Lecture 3: Demodulation with Noise

Scribe: 鄧傑方、黃安 Lecture Date: 3/15, 2017 Lecturer: I-Hsiang Wang

1 Outline

- Demodulation with noise (Detection)：在有雜訊的時候，收到的訊號會變形，要用 Detection 得出原來的訊號
- Basic Principles of Detection
  - Binary Detection in scalar
  - M-ary Detection
  - Minimum distance (MD)：雜訊是 Gaussian distribution，大小為零的機率最大，所以用最短距離來找還原後的訊號
  - Detection in Multi-Dimensional Space：解調原則一樣，只是比較複雜一點

2 Noise

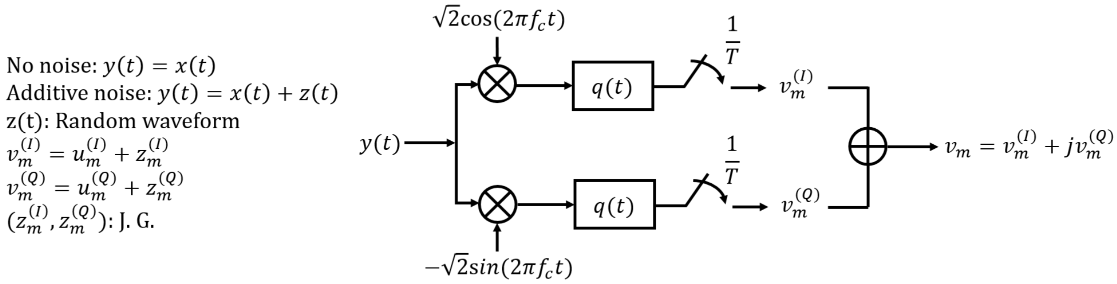

Figure 1: Basic Diagram of Received Signal

- Random Process $\{z(t), \forall t \in \mathbb{R}\}$
  - sampled at $t_1, t_2, \ldots, t_n, \ldots$
  - obtain: $z(t_1), z(t_2), \ldots, z(t_n), \ldots$
- How to specify the "probability distribution" of this random process?
- Answer: specify the joint distribution of $\{z(t) : t \in S \subseteq \mathbb{R}\}$ for all $|S|$ finite
  - e.g. need to specify $P_{z(t_1)}\ \forall t_1 \in \mathbb{R}, P_{z(t_1)z(t_2)}\ \forall t_1, t_2 \in \mathbb{R}$
- Def: (Gaussian Process) continuous time
  - $\{z(t) | t \in \mathbb{R}\}$ is a Gaussian Process if for all $n \in N = \{1, 2, 3, \ldots\}$ and any $\{t_1, t_2, \ldots, t_n\} \subseteq \mathbb{R}$
  - $(z(t_1), z(t_2), \ldots, z(t_n))$ is a multivariate Gaussian (jointly Gaussian)
- Def: (Joint Gaussianity) discrete time
  - $(z_1, z_2, \ldots, z_n)$: *J.G.* iff there exists iid N(0,1) $w_1, w_2, \ldots\ldots, w_m$, $\quad \underline{w} \triangleq \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix}$, $m \leq n$ and constant $\underline{B} \in \mathbb{R}^n$ and matrix $A \in \mathbb{R}^{n \times m}$
  - $\underline{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = A\underline{w} + \underline{B}$

  - 用 Gaussian distribution 表示雜訊分布: 干擾、熱擾動...... 各種獨立的雜訊疊加在一起，總和的效應呈 Gaussian distribution

- Def: (Additive White Gaussian Process)
  - Recall: Power Spectral Density
  - For a random process that is wide-sense stationary(WSS)
    $\begin{cases} \text{mean: } E[z(t)] = \mu \quad \forall t \in \mathbb{R} \\ \text{auto-correlation: } E[z(t_1)z(t_2)] = R_z(\tau) \quad \forall t_1, t_2, ..., t_1 - t_2 = \tau \text{ (只跟時間差有關)} \end{cases}$
  - Its P.S.D $S_z(f) \triangleq F\{R_z(\tau)\}$
  - $\{w(t)\}$ is a white Gaussian Process iff $R_w(\tau) = \frac{N_0}{2}\delta(\tau)$, $S_w(f) = \frac{N_0}{2}, \forall f$ (各種頻率的光都有等效貢獻)

- Claim: If $z(t)$ is a WGN, then $\{z_m^{(I)}\}$: iid Gaussian, $\{z_m^{(Q)}\}$: iid Gaussian, $\{z_m^{(I)}\}$ and $\{z_m^{(Q)}\}$ are independent.

