EC5.203 Communication Theory I (3-1-0-4):

Lectures 18-23:
Optimal Demodulation

24 March 2025 to 21 April 2025

INTERNATIONAL INSTITUTE OF
INFORMATION TECHNOLOGY
HYDERABAD

References

- Chap. 6 (Madhow)

Optimal Demodulation

- In 16 QAM, one of 16 passband waveforms corresponding to 16 symbols is sent, where each passband waveform is given by

$$s_i(t) = s_{b_c,c_s} = b_c p(t)\cos(2\pi f_c t) - b_s p(t)\sin(2\pi f_c t)$$

  where b_c, b_s each takes value in $\{\pm 1, \pm 3\}$

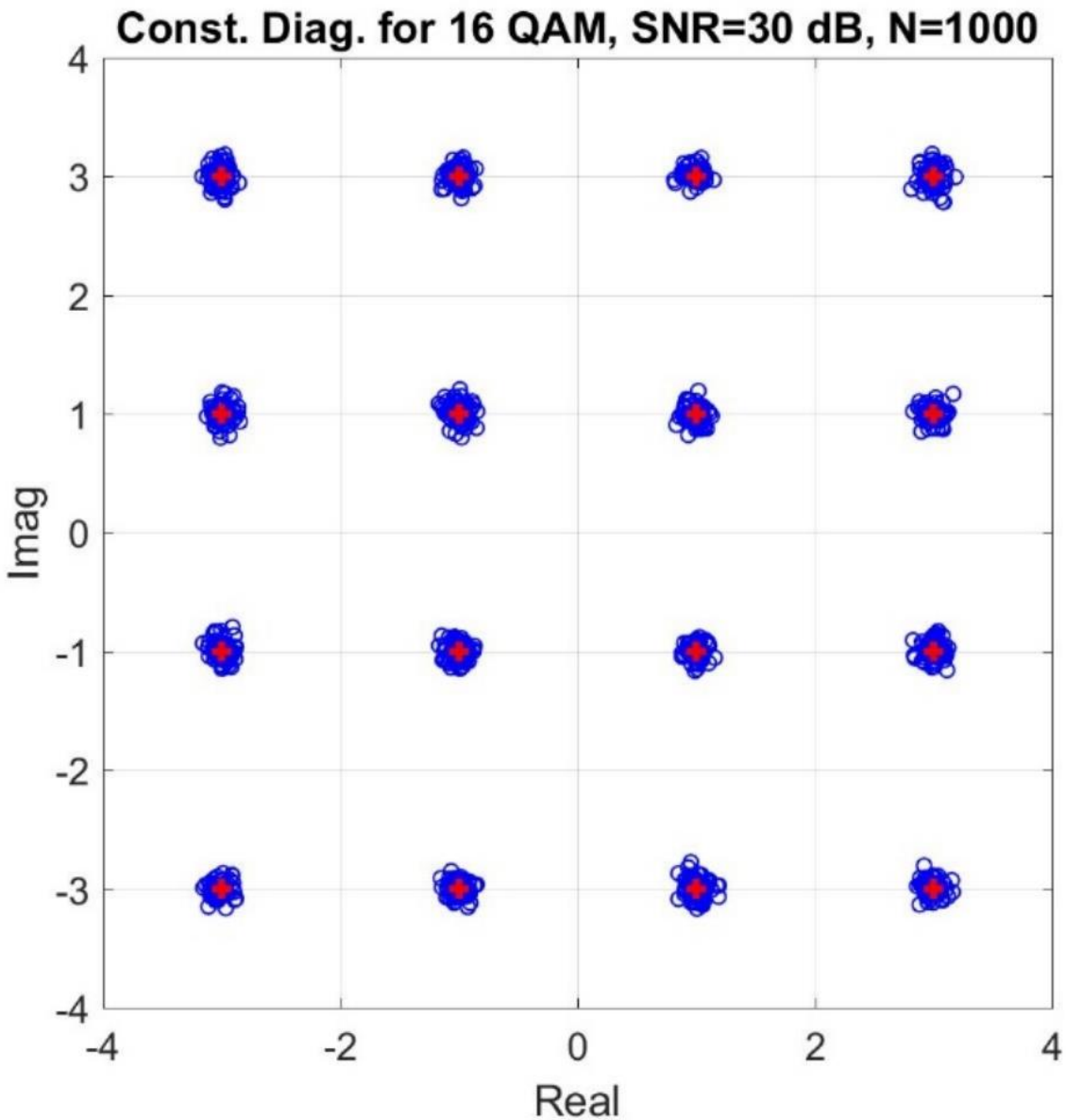
- At the receiver, we have noisy observations

$$y(t) = s_i(t) + n(t)$$

- Now, we are faced with a hypothesis testing problem at the receiver: we have M possible hypotheses about which signal was sent.

- Hypothesis: Possible cause of an event.

