5.1.5 A symbol-by-symbol detector for signals with memory

\[
q_i(s_{1:D}, s_1) = P(s_{1:D}, s_1) P(r_{1:D} | s_{1:D}, s_1) \\
= P(s_{1:D}, s_1) [P(r_{1:D} | s_{1:D}, s_1) P(r_{1:D} | s_{1:D} = \cdot^{D-L})] \\
= P(s_{1:D}, s_1) [P(r_{1:D} | s_{1:D} = \cdot^{D-L}) \sum_{s_{1:D-L}} P(r_{1:D} | s_{1:D-L}, s_1) ] \\
= P(s_{1:D}, s_1) \left[ P(r_{1:D} | s_{1:D-L}, s_1) \frac{\sum_{s_{1:D-L}} P(r_{1:D} | s_{1:D-L}, s_1) }{P(s_{1:D}, s_1)} \right] \\
= P(s_{1:D}, s_1) \left[ P(r_{1:D} | s_{1:D-L}, s_1) \right] \frac{\sum_{s_{1:D-L}} q_i(s_{1:D-L}, s_1) }{P(s_{1:D}, s_1)} \\
= P(r_{1:D} | s_{1:D-L}, s_1) \frac{P(s_{1:D}, s_1) \sum_{s_{1:D-L}} q_i(s_{1:D-L}, s_1) }{P(s_{1:D}, s_1)} \\
= P(r_{1:D} | s_{1:D-L}) \frac{P(s_{1:D}, s_1) \sum_{s_{1:D-L}} q_i(s_{1:D-L}, s_1) }{P(s_{1:D}, s_1)} \\
= P(r_{1:D} | s_{1:D-L}) \sum_{s_{1:D-L}} q_i(s_{1:D-L}, s_1) \\
\hat{s}_3 = \arg \max_{s_3 \in [-M,...,-1]} \sum_{s_1 \in [-M,...,-1]} \sum_{s_{1:D-2}} q_3(s_{1:D}, s_3) \\
q_3(s_{1:D}, s_3) = P(r_{1:D} | s_{1:D-2}, s_3) \sum_{s_{1:D-2}} q_3(s_{1:D-2}) \\
\hat{s}_4 = \arg \max_{s_4 \in [-M,...,-1]} \sum_{s_3 \in [-M,...,-1]} \sum_{s_{1:D-3}} q_4(s_{1:D}, s_4) \\
q_4(s_{1:D}, s_4) = P(r_{1:D} | s_{1:D-3}, s_4) \sum_{s_{1:D-3}} q_4(s_{1:D-3}) \\
\vdots \\
\hat{s}_k = \arg \max_{s_k \in [-M,...,-1]} \sum_{s_{k-1} \in [-M,...,-1]} \sum_{s_{1:D-k}} q_k(s_{1:D}, s_k) \\
q_k(s_{1:D}, s_k) = P(r_{1:D} | s_{1:D-k}, s_k) \sum_{s_{1:D-k}} q_k(s_{1:D-k}) \\
\vdots
5.1.5 A symbol-by-symbol detector for signals with memory

- Advantage of the above approach
  - Optimality (in MAP sense) under the constraint of delay = D.
  - Hence, it can be applied to the case of non-equal prior.
- Disadvantage of the above approach
  - Computational complexity, especially when \( M \) or \( L \) is large.

5.2 Performance of the optimum receiver for memoryless modulation

- Performance = Probability of error
- In this section, our goal is to derive the probability of error for memoryless modulation/demodulation.
5.2.1 Probability of error for binary modulation

**Antipodal signal**

*One-dimensional signal (Refer to slide 5-42)*

\[ r = \pm \sqrt{E} + n, \text{ where } n \text{ is zero - mean Gaussian distributed with variance } \frac{N_0}{2} \]

\[ \Rightarrow r \text{ is Gaussian distributed with mean } \pm \sqrt{E} \text{ and variance } N_0 / 2. \]

\[ d_{\text{Adj}}(r) = \begin{cases} \sqrt{E}, & \text{if } r < \frac{N_0}{4\sqrt{E}} \ln \frac{1-p}{p} \\ \sqrt{E}, & \text{otherwise} \end{cases} \]

\[ \Rightarrow P_{e,\text{BPAM}} = P(\sqrt{E})P(\text{error } | s = \sqrt{E}) + P(-\sqrt{E})P(\text{error } | s = -\sqrt{E}) \]

\[ = p \int_{-\infty}^{-\infty} \text{Normal}(\sqrt{E}, N_0 / 2) + (1-p) \int_{-\infty}^{\infty} \text{Normal}(-\sqrt{E}, N_0 / 2) \]

\[ = p \cdot \Phi \left( \frac{\tau - \sqrt{E}}{\sqrt{N_0 / 2}} \right) + (1-p) \cdot \left[ 1 - \Phi \left( \frac{\tau + \sqrt{E}}{\sqrt{N_0 / 2}} \right) \right] \]

\[ \text{SNR is usually defined as } (E / N_0). \quad P_{e,\text{BPAM}} = \Phi \left( \sqrt{2 \times \text{SNR}} \right) \quad (\text{SNR}_0 = \frac{E}{N_0}) \]

- The larger the SNR, the smaller the probability of error.

**Observation 2:** POE can also be expressed in terms of the distance between the two signals.

- The larger the distance among signals, the smaller the probability of error.

\[ P_{e,\text{BPAM}} = \Phi \left( \frac{-d_{\text{adj}}}{\sqrt{2N_0}} \right) \text{ where } d_{\text{adj}} = \sqrt{E} - (-\sqrt{E}) = 2\sqrt{E}. \]
### 5.2.1 Probability of error for binary modulation

\[
P_{e,\text{antipodal}} = \Phi(-\sqrt{2\gamma_b}) = 0.5\text{erfc}(10^{-\gamma_b/20})
\]

\[\gamma_b = \text{SNR per information bit}\]

On white noise

- **The power of white noise processes (an erroneous view).**

To derive the power of \( n(t) \), we need to first establish its autocorrelation function.

\[
\phi_n(\tau) = E[n^*(t)n(t+\tau)] = \begin{cases} 
N_0/2, & \text{if } \tau = 0 \\
0, & \text{otherwise}
\end{cases}
\]

\[
\Rightarrow \Phi_n(f) = N_0/2
\]

Recall that the average power of a WSS process is \( E[X_n^2] = \phi(0) \).

An alternative way to compute it is to use \( \phi(\tau) = \int_\tau^\infty \Phi(f)e^{i\omega \tau}df \Rightarrow \phi(0) = \int_\tau^\infty \Phi(f)df \)

Hence, \( \phi_n(0) = E[n^2(t)] = \int_\infty^\infty \Phi_n(f)df = \infty \) !!!
On white noise

\[ \square \text{Definition of white noise} \]
\[ \wedge \text{Discrete white noise} \]

- **Definition**
  A discrete random process \( \{X_n\} \) is said to be white if \( \Phi_X(f) = \text{constant for } -\frac{1}{2} \leq f < \frac{1}{2}. \)

- **Autocorrelation function of discrete white noise**
  \[ \phi_X(\tau) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \Phi_X(f) e^{j2\pi \tau f} df = \sigma_X^2 \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{j2\pi \tau f} df = \begin{cases} \sigma_X^2, & \text{if } \tau = 0 \\ 0, & \text{if } \tau = \pm 1, \pm 2, \ldots \end{cases} \]

- **Average power of discrete white noise**
  \[ \phi_X(0) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \Phi_X(f) df = \sigma_X^2. \]

---

On white noise

- A white process is often implicitly assumed zero-mean.

\[ \wedge \text{Advantage of adopting discrete white noise concept} \]

- It is well-defined and free of analytical difficulties.
**On white noise**

- **Continuous white noise**
  - **Definition**
    
    A continuous random process \( \{X_t\} \) is said to be white
    
    if \( \Phi_x(f) = \text{constant} \) for all \( f \in \mathbb{R} \)
  
  - **Autocorrelation function of discrete white noise**
    
    \[ \phi_x(\tau) = \int_{-\infty}^{\infty} \Phi_x(f) e^{i2\pi f \tau} df = \sigma_x^2 \delta(\tau). \]

  **Properties of \( \delta(\tau) \), Dirac delta function:**
  
  \begin{align*}
  (1) & \int_{-\infty}^{\infty} h(\tau) \delta(\tau) d\tau = h(0). \quad (\text{Hence, } \int_{-\infty}^{\infty} \delta(\tau) d\tau = 1.)
  
  (2) & \delta(0) = \infty.
  
  (3) & \delta(\tau) = 0 \text{ for } \tau \neq 0.
  \end{align*}

---

**On white noise**

- **Solution to the previous non-consistence.**

To derive the power of \( n(t) \), we need to first establish its autocorrelation function.

For continuous white noise, \( \phi(0) = \infty. \)

\[ \phi_n(\tau) = E[n^*(t)n(t+\tau)] = \frac{N_o}{2} \delta(\tau) = \begin{cases} \infty, & \text{if } \tau = 0 \\ 0, & \text{otherwise} \end{cases} \]

\[ \Rightarrow \Phi_n(f) = N_o/2 \]

Recall that the average power of a WSS process is \( E[X_n^2] = \phi(0). \)

An alternative way to compute it is to use \( \phi(\tau) = \int_{-\infty}^{\infty} \Phi(\tau) e^{i2\pi f \tau} df \Rightarrow \phi(0) = \int_{-\infty}^{\infty} \Phi(f) df \)

Hence, \( \phi_n(0) = E[n^2(t)] = \int_{-\infty}^{\infty} \Phi_n(f) df = \infty \) !!!
On white noise

Therefore, we say that the average power of a continuous white noise is infinite.

It should not be a problem for its analysis because the post-sampling noise power is finite!
On white noise

– So impractical, why the continuous white noise is so popular?
  • There do exist noise sources that have a flat power spectral density over a range of frequencies that is much larger than the bandwidths of subsequent filters or measurement devices.

