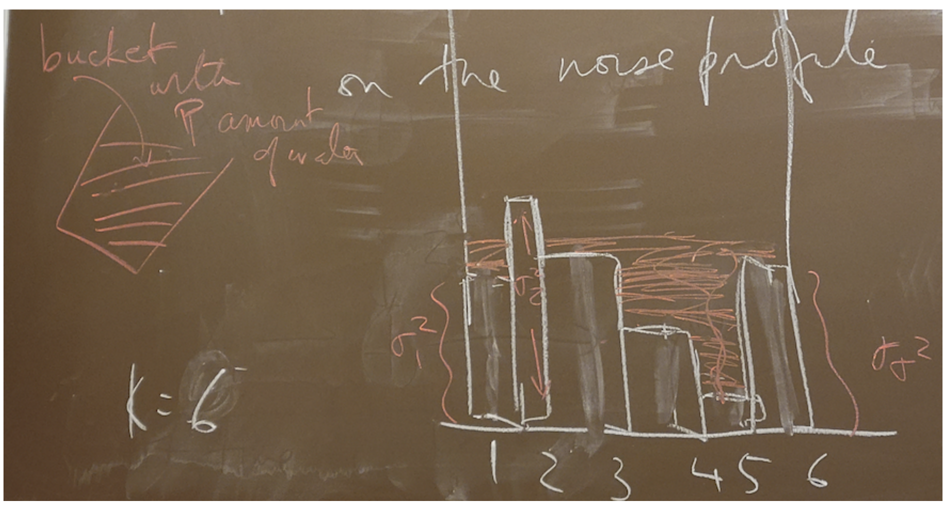
Then

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial P_k} &= (\log_2 e) \cdot \frac{1/\sigma_k^2}{2(1 + P_k/\sigma_k^2)} + \lambda \\
&= \frac{\log_2 e}{2} \cdot \frac{1}{\sigma_k^2 + P_k} + \lambda
\end{aligned}$$

We also need to bring in the non-negativity constraints. With these taken into account, the optimality criterion is that at the optimum, $\frac{\partial \mathcal{L}}{\partial P_k}$ must be ≤ 0 with strict inequality allowed only at $P_k^* = 0$. This leads to

$$\frac{\log_2 e}{2} \frac{1}{\sigma_k^2 + P_k^*} \leq -\lambda^*$$

for all k¡ with strict inequality only if $P_k^* = 0$. That is, $\sigma_k^* + P_k^* =$ constant, except possibly for k such that $P_k^* = 0$. This is waterfilling the available power $P = \sum_{k=1}^{L} P_k^*$ on the noise power. Imagine filling up the following bucket with water:



1.2 Power-constrained waveform channels with colored noise

What does this have to do with waveform channels in colored noise?

Definition 1.1. For a discrete time stationary process $(U_k, k \in \mathbb{Z})$, the **autocorrelation function** is

$$R_{U,U}(m, n) := \mathbb{E}[U_m U_n].$$

This is dependent only on $m-n$, and we may call it $R_{U,U}(m-n)$.

Definition 1.2. We call $(U_n, n \in \mathbb{Z})$ **wide sense stationary (WSS)** if $R_{U,U}(m,n)$ is dependent only on $m-n$ and if $\mathbb{E}[U_n]$ is constant.

Definition 1.3. The **power spectral density** of the process $(U_n, n \in \mathbb{Z})$ (assuming the sampling time is T) is

$$S_{U,U}(f) = \sum_{n=-\infty}^{\infty} R_{U,U}(n) e^{-i2\pi f n T},$$

which is periodic with period $2\pi/T$.

The coefficient of the autocorrelation function can be recovered as

$$\frac{1}{2W}\int_{-W}^{W} e^{-in\frac{\pi f}{W}} S_{U,U}(f)\, df. \quad \text{where} \quad W = \frac{\pi}{T}.$$

If the parallel Gaussian channel model is viewed as coming from quantizing the communication bandwidth into K levels with the noise power roughly flat over those levels, this leads to the capacity formula for power-constrained waveform channels with colored noise:

$$C = \int_{-W}^{W} \frac{1}{2}\log\left(1 + \frac{\max\{\nu - S_{U,U}(f), 0\}}{S_{U,U}(f)}\right) df,$$

where ν is chosen by waterfilling as the unique level with $\int \max\{\nu - S_{U,U}(f), 0\}\, df = P$.

Observe that if you consider the Toeplitz matrix

$$R_{U,U}^{(n)} = \begin{bmatrix} R_{U,U}(0) & R_{U,U}(1) & \cdots & R_{U,U}(n-1) \\ R_{U,U}(1) & \ddots & \ddots & \\ \vdots & & & R_{U,U}(1) \\ R_{U,U}(n-1) & \cdots & R_{U,U}(1) & R_{U,U}(0) \end{bmatrix},$$

then $w^\top R_{U,U}^{(n)} w = \mathbb{E}[(\sum_{\ell=0}^{n1} e_\ell U_\ell)^2]$. This matrix is positive semidefinite, so it has nonnegative, real eigenvalues $\tau_{n,1}, \ldots, \tau_{n,n}$.

Theorem 1.2 (Szegö)**.** *The fraction of these eigenvalues that lie in $(f_0, f_0+\varepsilon)$ for any f_0 converges to a limit in the sense that for any function $F : \mathbb{R}_+ \to \mathbb{R}$ that is continuous,*

$$\frac{1}{n}\sum_{k=1}^{n} F(\tau_{n,k}) \to \frac{T}{2\pi}\int_{-\pi/T}^{\pi/T} F(S(f))\, df.$$

Where does this theorem come from? Think about this in terms of eigenvalues:

$$R_{U,U}^{(n)} w_k^{(n)} = \tau_{n,k} w_k^{(n)}, \qquad w_k^{(n)} = \begin{bmatrix} w_{k,1}^{(n)} \\ \vdots \\ w_{k,n} \end{bmatrix}$$

normalized to make $\|w_k^{(n)}\|_2 = 1$. Associate to this

$$\psi_k^{(n)}(f) = \sum_{\ell=1}^{n} w_{k,\ell}^{(n)} e^{-i\frac{\ell\pi f}{W}},$$

which is a periodic function of period $2W$. Then

$$\int_{-W}^{W} |\psi_n^{(k)}(f)|^2 \, df = \|w_k^{(n)}\|_2^2 = 1,$$

and

$$\begin{aligned} \frac{1}{2W}\int_{-W}^{W} |\psi_k^{(n)}(f)|^2 S_W(f)\, df &= \frac{1}{2W}\int_{-W}^{W} \sum_{\ell=1}^{n}\sum_{j=1}^{n} w_{k,\ell}^{(n)} w_{k,j}^{(n)} e^{-i\frac{\pi f}{W}(j-\ell)} S_W(f)\, df \\ &= (w_k^{(n)})^\top R_{U,U}^{(n)} (w_k^{(n)}) \\ &= \tau_{n,k}. \end{aligned}$$