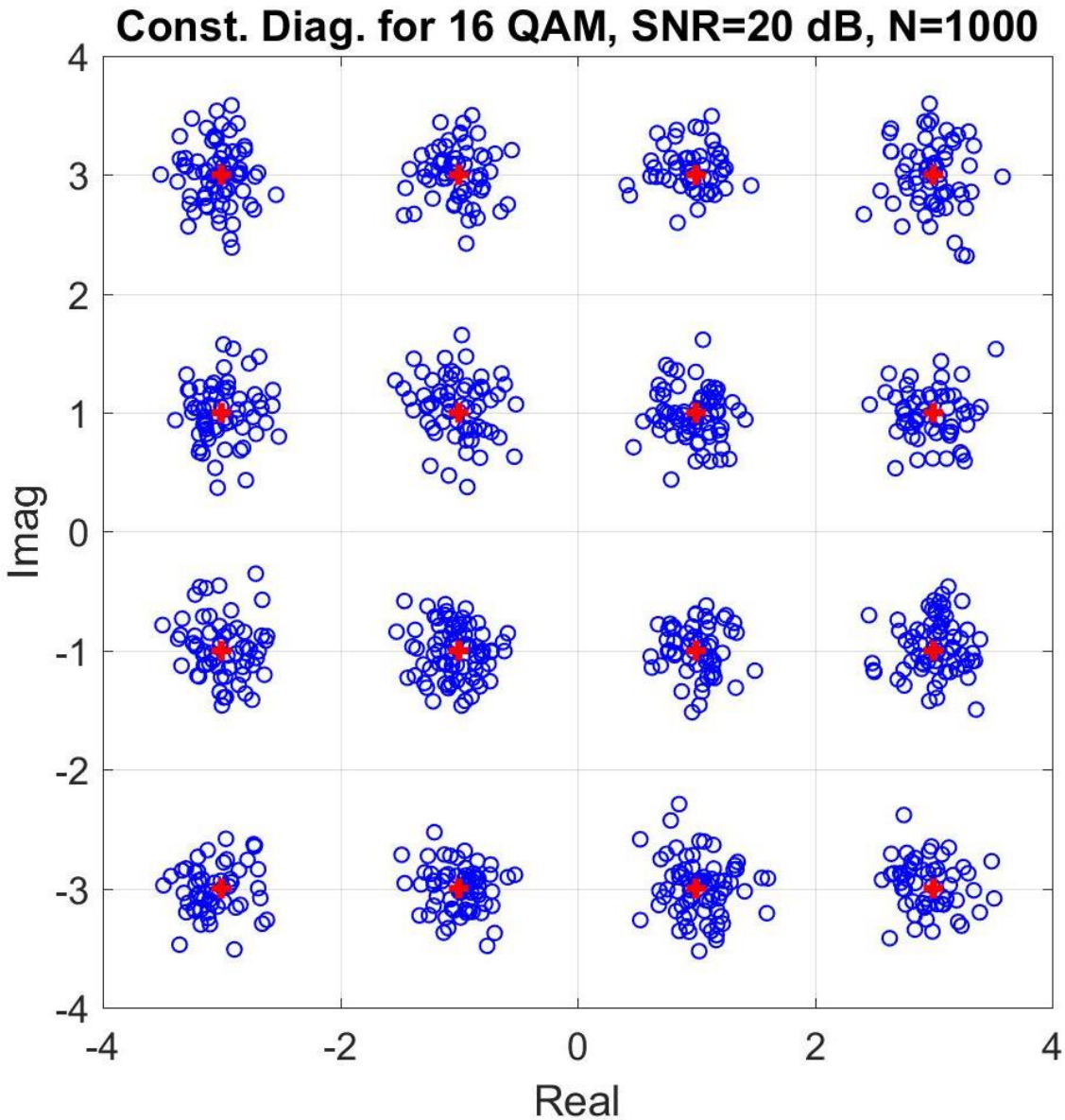
Received Data at SNR = 30 dB



$y(t) = s_i(t) + n(t)$

$n(t)$ is AWGN

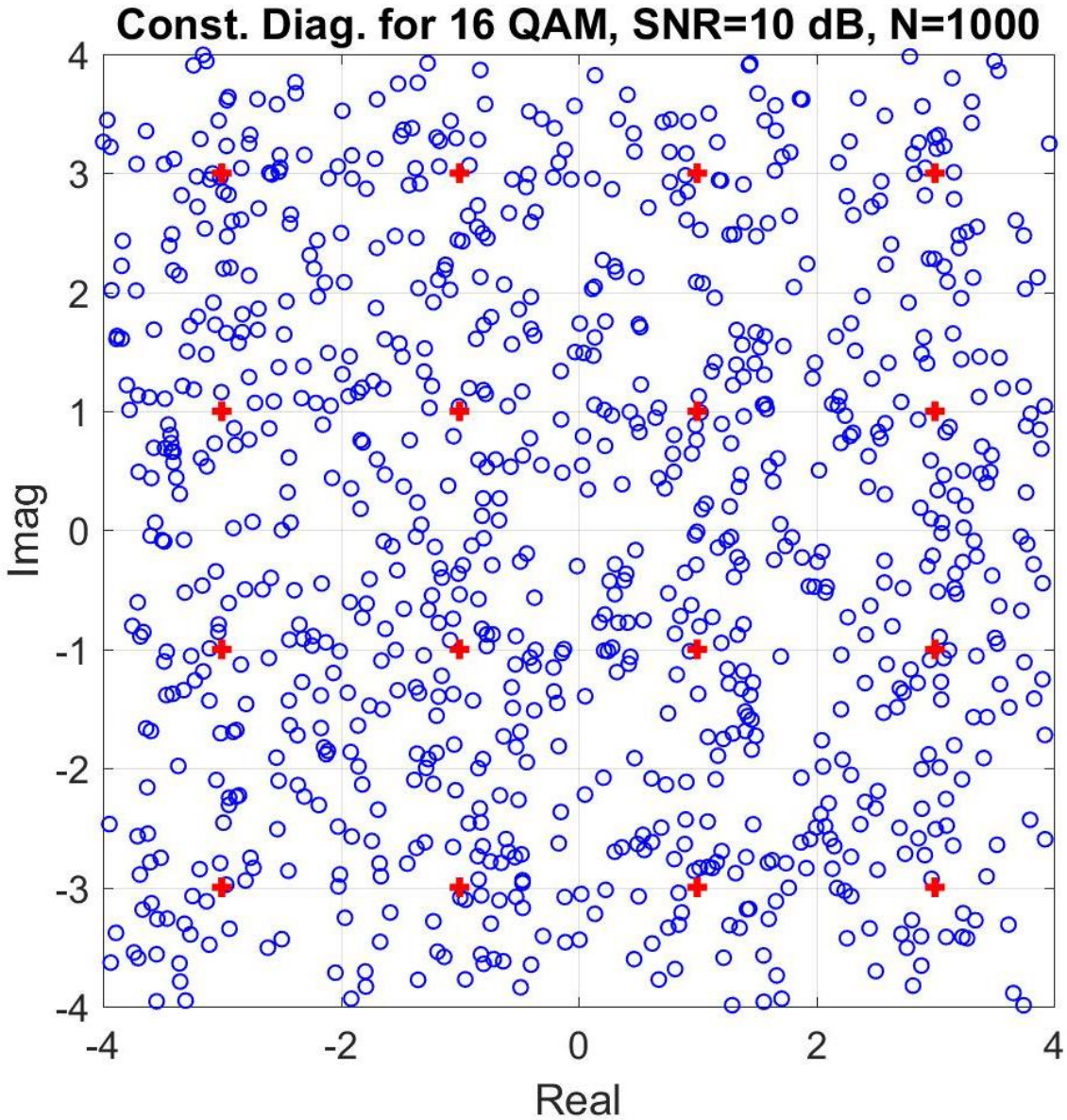
Received Data at SNR = 20 dB



$y(t) = s_i(t) + n(t)$

$n(t)$ is AWGN

Received Data at SNR=10dB



$y(t) = s_i(t) + n(t)$

$n(t)$ is AWGN

Optimal Demodulation

- In 16 QAM, one of 16 passband waveforms corresponding to 16 symbols is sent, where each passband waveform is given by

$$s_i(t) = s_{b_c,c_s} = b_c p(t)\cos(2\pi f_c t) - b_s p(t)\sin(2\pi f_c t)$$

  where b_c, b_s each takes value in $\{\pm 1, \pm 3\}$.

- At the receiver, we are faced with a hypothesis testing problem: we have M possible hypotheses about which signal was sent.

- Based on the observations

$$y(t) = s_i(t) + \boxed{n(t)} \quad \text{AWGN}$$

  we are interested in finding a decision rule to make a best guess which hypothesis was sent.

- For communications applications, performance criteria is to minimize the probability of error (i.e., the probability of making a wrong guess).

S&S Recap: Signal Energy

- The energy in a CT signal $x(t)$ over time interval (t_1, t_2)

$$\int_{t_1}^{t_2} |x(t)|^2 \, dt$$

where $x(t)$ is a complex signal.

- The energy in a DT signal $x[n]$ over sample interval $[n_1, n_2]$ is

$$\sum_{n_1}^{n_2} |x[n]|^2$$

- The energy in a CT $x(t)$ and a DT signal $x[n]$ over infinite time interval respectively are

$$E_\infty = \int_{-\infty}^{\infty} |x(t)|^2 \, dt \qquad \text{and} \qquad E_\infty = \sum_{-\infty}^{\infty} |x[n]|^2$$

S&S Recap: Signal Power

- The power in a continuous-time signal $x(t)$ over time interval (t_1, t_2)

$$\frac{1}{t_2 - t_1} \int_{t_1}^{t_2} |x(t)|^2 \, \mathrm{d}t$$

- The power in a discrete-time signal $x[n]$ over samples (n_1, n_2) is

$$\frac{1}{n_2 - n_1 + 1} \sum_{n_1}^{n_2} |x[n]|^2$$

- The power in $x(t)$ and $x[n]$ over infinite time interval respectively are

$$P_\infty = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} |x(t)|^2 \, dt \qquad \text{and} \qquad P_\infty = \lim_{N \to \infty} \frac{1}{2N} \sum_{-N}^{N} |x[n]|^2$$

S&S Recap: Note on Dimension of Energy and Power Definitions

- Consider an example where $v(t)$ and $i(t)$ are the instantaneous voltage and current across a resistor R, then the instantaneous power across the resistor is

$$p(t) = v(t)i(t) = \frac{v^2(t)}{R}$$

- The total energy dissipated over the time interval (t_1, t_2) is

$$E = \int_{t_1}^{t_2} p(t)\ dt$$

- The average power dissipated over the time interval (t_1, t_2) is

$$P = \frac{E}{t_2 - t_1}$$

The definitions used earlier are generic and may have wrong dimensions and scaling. The advantage is the convenience and wide applicability irrespective of where the signal is coming from.

Example 5.6.3

- **Binary on-off keying in Gaussian noise**

$$Y = m + n \quad \text{if 1 is sent}$$
$$Y = n \quad \text{if 0 is sent}$$

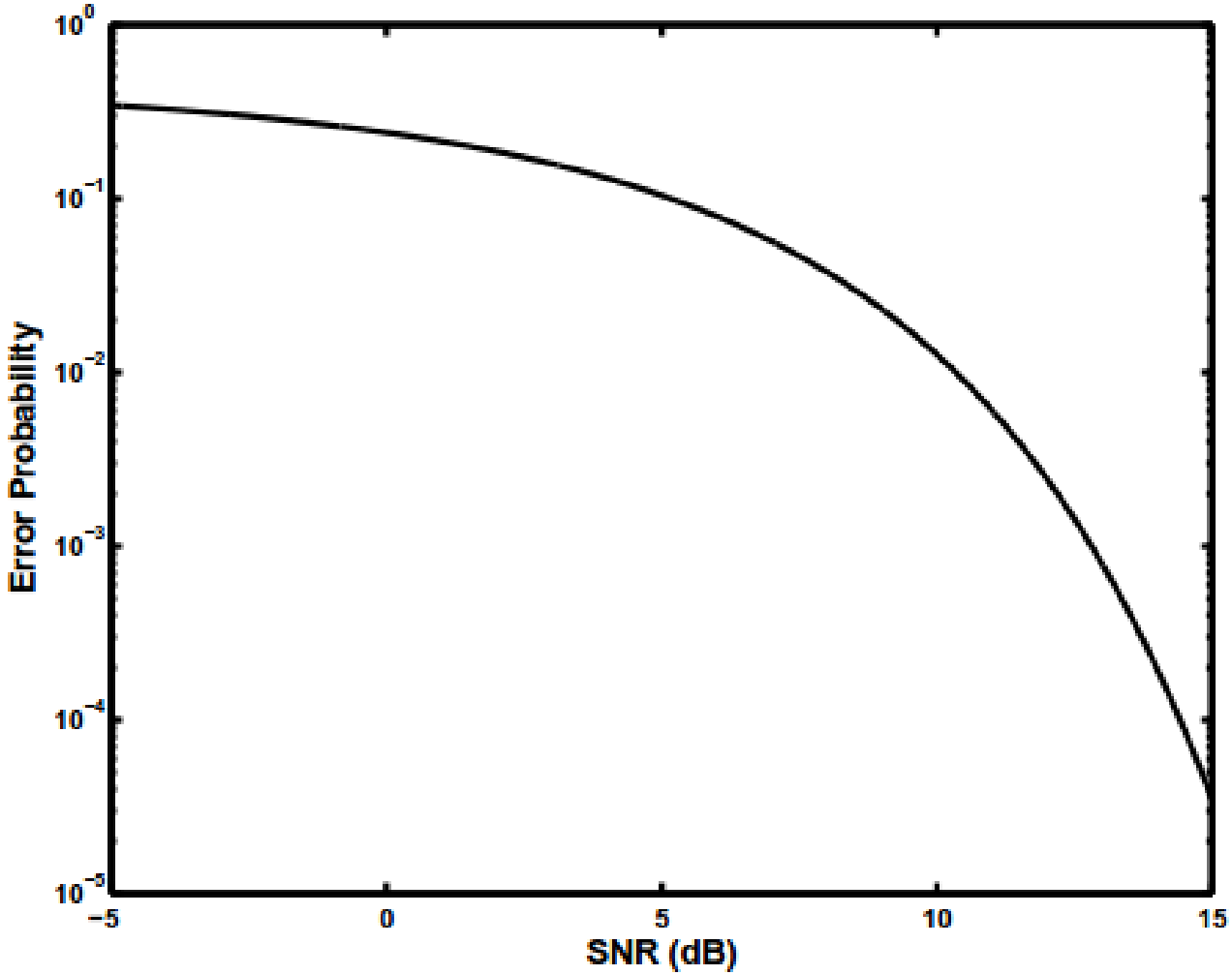
Here Y is the received sample, $m > 0$ is some constant and n is AWGN sample with $\mathcal{N}(0, v^2)$.

- At the receiver, the detection strategy is

$$Y > m/2 \quad \text{Decide 1 is sent}$$
$$Y \leq m/2 \quad \text{Decide 0 is sent}$$

- Assuming that both 0 and 1 are equally likely,
  - Find the average signal power
  - Find the conditional probability of error conditioned on 0 being sent
  - Find the conditional probability of error conditioned on 1 being sent
  - Find average error probability
  - Find the probability of error for SNR of 13 dB?