Based on physically measurement, some have shown that the autocorrelation function of the noise has the form

\[ \phi_t(t) = kTRe^{-\alpha t}, \]

where \( k \) is the Boltzmann's constant,

\( T \) is the absolute temperature,

\( R \) and \( \alpha \) are the parameters of the physical medium.

Its power spectrum density is

\[ \Phi(f) = kTR \frac{2\alpha^2}{\alpha^2 + (2\pi f)^2} \approx 2kTR \text{, when } f \ll \alpha. \]

Hence, the noise process looks like white over a large bandwidth, a fact that has been verified closely by experiment.

On white noise

– When the physical noise is introduced to a filter, the output power spectrum density will be approximately the same as it would be if noise is assumed ideally to be white.

\[ \Phi_o(f) = |H(f)|^2 \Phi_t(f) \]

– Commonly, the continuous-time white noise is represented as being white with flat spectral density equal to \( N_0/2 \).

  • The factor of 2 is included because of the “two-sided” nature of filter transfer functions.

– Conclusion:
  • Some researchers think that such a white noise makes sense only if one sees it through a filter.
On white noise

- Properties of continuous white noise
  - No matter how close in time, the noise samples are uncorrelated.
  - If Gaussian, no matter how close in time, the noise samples are independent.
  - Physical interpretation: The adjacent noise samples are uncorrelated at time greater than the shortest time delays in subsequent filtering.

- If the adjacent noise samples at time period smaller than the shortest time delays in subsequent filtering, then the noise samples become correlated in nature.

On white noise

- Research on white noise seen through a filter.

Spectral factorization theorem:

\[
\text{if } \int_{-\infty}^{\infty} \ln \Phi_s(f) \, df > -\infty,
\]

then \( \Phi_s(f) = |H(f)|^2 \left( \frac{N_o}{2} \right) \) for a causal linear stable time-invariant filter.

A process satisfying spectral factorization theorem is physically realizable and possibly seen in practice.

- This is the end of the discussion on white noises.
5.2.1 Probability of error for binary modulation

- Orthogonal two-dimensional signals

\[
\begin{bmatrix}
  s_1 \\
  s_2 
\end{bmatrix} = \begin{bmatrix}
  \sqrt{E} & 0 \\
  0 & \sqrt{E}
\end{bmatrix}
\begin{bmatrix}
  r_1 \\
  r_2 
\end{bmatrix}
\]

\[\|s_1 - s_2\| = \sqrt{2E}\]

\[r = [\sqrt{E} + n_1, n_2], \text{if } s(t) \text{ is transmitted.}\]

---

5.2.1 Probability of error for binary modulation

- Determine the optimum MAP detector under AWGN with power spectrum density \(N_0/2\) and \(M = 2\).

\[r = [\sqrt{E} + n_1, n_2] \text{ or } [n_1, \sqrt{E} + n_2], \text{ where } n_1 \text{ and } n_2 \text{ are independent Gaussians with zero mean and marginal variance } N_0/2.\]

\[P(s_1) = p \text{ and } P(s_2) = 1 - p.\]

\[d_{MAP}(r) = \arg \max_{s_1, s_2} p[r | s_m]P(s_m) = \arg \max \left\{ p e^{-\frac{(r_1 - \sqrt{E} + r_2)^2}{2N_0}}, e^{-\frac{(r_1 + \sqrt{E} + r_2)^2}{2N_0}} \right\}\]

\[= \begin{cases}
  s_1, & r_2 - r_1 \leq \tau = \frac{N_0}{2\sqrt{E}} \ln \frac{p}{1 - p} \\
  s_2, & \text{otherwise.}
\end{cases}\]
5.2.1 Probability of error for binary modulation

\[ \begin{align*}
  r_2 - r_1 &= \text{Normal}(-\sqrt{E}, N_0), \text{if } s_1 \text{ is transmitted.} \\
  r_2 - r_1 &= \text{Normal}(\sqrt{E}, N_0), \text{if } s_2 \text{ is transmitted.}
\end{align*} \]

\[ P_{\text{e,BFSK}} = P(s_1)P(\text{error} | s_1) + P(s_2)P(\text{error} | s_2) \]

\[ = p\int_{-\sqrt{E}}^{\infty} \text{Normal}(-\sqrt{E}, N_0) + (1 - p)\int_{-\sqrt{E}}^{\infty} \text{Normal}(\sqrt{E}, N_0) \]

\[ = p \cdot \left[ 1 - \Phi \left( \frac{\sqrt{E}}{\sqrt{N_0}} \right) \right] + (1 - p) \cdot \Phi \left( \frac{-\sqrt{E}}{\sqrt{N_0}} \right) \]

---

5.2.1 Probability of error for binary modulation

\[ \def\thetitle#1\end{title}{\textbf{Observation 1}} \]

For equal prior, \( \tau = 0. \)

\[ \Rightarrow P_{\text{e,BFSK}} = \frac{1}{2} \left[ 1 - \Phi \left( \frac{\sqrt{E}}{\sqrt{N_0}} \right) \right] + \frac{1}{2} \cdot \Phi \left( \frac{-\sqrt{E}}{\sqrt{N_0}} \right) = \Phi \left( \frac{E}{\sqrt{N_0}} \right) \]

Since \( \text{SNR} = \frac{E}{N_0}, \ P_{\text{e,BFSK}} = \Phi \left( \frac{-\sqrt{\text{SNR}}}{\sqrt{N_0}} \right) \)

\( \text{– The larger the SNR, the smaller the probability of error.} \)

\[ \def\thetitle#1\end{title}{\textbf{Observation 2}} \]

\( \text{– The larger the distance among signals, the smaller the probability of error (Cf. slide 5-90).} \)

\[ P_{e,BFSK} = \Phi \left( \frac{-d_{12}}{\sqrt{2N_0}} \right), \text{ where } d_{12} = \sqrt{2E}. \]
5.2.1 Probability of error for binary modulation

Comparison between binary antipodal signals (BPAM) and binary orthogonal signals (BFSK)

- Binary orthogonal signal requires twice the transmitted power than antipodal signal to achieve the same probability of error.
  \[ P_{e, \text{antipodal}} = \Phi\left( -\sqrt{2 \times \text{SNR}} \right) \]
  \[ P_{e, \text{BOrthogonal}} = \Phi\left( -\sqrt{\text{SNR}} \right) \]

- Binary orthogonal signal is 3dB poorer than antipodal signal since \(10\log_{10}2 = 3\) dB.

\[ \Phi(\sqrt{\gamma}) = Q\left( \sqrt{\gamma}\right) = Q\left(\frac{10^{\gamma_{\text{dB}}}/20}{\sqrt{2}} \right) = 0.5\text{erfc}\left(\frac{10^{\gamma_{\text{dB}}}/20}{\sqrt{2}} \right) \]

\[ P_{e, \text{antipodal}} = \Phi\left( -\sqrt{2\gamma} \right) = Q\left(\sqrt{\gamma}\right) = Q\left(\frac{10^{\gamma_{\text{dB}}}/20}{\sqrt{2}} \right) = 0.5\text{erfc}\left(\frac{10^{\gamma_{\text{dB}}}/20}{\sqrt{2}} \right) \]

\[ P_{e, \text{BOrthogonal}} = \Phi\left( -\sqrt{\gamma} \right) = Q\left(\sqrt{\gamma}\right) = Q\left(\frac{10^{\gamma_{\text{dB}}}/20}{\sqrt{2}} \right) = 0.5\text{erfc}\left(\frac{10^{\gamma_{\text{dB}}}/20}{\sqrt{2}} \right) \]
5.2.2 Probability of error for M-ary orthogonal signals

We now turn to M-ary signals.

**Probability of error under ML detection and equal prior for M-ary equal-power orthogonal signals**

**ML detector**

\[ d_{ML}(\vec{r}) = \arg \max_{1 \leq m \leq M} \left( \langle \vec{r}, \vec{s}_m \rangle - \frac{\|\vec{s}_m\|^2}{2} \right) \]

\[ = \arg \max_{1 \leq m \leq M} \langle \vec{r}, \vec{s}_m \rangle \]

\[ = \arg \max_{1 \leq m \leq M} C^*(\vec{r}, \vec{s}_m) \]

**M-ary equal-power orthogonal signals**

\[
\begin{bmatrix}
\vec{s}_1 \\
\vec{s}_2 \\
\vdots \\
\vec{s}_M
\end{bmatrix} =
\begin{bmatrix}
\sqrt{E} & 0 & 0 & \cdots & 0 & 0 \\
0 & \sqrt{E} & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & \sqrt{E}
\end{bmatrix}
\]

---

5.2.2 Probability of error for M-ary orthogonal signals

**Outputs of M correlators**

\( \vec{r} = [\sqrt{E} + n_1, n_2, \ldots, n_M], \) if \( s_i(t) \) is transmitted.

\[ C^*(\vec{r}, \vec{s}_1) = [\sqrt{E} + n_1, n_2, \ldots, n_M] \cdot [\sqrt{E}, 0, \ldots, 0] = \sqrt{E}(\sqrt{E} + n_1) = \sqrt{E}r_1 \]

\[ C^*(\vec{r}, \vec{s}_2) = [\sqrt{E} + n_1, n_2, \ldots, n_M] \cdot [0, \sqrt{E}, \ldots, 0] = \sqrt{E}n_2 = \sqrt{E}r_2 \]

\[ \vdots \]

\[ C^*(\vec{r}, \vec{s}_M) = [\sqrt{E} + n_1, n_2, \ldots, n_M] \cdot [0, 0, \ldots, \sqrt{E}] = \sqrt{E}n_M = \sqrt{E}r_M \]

\[ d_{ML}(\vec{r}) = \arg \max_{1 \leq m \leq M} C^*(\vec{r}, \vec{s}_m) \]

The decision is correct only when

\[ \sqrt{E}r_i \geq \max_{2 \leq m \leq M} \sqrt{E}r_m, \]

or equivalently, \( r_i \geq \max_{2 \leq m \leq M} r_m \)
5.2.2 Probability of error for M-ary orthogonal signals

