ECE253/CSE208 Introduction to Information Theory

Lecture 14: AWGN Channel

Dr. Yu Zhang

ECE Department
University of California, Santa Cruz

- Chap 9 of *Elements of Information Theory (2nd Edition) by Thomas Cover & Joy Thomas*.

Additive White Gaussian Noise (AWGN) Channels

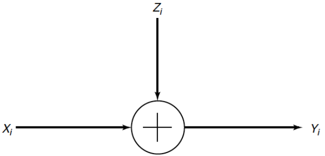


Figure: Gaussian channel: Infinite capacity if no noise or unbounded transmit power.

A discrete-time Gaussian (continuous-alphabet) channel is of the form:

$$Y_i = X_i + Z_i,$$

where at time i, the Gaussian noise $Z_i \sim \mathcal{N}(0, N)$ is assumed to be independent of the input signal X_i whose transmit power is bounded: $\mathrm{E}(X_i^2) \leq P$.

- **Additive:** noise is added to the system.
- **White:** a constant power spectral density across the frequency band[1].
- **Gaussian:** normal distribution of the noise amplitude. The Gaussian assumption is valid in a large number of situations due to the CLT.

This simple and tractable model is useful for gaining vital insight for many practical channels with fading, frequency selectivity, interference, nonlinearity or dispersion.

[1] `https://sound.softdb.com/what-is-white-noise/`

Shannon Capacity

Theorem (AWGN channel capacity)

The information capacity of a Gaussian channel with power constraint P and noise variance N is given as

$$\boxed{C = \max_{f(x):\mathrm{E}(X^2)\leq P} I(X;Y) = \frac{1}{2}\log\left(1+\frac{P}{N}\right)} \text{ bits per transmission.}$$

Proof: $\mathrm{E}(Y^2) = \mathrm{E}(X+Z)^2 = \mathrm{E}(X^2) + \mathrm{E}(Z^2) + 2\mathrm{E}(X)\mathrm{E}(Z) \leq P + N.$

$$\begin{aligned}
I(X;Y) = h(Y) - h(Y|X) &= h(Y) - h(Z) && \text{(1a)}\\
&\leq \frac{1}{2}\log[2\pi e\mathrm{Var}(Y)] - \frac{1}{2}\log(2\pi eN) && \text{(1b)}\\
&\leq \frac{1}{2}\log[2\pi e(P+N)] - \frac{1}{2}\log(2\pi eN) && \text{(1c)}\\
&= \frac{1}{2}\log(1+\frac{P}{N}), && \text{(1d)}
\end{aligned}$$

where the maximum is attained when $X \sim \mathcal{N}(0,P)$. Note that the capacity is infinite if we have zero-variance noise or unconstrained input.

Infinity Capacity

Operational-wise, how can we achieve infinity capacity if $P \to +\infty$ or $N \to 0$?

If transmit 1 bit info. Set $X(0) = -\sqrt{P}$ and $X(1) = +\sqrt{P}$. Decoding: $\hat{X} = \mathbb{1}_{\{Y>0\}}$.

Then,

$$\begin{aligned} P_e &= \frac{1}{2}\Pr(Y < 0 | X = +\sqrt{P}) + \frac{1}{2}\Pr(Y > 0 | X = -\sqrt{P}) \\ &= \Pr(Z > \sqrt{P}) = 1 - \Phi(\sqrt{P/N}). \end{aligned}$$

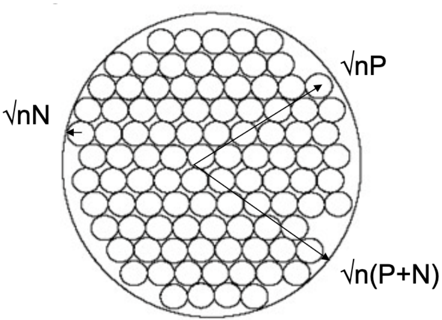
$P_e \to 0$ if $P \to +\infty$ or $N \to 0$. So a rate of 1 is achievable. We can increase the rate from 1 to 2 by encoding two info bits at a time:

$$X(b_1, b_2) = \begin{cases} \sqrt{P}, \text{ if } (b_1, b_2) = (0,0) \\ \frac{\sqrt{P}}{2}, \text{ if } (b_1, b_2) = (0,1) \\ -\sqrt{P}, \text{ if } (b_1, b_2) = (1,0) \\ -\frac{\sqrt{P}}{2}, \text{ if } (b_1, b_2) = (1,1) \end{cases}$$

Follow the same procedure, we have $P_e = \Pr\left(Z > \frac{\sqrt{P}}{2}\right) = 1 - \Phi(\sqrt{P/4N})$.