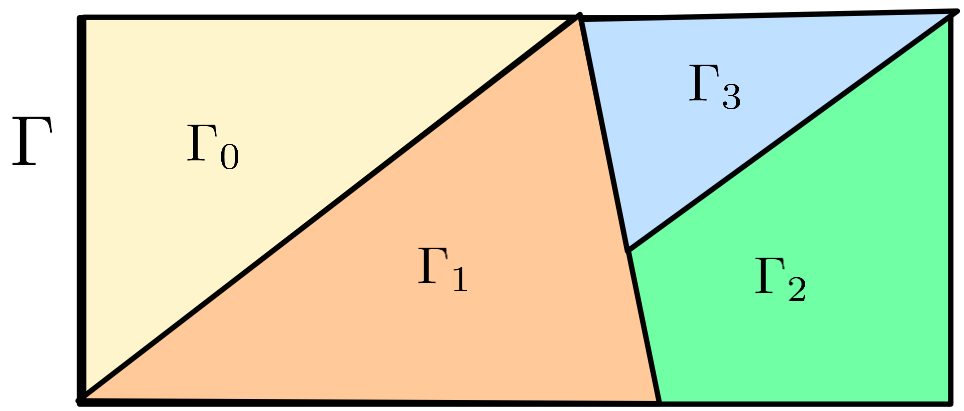
(Bit) Error Probability vs SNR for Example



Ingredients of Hypothesis Testing Framework

- Hypotheses $H_0, H_1, \ldots, H_{M-1}$
- Observation $Y \in \Gamma$
- Conditional densities $p(y|i)$ for $i = 0, 1, \ldots, M-1$
- Prior probabilities $\pi_i = P(H_i)$ with $\sum_i \pi_i = 1$
- Decision rule $\delta : \Gamma \to \{0, 1, M-1\}$
- Decision region $\Gamma_i : \{y \in \Gamma_i : \delta(y) = i\}$ for $i = 0, 1, M-1$

Example of Decision regions for M=4



Error Probabilities

- Conditional error probabilities, conditioned on H_i, is

$$\begin{aligned} P_{e|i} &= P(\text{decide } j \text{ for some } j \neq i | H_i \text{ is true}) \\ &= \sum_{j \neq i} P(Y \in \Gamma_j | H_i) \\ &= 1 - P(Y \in \Gamma_i | H_i) \end{aligned}$$

- Conditional probabilities of correct detection, conditioned on H_i, is

$$\begin{aligned} P_{c|i} &= P(Y \in \Gamma_i | H_i) \\ &= 1 - P_{e|i} \end{aligned}$$

- Average error probability

$$P_e = \sum_{i=1}^{M} \pi_i P_{e|i}$$

- Average probability of correct detection

$$P_c = \sum_{i=1}^{M} \pi_i P_{c|i}$$

Ingredients of Hypothesis Testing Framework

- Hypotheses $H_0, H_1, \ldots, H_{M-1}$
- Observation $Y \in \Gamma$
- Conditional densities $p(y|i)$ for $i = 0, 1, \ldots, M-1$
- Prior probabilities $\pi_i = P(H_i)$ with $\sum_i \pi_i = 1$
- Decision rule $\delta : \Gamma \to \{0, 1, M-1\}$
- Decision region $\Gamma_i : \{y \in \Gamma_i : \delta(y) = i\}$ for $i = 0, 1, M-1$

- In earlier example
  - Hypotheses H_0, H_1, Observation $Y \in \Gamma = \mathcal{R}$
  - Conditional densities $p(y|0)$ and $p(y|1)$
  - Prior probabilities π_0 and π_1
  - Decision rule δ:
  $$\delta(y) = \begin{cases} 0, & y \leq m/2 \\ 1, & y > m/2 \end{cases}$$
  - Decision regions: $\Gamma_0 = (-\infty, m/2]$ and $\Gamma_1 = (m/2, \infty)$

MAP rule

- Definitions:
  - *A priori* probability: Before the data is observed : $P(H_i) = \pi_i$
  - *A posteriori* probability: After the data is observed: $P(H_i|y)$
- Maximum *a posteriori* probability (MAP) rule:

$$\delta_{\text{MAP}}(y) = \arg\max_i P(H_i|Y = y)$$

where $i = 0, 1, \ldots, M-1$

- Using Bayes rule $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$, MAP rule can be rewritten as

$$\begin{aligned}
\delta_{\text{MAP}}(y) &= \arg\max_i \frac{P(Y = y|H_i)P(H_i)}{P(Y = y)} \\
&= \arg\max_i \frac{p(y|i)\pi_i}{p(y)} \\
&= \arg\max_i p(y|i)\pi_i \\
&= \arg\max_i \log \pi_i + \log p(y|i)
\end{aligned}$$

Optimality of MAP (or MPE) rule

- Optimality of MAP rule: The MAP rule minimizes the probability of error. Proof!

ML rule

- Definitions:
  - Likelihood function: $p(y|i) = P(Y = y|H_i)$
- Maximum likelihood (ML) rule

$$\delta_{\mathrm{ML}}(y) = \arg\max_i p(y|i)$$

where $i = 0, 1, \ldots, M-1$

- Equivalently,

$$\delta_{\mathrm{ML}}(y) = \arg\max_i \log p(y|i)$$

- ML is equivalent of MAP for equal prior probabilities, i.e., $\pi_i = \frac{1}{M}$, we have

$$\begin{aligned}\delta_{\mathrm{MAP}}(y) &= \arg\max_i \log \pi_i + \log p(y|i) \\ &= \arg\max_i \log \frac{1}{M} + \log p(y|i) \\ &= \arg\max_i \log p(y|i)\end{aligned}$$

Binary Hypothesis Testing Problem

- For two hypotheses case, ML decision rule is

$$\begin{aligned}\delta_{\mathrm{ML}}(y) &= \arg\max_i p(y|i) \\ &= \arg\max_i \{p(y|0), p(y|1)\}\end{aligned}$$

- Equivalently,

$$\begin{aligned} p(y|0) > p(y|1) \quad &\rightarrow \quad \delta_{\mathrm{ML}}(y) = 0 \\ p(y|1) > p(y|0) \quad &\rightarrow \quad \delta_{\mathrm{ML}}(y) = 1 \end{aligned}$$

- This can be written as

$$p(y|1) \underset{H_0}{\overset{H_1}{\gtrless}} p(y|0)$$

- Similarly, MAP or MPE rule for binary hypothesis testing problem can be written as

$$\pi_1 p(y|1) \underset{H_0}{\overset{H_1}{\gtrless}} \pi_0 p(y|0)$$

$$P(H_1|Y = y) \underset{H_0}{\overset{H_1}{\gtrless}} P(H_0|Y = y)$$

Example of exponential distribution

- A binary hypothesis testing problem is specified as follows

$$H_0 : Y \sim \mathcal{E}(1)$$
$$H_1 : Y \sim \mathcal{E}(1/4)$$

  where $\mathcal{E}(\mu)$ denotes an exponential density $\mu e^{-\mu y}$ and CDF $1 - e^{-\mu y}$ where $y \geq 0$. Note that the mean of $\mathcal{E}(\mu)$ is $1/\mu$.

- Find the ML rule and the corresponding error probabilities.

- Find the MAP rule when the prior probability of H_1 is $1/5$. Also find the conditional and average error probabilities.

Likelihood Ratio

- For two hypotheses case, ML decision rule can be written as

$$p(y|1) \underset{H_0}{\overset{H_1}{\gtrless}} p(y|0)$$

- Equivalently,

$$\frac{p(y|1)}{p(y|0)} \underset{H_0}{\overset{H_1}{\gtrless}} 1$$

  This ratio of likelihood functions is called likelihood ratio(LR) and denoted by $L(y)$ and the test is called as likelihood ratio test (LRT)

- Taking log of both sides

$$\log \frac{p(y|1)}{p(y|0)} \underset{H_0}{\overset{H_1}{\gtrless}} 0$$

  The test statistic in this case is Log of LR and is called log-likelihood ratio (LLR) while the test is called as LLRT

Likelihood Ratio: MAP

- For two hypotheses case, MAP decision rule is

$$\pi_1 p(y|1) \underset{H_0}{\overset{H_1}{\gtrless}} \pi_0 p(y|0)$$

- In terms of LR, the LRT is

$$L(y) = \frac{p(y|1)}{p(y|0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\pi_0}{\pi_1}$$

- Taking log of both sides, the test is given in terms of LLR

$$\log L(y) \underset{H_0}{\overset{H_1}{\gtrless}} \log \frac{\pi_0}{\pi_1}$$

Signal Space Concepts

Receiver design as hypothesis testing

- Consider the multiple hypothesis testing

$$H_i \ : \ y(t) = s_i(t) + n(t) \quad i = 0, 1, \ldots, M-1$$

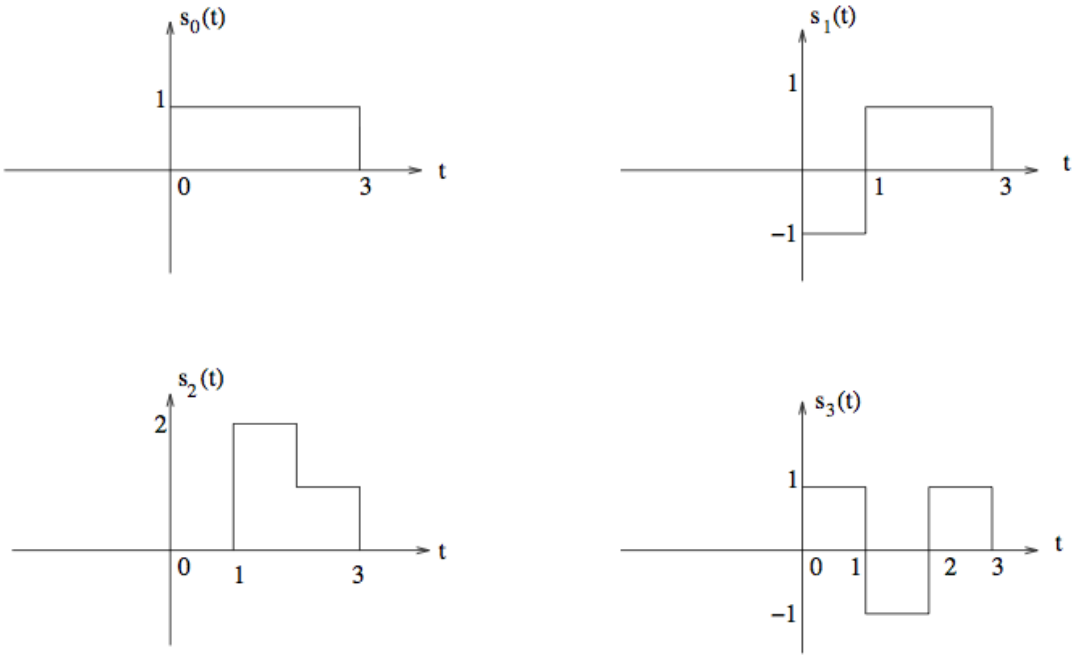
  where $n(t)$ is WGN with $S_n(f) = \frac{N_0}{2} = \sigma^2$

- Strategy:
  - Show that we can reduce the continuous time received signal to a finite-dimensional vector without losing information
  - Derive the optimal receiver based on the finite-dimensional vector observation
  - Map the optimal receiver back to continuous time

- This approach is based on signal space concepts: even though the received signal lives in an infinite-dimensional space, we can restrict attention to the subspace spanned by the signals that could have been transmitted