- It is sometimes convenient to first derive the probability that the detector makes a correct decision.

\[
\Pr(\text{correct} \mid \tilde{x}_n) = \Pr \left( R_n > \max_{m=1,\ldots,M} R_m \right) = \Pr \left( R_n > R_1 \cap \cdots \cap R_n > R_M \right) \\
= \prod_{i=1}^{M} \Pr(R_i > R_n) \Pr(R_i = r_i | R_i = \tilde{x}_n) \Pr(R_i = r_i) \\
= \int_{-\infty}^{\infty} \left[ \Pr(R_i < \tilde{x}_n | R_i = r_i) \right]^{M-1} \Phi^{-1} \left( \frac{r_i}{\sqrt{N_0/2}} \right) e^{-\frac{r_i^2}{2N_0}} \, dr_i \\
= \int_{-\infty}^{\infty} \Phi^{-1} \left( \frac{r_i}{\sqrt{N_0/2}} \right) e^{-\frac{r_i^2}{2N_0}} \, dr_i \\
\]

where \( N_{(\mu,\sigma^2)} \) is the Gaussian pdf with mean \( \mu \) and variance \( \sigma^2 \).

\[
\Pr(\text{correct}) = \sum_{m=1}^{M} \Pr(\text{correct} \mid \tilde{x}_n) = \Pr(\text{correct} \mid \tilde{x}_n)
\]

5-111

--- Po-Ning Chen ---
5.2.2 Probability of error for M-ary orthogonal signals

- Probability of error (POE) versus SNR(-per-bit)
  - In previous section, we establish the relation between POE and SNR for binary antipodal signals (PAM) and binary orthogonal signals:
    \[
    \begin{align*}
    P_e,\text{antipodal} & = \Phi\left(-\sqrt{2\gamma_b}\right) = \mathcal{Q}\left(\sqrt{2\gamma_b}\right) = \mathcal{Q}\left(10^{\gamma_{\text{dB}}/20}\sqrt{2}\right) = 0.5\text{erfc}\left(10^{\gamma_{\text{dB}}/20}/\sqrt{2}\right) \\
    P_e,\text{orthogonal} & = \Phi\left(-\sqrt{2\gamma_b}\right) = \mathcal{Q}\left(\sqrt{2\gamma_b}\right) = \mathcal{Q}\left(10^{\gamma_{\text{dB}}/20}\sqrt{2}\right) = 0.5\text{erfc}\left(10^{\gamma_{\text{dB}}/20}/\sqrt{2}\right)
    \end{align*}
    \]

- A fair comparison must be made based on the POE-per-bit and SNR-per-bit.

5.2.2 Probability of error for M-ary orthogonal signals

- SNR-per-bit for M-ary orthogonal signals
  
  For \( M \) channel symbols, \( E(=E_{\text{symbol}} = kE_{\text{bit}}) = kE_s \), if \( k = \log_2 M \).
  
  Noise power is still \( N_o \).
  
  \[
  P_{\text{MFSK}} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[1 - \Phi^{M-1}(x)\right] e^{-i\sqrt{2N_o}y^2/2} \, dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[1 - \Phi^{M-1}(x)\right] e^{-i\sqrt{2N_o}y^2/2} \, dx
  \]

- POE-per-bit for M-ary orthogonal signals
  
  When transmitting one out of \( 2^M \) symbols,
  
  \[
  \begin{align*}
  \text{Correct, if decision = the transmitted symbol;} \\
  \text{Error, if decision = the other (}2^M - 1\text{) symbols.} \\
  \text{1 symbol with probability } (1-P_{\text{MFSK}}) \\
  \text{(Assume equal contribution to the error probability} \\
  \text{based on the symmetry of orthogonal signals.)}
  \end{align*}
  \]

5-113
5.2.2 Probability of error for M-ary orthogonal signals

Example. $k = 2$.

- $00 \rightarrow 00$ with probability $(1 - P_{\text{MFSK}})$, no bit error = no symbol error
- $00 \rightarrow 01$ with probability $P_{\text{MFSK}}/(2^k - 1)$, one bit error = one symbol error
- $00 \rightarrow 10$ with probability $P_{\text{MFSK}}/(2^k - 1)$, one bit error = one symbol error
- $00 \rightarrow 11$ with probability $P_{\text{MFSK}}/(2^k - 1)$, two bit error = one symbol error

Hence, $E[\text{number of error bit}] = \frac{P_{\text{MFSK}}}{2^k - 1} \sum_{i=1}^{k} \binom{k}{i} = k \frac{2^{k-1}}{2^k - 1} P_{\text{MFSK}}$.

Also it usually assumes that $E[\text{number of error bit}] = k \times P_{\text{MFSK}}$.

\[ P_{\text{MFSK}} = \frac{2^{k-1}}{2^k - 1} \approx \frac{1}{2} P_{\text{MFSK}}, \]

\[ P_{\text{MFSK}} = \frac{1}{\sqrt{2\pi}} \frac{2^{k-1}}{2^k - 1} \int_{-\infty}^{\infty} \left[ 1 - \Phi^{-1}(x) \right] e^{-i(2\pi f_{\text{MFSK}}) t} dt \]

\[ = \left( \frac{2^{k-1}}{2^k - 1} \right) \left( \frac{1}{\sqrt{2\pi}} \right) \left( \frac{1}{2} P_{\text{MFSK}} \right) \]

\[ \approx \frac{1}{2} P_{\text{MFSK}} \]

5.2.2 Probability of error for M-ary orthogonal signals

- **Concluding remarks**
  - The larger the $M$ is, the better the system performance.
  - For example, to achieve bit-POE = $10^{-5}$, one needs $\gamma_b = 12$ dB for $M = 2$; but only requires $\gamma_b = 6$ dB for $M = 64$, a 6 dB save in transmission power !!!
5.2.2 Probability of error for M-ary orthogonal signals

- Since bit-POE is decreasing with respect to $M$, one may question that whether $\lim_{M \to \infty} P_{eb,MFSK} = 0$?
  - Answer: Not necessary.

- **Union Bound** on the probability of error
  - The previous formula is too complicated to evaluate its ultimate behavior with respect to $M$. We therefore develop the (famous) **Union bound**.

---

Recall that $\Pr(\text{correct } | \bar{x}) = \Pr(R_i > \max_{i=2,..,M} R_{\bar{x}} | \bar{x})$

\[
P_{e,MFSK} = \Pr(\text{error } | \bar{x}) \\
= \Pr(R_i > \max_{i=2,..,M} R_{\bar{x}} | \bar{x}) \\
= \Pr(R_i \leq R_1) \lor \cdots \lor \Pr(R_i \leq R_\bar{x} | \bar{x}) \\
= \int_{-\infty}^{\infty} \Pr(R_i \leq R_1) \cdots \Pr(R_i \leq R_\bar{x} | \bar{x}) \Pr(R_i = r_i | \bar{x}) dr_i \\
\leq \int_{-\infty}^{\infty} (M-1) \Pr(R_i \geq r_i | \bar{x}), \Pr(R_i = r_i | \bar{x}) dr_i \\
= (M-1) \Pr(R_i \geq r_i | \bar{x}), \Pr(R_i = r_i | \bar{x}) dr_i \\
= (M-1) P_{eb,MFSK}
\]

For equal prior, $P_{e,BFSK} = \Phi \left( -\frac{E}{\sqrt{N_0}} \right)$
5.2.2 Probability of error for M-ary orthogonal signals

A famous approximation formula for Gaussian cdf

\[ \Phi(-u) = \frac{1}{\sqrt{2\pi u}} e^{-u^2/2} \left(1 - \frac{1}{u^2} + \frac{1}{u^4} - \frac{1}{u^6} + \ldots\right) \]

where the approximation error is less than the last term used.

\[ \Rightarrow (\forall u \geq 0) \quad \Phi(-u) \leq \frac{1}{\sqrt{2\pi u}} e^{-u^2/2}. \]

\[ \Rightarrow (\forall u \geq 0) \quad \Phi(-u) \geq \frac{1}{\sqrt{2\pi u}} \left(1 - \frac{1}{u^2}\right) e^{-u^2/2}. \]

5.2.2 Probability of error for M-ary orthogonal signals

\[ P_{e,\text{MAX}} \leq (M - 1) P_{e,\text{MAX}} \]
\[ = (M - 1) \Phi\left(-\sqrt{\frac{E}{N_0}}\right) \]
\[ \leq (M - 1) \frac{1}{\sqrt{2\pi \frac{E}{N_0}}} \exp\left(-\frac{E}{2N_0}\right) \]
\[ = (2^t - 1) \frac{1}{\sqrt{2\pi \frac{E_0}{N_0}}} \exp\left(-\frac{E_0}{2N_0}\right) \]
\[ \leq 2^t \quad \frac{1}{\sqrt{2\pi \gamma_0}} \exp\left[-k\gamma_0 / 2\right] \]
\[ = \frac{1}{\sqrt{2\pi \gamma_0}} \exp\left[k \log(2)\right] \exp\left[-k\gamma_0 / 2\right] \]
\[ = \frac{1}{\sqrt{2\pi \gamma_0}} \exp\left[-k\gamma_0 - 2\log(2) / 2\right] \quad \text{(Union bound)} \]
### 5.2.2 Probability of error for M-ary orthogonal signals

Therefore, if \( \gamma_s > 2 \ln(2) = 1.42 \text{dB} \), then

\[
\lim_{M \to \infty} P_{e,M} \leq \lim_{k \to \infty} \frac{1}{\sqrt{2\pi\gamma_s}} e^{-\frac{1}{2}k^2\gamma_s} = 0
\]

\[
\lim_{M \to \infty} P_{e,M} = \lim_{M \to \infty} \frac{M/2}{M-1} P_{e,M-1} \leq \lim_{k \to \infty} \frac{2^{k-1}}{2^{k-1} - 1} \frac{1}{\sqrt{2\pi\gamma_s}} e^{-\frac{1}{2}k^2\gamma_s} = 0
\]

Error approaches zero as \( M \) grows.
5.2.2 Probability of error for M-ary orthogonal signals

- What if $\gamma_k \leq 2 \log(2) = 1.42\, \text{dB}$?
  - Answer:
    $\gamma_k \leq 1.42\, \text{dB}$ does not guarantee $\lim_{M \to \infty} P_{e,MFSK} > 0$?
  - In other words, 1.42 dB is not a tight bound!!
- In fact, it can be shown that
  $P_{e,MFSK} \leq e^{-(\gamma_k - \sqrt{\log(2)})}$.