3 Basic Principles of Detection

- Binary Detection in Scalar
  model: $y = x + z$, $x \in \{\pm A\} \quad z \sim N(0, \sigma^2)$
- BPSK/Binary PAM:

Figure 2: Block diagram of BPSK/Binary PAM

- Goal: Decide $x = +A$ or $-A$ (equivalently, $\hat{c} = 1$ or 0) such that "Probability of error" is minimized.
- Problem Formulation:
  - Given
    1. Prior distribution $Pr\{c = 3\}$ and $Pr\{c = 1\}$
    2. $P_{Y|C}(y|c)$: likelihood function
  - Find a rule, $\hat{c}(y)$, such that $P_e \triangleq Pr\{c \neq \hat{c}(Y)\}$ is minimized.
- Deeper Look into P_e:
  - $P_e = Pr\{c \neq \hat{c}(Y)\}$
    $\hat{c}$ is a mapping from $\mathbb{R}$ to $\{0, 1\}$ and is equivalent to "Decision Region"
  - D_0 and D_1 :
    * $D_0 \triangleq \{y \in \mathbb{R} : \hat{c}(y) = 0\}$
    * $D_1 \triangleq \{y \in \mathbb{R} : \hat{c}(y) = 1\}$
  - $1 - P_e = Pr\{c = \hat{c}(Y)\} = \sum_{i=0,1} Pr\{c = i \text{ and } \hat{c}(Y) = i\}$
    $= \sum_{i=0,1} Pr\{c = i|Y \in D_i\}Pr\{Y \in D_i\} = \sum_{i=0,1} \int_{D_i} P_{C|Y}(c = i|y)f_Y(y)\,dy$
- Q: Given y, should it belongs to D_0 or D_1?
  - $P_{C|Y}(c = 0|y)$, $P_{C|Y}(c = 1|y)$: 事後機率 (a posterior probability)
  - general: $D_i = \{y|P_{C|Y}(c = i|y)$ is larger than $P_{C|Y}(c = j|y) \forall j \neq i\}$
    * In D_0: $P_{C|Y}(c = 0|y)$ should be larger than $P_{C|Y}(c = 1|y)$
    * In D_1: $P_{C|Y}(c = 1|y)$ should be larger than $P_{C|Y}(c = 0|y)$

- Answer: (maximum a Posterier Probability Rule) "MAP"

$$\begin{aligned}\hat{c}_{MAP}(y) &= \underset{i\in M}{\operatorname{argmax}}\{P_{C|Y}(c=i|y)\} \quad (M=\{0,1\} \text{ in the binary case})\\ &= \underset{i\in M}{\operatorname{argmax}}\{f_{Y|C}(y|c=i)\cdot Pr\{c=i\}\}\\ &= \underset{i\in M}{\operatorname{argmax}}\{f_{Y|C}(y|c=i)\} \quad (\text{when } Pr\{c=i\}=\frac{1}{M} \text{ is uniform prior})\end{aligned}$$

- Look into a posterier probability

$$P_{C|Y}(c=i|y)=\frac{f_{Y|C}(y|c=i)Pr\{c=i\}}{f_Y(y)} \propto f_{Y|C}(y|c=i)Pr\{c=i\} \quad \text{for fixed y}$$

- Maximum Likelihood Rule

$$\hat{c}_{ML}(y)=\underset{i\in M}{\operatorname{argmax}}\{f_{Y|C}(y|c=i)\} \equiv \hat{c}_{MAP} \quad \text{when prior is uniform}$$

- Summary:
  - MAP detection: $\hat{c}_{MAP}(y)=\underset{i\in M}{\operatorname{argmax}}\{f_{Y|C}(y|c=i)Pr\{c=i\}\}$
  - ML detection: $\hat{c}_{ML}(y)=\underset{i\in M}{\operatorname{argmax}}\{f_{Y|C}(y|c=i)\}$
  - MAP is the optimal detector
  - ML≡MAP when prior is uniform, i.e., $Pr\{c=i\}=\frac{1}{M} \quad \forall i\in\mu,\ M=|\mu|$
- Specialization to the additive Gaussian noise case

$$\begin{aligned}&c\in M \leftrightarrow x\in A \qquad A: \text{constellation set}\\ &A=\{a_1,a_2,\ldots,a_M\},\quad M=2^l\\ &y=x+z,\ z\sim N(0,\sigma^2) \Rightarrow Y|C=i \quad \sim N(a_i,\sigma^2)\\ &\qquad\qquad \Rightarrow f_{Y|C=i}(y|c=i)=\frac{1}{\sqrt{2\pi\sigma^2}}exp\left(\frac{-(y-a_i)^2}{2\sigma^2}\right)\\ &(\text{If prior is uniform})\hat{c}_{MAP}=\hat{c}_{ML}(y)=\underset{i\in M}{\operatorname{argmin}}\{(y-a_i)^2\}=\text{Minimum Distance Detection}\end{aligned}$$