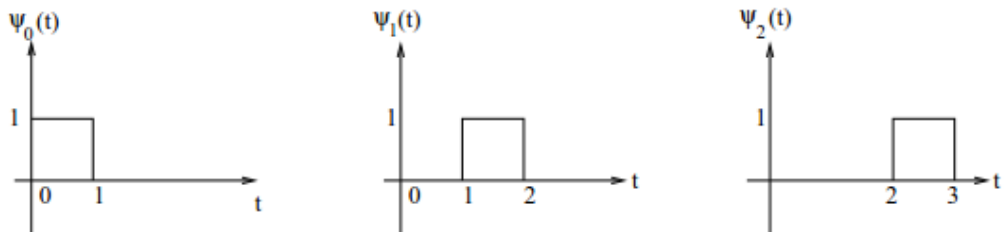
Signal space concepts: an example

- Four possible transmitted signals living in a 3-dimensional space

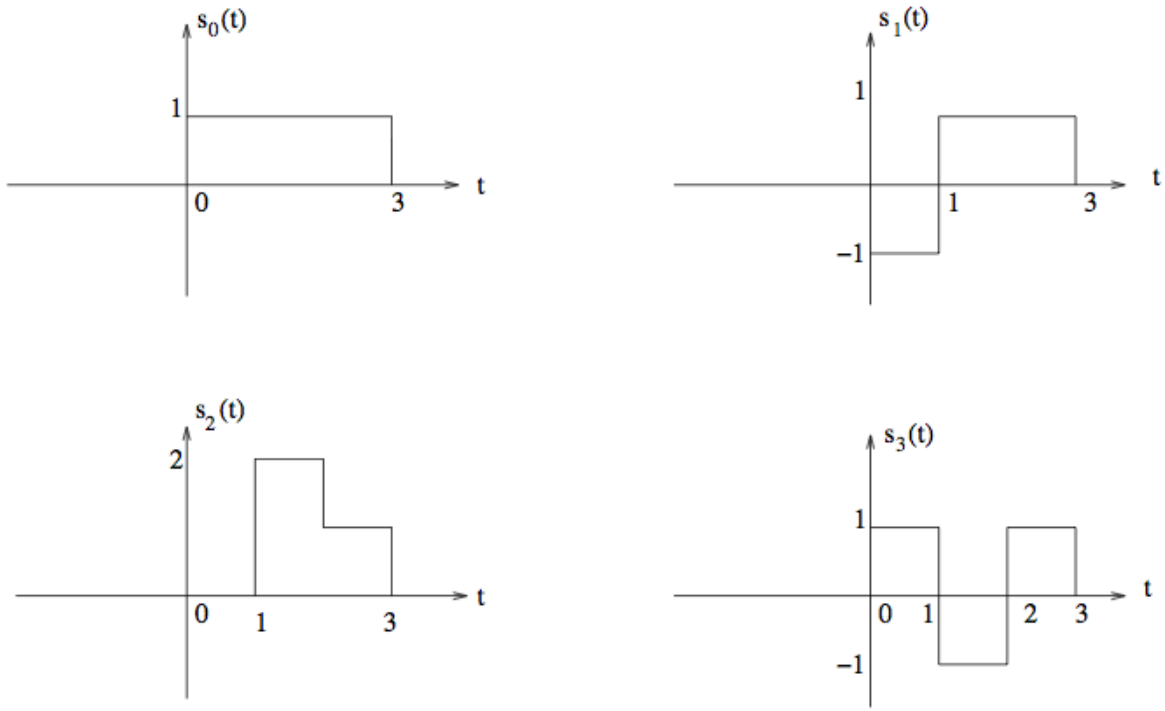


- Orthonormal basis (by inspection)

$$\psi_0(t) = I_{[0,1]}(t),\ \psi_1(t) = I_{[1,2]}(t),\ \psi_2(t) = I_{[2,3]}(t),$$



Example (continued)



- Expand with respect to basis $s_i(t) = \sum_{k=0}^{2} s_i[k]\psi_k(t)$
- Vectors of basis coefficient

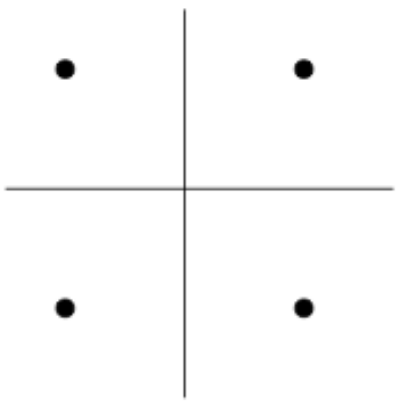
$$\mathbf{s}_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}; \quad \mathbf{s}_1 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}; \quad \mathbf{s}_2 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}; \quad \mathbf{s}_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix};$$

Another example: 2D modulation

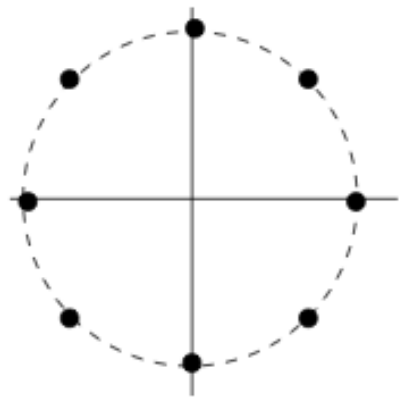
- 2D constellation for passband signaling: signal transmitted over a single symbol interval

$$s_i(t) = s_{b_c,b_s}(t) = b_c p(t) \cos 2\pi f_c t - b_s p(t) \sin 2\pi f_c t$$

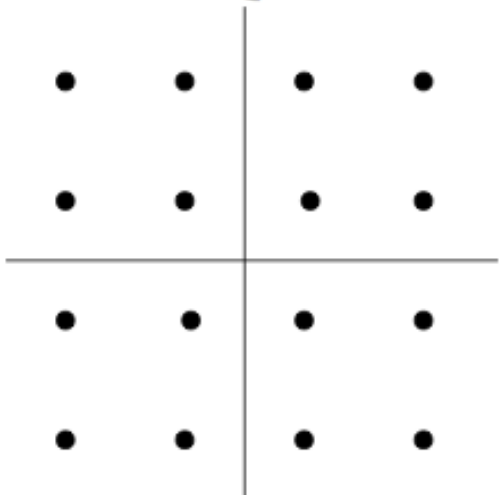
QPSK/4PSK/4QAM



8PSK



16QAM



- Signal is spanned by the following two signals

$$\phi_c(t) = p(t) \cos 2\pi f_c t; \quad \phi_s(t) = -p(t) \sin 2\pi f_c t$$

Another example: 2D modulation…

- Signal is spanned by the following two orthogonal signals

$$\phi_c(t) = p(t)\cos 2\pi f_c t; \quad \phi_s(t) = -p(t)\sin 2\pi f_c t$$

- Noting that $||\phi_c||^2 = ||\phi_s||^2 = \frac{1}{2}||p||^2$, corresponding orthonormal basis functions are

$$\psi_c(t) = \frac{\phi_c(t)}{||\phi_c||}; \quad \psi_s(t) = \frac{\phi_s(t)}{||\phi_s||};$$

- Expansion with respect to orthonormal basis is

$$s_{b_c,b_s}(t) = \frac{1}{\sqrt{2}}||p||b_c\psi_c(t) + \frac{1}{\sqrt{2}}||p||b_s\psi_s(t)$$

Basis expansion ⇔ constellation

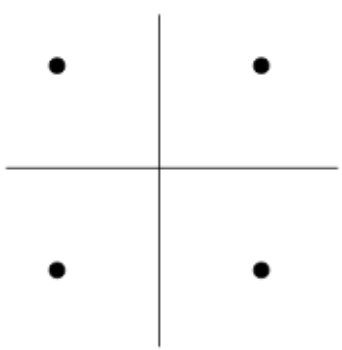
- 2D constellation for passband signaling: signal transmitted over a single symbol interval

$$s_i(t) = s_{b_c,b_s}(t) = b_c p(t) \cos 2\pi f_c t - b_s p(t) \sin 2\pi f_c t$$

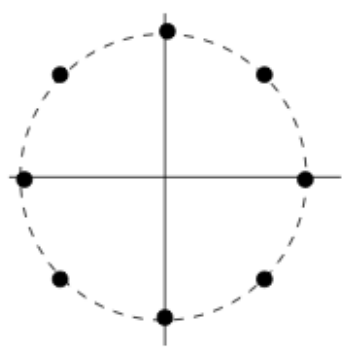
- Vector basis coefficients are given by

$$\mathbf{s}_i = \frac{1}{\sqrt{2}} ||p|| \begin{pmatrix} b_c \\ b_s \end{pmatrix}$$

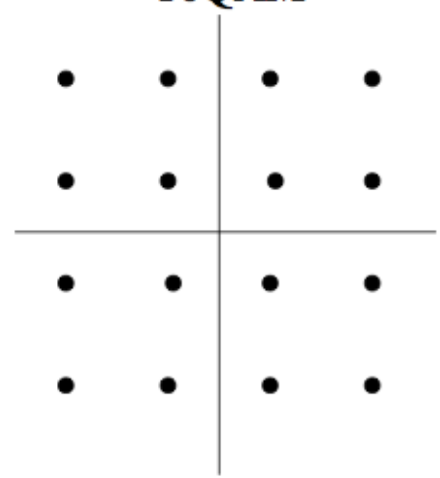
QPSK/4PSK/4QAM



8PSK



16QAM



From Signals to Vectors

- Signal space consists of all possible linear combinations of $s_0(t), s_1(t), \ldots,$ $s_{M-1}(t)$.