Therefore if $\gamma_k > \log(2) = -1.6\, \text{dB}$, then bit-POE approaches zero as $M$ tends to infinity.

\[ -1.6\, \text{dB} \text{ is called the Shannon limit for M-ary orthogonal signals under AWGN channel. This is the minimum bit-SNR to achieve arbitrarily small bit-POE for M-ary orthogonal signals.} \]

7.1.2 Channel capacity (with orthogonal signals for AWGN channels)

\[ \begin{bmatrix} M \\ 1/2T \quad 1/2T \quad 1/2T \quad 1/2T \end{bmatrix} \quad \Rightarrow \quad W = \frac{M}{2T} \quad \text{or equivalently} \quad \frac{M}{2} = WT. \]

Shannon(1948) proved the upper limit of bits per symbol for arbitrary small bit-POE is:

\[ I(X^n;Y^n) \text{ bits/symbol} = M \times I(X;Y) \]
\[ = M \times \frac{1}{2} \log_2 \left( 1 + \frac{P_v}{N_0} \right) \]
\[ = WT \log_2 \left( 1 + \frac{P_v}{WN_0} \right) \]
7.1.2 Channel capacity (with orthogonal signals for AWGN channels)

Shannon proved that:
if \( R > C \),
then the bit-POE of any transmission scheme is bounded away from zero;
if \( R < C \),
then for any given \( \varepsilon > 0 \), there exists one transmission scheme (such as by letting \( M \) large) whose bit-POE < \( \varepsilon \).

\[
C = \frac{I(X^M;Y^M)}{T} \text{bits/symbol} = W \log_2 \left( 1 + \frac{P_{av}}{W N_0} \right) \text{bits/second.}
\]

Assume that the transmission rate is \( R \) bits/second.

\[\Rightarrow P_{av} \text{ Joule/second} = R \text{ bits/second} \times E_b \text{ Joule/bit}\]

\[
R > C = W \log_2 \left( 1 + \frac{P_{av}}{W N_0} \right) = W \log_2 \left( 1 + \frac{R E_b}{W N_0} \right) = W \log_2 \left( 1 + \frac{R}{W} \gamma_s \right)
\]

\[
\Rightarrow \frac{R}{W} > \log_2 \left( 1 + \frac{R}{W} \gamma_s \right)
\]

\[
\Rightarrow \gamma_s < 2^{\frac{W}{R}} - 1 \text{ implies that bit-POE is bounded away from zero.}
\]
7.1.2 Channel capacity (with orthogonal signals for AWGN channels)

Suppose $T$ seconds/symbol is fixed. Observe that $R$ bits/sec $\times T$ seconds/symbol $= k$ bits/symbol and $M = 2^k$. Also, recall that $W = \frac{M}{2T}$, so $R = \frac{1}{W} R = \frac{2T}{M} R = \frac{2k}{2^k}$.

Then $r_c < \frac{2^{k/2^k} - 1}{R/W} = r_c < \frac{2^{k/2^k} - 1}{2k/2^k}$.
7.1.2 Channel capacity (with orthogonal signals for AWGN channels)

- Conclusion for the above discussion.

\[ \text{If } \gamma_b < \log(2), \text{ then } (\exists \text{ an absolute constant } b > 0) \text{ bit - POE} > b \text{ for every } k. \]

- Now, we need to show that if \( \gamma_b > \log(2) \), then we can make bit-POE arbitrarily small by taking \( k \) large enough, i.e.,

\[ \text{If } \gamma_b > \log(2), \text{ then } (\forall \varepsilon > 0) (\exists k_b) \text{ bit - POE} < \varepsilon \text{ for } k > k_b. \]

7.1.3 Achieving channel capacity with orthogonal signals

- The Union Bound is loose. This can be seen from that

\[
P_{\text{err}} \leq (M-1)P_{\text{err}}
= (M-1)\phi\left(-\sqrt{E/N_o}\right)
\leq (M-1)\frac{1}{\sqrt{2\pi E/N_o}} \exp\left[-\frac{E}{(2N_o)}\right]
\]

does not provide a meaningful bound when

\[
(M-1)\frac{1}{\sqrt{2\pi E/N_o}} \exp\left[-\frac{E}{(2N_o)}\right] > 1, \text{ or equivalently, } (E/N_o) \text{ is small.} \]
7.1.3 Achieving channel capacity with orthogonal signals

Recall that we have the close form for $P_{e,MFSK}$.

$$P_{e,MFSK} = 1 - \text{Pr(correct)}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[ 1 - \Phi^{-1}(x) \right] e^{-\frac{x^2}{2}} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[ 1 - \Phi^{-1}(x) \right] e^{-\frac{x^2}{2}} dx + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[ 1 - \Phi^{-1}(x) \right] e^{-\frac{x^2}{2}} dx$$

$$\leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} [(M-1)\Phi(-x)] e^{-\frac{x^2}{2}} dx, \text{ for } \gamma_0 > 0.$$  

Since $(\forall x) \left[ 1 - \Phi^{-1}(x) \right] \leq 1$ and $(\forall x \geq 0) \left[ 1 - \Phi^{-1}(x) \right] \leq (M-1)\Phi(-x)$,

$$\leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx + \frac{M}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx$$

$$\langle y_0 \geq \frac{1}{\sqrt{2\pi}} \rangle$$


7.1.3 Achieving channel capacity with orthogonal signals

$$P_{e,MFSK} \leq \min_{y_0 \geq \frac{1}{\sqrt{2\pi}}} \left[ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx + \frac{2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} e^{-\frac{y_0^2}{2}} dx \right]$$

$$\frac{\partial}{\partial y_0} \left( \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx + \frac{2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} e^{-\frac{y_0^2}{2}} dx \right)$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{y_0^2}{2}} - \frac{2}{\sqrt{2\pi}} e^{-\frac{y_0^2}{2}} = 0$$

$$\Rightarrow y_0^* = 2k \log(2) > \frac{1}{\sqrt{2\pi}} \text{ if } k \geq \frac{1}{4\pi \log(2)} = 0.114806.$$  

$$y_0^* = \sqrt{2k \log(2)}$$
7.1.3 Achieving channel capacity with orthogonal signals

\[ P_{e,\text{MFSK}} \leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\gamma_s} e^{-x^2/2} dx + \frac{2^e}{\sqrt{2\pi}} \int_{\gamma_s}^{\infty} e^{-x^2/2} dx \]

\[ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\gamma_s} e^{-x^2/2} dx + \frac{2^e}{\sqrt{2\pi}} \int_{\gamma_s}^{\infty} e^{-x^2/2} dx \]

\[ < \Phi\left(\frac{\gamma_s - \sqrt{2\log(2)}}{\sqrt{0.5}}\right) + 2^e e^{-\gamma_s^2/2} \Phi\left(\frac{\gamma_s - \sqrt{4\log(2)}}{\sqrt{0.5}}\right) \]

when \( \log(2) < \gamma_s < 4 \log(2) \).

\[ P_{e,\text{MFSK}} \leq \frac{1}{\sqrt{2\pi}} \left[ 2\pi \left( \gamma_s - \sqrt{2\log(2)} \right) \right] e^{-\gamma_s^2/2 + \epsilon^{-\left(\gamma_s - \sqrt{2\log(2)}\right)^2}} \]

\[ + \frac{2^e}{\sqrt{2\pi}} \left[ 2\pi \left( 4\log(2) - \gamma_s \right) \right] e^{-4\log(2)^2/2 + \epsilon^{-\left(4\log(2) - \gamma_s\right)^2}} \]

when \( 4 \log(2) \leq \gamma_s \).

\[ P_{e,\text{MFSK}} \leq \frac{1}{\sqrt{2\pi}} \left[ 2\pi \left( \gamma_s - \sqrt{2\log(2)} \right) \right] e^{-\gamma_s^2/2 + \epsilon^{-\left(\gamma_s - \sqrt{2\log(2)}\right)^2}} \]

when \( \log(2) < \gamma_s < 4 \log(2) \).

\[ P_{e,\text{MFSK}} \leq \frac{2^e}{\sqrt{2\pi}} \left[ 2\pi \left( 4\log(2) - \gamma_s \right) \right] e^{-4\log(2)^2/2 + \epsilon^{-\left(4\log(2) - \gamma_s\right)^2}} \]

when \( 4 \log(2) \leq \gamma_s \).

\[ \Rightarrow \lim_{\gamma_s \to \log(2)} P_{e,\text{MFSK}} = 0. \]
5.2.3 Probability of error for M-ary biorthogonal signals

- Summary for the POE study
  - Binary bi-orthogonal performs better than binary orthogonal in bit-POE. (The latter requires twice bit-SNR to achieve the same bit-POE.)
  - M-ary orthogonal performs better with larger M.
  - There is a Shannon-Limit for minimum bit-SNR for M-ary orthogonal to achieve arbitrarily small bit-POE.

- How about M-ary bi-orthogonal? How good can it perform over M-ary orthogonal?

