A minor correction to a slide

- IDC 5-104: Please replace "Details theoretical background" by "Detail theoretical background".

1. (Optimality of Maximal-ratio combiner) Suppose (after performing inner products) the receiver receives

$$r = \pm\alpha\sqrt{\mathcal{E}} + z,$$

and assume α can be perfectly known to the receiver. Let z be real-valued zero-mean Gaussian distributed with variance $E[z^2] = N_0$. With L-diversity, we have

$$\begin{aligned} r_1 &= \pm\alpha_1\sqrt{\mathcal{E}} + z_1 \\ r_2 &= \pm\alpha_2\sqrt{\mathcal{E}} + z_2 \\ &\vdots \\ r_L &= \pm\alpha_L\sqrt{\mathcal{E}} + z_L \end{aligned}$$

where $\{z_k\}_{k=1}^L$ are zero-mean i.i.d. with variance σ^2 and $\{\alpha_k\}_{k=1}^L$ are assumed to be perfectly estimated. Find the optimal **linear** combiner

$$r = \sum_{k=1}^{L} w_k \cdot r_k$$

that maximizes the output signal-to-noise ratio (SNR).

Hint: Find the weights $\{w_k\}_{k=1}^L$ such that the output SNR is maximized by using the Cauchy-Schwartz inequality.

Solution. We derive

$$\begin{aligned} r &= \sum_{k=1}^{L} w_k r_k \\ &= \sum_{k=1}^{L} w_k(s + z_k) \\ &= \sum_{k=1}^{L} w_k\alpha_k s + \sum_{k=1}^{L} w_k z_k \\ &= s\sum_{k=1}^{L} w_k\alpha_k + \sum_{k=1}^{L} w_k z_k, \end{aligned}$$

where $s \in \{\pm\sqrt{\mathcal{E}}\}$. Under the assumption that $\{\alpha_k\}_{k=1}^L$ can be perfectly estimated by the receiver, the output SNR satisfies

$$\text{SNR} = \frac{\left(\sum_{k=1}^L w_k\alpha_k\right)^2 E[s^2]}{\sum_{k=1}^L w_k^2 E[z_k^2]} \le \frac{\left(\sum_{k=1}^L w_k^2\right)\left(\sum_{k=1}^L \alpha_k^2\right) E[s^2]}{\sigma^2 \sum_{k=1}^L w_k^2} = \frac{\left(\sum_{k=1}^L \alpha_k^2\right) E[s^2]}{N_0},$$

where the inequality follows the Cauchy-Schwartz inequality. The optimal $\{w_k\}_{k=1}^L$ are thus the ones that equate the Cauchy-Schwartz inequality, i.e., $w_k = C \cdot \alpha_k$ for some constant C.

Note: This problem shows that as long as $\{\alpha_k\}_{k=1}^L$ can be perfectly estimated, and $\{z_k\}_{k=1}^L$ is i.i.d., the maximal ratio combiner is optimal (in the sense of maximizing the output SNR). No Rayleigh or Gaussian assumptions are actually needed.

2. In an antenna arrays system, suppose the first user equipment (UE) has two antennas, the second UE has one antennas, the third UE has two antennas, and the base station has eight antennas. Thus, the base station receives

$$\underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{bmatrix}}_{=\boldsymbol{x}} = \begin{bmatrix} \boldsymbol{c}_1 & \boldsymbol{c}_2 & \boldsymbol{c}_3 & \boldsymbol{c}_4 & \boldsymbol{c}_5 \end{bmatrix} \begin{bmatrix} m_1 \\ m_1 \\ m_2 \\ m_3 \\ m_3 \end{bmatrix} + \underbrace{\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \\ v_8 \end{bmatrix}}_{\boldsymbol{v}},$$

where we suppose $\{\boldsymbol{c}_i\}_{i=1}^8$ can be perfectly pre-measured, and $\{v_k\}_{k=1}^8$ are zero-mean i.i.d. with variance σ^2. Let m_1 be the message of interest. The receiver then uses a **linear** combiner with weights $\{w_k\}_{k=1}^8$ to determine m_1. Find the weights that maximize the output SNR, subject to that $\boldsymbol{w}$ is orthogonal to $\boldsymbol{c}_3$ and $\boldsymbol{c}_{45} \triangleq \boldsymbol{c}_4 + \boldsymbol{c}_5$, where

$$\boldsymbol{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ w_6 \\ w_7 \\ w_8 \end{bmatrix}.$$

Hint: Denote for convenience $\boldsymbol{c}_{12} = \boldsymbol{c}_1 + \boldsymbol{c}_2$.

Solution. By using a dot to denote the inner product, the receiver performs

$$\begin{aligned} \boldsymbol{w} \cdot \boldsymbol{x} &= \boldsymbol{w} \cdot \begin{bmatrix} \boldsymbol{c}_1 & \boldsymbol{c}_2 & \boldsymbol{c}_3 & \boldsymbol{c}_4 & \boldsymbol{c}_5 \end{bmatrix} \begin{bmatrix} m_1 \\ m_1 \\ m_2 \\ m_3 \\ m_3 \end{bmatrix} + \boldsymbol{w} \cdot \boldsymbol{v} \\ &= \boldsymbol{w} \cdot \begin{bmatrix} \boldsymbol{c}_1 + \boldsymbol{c}_2 & \boldsymbol{c}_3 & \boldsymbol{c}_4 + \boldsymbol{c}_5 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} + \boldsymbol{w} \cdot \boldsymbol{v} \\ &= \begin{bmatrix} \boldsymbol{w} \cdot \boldsymbol{c}_{12} & 0 & 0 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} + \boldsymbol{w} \cdot \boldsymbol{v} \\ &= m_1 \boldsymbol{w} \cdot \boldsymbol{c}_{12} + \boldsymbol{w} \cdot \boldsymbol{v}. \end{aligned}$$