- We can always find an orthonormal basis for the signal space, i.e., $\psi_0(t)$, $\psi_1(t), \ldots, \psi_{n-1}(t)$ where $n \leq M$

- We can express each signal as a vector of basic coefficients

$$s_i(t) = \sum_{k=0}^{n-1} s_i[k]\psi_k(t)$$

  where $s_i[k] =< s_i, \psi_k >$

- Using $s_i[k]$, we can now express $s_i(t)$ as a $n \times 1$ vector $\mathbf{s}_i = [s_i[0]\ s_i[1]\ \ldots s_i[n-1]]^T$

- Note that finite-dimensional basis always exists

- If it is not possible to find natural basis by inspection, then we can always use Gram-Schmidt orthoginalization procedure

Gram-Schmit Orthogonalization Process

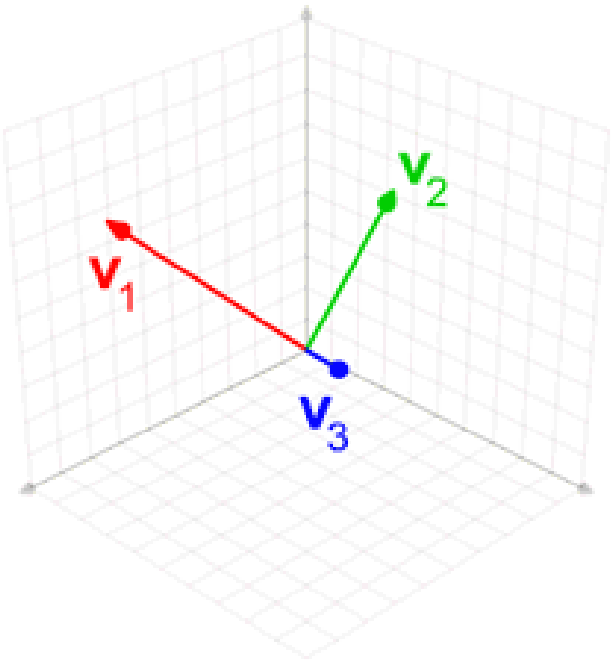
Orthogonal Vectors	Orthonormal Vectors
$\mathbf{u}_1 = \mathbf{v}_1,$	$\mathbf{e}_1 = \dfrac{\mathbf{u}_1}{\|\mathbf{u}_1\|}$
$\mathbf{u}_2 = \mathbf{v}_2 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_2),$	$\mathbf{e}_2 = \dfrac{\mathbf{u}_2}{\|\mathbf{u}_2\|}$
$\mathbf{u}_3 = \mathbf{v}_3 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_3) - \mathrm{proj}_{\mathbf{u}_2}(\mathbf{v}_3),$	$\mathbf{e}_3 = \dfrac{\mathbf{u}_3}{\|\mathbf{u}_3\|}$
$\mathbf{u}_4 = \mathbf{v}_4 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_4) - \mathrm{proj}_{\mathbf{u}_2}(\mathbf{v}_4) - \mathrm{proj}_{\mathbf{u}_3}(\mathbf{v}_4),$	$\mathbf{e}_4 = \dfrac{\mathbf{u}_4}{\|\mathbf{u}_4\|}$
$\vdots$	$\vdots$
$\mathbf{u}_k = \mathbf{v}_k - \sum_{j=1}^{k-1} \mathrm{proj}_{\mathbf{u}_j}(\mathbf{v}_k),$	$\mathbf{e}_k = \dfrac{\mathbf{u}_k}{\|\mathbf{u}_k\|}.$

$$\mathrm{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}\mathbf{u},$$

This operator projects the vector **v** orthogonally onto the line spanned by vector **u**.

Source: Wikipedia http://en.wikipedia.org/wiki/Gram%E2%80%93Schmidt_process

Gram-Schmit Orthogonalization Process



$$\mathrm{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u},$$

Source: Wikipedia http://en.wikipedia.org/wiki/Gram%E2%80%93Schmidt_process

Gram-Schmidt Orthogonalization Process

- **Step 0 (Initialization)**: Let $\phi_0 = s_0$. If $\phi_0 \neq 0$, then set $\psi_0 = \frac{\phi_0}{||\phi_0||}$. Note that ψ_0 provides a basis function for $\mathcal{S}_0$.

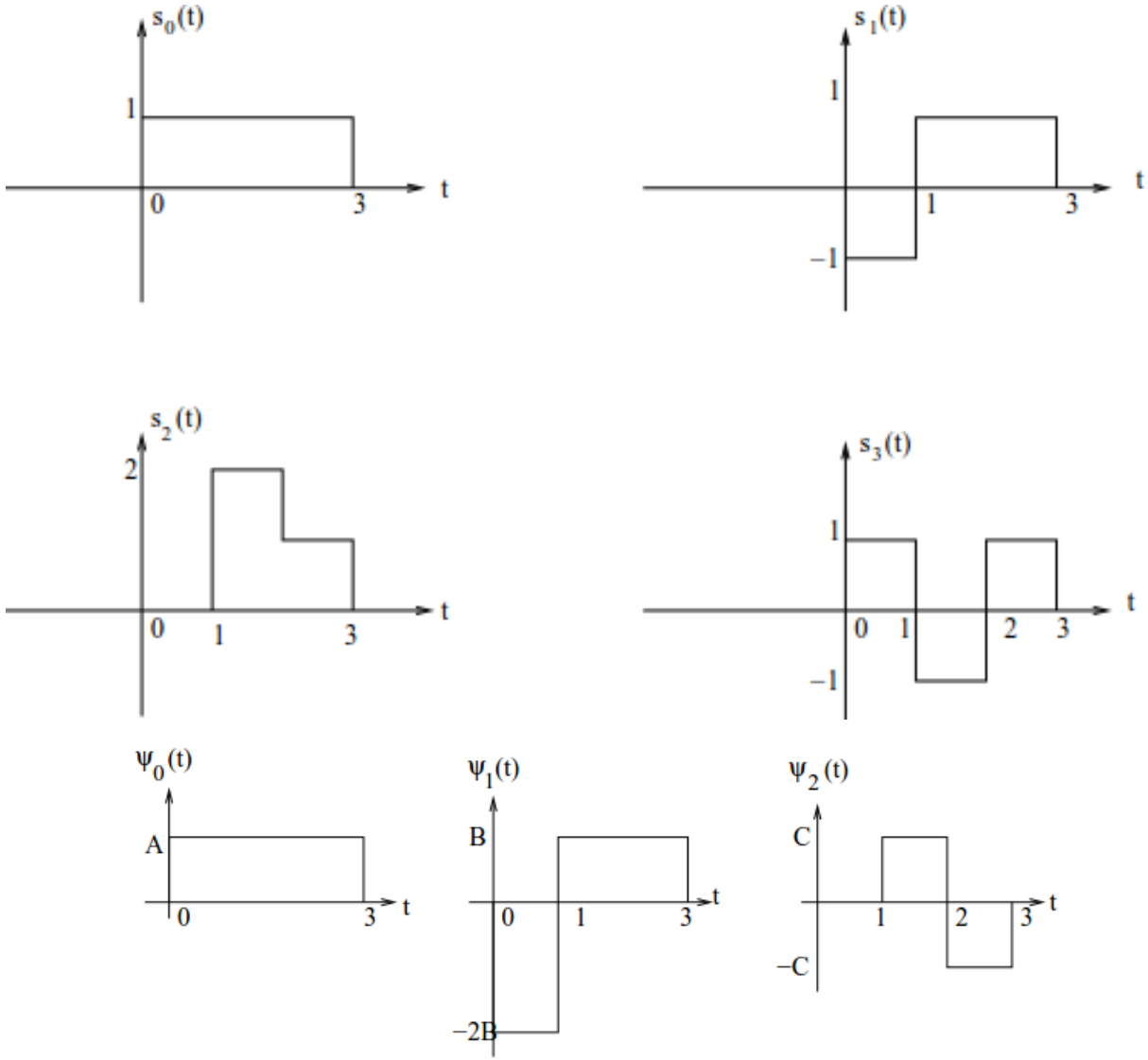
- **Step k**: Suppose that we have constructed an orthornormal basis $\mathcal{B}_{k-1} = \{\psi_0, \ldots, \psi_{m-1}\}$ for subspace $\mathcal{S}_{k-1}$ spanned by the first k signals $s_0, \ldots, s_{k-1}$. Note that $m \leq k$. Define

$$\phi_k(t) = s_k(t) - \sum_{i=0}^{m-1} < s_k, \psi_i > \psi_i(t)$$

- The signal $\phi_k(t)$ is the component of $s_k(t)$ orthogonal to the subspace $\mathcal{S}_{k-1}$. If $\phi_k \neq 0$, define a new basis function $\psi_m = \frac{\phi_k(t)}{||\phi_k||}$ and the basis as $\mathcal{B}_k = \{\psi_0, \ldots, \psi_m\}$. If $\phi_k = 0$, then $s_k \in \mathcal{S}_{k-1}$ and it is not necessary to update the basis. In this case $\mathcal{B}_k = \mathcal{B}_{k-1} = \{\psi_0, \ldots, \psi_{m-1}\}$.

Example: Gram-Schmidt Orthogonalization

- Find orthonormal basis set for these signals.



From signals to vectors

- Signal space consists of all possible linear combinations of $s_0(t), s_1(t), \ldots, s_{M-1}(t)$.
- We can always find an orthonormal basis for the signal space, i.e., $\psi_0(t)$, $\psi_1(t), \ldots, \psi_{n-1}(t)$ where $n \leq M$

We can express each signal as a vector of basis coefficients

$$s_i(t) = \sum_{l=0}^{n-1} s_i[k]\psi_k(t), \text{ where } s_i[k] = \langle s_i, \psi_k \rangle \quad \leftrightarrow \mathbf{s}_i = \begin{pmatrix} s_i[0] \\ s_i[1] \\ . \\ . \\ . \\ s_i[n-1] \end{pmatrix}$$

Inner products are preserved

- Performance of M-ary signaling in AWGN depends only on the inner products between the signals if the noise PSD is fixed.
- When mapping the CT hypothesis testing problem to DT, it is important to check that these inner products are preserved when projecting them to signal space.
- Check the following

$$\langle s_i, s_j \rangle = \int s_i(t) s_j(t) dt = \sum_{k=0}^{n-1} s_i[k] s_j[k] = \langle \mathbf{s}_i, \mathbf{s}_j \rangle$$

Inner products are preserved from CT to DT and vice versa
➔ norms, energies, distances are preserved