5-135

5.2.3 Probability of error for M-ary biorthogonal signals

- M-ary bi-orthogonal signals
  \[ \tilde{r} = [\sqrt{E} + n_1, n_2, ..., n_{M/2}] \] if \( s(t) \) is transmitted.

  \[ C^r(\tilde{r}, \tilde{s}_{M/2}) = [\sqrt{E} + n_1, n_2, ..., n_{M/2}] \cdot [0, 0, ..., -\sqrt{E}] = -\sqrt{E}n_{M/2} = -\sqrt{E}r_{M/2} \]

  \[ \vdots \]

  \[ C^r(\tilde{r}, \tilde{s}_2) = [\sqrt{E} + n_1, n_2, ..., n_{M/2}] \cdot [0, -\sqrt{E}, ..., 0] = -\sqrt{E}n_2 = -\sqrt{E}r_2 \]

  \[ C^r(\tilde{r}, \tilde{s}_4) = [\sqrt{E} + n_1, n_2, ..., n_{M/2}] \cdot [-\sqrt{E}, 0, ..., 0] = -\sqrt{E}(\sqrt{E} + n_1) = -\sqrt{E}r_1 \]

  \[ C^r(\tilde{r}, \tilde{s}_8) = [\sqrt{E} + n_1, n_2, ..., n_{M/2}] \cdot [\sqrt{E}, 0, ..., 0] = \sqrt{E}(\sqrt{E} + n_1) = \sqrt{E}r_1 \]

  \[ C^r(\tilde{r}, \tilde{s}_{16}) = [\sqrt{E} + n_1, n_2, ..., n_{M/2}] \cdot [0, \sqrt{E}, ..., 0] = \sqrt{E}n_2 = \sqrt{E}r_2 \]

  \[ \vdots \]

  \[ C^r(\tilde{r}, \tilde{s}_{M/2}) = [\sqrt{E} + n_1, n_2, ..., n_{M/2}] \cdot [0, 0, ..., \sqrt{E}] = \sqrt{E}n_{M/2} = \sqrt{E}r_{M/2} \]

5-136
5.2.3 Probability of error for M-ary biorthogonal signals

To minimize error probability under ML detection and equal prior for M-ary equal-power bi-orthogonal signals

Hence, $d_{ML}(\tilde{r}) = \arg \max_{m \in M} C'(\tilde{r}, \tilde{s}_m)$

Since $C'(\tilde{r}, \tilde{s}_m) = -C'(\tilde{r}, \tilde{s}_m)$,

$u = \arg \max_{|r| \leq \sqrt{E}} |C'(\tilde{r}, \tilde{s}_m)|$

$= \arg \max_{|r| \leq \sqrt{E}} \sqrt{E} r_m$

$= \arg \max_{|r| \leq \sqrt{E}} |r_m|$

$d_{ML}(\tilde{r}) = \begin{cases} u, & \text{if } r_u \geq 0, \\ -u, & \text{if } r_u < 0. \end{cases}$

The decision is correct for transmitting $\tilde{s}_i$ when $\tilde{r}_i \geq 0$ and $\tilde{r}_i \geq \max_{|r| \leq \sqrt{E}} |r_m|$.

5.2.3 Probability of error for M-ary biorthogonal signals

$$P(\text{correct} | \tilde{s}_i) = P\left( \left\{ \tilde{r}_{i_1}, \ldots, \tilde{r}_{i_M} : \tilde{r}_{i_j} \geq 0 \text{ and } \tilde{r}_{i} > \max_{|r| \leq \sqrt{E}} |r_m| \right\} | \tilde{s}_i \right)$$

$$= \int \prod \left[ P(\tilde{r}_{i_1}, \ldots, \tilde{r}_{i_M} : \tilde{r}_{i_j} \geq \tilde{r}_j \text{ and } \tilde{r}_{i} > \tilde{r}_m) \right] \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{(x-\tilde{r}_m)^2}{2\sigma^2}} dx$$

since given $\tilde{r}_i$ and $\tilde{s}_i$, $\tilde{r}_i$, $r_m$, are independent;

$$= \int \prod \left[ N(\tilde{r}_i, \sigma^2) \right] dx$$

$$= \int \prod \left[ \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{(x-\tilde{r}_i)^2}{2\sigma^2}} dx \right]$$

$$= \int \left[ 1 - 2\Phi \left( \frac{-x}{\sqrt{\sigma^2}} \right) \right] dx$$

$$= 1 - \frac{1}{M} \sum_{m=1}^{M} \Pr(\text{correct} | \tilde{s}_m)$$

where $\gamma = E/N_0.$
5.2.3 Probability of error for M-ary biorthogonal signals

- Re-formulating the formula using bit-SNR instead of symbol-SNR, we obtain:
  - Larger M, better performance except M = 2 and M = 4.
  - Note that symbol-POE comparison does not really tell the winner in performance.
  - Symbol-POE BPSK (M = 2) < symbol-POE QPSK (M = 4), but bit-POE BPSK (M = 2) = bit-POE QPSK (M = 4),
  - The Shannon-limit remains the same.

5.2.4 Probability of error for simplex signals

- Simplex signals from orthogonal signals

  Given the vector representations of (orthogonal and equal - power) channel symbols
  \[ s_m = [a_{m1}, a_{m2}, ..., a_{mk}] \] for \( m = 1, ..., M \), its center is

  \[ \bar{s} = \left[ \frac{1}{M} \sum_{m=1}^{M} a_{m1}, \frac{1}{M} \sum_{m=1}^{M} a_{m2}, ..., \frac{1}{M} \sum_{m=1}^{M} a_{mk} \right] \]

  Define new channel symbols as \( s'_m = s_m - \bar{s} \). Then \( \{s'_1, s'_2, ..., s'_M\} \) is called the simplex signal.
5.2.4 Probability of error for simplex signals

- Transmitted energy of simplex signals is reduced.

\[
E' = \left(1 - \frac{1}{M}\right)E
\]

\[
\Rightarrow (10\log_{10} E)\ dB = 10\log_{10} \left(\frac{M-1}{M}\right)\ dB + (10\log_{10} E)\ dB
\]

save \[10\log_{10} \left(\frac{M-1}{M}\right)\ dB. \quad (\text{\textasciitilde 3 dB for } M = 2)
\]

- Since the inter-symbol distances (as well as their relative positions) remain the same, the POE of the simplex signal is the same as M-ary orthogonal signals.

5.2.5 Probability of error for M-ary binary-coded signals

□ Example. Multi-dimensional BPSK.

Channel symbols: \( s_m = \{s_m, s_{m+1}, \ldots, s_{M}\}, 1 \leq m \leq M. \)

where \( s_{my} = \pm \sqrt{E/N_0} \)

(Here, we equivalently assume that the noise variance is one.)

- Possibly \( M < 2^K \).

\[
C'(\vec{r}, \vec{s}_m) = [r_1, r_2, \ldots, r_K] \cdot [\pm \sqrt{E/N_0}, \pm \sqrt{E/N_0}, \ldots, \pm \sqrt{E/N_0}]
\]

\[
= \sqrt{E/N_0} (\pm r_1 \pm r_2 \cdots \pm r_K)
\]

\[
d_M(\vec{r}) = \arg \max_{(a_1, a_2, \ldots, a_K) \in \mathbb{Z}^K} \left\{ -\sqrt{E/N_0}, \sqrt{E/N_0}\right\} (a_1 r_1 + a_2 r_2 + \cdots + a_K r_K) \]
5.2.5 Probability of error for M-ary binary-coded signals

\[ P_{d_{	ext{min}}(w|\bar{x})} = P\left(N: s_{1}t_{1} + \cdots + s_{M}t_{M} \leq \max_{(a_{1}, \ldots, a_{M})} a_{1}t_{1} + \cdots + a_{M}t_{M} \right) \]

\[ = \int_{N} P\left(\sum_{i=1}^{M} s_{i}t_{i} \leq a_{1}t_{1} + \cdots + a_{M}t_{M} \right) d\bar{x} \]

\[ = \int_{N} \sum_{(a_{1}, \ldots, a_{M})} P\left(\sum_{i=1}^{M} s_{i}t_{i} \leq a_{1}t_{1} + \cdots + a_{M}t_{M} \right) d\bar{x} \]

\[ \leq \int_{N} \sum_{(a_{1}, \ldots, a_{M})} P\left(\sum_{i=1}^{M} s_{i}t_{i} \leq a_{1}t_{1} + \cdots + a_{M}t_{M} \right) d\bar{x} \]

\[ \leq \int_{N} \left(\sum_{(a_{1}, \ldots, a_{M})} P\left(\sum_{i=1}^{M} s_{i}t_{i} \leq a_{1}t_{1} + \cdots + a_{M}t_{M} \right) d\bar{x} \right) \]

\[ \leq M \int_{N} P\left(\sum_{i=1}^{M} s_{i}t_{i} \leq a_{1}t_{1} + \cdots + a_{M}t_{M} \right) d\bar{x} \]

where \((a_{1}, \ldots, a_{M})\) is the transmitted signal that has the minimum Euclidean distance to \(\bar{x}\).
The performance of M-ary binary-coded signals will depend on the minimum Euclidean distance among transmitted signals (codewords), which is determined at the design stage of the code book.

5.2.5 Probability of error for M-ary binary-coded signals

- Symbol-POE under equal prior
  - Vectorized channel symbols: \( s_m = \sqrt{E} A_m \)
  - where \( A_m = (2m - 1 - M)d \), \( m = 1, 2, ..., M \).
  - \( C(\vec{r}, \vec{s}_m) = \| r - s_m \|^2 \)
  - \( d_{\text{ML}}(r) = \arg \min_{1 \leq m \leq M} \| r - s_m \|^2 \)
  - \( \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_M \end{bmatrix} \)
  - Error happens when \( |r - s_m| > d\sqrt{E}/2 \) for \( 2 \leq m \leq M - 1 \) given \( s_m \).
  - \( r - s_1 > d\sqrt{E}/2 \) given \( s_1 \)
  - \( r - s_M < -d\sqrt{E}/2 \) given \( s_M \)
5.2.6 Probability of error for M-ary PAM

\[
P_{e,M\text{-PAM}} = \frac{1}{M} \left( \Pr \{ r - s_i > d \sqrt{E/2} | s_i \} + \sum_{m=1}^{M} \Pr \{ r - s_{m} > d \sqrt{E/2} | s_{m} \} + \Pr \{ r - s_{M} < -d \sqrt{E/2} | s_{M} \} \right)
\]

\[
= \frac{M-1}{M} \Pr \left( r - s_{M} > d \sqrt{E/2} \right) 
\]

\[
= \frac{M-1}{M} 2\Phi \left( \frac{-d \sqrt{E/2} - 0}{\sqrt{N_0/2}} \right) = \frac{2(M-1)}{M} \Phi \left( -d \sqrt{ \frac{E}{N_0} } \right)
\]

\[\square\] Average symbol energy of (unequal-energy) PAM signals

\[
E_m = \frac{E}{2} A_m^2 
\]

\[\Rightarrow E_{av} = \frac{1}{M} \sum_{m=1}^{M} E_m = \frac{1}{M} \frac{E}{2} d^2 \sum_{m=1}^{M} (2m - 1 - M)^2 = \frac{1}{6} d^2 E (M^2 - 1)
\]

The larger the M is, the worse the symbol performance!!!