Hence, without noise or a bound on transmit power, we can squeeze as many info bits as we want that achieves an arbitrarily small P_e yielding $+\infty$ capacity.

Sphere Packing



By LLN, as $n \to \infty$, we have the following:

- A received vector $\mathbf{y}^n \sim \mathcal{N}(\mathbf{x}^n, N\mathbf{I})$ is contained in a **small** sphere of radius $r_{\text{small}} = \sqrt{nN}$ around the sent codeword (with high probability).
- Most received vectors $\{\mathbf{y}^n\}$ lie inside the n-dimensional **large** sphere of radius $r_{\text{large}} = \sqrt{n(P+N)}$.

Sphere Packing (cont'd)

- Decoding: assign everything in this sphere to the given sent codeword. An error happens only if $\mathbf{y}^n$ falls outside the sphere with low probability. Similarly, we can choose other codewords and their corresponding decoding spheres.
- Maximum number of non-overlapping spheres = Maximum number of codewords that can be reliably transmitted.

Q: How many non-overlapping spheres can be packed into the large sphere?
A: The n-dimensional volume of a sphere with radius r is $V_n(r) = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} r^n \implies$.

$$2^{nR} \leq \left(\frac{r_{\text{large}}}{r_{\text{small}}}\right)^n = \left(\frac{\sqrt{n(P+N)}}{\sqrt{nN}}\right)^n \implies R \leq \frac{1}{2}\log\left(1+\frac{P}{N}\right).$$

This sphere-packing argument indicates that we cannot hope to send at rates greater than C with a low probability of error. However, we can indeed do almost as well as this, as is proved next.

Proof of Achievability and Converse Statement

The proof is very similar to that of DMC. The key difference is changing from an arbitrary random coding to Gaussian coding while considering the power constraint[2].

- To form codewords $X^n(1), X^n(2), \ldots, X^n(2^{nR}) \in \mathbb{R}^n$, we draw $X_i(w)$ from i.i.d. $\mathcal{N}(0, P-\epsilon)$, for $i = 1, 2, \ldots, n$ and $w = 1, 2, \ldots, 2^{nR}$.
- A new type of error: Power constraint violation $E_0 = \left\{ \frac{1}{n} \sum_{i=1}^{n} X_i^2(1) > P \right\}$.
- For the average probability of error, we have

$$
\begin{aligned}
\Pr(\mathcal{E}|W=1) &\le \Pr\left(E_0|W=1\right) + \Pr\left(E_1^c|W=1\right) + \sum_{i=2}^{2^{nR}} \Pr\left(E_i|W=1\right) &\text{(2a)}\\
&\le \epsilon + \epsilon + (2^{nR}-1) \times 2^{-n(I(X;Y)-3\epsilon)} &\text{(2b)}\\
&\le 3\epsilon, &\text{(2c)}
\end{aligned}
$$

  if n is sufficiently large and $R < I(X;Y) - 3\epsilon$.

- For the converse, we use Fano's inequality again.

[2]See details on pages 266-270 of Cover's book.

Band-limited AWGN

A more realistic channel model is the band-limited continuous AWGN

$$Y(t) = (X(t) + Z(t)) * h(t),$$

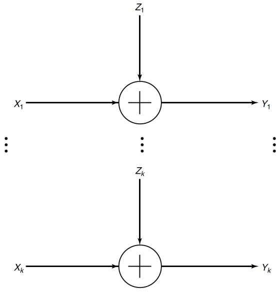
where $h(t)$ is the impulse response of an ideal bandpass filter with $f_{\max} = W$. Consider the channel being used over the interval $[0, T]$. We have

- $2W$ samples per second: this is sufficient for non-aliasing reconstruction reconstruction thanks to the *Nyquist-Shannon sampling theorem*.
- Signal energy per sample is $\frac{PT}{2WT}$.
- Noise variance per sample is $\frac{N_0 WT}{2WT}$.
  We assume double-sided noise power spectral density is $\frac{N_0}{2}$.
- Channel capacity per sample is $C = \frac{1}{2}\log(1 + \frac{P}{N_0 W})$.

Therefore, the capacity of a band-limited AWGN is given as

$$\boxed{C = W \log\left(1 + \frac{P}{N_0 W}\right) \xrightarrow{W \to \infty} \frac{P}{N_0} \log_2 e \text{ bits per second}}$$

Parallel AWGN



Consider a set of AWGN channels in parallel:

- channel output $Y_i = X_i + Z_i, \quad i = 1, \ldots, k.$
- independent noise $Z_i \sim \mathcal{N}(0, N_i)$.
- coupling power budget: $\mathrm{E}\left(\sum_i X_i^2\right) \leq P$.
- mutual information:

$$\begin{aligned} I(X^k; Y^k) = h(Y^k) - h(Y^k|X^k) = h(Y^k) - h(Z^k) \\ \leq \sum_i h(Y_i) - h(Z_i) \\ \leq \sum_i \frac{1}{2} \log\left(1 + \frac{P_i}{N_i}\right), \end{aligned}$$

where "=" is achieved by independent $X_i \sim \mathcal{N}(0, P_i)$.