Modeling WGN in signal space

- What about projection of noise on the signal subspace?
- The noise projection onto the ith basis function as

$$N[i] = \langle \boxed{n}, \psi_i \rangle = \int n(t)\psi_i(t)$$

Noise

for $i = 0, 1, \ldots, \boxed{n} - 1$

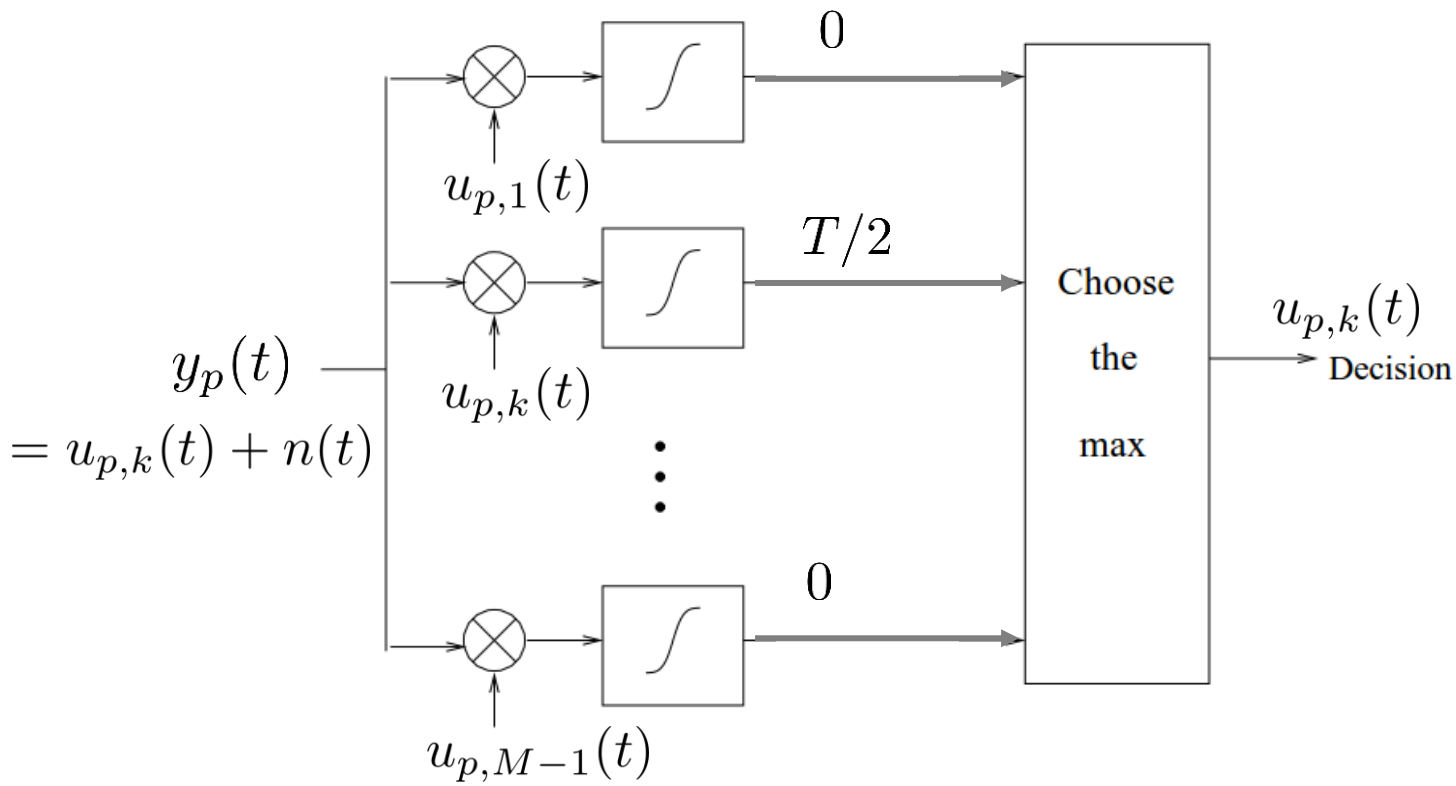
Dimension of signal space

- We can write noise $n(t)$ as follows

$$n(t) = \boxed{\sum_{i=0}^{n-1} N[i]\psi_i(t)} + \boxed{n^{\perp}(t)}$$

where $n^{\perp}(t)$ is the projection of noise orthogonal noise subspace.

Recap: *FSK coherent demodulation*



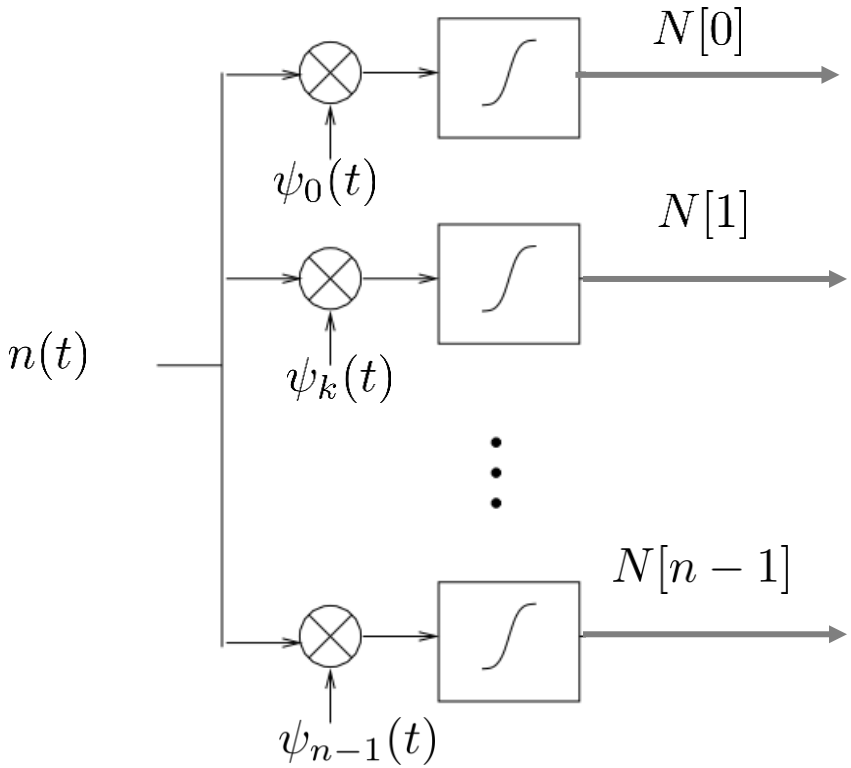
- Correlate the incoming signal with all M reference sinusoidal tones (passband signal) given by

$$u_{p,k} = \cos(2\pi(f_c + k\Delta f)t) \quad 0 \leq t \leq T$$

for $k = 0, 1, \ldots, M-1$.

- Choose $u_{p,m}$ for which the output is maximum among all the correlated outputs.

Noise projection on Signal Subspace



- Prove that noise projection on signal space is discrete-time WGN.
- For this we will need theorem on the next slide!

Noise projection on Signal Subspace

- **WGN through correlators** The random variables $Z_1 = \langle n, u_1 \rangle$ and $Z_2 = \langle n, u_2 \rangle$ are zero-mean, jointly Gaussian, with

$$\mathrm{cov}(Z_1, Z_2) = \sigma^2 \langle u_1, u_2 \rangle$$

so that $\mathbf{Z} = (Z_1, Z_2)^T \sim \mathcal{N}(\mathbf{0}, \mathbf{C})$ with covariance matrix

$$\mathbf{C} = \begin{pmatrix} \sigma^2 ||u_1||^2 & \sigma^2 \langle u_1, u_2 \rangle \\ \sigma^2 \langle u_1, u_2 \rangle & \sigma^2 ||u_2||^2 \end{pmatrix}$$

Proof!

- Prove that noise projection on signal space is discrete-time WGN.

Geometric interpretation of WGN

- The projection of WGN in any direction in signal space is an $\mathcal{N}(0, \sigma^2)$ random variable
- Projections in orthogonal directions are independent
- Projections along an orthonormal basis are iid $\mathcal{N}(0, \sigma^2)$ random variables

Hypothesis testing in signal space

- For the H_i hypothesis, the received signal is given by

$$y(t) = s_i(t) + n(t)$$

for $i = 0, 1, \ldots, M-1$

- By projecting this on signal space, we get

$$Y[k] = \langle y, \psi_k \rangle = s_i[k] + N[k] \quad k = 0, 1, \ldots, n-1$$

- On collecting these into an $n-$dimensional vector, we get

$$\mathbf{Y} = \mathbf{s}_i + \mathbf{N}$$

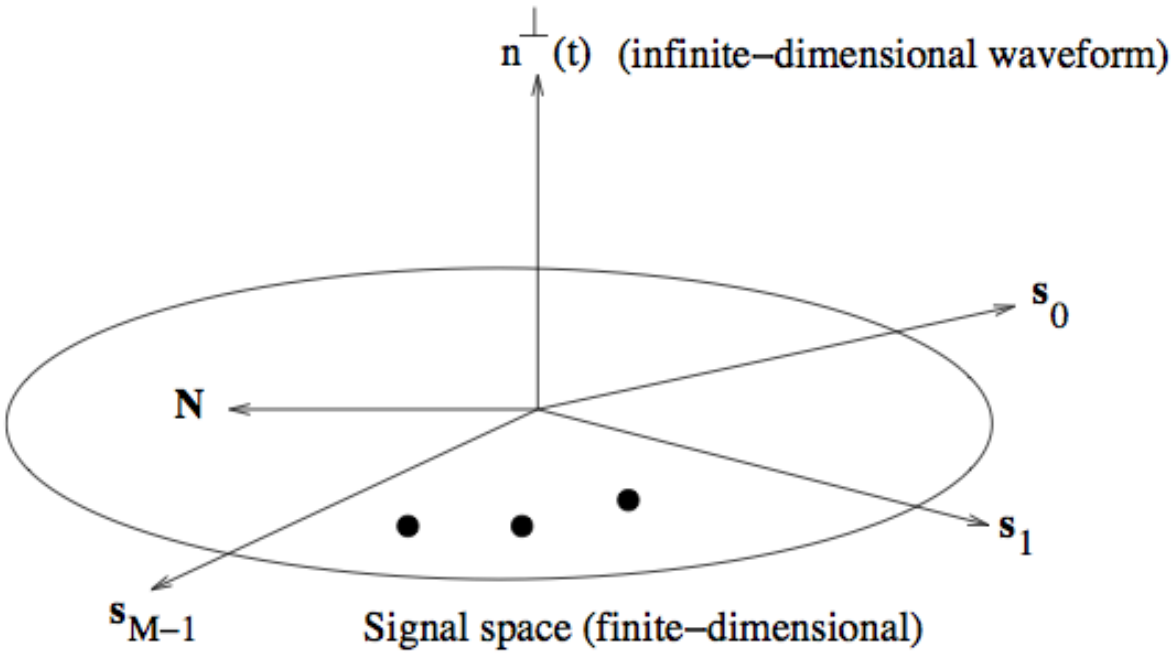
with $\mathbf{N} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$

- Note that the vector $\mathbf{Y}$ completely describes the component of the received signal $y(t)$ in the signal space given by

$$y_{\mathcal{S}} = \sum_{j=0}^{n-1} \langle y, \psi_j \rangle \psi_j(t) = \sum_{j=0}^{n-1} Y[j] \psi_j(t)$$