Recall that for M-ary orthogonal signals, the larger the M is, the better the symbol performance (approaching Shannon Limit).
5.2.7 Probability of error for M-ary PSK

**M-ary PSK**

Channel symbol: \( s_m(t) = g(t) \cos(2\pi f t + \theta_m) \)

where \( \theta_m = 2\pi (m-1)/M \) and \( m = 1, 2, \ldots, M \).

**Example.** \( M = 4 \).

\[
\begin{align*}
  s_1(t) &= g(t) \cos(2\pi f t) \\
  s_2(t) &= g(t) \cos(2\pi f t + \pi/2) \\
  s_3(t) &= g(t) \cos(2\pi f t + \pi) \\
  s_4(t) &= g(t) \cos(2\pi f t + 3\pi/2)
\end{align*}
\]

**Vectorization of M-ary PSK**

\[
\begin{align*}
  u_1(t) &= \frac{g(t)}{g(t)} \sqrt{2} \cos(2\pi f t) \\
  u_2(t) &= -\frac{g(t)}{g(t)} \sqrt{2} \sin(2\pi f t)
\end{align*}
\]

\[\Rightarrow \bar{s}_m = \left[ \frac{g(t)}{\sqrt{2}} \cos(\theta_m), \frac{g(t)}{\sqrt{2}} \sin(\theta_m) \right] \]

5.2.7 Probability of error for M-ary PSK

**ML decision maker**

\[
d_{ML}(\bar{r}) = \arg \max_{1 \leq m \leq M} C(\bar{r}, \bar{s}_m)
\]

\[= \arg \max_{1 \leq m \leq M} \left( \langle \bar{r}, \bar{s}_m \rangle - \frac{1}{2} \| \bar{s}_m \|^2 \right)\]

\[= \arg \max_{1 \leq m \leq M} \langle \bar{r}, \bar{s}_m \rangle\]

\[\bar{s}_m = \left[ \frac{g(t)}{\sqrt{2}} \cos(\theta_m), \frac{g(t)}{\sqrt{2}} \sin(\theta_m) \right]\]

\[
d_{ML}(\bar{r}) = \arg \max_{1 \leq m \leq M} \left( r_1 \cos(\theta_m) + r_2 \sin(\theta_m) \right)
\]

\[= \arg \max_{1 \leq m \leq M} \left( \cos(\theta) \cos(\theta_m) + \sin(\theta) \sin(\theta_m) \right), \text{ where } \theta \text{ is equal to } \tan^{-1} \left( \frac{r_2}{r_1} \right)\]

\[= \arg \max_{1 \leq m \leq M} \cos(\theta - \theta_m)\]
5.2.7 Probability of error for M-ary PSK

Suppose \( \theta_m = 0 \) is transmitted.

\[
\Rightarrow \hat{\mathbf{s}}_m = \left[ \sqrt{E}, 0 \right]
\]

\[
\Rightarrow \hat{\mathbf{r}} = [r_1, r_2] = \left[ \sqrt{E + n_1, n_2} \right]
\]

\[
p(\tau_1, \tau_2) = \frac{1}{\pi N_0} \exp \left[ - \frac{(\tau_1 - \sqrt{E})^2 + r_2^2}{N_0} \right]
\]

\[
p(v, \theta) = \det \begin{bmatrix}
\frac{\partial \tau_1}{\partial v} & \frac{\partial \tau_2}{\partial v} \\
\frac{\partial \tau_1}{\partial \theta} & \frac{\partial \tau_2}{\partial \theta}
\end{bmatrix} \times \frac{1}{\pi N_0} \exp \left[ - \frac{v^2 - 2 \sqrt{E} \cos(\theta) + E}{N_0} \right]
\]

\[
= \frac{v}{\pi N_0} \exp \left[ - \frac{v^2 - 2 \sqrt{E} \cos(\theta) + E}{N_0} \right]
\]

where \( \tau_1 = v \cos \theta \) and \( \tau_2 = v \sin \theta \).

\[\text{E} = \frac{||\mathbf{g}(t)||^2}{2}\]

\[\text{5-151}\]

5.2.7 Probability of error for M-ary PSK

\[
p(\theta) = \int p(v, \theta) dv
\]

\[
= \frac{1}{\pi N_0} \int_0^{\pi} \int_0^\infty \frac{v^2 - 2 \sqrt{E} \cos(\theta) + E}{N_0} \cdot \frac{1}{\pi N_0} \exp \left[ - \frac{v^2 - 2 \sqrt{E} \cos(\theta) + E}{N_0} \right] dv
\]

\[
= \frac{1}{\pi N_0} \int_0^{\pi} \int_0^\infty \frac{v}{\pi N_0} \exp \left[ - \frac{v^2 - 2 \sqrt{E} \cos(\theta) + E}{N_0} \right] dv
\]

\[
= \frac{1}{2\pi} e^{-\gamma \sin^2(\theta)} \int_0^{\gamma} \int_0^\infty \frac{v^2 - 2 \sqrt{E} \cos(\theta) + E}{N_0} \cdot \frac{1}{\pi N_0} \exp \left[ - \frac{v^2 - 2 \sqrt{E} \cos(\theta) + E}{N_0} \right] dv
\]

by letting \( \tilde{v} = v/\sqrt{N_0}/2 \) and \( \gamma = E/N_0 \).

\[\text{5-152}\]
5.2.7 Probability of error for M-ary PSK

\[ p(\theta) = \frac{1}{2\pi} e^{-\frac{\pi^2}{2} \theta^2} - \int_0^{\frac{\pi}{2}} \frac{1}{\sqrt{\pi}} e^{-\frac{\pi^2}{2} \theta^2} d\theta \]

- Observations.
  - The larger the SNR is, the narrower the pdf.
  - \( P_{error,PSK} = P(error | \theta_m = 0) = 1 - \int_{-\pi/2}^{\pi/2} p(\theta)d\theta \).
  - The larger the SNR is, the smaller the symbol-POE.

- Similar to M-ary PAM, the difference between \( M = 4 \) and \( M = 8 \) is around 4 dB. (See slide 5-148.)
- The same as M-ary PAM, doubling \( M \) requires an additional 6 dB/bit for large \( M \).
- Exception: Difference between \( M = 2 \) and \( M = 4 \).
5.2.7 Probability of error for M-ary PSK

- In general, the POE of M-ary PSK does not reduce to a simple form, except for $M = 2$ (binary antipodal) and $M = 4$ (whose POE is the same as $\pi/4$-QPSK since they have the same signal space topology.)

$$ P_{\text{err,PSK}} = \Phi\left( -\frac{2E_b}{N_r} \right) $$

Assume that no interference between the signals on the two quadrature carriers.

$$ P_{\text{err,QPSK}} = 1 - \left[ 1 - P_{\text{err,PSK}} \right] = 1 - \left[ 1 - \Phi\left( -\frac{2E_b}{N_r} \right) \right]^2 $$

$$ = \Phi\left( -\frac{2E_b}{N_r} \right) \left[ 2 - \Phi\left( -\frac{2E_b}{N_r} \right) \right] $$

---

5.2.7 Probability of error for M-ary PSK

- We now have simple formulas for BPSK and QPSK.

$$ P_{e,\text{BPSK}} = \Phi\left(-\sqrt{2\gamma_s}\right) $$

$$ P_{e,QPSK} = \Phi\left(-\sqrt{2\gamma_s}\right) \left[ 2 - \Phi\left(-\sqrt{2\gamma_s}\right) \right] $$

- But, it seems that no simple formulas for M-ary PSK at $M > 4$.

- Question: “Can we obtain a simple approximate formula for higher dimensional PSK?”
5.2.7 Probability of error for M-ary PSK

Approximation of \( p(\theta) \) for \( |\theta| < \pi/M \) (and \( M > 4 \)).

\[
p(\theta) = \frac{1}{2\pi} e^{-\gamma u^2/2} \int_{-\infty}^{\infty} \overline{\left[ x + \sqrt{2} \gamma \cos(\theta) \right] e^{xt} dx},
\]
by letting \( x = \sqrt{2} \gamma \cos(\theta) \).

Thus,

\[
p(\theta) = \frac{1}{2\pi} e^{-\gamma u^2/2} \left[ x x e^{x^2/2} dx + \sqrt{2\gamma} \cos(\theta) \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \right]
\]

\[
= \frac{1}{2\pi} e^{-\gamma u^2/2} \left[ e^{-\gamma \cos(\theta)} + \sqrt{2\gamma} \cos(\theta) \left[ 1 - \frac{1}{\sqrt{2\gamma} \cos(\theta)} e^{-\gamma \cos(\theta)} \right] \right]
\]

\[
(\forall u \geq 0, \Phi(-u) < \frac{1}{\sqrt{2\pi} u} e^{-u^2/2})
\]

Here, "\( \infty \)" becomes "\( \pi \)"
when \( \sqrt{2\gamma} \cos(\pi/M) >> 1 \).

Thus,

\[
p(\theta) \approx \sqrt{\frac{\gamma}{\pi}} \cos(\theta) e^{-\gamma u^2/2}
\]
when \( \gamma \approx 1 \). (From \( \sqrt{2\gamma} \cos(\pi/M) >> 1 \) and fixed \( M > 4 \).)