- Detection in Vector: scalar → vector: straight forward extension!
  - Model:

$$\begin{aligned}&\underline{y}=\underline{x}+\underline{z},\ \underline{x}\in A\subseteq\mathbb{R}^n,\ |A|=M \quad c\in\mu\leftrightarrow\underline{x}\in A \quad A=\{\underline{a_1},\underline{a_2},\ldots,\underline{a_M}\},\quad M=2^l\\ &\underline{y}=\underline{x}+\underline{z} \quad \underline{z}=\begin{bmatrix}z_1\\ \vdots\\ z_n\end{bmatrix} \quad z_1,z_2,\ldots,z_n\text{: iid } N(0,\sigma^2)\\ &\Rightarrow \underline{Y}|C=i \quad \sim N(\underline{a_i},\sigma^2 I_n)\\ &\Rightarrow f_{\underline{Y}|C=i}(y|c=i)=\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n\prod_{k=1}^{n}exp\left(\frac{-(y_k-a_{ik})^2}{2\sigma^2}\right)\\ &=\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n exp\left(\frac{-1}{2\sigma^2}\sum_{k=1}^{n}|y_k-a_{ik}|^2\right)=\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n exp\left(\frac{-1}{2\sigma^2}|y-a_i|^2\right)\end{aligned}$$

- Example:

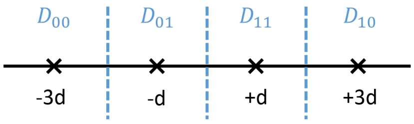

Figure 3: 4-PAM

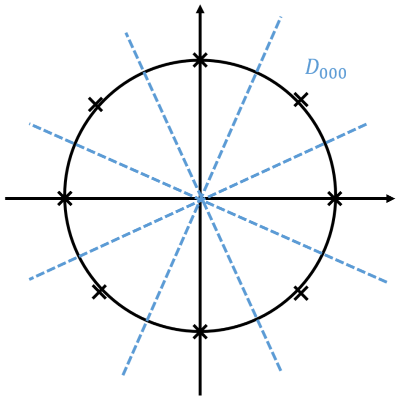

Figure 4: 8-PSK

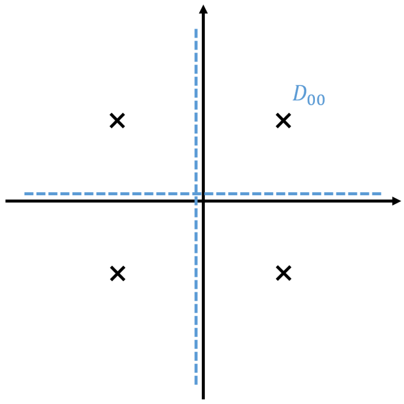

Figure 5: 4QAM

- Probability of Error and SNR:
  Now let's compare the performance of MD detector under iid Gaussian noise and see the performance of detection.

- Binary PAM
  - $y = x + z$, $z \sim N(0, \sigma^2) \leftarrow$ prior: uniform

$$\begin{aligned} P_e &= Pr\{\hat{c}(Y) \neq c\} \\ &= Pr\{c = 0\}Pr\{\hat{c}(Y) \neq 0|c = 0\} + Pr\{c = 1\}Pr\{\hat{c}(Y) \neq 1|c = 1\} \end{aligned}$$

$$\begin{aligned} Pr\{\hat{c}(Y) \neq 0|c = 0\} &= Pr\{Y \in D_1|c = 0\} \\ &= Pr\{Y > 0|c = 0\} \\ &= Pr\{N(-d, \sigma^2) > 0\} \\ &= Pr\{N(0, \sigma^2) > d\} \\ &= Pr\{N(0, 1) > \frac{d}{\sigma}\} \Rightarrow Q(\frac{d}{\sigma}) \end{aligned}$$

- Def: (Q-function) $Q(x) = Pr\{N(0, 1) > x\} = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\, dt$

$$\begin{aligned} P_e &= \frac{1}{2}Q(\frac{d}{\sigma}) + \frac{1}{2}Q(\frac{d}{\sigma}) = Q(\frac{d}{\sigma}) \\ &\Rightarrow -log\, P_e \sim \frac{d^2}{\sigma^2} \Rightarrow P_e = Q(\sqrt{SNR}) \end{aligned}$$

$$P_e \text{ increases when } \frac{d}{\sigma} \text{ decreases; } P_e \text{ decreases when } \frac{d}{\sigma} \text{ increases.}$$

  - Bound on the Q-function: $Q(x) \triangleq \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\, dt \leq \frac{1}{2}e^{-\frac{1}{2}x^2}, \forall x$

- Signal-to-Noise ratio
  - Noise variance: σ^2
  - Signal energy: $E_s = \frac{1}{2}|+d|^2 + \frac{1}{2}|-d|^2 = d^2$
  - $SNR = \frac{\text{signal energy}}{\text{noise variance}} = \frac{E_s}{\sigma^2} = \frac{d^2}{\sigma^2} \Rightarrow P_e = Q(\frac{d}{\sigma}) = Q(\sqrt{SNR})$