Irrelevance of orthogonal-noise component

- The noise component orthogonal to signal space, denoted by $n^{\perp}$,
  - does not give any information about signal
  - does not give any information about noise components in the signal space



Optimal Reception in AWGN

- Prove that the MAP/MPE rule for optimum demodulation in AWGN is

$$\delta_{\mathrm{MAP}}(\mathbf{y}) = \arg\max_i ||\mathbf{y} - \mathbf{s}_i||^2 - 2\sigma^2 \log \pi_i$$

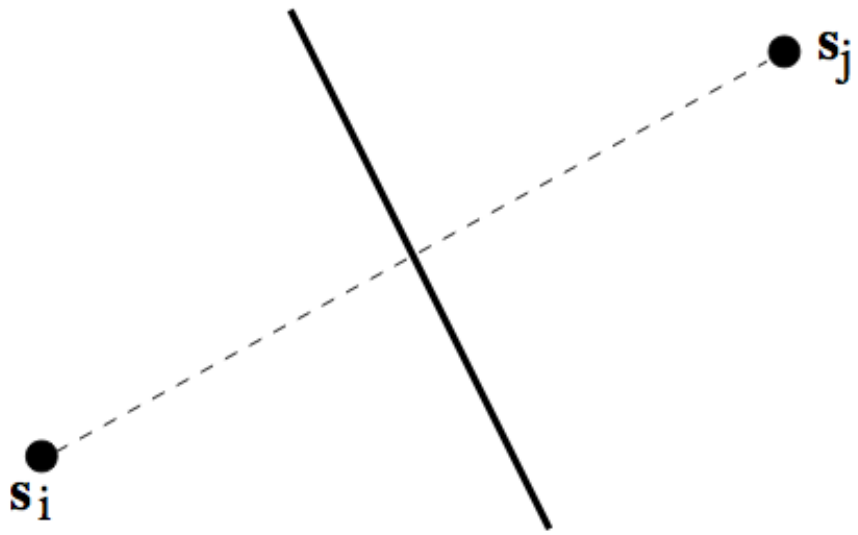
- For equiprobable prior probabilities, MAP is equivalent to ML

$$\begin{aligned}\delta_{\mathrm{ML}}(\mathbf{y}) &= \delta_{\mathrm{MAP}}(\mathbf{y}) \\ &= \arg\min_i ||\mathbf{y} - \mathbf{s}_i||^2 - 2\sigma^2 \log \frac{1}{M} \\ &= \arg\min_i ||\mathbf{y} - \mathbf{s}_i||^2\end{aligned}$$

This is **minimum distance rule**!

Min distance rule between 2 signals

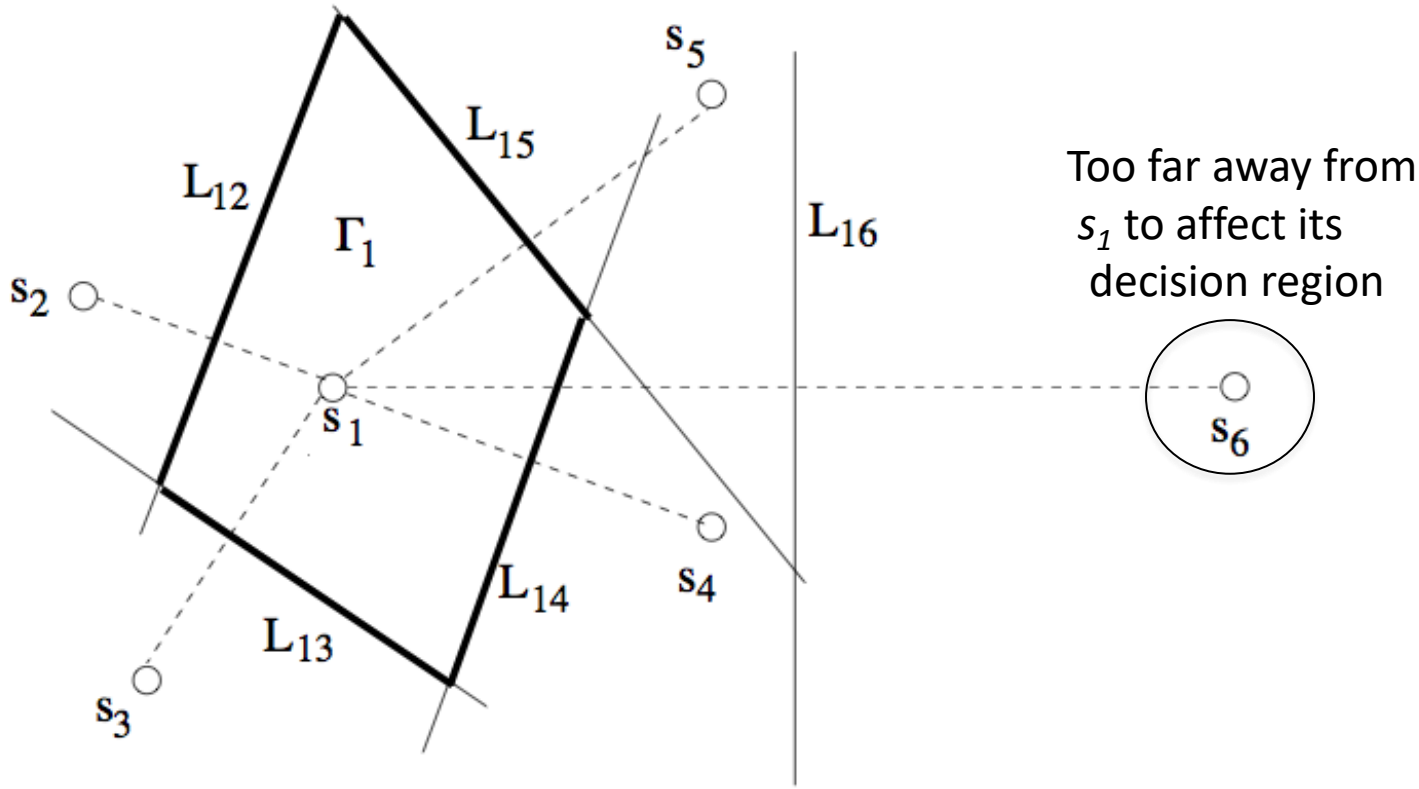
- Draw line between signal points.
- The ML decision boundary between them is the line bisecting this in 2D
- For two signal points in n dimensions, the ML decision boundary is $(n-1)$ dimensional hyperplane bisecting the line connecting the two points



ML decision boundary

is an (n−1) dimensional hyperplane

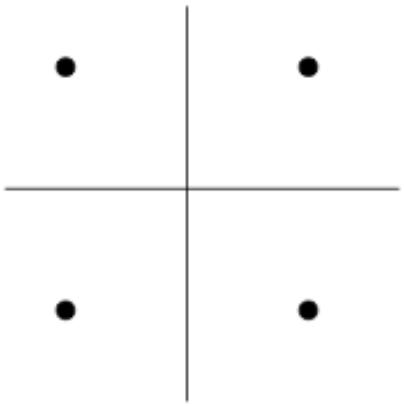
Geometry of min distance rule



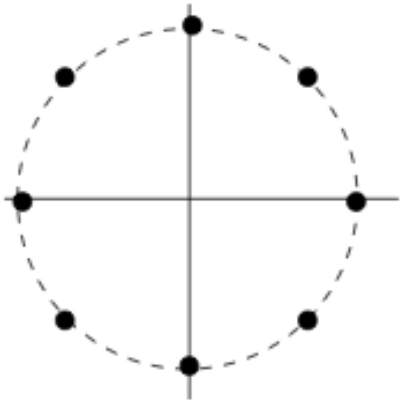
- Decision region for a signal is where it wins compared to all other signals
- Intersect the half-planes corresponding to all the pairwise comparisons
- Note that the decision region is typically determined by nearby neighbors

2 Dimensional Examples

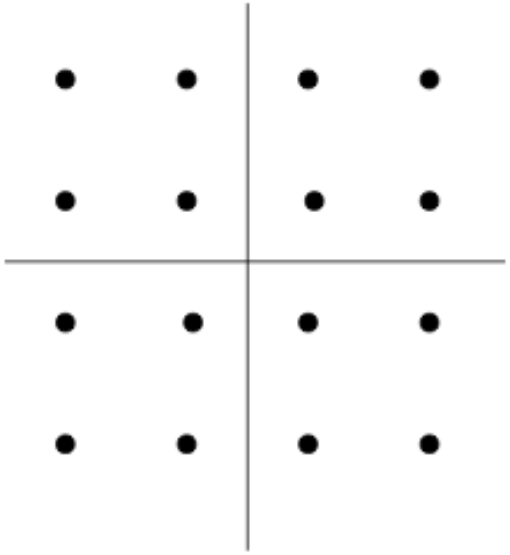
QPSK/4PSK/4QAM



8PSK



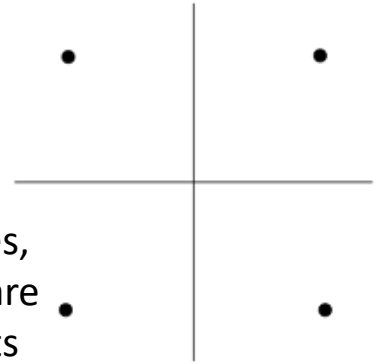
16QAM



Solve: Find decision regions for these constellations for ML demodulation

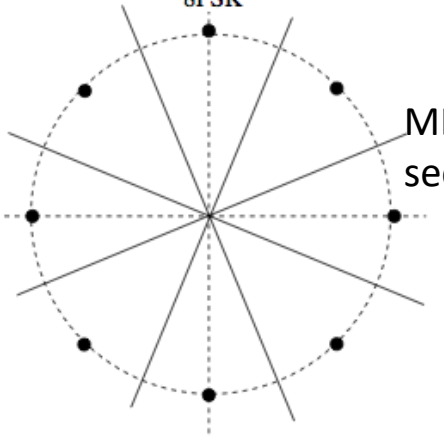
ML decisions for common 2D constellations

QPSK



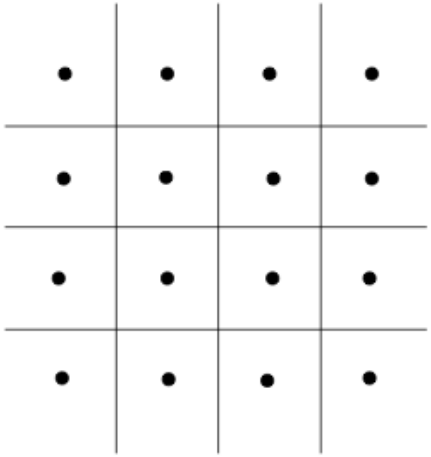
ML boundaries parallel to I/Q axes, decision regions are the four quadrants