The output SNR is thus given by

$$\text{SNR} = \frac{E[|m_1 \boldsymbol{w} \cdot \boldsymbol{c}_{12}|^2]}{E[|\boldsymbol{w} \cdot \boldsymbol{v}|^2]} = \frac{|\boldsymbol{w} \cdot \boldsymbol{c}_{12}|^2 \, E[|m_1|^2]}{\|\boldsymbol{w}\|^2 \sigma^2}.$$

Note that $\boldsymbol{w}$ must be orthogonal to the space $\mathcal{I}$ spanned by $\boldsymbol{c}_3$ and $\boldsymbol{c}_{45}$. Decompose

$$\boldsymbol{c}_{12} = \boldsymbol{u}^\perp + \boldsymbol{u},$$

where $\boldsymbol{u}$ is on the space $\mathcal{I}$ (so, $\boldsymbol{w} \cdot \boldsymbol{u} = 0$) and $\boldsymbol{u}^\perp$ is orthogonal to the space $\mathcal{I}$. The output SNR can the be rewritten as

$$\text{SNR} = \frac{\boldsymbol{w} \cdot (\boldsymbol{u}^\perp + \boldsymbol{u}) E[|m_1|^2]}{\|\boldsymbol{w}\|^2 \sigma^2} = \frac{\boldsymbol{w} \cdot \boldsymbol{u}^\perp E[|m_1|^2]}{\|\boldsymbol{w}\|^2 \sigma^2} \le \frac{\|\boldsymbol{w}\|^2 \|\boldsymbol{u}^\perp\|^2 E[|m_1|^2]}{\|\boldsymbol{w}\|^2 \sigma^2} = \frac{\|\boldsymbol{u}^\perp\|^2 E[|m_1|^2]}{\sigma^2}$$

where we use the Cauchy-Schwartz inequality. The optimal $\boldsymbol{w}$ is thus proportional to $\boldsymbol{u}^\perp$.

Note: The L-diversity communication in Problem 1 can be regarded as a special case of Problem 2, where $\boldsymbol{c}_3 = \boldsymbol{c}_4 = \boldsymbol{c}_5 = \boldsymbol{0}$, $m_1 = s \in \{\pm\sqrt{\mathcal{E}}\}$, and

$$\boldsymbol{x} = \begin{bmatrix} r_1 \\ \vdots \\ r_L \end{bmatrix}, \quad \boldsymbol{u} = \boldsymbol{c}_{1,2} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_L \end{bmatrix}.$$

Thus, the antenna array system can be considered a further generalization of L-diversity.

3. After presuming that "information measure" is a function of event probability, three axioms have been proposed by Shannon. What are these three axioms? Also, give their corresponding mathematical formulations.

   Solution. See Slides IDC 6-6 $\sim$ 6-8.

   Note: In fact, Shannon mentioned four axioms. The last axiom is that **the information measure should be non-negative**. There is no reason for one to make a new observation but gain "negative" information amount due to it. But the function of information measure that satisfies the first three axioms turns out to automatically satisfy the non-negativity axiom. So, this axiom can be ignored and removed.

4. (a) Compute the entropy $H(X)$ of random variable X, where

$$\Pr[X = 0] = \frac{1}{8}, \quad \Pr[X = 1] = \frac{1}{8}, \quad \Pr[X = 2] = \frac{1}{4}, \quad \text{and} \quad \Pr[X = 3] = \frac{1}{2}.$$

   (b) Compute the entropy $H(Y)$ of random variable Y, where

$$\Pr[Y = i] = \frac{1}{D} \quad \text{for } i = 1, 2, \ldots, D.$$

   Solution

   (a)

$$\begin{aligned} H(X) &= \frac{1}{8}\log_2 \frac{1}{(\frac{1}{8})} + \frac{1}{8}\log_2 \frac{1}{(\frac{1}{8})} + \frac{1}{4}\log_2 \frac{1}{(\frac{1}{4})} + \frac{1}{2}\log_2 \frac{1}{(\frac{1}{2})} \\ &= \frac{1}{8} \times 3 + \frac{1}{8} \times 3 + \frac{1}{4} \times 2 + \frac{1}{2} \times 1 \\ &= \frac{7}{4} \text{ bits} \end{aligned}$$

(b)

$$H(Y) = \sum_{i=1}^{D} \frac{1}{D} \log_2 \frac{1}{(\frac{1}{D})} = \log_2(D) \text{ bits.}$$

Note: Rudolph Clausius had introduced entropy as a concept in thermodynamics in 1854 and defined entropy from a macroscopic point of view. As an alternative, Boltzmann gave a microscopic description of entropy in 1877:

$$S = k_B \log D,$$

where S is entropy, D is the number of energy states in the system that can be randomly filled with energy, and K_B is the Boltzmann constant. In comparison with Shannon's entropy formula, we know that Boltzmann's formula assumes uniform distribution over all energy states, as exemplified in this subproblem. Thus, it is reasonable to name the "uncertainty" formula developed from the three axioms proposed by Shannon the *entropy* of the source.