Thus, \( P_{\text{MPSK}} = 1 - \int_{-\infty}^{\infty} p(\theta) d\theta \)

\[
\leq 1 - 2\int_{0}^{\gamma} \sqrt{\frac{\gamma}{\pi}} \cos(\theta) e^{-\gamma u^2/2} d\theta
\]

\[
= 1 - 2\int_{0}^{\gamma} \sqrt{\frac{\gamma}{\pi}} e^{-\gamma u^2/2} du, \text{ by letting } u = \sqrt{2\gamma} \sin(\theta)
\]

\[
= 2 \Phi\left(-\sqrt{2\gamma} \sin(\pi/M)\right)
\]

When \( \sqrt{2\gamma} \cos(\pi/M) >> 1 \), the above approximation is indeed accurate!
5.2.7 Probability of error for M-ary PSK

- **Exact formulas.**
  \[
  P_{e,\text{BPSK}} = \Phi\left(-\sqrt{2\gamma_b}\right) \\
  P_{e,\text{QPSK}} = \Phi\left(-\sqrt{2\gamma_b}\right) + 2\Phi\left(-\sqrt{2\gamma_b}\right) - \Phi\left(-\sqrt{2\gamma_b}\right) - \Phi\left(-\sqrt{2\gamma_b}\right)
  \]

- **Approximated formulas.**
  \[
  P_{e,\text{BPSK}} = 2\Phi\left(-\sqrt{2\gamma_b}\right) \\
  P_{e,\text{QPSK}} = 2\Phi\left(-\sqrt{2\gamma_b}\right) = \Phi\left(-\sqrt{4\gamma_b}\right) + \Phi\left(-\sqrt{4\gamma_b}\right) = 2\Phi\left(-\sqrt{2\gamma_b}\right)
  \]

When \(\sqrt{2\gamma \cos(\pi / M)} \gg 1\), the above approximation is indeed accurate! Hence for fixed \(M\), the larger SNR is, the more accurate the approximation is.

5.2.7 Probability of error for M-ary PSK

- **Another viewpoint on the relationship of bit-POE and symbol-POE (for M-ary PSK)**

  Recall that \(P_{e,\text{BPSK}} = \frac{M}{2(M-1)} P_{e,\text{QPSK}}\) (See slide 5-115)

  based on the assumption that the error probability is equally contributed by the other \((M-1)\) symbols, which is a reasonable assumption for QPSK signals.

  This assumption, however, may not be applicable to MPSK signals since the "symmetry" on signal error contribution is no longer valid.
These two symbols contribute more in error probability when 000 is transmitted.

\[ M = 2 \]

\[ M = 4 \]

\[ M = 8 \]

--- Po-Ning Chen ---
NCTU/CM

5.2.7 Probability of error for M-ary PSK

\[ P_{e,\text{MPSK}} = 1 - P_{e,\text{MPSK}} \]
\[ \approx 1 - \left( P_{e,b,\text{MPSK}} \right)^k \]
\[ = 1 - \left[ 1 - P_{e,b,\text{MPSK}} \right] \]
\[ \approx k \cdot P_{e,b,\text{MPSK}} \cdot (\text{By Taylor’s expansion}) \]
5.2.8 Differential PSK (DPSK) and its performance

- Phase ambiguity (due to frequency shift) of M-ary PSK.

\[
\begin{align*}
\text{ideal case:} & \quad \begin{cases} 
\text{receive } \cos(2\pi f_c t + \theta) \\
\text{estimate it in terms of } f_c 
\end{cases} \\
\text{ambiguous case:} & \quad \begin{cases} 
\text{receive } \cos(2\pi f_{\text{wrong}} t + \theta) \\
\text{estimate it in terms of } f_{\text{wrong}} 
\end{cases}
\Rightarrow \begin{cases} 
\text{receive } \cos(2\pi f_{\text{wrong}} t + [2\pi(f_c - \tilde{f}_c)t] + \theta) \\
\text{estimate it in terms of } \tilde{f}_c 
\end{cases} \\
\Rightarrow \hat{\theta} = [2\pi(f_c - \tilde{f}_c)t] + \theta.
\end{align*}
\]

5.2.8 Differential PSK (DPSK) and its performance

- Differential Encoding

\[
\begin{align*}
\text{DBPSK:} & \quad \begin{cases} 
\text{Shift the phase of previous symbol by 0 degree, if } Input = 0 \\
\text{Shift the phase of previous symbol by 180 degree, if } Input = 1 
\end{cases} \\
\text{DQPSK:} & \quad \begin{cases} 
\text{Shift the phase of previous symbol by 0 degree, if } Input = 00 \\
\text{Shift the phase of previous symbol by 90 degree, if } Input = 01 \\
\text{Shift the phase of previous symbol by 180 degree, if } Input = 11 \\
\text{Shift the phase of previous symbol by -90 degree, if } Input = 10 
\end{cases}
\end{align*}
\]
5.2.8 Differential PSK (DPSK) and its performance

- POE of DifferentiallyEncoded PSK
  - The current bit-error will affect the correctness of the decision for the next bit.

\[
\begin{align*}
\text{input} &= 010101010101... \\
\text{DE code} &= 011001100110... \\
&\downarrow \text{channel} \\
\text{receive code} &= 01101100110... \\
\text{output} &= 0101010101... \\
\end{align*}
\]

- One-bit error in the received codeword will induce two-bit error in decision.

\[
\begin{align*}
C_{k+1}^{\text{wrong}} &\rightarrow \\
I_k &= C_k^{\text{wrong}} \oplus C_{k-1}^{\text{wrong}} \\
I_{k+1} &= C_{k+1} \oplus C_k^{\text{wrong}}
\end{align*}
\]

- One-bit error in the received codeword will induce two-bit error in decision.
5.2.8 Differential PSK (DPSK) and its performance

- **Demodulation without phase estimation**
  At time $k$, the received vector is
  \[ \vec{r}_k = \sqrt{E} \cos(\theta_k - \phi) + n_{1k} + \sqrt{E} \sin(\theta_k - \phi) + n_{2k}, \]
  \[ \Rightarrow (r_{1k} + j r_{2k}) = \sqrt{E} e^{j(\theta_k - \phi)} + (n_{1k} + j n_{2k}). \]
  At time $k - 1$, the received vector is
  \[ \Rightarrow (r_{1k-1} + j r_{2k-1}) = \sqrt{E} e^{j(\theta_{k-1} - \phi)} + (n_{1k-1} + j n_{2k-1}). \]

\[ (r_{1k} + j r_{2k}) \cdot (r_{1k-1} + j r_{2k-1})^* \]
\[ = (\sqrt{E} e^{j(\theta_k - \phi)} + (n_{1k} + j n_{2k})) \cdot (\sqrt{E} e^{j(\theta_{k-1} - \phi)} + (n_{1k-1} + j n_{2k-1})) \]
\[ = \sqrt{E} \cos(\phi) + (n_{1k} + j n_{2k}) \cdot (\sqrt{E} \sin(\phi) + (n_{1k-1} + j n_{2k-1})) \]

(Note that the phase offset $\phi$ has been cancelled in the first term.)
\[ \Rightarrow E[(r_{1k} + j r_{2k}) \cdot (r_{1k-1} + j r_{2k-1})^*] = E[e^{j(\phi - \phi)}]. \]

- **POE of DPSK**
  - Hard to derive directly in its original form.
  - Instead, approximation is applied.

\[ (r_{1k} + j r_{2k}) \cdot (r_{1k-1} + j r_{2k-1})^* \]
\[ = \sqrt{E} \cos(\phi) + (n_{1k} + j n_{2k}) \cdot (\sqrt{E} \sin(\phi) + (n_{1k-1} + j n_{2k-1})) \]
\[ = \sqrt{E} \cos(\phi) + (\hat{a}_{k-1} + j \hat{a}_{k-1}) \cdot (\hat{a}_{1k} + j \hat{a}_{2k}) + \frac{1}{\sqrt{E}} (n_{1k} + j n_{2k}) (n_{1k-1} + j n_{2k-1}) \]

where \((\hat{a}_{k-1} + j \hat{a}_{k-1})\) has the same distribution as \((n_{1k} + j n_{2k})\), and
\((\hat{a}_{1k} + j \hat{a}_{2k})\) has the same distribution as \((n_{1k-1} + j n_{2k-1})\).

Problem 2-5.
Given $Y = e^{\phi} \hat{X}$ and $\hat{X}$ is Gaussian with zero-mean vector and diagonal covariance matrix with identical marginal variance. Then
\[ Y \] has the same distribution as $\hat{X}$. 