Water-filling Optimal Resource Allocation

The following convex problem yields the channel capacity:

$$\max_{\{P_i \geq 0\}_{i=1}^{k}} \quad \sum_{i=1}^{k} \frac{1}{2} \log\left(1 + \frac{P_i}{N_i}\right) \tag{3a}$$

$$\text{subject to} \quad \sum_{i=1}^{k} P_i = P \tag{3b}$$

- Lagrangian relaxation: $L(\{P_i\}; \lambda) = \sum_i \frac{1}{2} \log\left(1 + \frac{P_i}{N_i}\right) + \lambda(\sum_i P_i - P)$
- Setting partial derivatives equal to zero: $\frac{1}{2\ln 2} \frac{1}{P_i + N_i} + \lambda = 0.$
- The optimal solution is obtained by the *water-filling* scheme:

$$P_i^* = (\nu - N_i)^+, \; i = 1, 2, \ldots k,$$

where $(\cdot)^+$ is the ReLU function (projection to the non-negative reals). The water-filling level $\nu = \frac{-1}{\lambda \ln 4}$ is obtained by solving the equation $\sum_i (\nu - N_i)^+ = P$.

Water-filling Algorithm (cont'd)

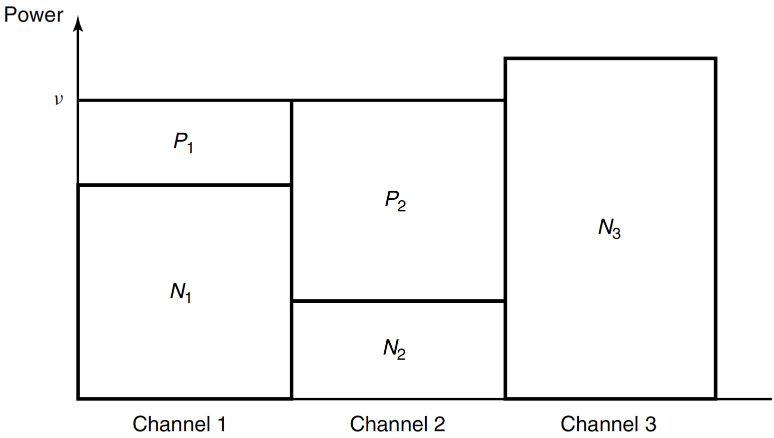


Figure: Water-filling solution: **allocate more power in less noisy channels**. Imaging each channel is a 'bathtub' with a height of N_i. In this example, we will have the solution $P_2 > P_1 > P_3 = 0$ and $P_1 + P_2 = P$. No power allocation and communication over the worst channel (#3).

Colored Gaussian Noise

Q: How about the case when the channel has Gaussian noise with memory?

A: We can consider a block of n consecutive uses of the channel as n channels in parallel with colored (or dependent) noise.

- Colored noise: noise values are correlated and its covariance matrix $\mathbf{K}_Z$ can be any arbitrary positive semi-definite matrix (but not diagonal)[3].
- Input power constraint: $\frac{1}{n}\sum_{i=1}^{n} \mathrm{E}X_i^2 \le P \Longleftrightarrow \mathrm{Tr}(\mathbf{K}_X) = nP$.
- Output signal covariance is $\mathbf{K}_Y = \mathbf{K}_X + \mathbf{K}_Z$ because $X \perp\!\!\!\perp Z$.
- Maximizing $I(X^n; Y^n) = h(Y^n) - h(Z^n)$ can be formulated as

$$\max_{\mathbf{K}_X : \mathrm{Tr}(\mathbf{K}_X) = nP} \frac{1}{2} \log \left((2\pi e)^n |\mathbf{K}_X + \mathbf{K}_Z| \right).$$

  That is, find $\mathbf{K}_X$ to maximize $|\mathbf{K}_X + \mathbf{K}_Z|$, subject to the trace constraint.