5.2.8 Differential PSK (DPSK) and its performance

If $2 \cdot \text{SNR} >> 1$, then

\[
\text{Var}[\zeta_{n+1} + \zeta_n] >> \text{Var}[n_{x,1} + n_{x,2} + n_{x,3}/\sqrt{E}]
\]

\[
\text{Var}[\zeta_{n+1} + \zeta_n] = N_0 >> \text{Var}[n_{x,1} + n_{x,2}+ n_{x,3}/\sqrt{E}] = N_0^2/(2E) = N_0/(2 \cdot \text{SNR})
\]

Therefore,

\[
(r_{x,1} + jn_{x,1}) - (r_{x,2} + jn_{x,2}) \frac{1}{\sqrt{E}}
\]

\[
= \sqrt{E} e^{j[\beta - \mu]} \left[ (\tilde{r}_{n+1} + j\tilde{n}_{n+1}) + (\tilde{r}_{n+2} + j\tilde{n}_{n+2}) + \frac{1}{\sqrt{E}} (n_{x,1} + jn_{x,2}) (n_{x,1} + jn_{x,2}) \right]
\]

\[
= \sqrt{E} e^{j[\beta - \mu]} \left[ (\tilde{r}_{n+1} + j\tilde{n}_{n+1}) + (\tilde{r}_{n+2} + j\tilde{n}_{n+2}) \right]
\]

5.2.8 Differential PSK (DPSK) and its performance

Let $(X,Y) = \left[ \sqrt{E} \cos(\theta_{x} - \theta_{n-1}) + n_{x,1,1} + \sqrt{E} \sin(\theta_{x} - \theta_{n-1}) - \hat{\eta}_{x,1,2} + \hat{n}_{x,2} \right]$

Then $(X,Y) = \text{Normal} \left[ \left[ \sqrt{E} \cos(\theta_{x} - \theta_{n-1}) \right] \left[ \begin{array}{c} N_0 \\ 0 \end{array} \right] \right]$

Our goal is to estimate:

\[
(\theta_{x} - \theta_{n-1}) = \tan^{-1} \frac{Y}{X}
\]
5.2.8 Differential PSK (DPSK) and its performance

Comparison of POE between PSK and DPSK

PSK: when $\theta = 0 \Rightarrow \hat{r} = [r_i, r_f] = \left[ \sqrt{E} + n_i, n_f \right]

\Rightarrow \hat{\theta} = \tan^{-1} \frac{n_f}{r_f} \quad \text{and} \quad [r_i, r_f] = \text{Normal} \left( \begin{bmatrix} \sqrt{E} & 0 \\ 0 & N_0/2 \end{bmatrix} \right)

DPSK: when $\theta_d = 0$

\Rightarrow \hat{\theta}_d = \tan^{-1} \frac{Y}{X} \quad \text{and} \quad (X, Y) = \text{Normal} \left( \begin{bmatrix} \sqrt{E} & 0 \\ 0 & N_0/2 \end{bmatrix} \right)

\text{Observe that the noise variance of DPSK is twice as large as that of PSK. Hence, the POE of DPSK is larger than that of PSK.}

\text{The above formula of DPSK performance is only an approximation. (Cf. Appendix B of the textbook.)}

--- Po-Ning Chen ---

\begin{align*}
\text{The POE degradation of DPSK to PSK is better than “double” or 3 dB.} \\
\text{When symbol-POE is less than } 10^{-5}, \quad \text{the difference in bit-SNR for binary cases is only 1 dB.}
\end{align*}
5.2.8 Differential PSK (DPSK) and its performance

The good performance of DBPSK is due to that \((\theta_k - \theta_{k-1})\) is either 0 or \(\pi\). Hence, we only need to consider the real part of

\[
(r_{k,1} + jr_{k,2}) \cdot (r_{k-1,1} + jr_{k-1,2})^* / \sqrt{E}
\]

i.e.,

\[
\text{Re}\{r_{k,1} + jr_{k,2}) \cdot (r_{k-1,1} + jr_{k-1,2})^* / \sqrt{E}\}
\]

Then, the optimal decision maker is:

\[
\text{Re}\{r_{k,1} + jr_{k,2}) \cdot (r_{k-1,1} + jr_{k-1,2})^* \} \geq 0. \quad (A \text{ special case of Appendix B})
\]

Appendix B

What is the probability of \(P_b = \Pr\{D < 0\}\), where

\[
D = \sum_{i=1}^{\infty} \left( A |X_i|^2 + B |Y_i|^2 + C X_i^* Y_i + C^* X_i Y_i \right),
\]

\(\{X_i, Y_i\}_{i=1}^{\infty}\) independent complex-valued Gaussian random variables with common covariance matrix \((X_i, Y_i)\) may be correlated, and may have different means for different \(k\), and \(A, B, C\) are complex-valued constants satisfying \(|C| - AB > 0\).

Denote \(m_{x,i} = E[X_i], m_{y,i} = E[Y_i]\),

\[
\mu_x = \frac{1}{2} E[|X_i - m_{x,i}|^2], \mu_y = \frac{1}{2} E[|Y_i - m_{y,i}|^2], \mu_{xy} = \frac{1}{2} E[(X_i - m_{x,i})(Y_i - m_{y,i})^*].
\]
Appendix B

\[ Q(a,b) - L_{0}(ab, \exp(-\frac{1}{2}(a^2 + b^2))) \cdot L_{1}(ab, \exp(-\frac{1}{2}(a^2 + b^2))/2) \cdot \frac{2L}{k} \cdot \frac{1}{\nu} \]
\[ P_\gamma = \begin{cases} \sum \sum & \text{if } L = 1 \end{cases} \]
\[ Q(a,b) - \frac{v_\gamma}{1 + v_\gamma} L_{1}(ab, \exp(-\frac{1}{2}(a^2 + b^2))) \]
\[ Q(a,b) - \frac{v_\gamma}{1 + v_\gamma} L_{0}(ab, \exp(-\frac{1}{2}(a^2 + b^2))) \]

where \( a_0 = 2(C^\dagger - DB) || m_{0} || m_{0} ^* + m_{0} ^* || m_{0} - m_{0} ^* || m_{0} ^* \), \( a_{n} = d(m_{n} ^\dagger + B) | m_{n} ^\dagger + Cm_{n} ^\dagger | m_{n} + Cm_{n} ^\dagger m_{n} ^\dagger \), \( Q(a,b) = e^{-2w^\dagger x} \sum \frac{x^{j}}{j!} J_{0}(ab), b > a > 0 \) Marcum Q-function

\[ f(x) = \sum \frac{e^{x^2}}{x^{2j+1}} \] is the \( j \)th order modified Bessel function of the first kind

Applying the Appendix B to DBPSK for \( (\theta_k - \theta_{k-1}) = \pi \):

\[ X_l = \sqrt{E} \exp(j\theta_l) + (n_{0} + jn_{0}^*) \]
\[ Y_l = \sqrt{E} \exp(j\theta_l) + (n_{0} + jn_{0}^*) \]
\[ L = 1, A = B = 0, C = 1, D = X_l \exp(j\theta_l) + Y_l \exp(j\theta_l) = 2 \text{Re}[X_l Y_l^\dagger] \]
\[ m_n = \sqrt{E} \exp(j\theta_l), m_{n}^* = \sqrt{E} \exp(j\theta_l^*), \mu_n = \mu_{n}^* = \frac{N_0}{2}, \mu_n = 0 \]
\[ a = 0, b = \frac{2E}{N_0}, \nu_1 = v_2 = \frac{1}{N_0} \]
\[ \implies a_t = 2EN_0, a_t = E \exp(j\theta_{0} - \theta_{1}) + \exp(j\theta_{1} - \theta_{0}) = 2E \] (Because error occurs when \( (\theta_l - \theta_{l-1}) = 0 \))
\[ Q(0, 2E / N_0) = 0 \lim_{x \to a} J_x(x) = 1 \]
\[ \implies P_a = \frac{1}{2} \exp\left(-\frac{E}{N_0}\right) \]
5.2.8 Differential PSK (DPSK) and its performance

- **DBPSK** is in general 1 dB inferior in bit-POE to BPSK/QPSK.
- **DQPSK** (see the formula in the next slide) is in general 2.3 dB inferior in bit-POE to DBPSK.

\[ \theta_k - \theta_{k-1} \]

The performance of DQPSK can be derived as (\(\theta_k - \theta_{k-1}\)) is either 0 or \(\pi/2\) or \(\pi\) or \(3\pi/2\). Hence, we only need to consider:

\[ n_k^* \sqrt{E} = \sqrt{E} e^{i \theta_k} + \left[ (n_{k-1} + j m_{k-1}) + (n_k + j m_k) + \frac{1}{\sqrt{E}} (n_{k-1} + j m_{k-1})(n_k + j m_k) \right] \]

Then, the optimal decision maker is:

- for the first bit, \(\text{Re}[n_k^*] + \text{Im}[n_k^*] \geq 1\)
- for the second bit, \(\text{Re}[n_k^*] - \text{Im}[n_k^*] \geq 1\)
5.2.8 Differential PSK (DPSK) and its performance

Take the second bit as an example for \((\theta_k - \theta_{k-1}) = \pi/2 \text{ or } \pi\).

\[
2 \text{Re}\{r_{k-1}^r r_k^*\} - 2 \text{Im}\{r_{k-1}^r r_k^*\} = (1+j)r_{k-1}^r r_k^* + (1-j)(r_{k-1}^r r_k^*)^*
\]

\[
X_k = \eta_k = \sqrt{E_b}e^{j(\theta_k - \theta_{k-1})} + j\eta_{k-1} \quad Y_k = t_{k-1} = \sqrt{E_b}e^{j(\theta_k - \theta_{k-1})} + j\eta_{k-1}
\]

\[
L = 1, A = B = C = 1 + j, D = (1 + j)X_k Y_k^* + (1 - j)X_k Y_k \\
\{m_y = \sqrt{E_b}e^{j(\theta_k - \theta_{k-1})}, m_x = \sqrt{E_b}e^{j(\theta_{k-1} - \theta_{k-2})}, \mu_y = \mu_x = N_b/2, \mu_{yy} = 0
\]

\[
a = \left[\frac{2 - \sqrt{2} E}{N_b}\right]^{1/2}, \quad b = \left[\frac{2 + \sqrt{2} E}{N_b}\right]^{1/2}
\]

\[
\Rightarrow \quad v_1 = v_2 = \frac{1}{\sqrt{2}N_b}, \quad \alpha_1 = 4EN_b, \quad \alpha_2 = \sqrt{2E}\left(e^{-(\theta_{k-1} - \theta_{k-2}) + j\pi/4} + e^{j(\theta_{k-1} - \theta_{k-2}) + j\pi/4}\right) = 2E
\]

(Because error occurs when \((\theta_k - \theta_{k-1}) = 0 \text{ or } 3\pi/2\).)

Therefore,

\[
P_{e,DPSK} = Q_1(a, b) - \frac{1}{2} I_0(ab) \exp\left\{-\frac{1}{2}(a^2 + b^2)\right\}
\]

where \(a = \sqrt{2}\gamma_b\left(1 - 1/\sqrt{2}\right)\) and \(b = \sqrt{2}\gamma_b\left(1 + 1/\sqrt{2}\right)\).