[3] https://en.wikipedia.org/wiki/White_noise

Colored Gaussian Noise (cont'd)

Eigenvalue decomposition: $\mathbf{K}_Z = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top$, where $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ collect all eigenvalues and $\mathbf{U}\mathbf{U}^\top = \mathbf{I}$.

$$\begin{aligned} |\mathbf{K}_X + \mathbf{K}_Z| &= |\mathbf{U}(\underbrace{\mathbf{U}^\top \mathbf{K}_X \mathbf{U}}_{\hat{\mathbf{K}}_X} + \mathbf{\Lambda})\mathbf{U}^\top| \\ &= |\hat{\mathbf{K}}_X + \mathbf{\Lambda}| \\ &\le \prod_i [(\hat{\mathbf{K}}_X)_{ii} + \lambda_i], \end{aligned}$$

with equality iff $\hat{\mathbf{K}}_X$ is diagonal due to Hadamard's inequality.

- Recall the constraints: $\mathrm{Tr}(\hat{\mathbf{K}}_X) \le nP$ and $(\hat{\mathbf{K}}_X)_{ii} \ge 0$. Then, the optimal solution is obtained by *water-filling in the eigen-space* of $\mathbf{K}_Z$:

$$\mathbf{K}_X^{\mathrm{opt}} = \mathbf{U}\mathrm{diag}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_n)\mathbf{U}^\top,$$

  where $(\hat{\mathbf{K}}_X)_{ii} := \tilde{\lambda}_i = (\nu - \lambda_i)^+$ and $\sum_i \tilde{\lambda}_i = nP$ (to determine ν).
- Consequently, the capacity becomes

$$\boxed{C = \frac{1}{2n} \sum_{i=1}^{n} \log \left(1 + \frac{(\nu - \lambda_i)^+}{\lambda_i}\right)}$$

Colored Gaussian Noise (cont'd)

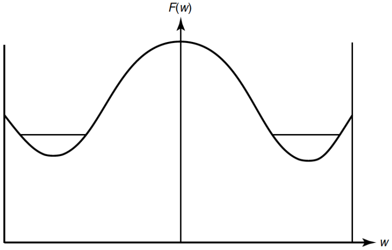


Figure: Water-filling in the spectral domain.

- If noise Z is a stationary stochastic process, it translates to water-filling in the spectral domain.
- An alternative approach: pre-whitening the noise:

$$\mathbf{y}^n = \mathbf{x}^n + \mathbf{z}^n \implies \underbrace{\mathbf{U}^\top \mathbf{y}^n}_{\tilde{\mathbf{y}}^n} = \underbrace{\mathbf{U}^\top \mathbf{x}^n}_{\tilde{\mathbf{x}}^n} + \underbrace{\mathbf{U}^\top \mathbf{z}^n}_{\tilde{\mathbf{z}}^n}$$

Gaussian Channels with Feedback

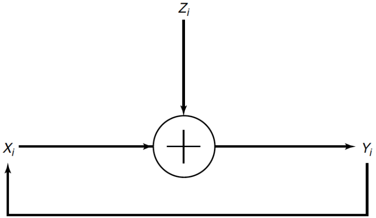


Figure: Gaussian channel with feedback: Because of the feedback, $\mathbf{x}^n$ and $\mathbf{z}^n$ are not independent; X_i is casually dependent on the past values of Z.

- For a DMC, feedback does not increase the capacity, although it may help greatly in reducing the complexity of encoding or decoding.
- The same is true of memoryless AWGN channels. However, for channels with memory, where the noise is correlated across time, feedback does increase capacity.
- However, there is no simple explicit characterization of the capacity with feedback.

Gaussian Channels with Feedback (cont'd)

Capacity with feedback and its upper bounds[4].

$$\boxed{C_{n,\mathrm{FB}} = \max_{\mathbf{K}_Z:\mathrm{Tr}(\mathbf{K}_X)\le nP} \frac{1}{2n}\log\frac{|\mathbf{K}_{X+Z}|}{|\mathbf{K}_Z|} \le \min\{C_n + \frac{1}{2},\, 2C_n\}.}$$

Gaussian channel capacity is not increased by more than half a bit or by more than a factor of 2 when we have feedback.

The above two upper bounds leverage the following results:

- $\mathbf{K}_{X+Z} + \mathbf{K}_{X-Z} = 2(\mathbf{K}_X + \mathbf{K}_Z)$
- $|\mathbf{K}_{X+Z}| \le 2^n |\mathbf{K}_X + \mathbf{K}_Z|$
- $|\mathbf{K}_{X-Z}| \ge |\mathbf{K}_Z|$ when sequences X and Z are causally related.
- $|\mathbf{K}_X + \mathbf{K}_Z| \ge \sqrt{|\mathbf{K}_{X+Z}||\mathbf{K}_{X-Z}|} \ge \sqrt{|\mathbf{K}_{X+Z}||\mathbf{K}_Z|}$

[4]See details in Sec 9.6 of Cover's book.

Thank You!

Email: `<zhangy@ucsc.edu>`

Homepage: `https://people.ucsc.edu/~yzhan419/`