8PSK



ML regions are sectors

16QAM



ML boundaries parallel to I/Q axes

MAP Demodulation In CT

- Remember that optimal modulation for signaling in DT AWGN is the MAP/MPE rule given by

DT(Finite Dimensional)

$$\delta_{\text{MAP}}(\mathbf{y}) = \arg\min_i ||\mathbf{y} - \mathbf{s}_i||^2 - 2\sigma^2 \log \pi_i$$
$$= \arg\max_i \langle \mathbf{y}, \mathbf{s}_i \rangle - ||\mathbf{s}_i||^2/2 + 2\sigma^2 \log \pi_i$$

- Equivalently, optimal demodulation for signaling in CT AWGN is

$$\delta_{\text{MAP}}(\mathbf{y}) = \arg\max_i \langle y, s_i \rangle - ||s_i||^2/2 + \sigma^2 \log \pi_i$$
$$= \arg\max_i \langle y_{\mathcal{S}}, s_i \rangle - ||s_i||^2/2 + \sigma^2 \log \pi_i$$

CT(Infinite Dimensional)

Recap: Receiver design as hypothesis testing

- Consider the multiple hypothesis testing

$$H_i \ : \ y(t) = s_i(t) + n(t) \quad i = 0, 1, \ldots, M-1$$

  where $n(t)$ is WGN with $S_n(f) = \frac{N_0}{2} = \sigma^2$

- Strategy:
  - Show that we can reduce the continuous time received signal to a finite-dimensional vector without losing information
  - Derive the optimal receiver based on the finite-dimensional vector observation
  - Map the optimal receiver back to continuous time

- This approach is based on signal space concepts: even though the received signal lives in an infinite-dimensional space, we can restrict attention to the subspace spanned by the signals that could have been transmitted

ML Demodulation In CT

- For equiprobable prior probabilities, MAP is equivalent to ML

$$\delta_{\mathrm{ML}}(\mathbf{y}) = \arg\min_i ||\mathbf{y} - \mathbf{s}_i||^2$$

DT(Finite Dimensional)

$$= \arg\max_i \langle y, s_i \rangle - ||s_i||^2/2$$

CT(Infinite Dimensional)

- Here we have used

$$\begin{aligned}
||\mathbf{s}_i||^2 &= ||s_i||^2 \\
\langle y, s_i \rangle &= \langle y_{\mathcal{S}} + y^{\perp}, s_i \rangle \\
&= \langle y_{\mathcal{S}}, s_i \rangle + \langle y^{\perp}, s_i \rangle \\
&= \langle y_{\mathcal{S}}, s_i \rangle \\
&= \langle \mathbf{y}, \mathbf{s}_i \rangle
\end{aligned}$$

- Note that we don't have minimum distance rule in CT as the squares of distance in CT will be

$$\begin{aligned}
||y - s_i||^2 &= ||y_{\mathcal{S}} - s_i||^2 + ||y^{\perp}||^2 \\
&= ||y_{\mathcal{S}} - s_i||^2 + \boxed{||n^{\perp}||^2}
\end{aligned}$$

Infinite Power

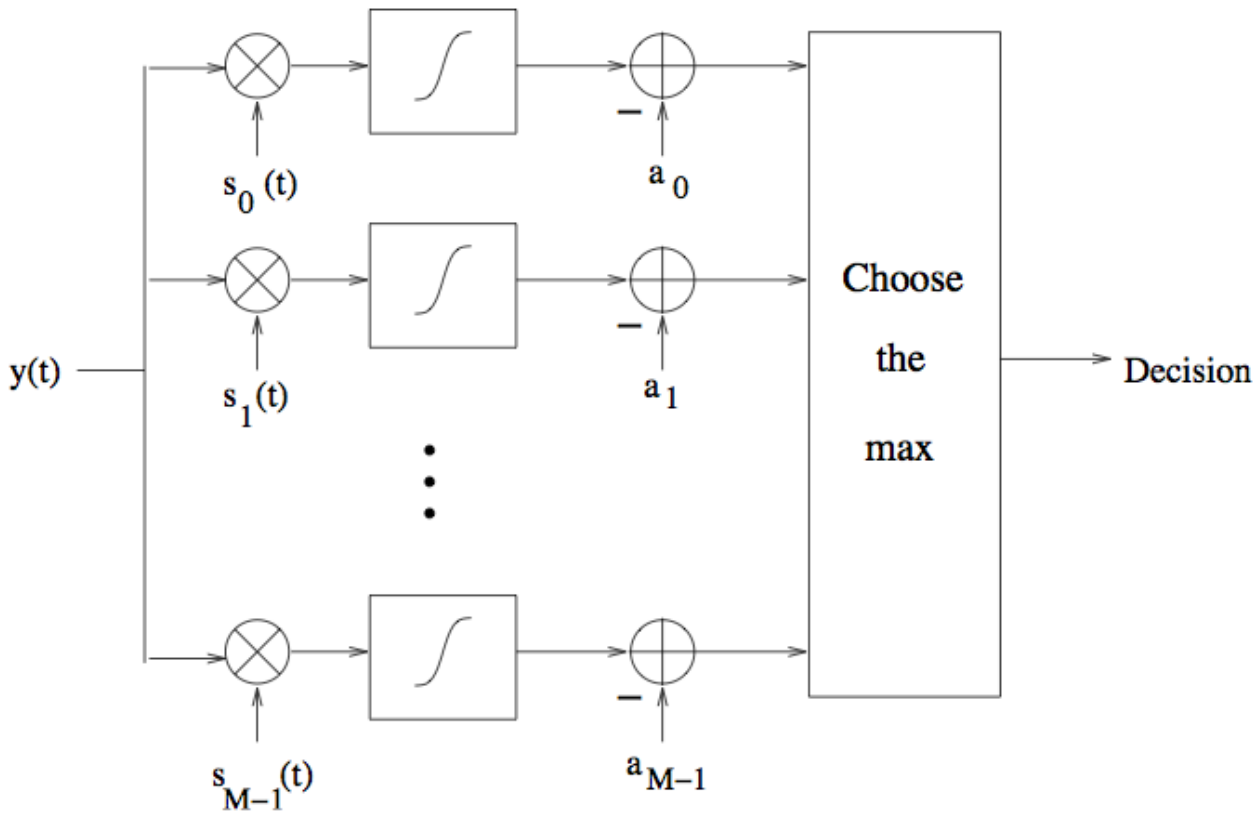
Implementation using correlation

- For equiprobable prior probabilities, MAP is equivalent to ML

$$\delta_{\mathrm{ML}}(\mathbf{y}) = \arg\min_i ||\mathbf{y} - \mathbf{s}_i||^2$$ DT(Finite Dimensional)

$$= \arg\max_i \langle y, s_i \rangle - ||s_i||^2/2$$

CT(Infinite Dimensional)



Equivalent Match Filter Implementation

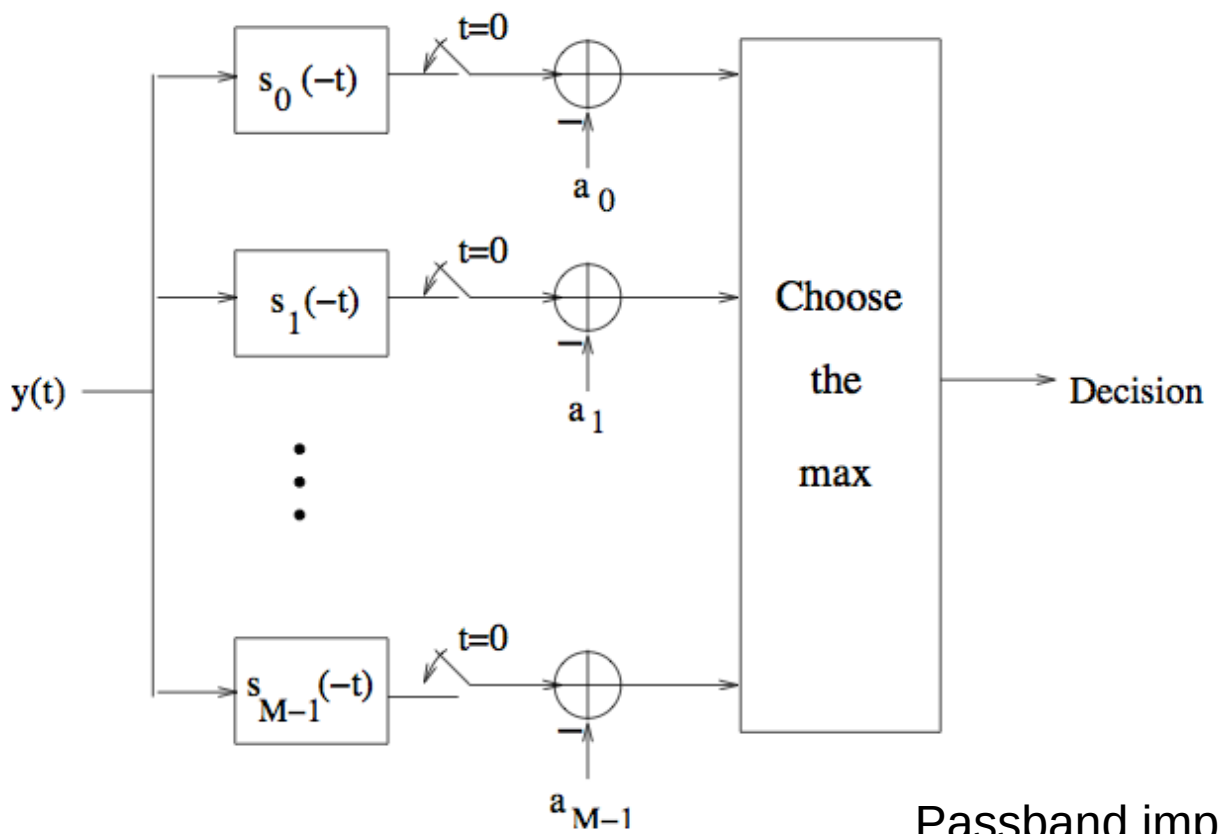
- For equiprobable prior probabilities, MAP is equivalent to ML

$$\delta_{\mathrm{ML}}(\mathbf{y}) = \arg\min_i ||\mathbf{y} - \mathbf{s}_i||^2$$

DT(Finite Dimensional)

$$= \arg\max_i \langle y, s_i \rangle - ||s_i||^2/2$$

CT(Infinite Dimensional)



Passband implementation

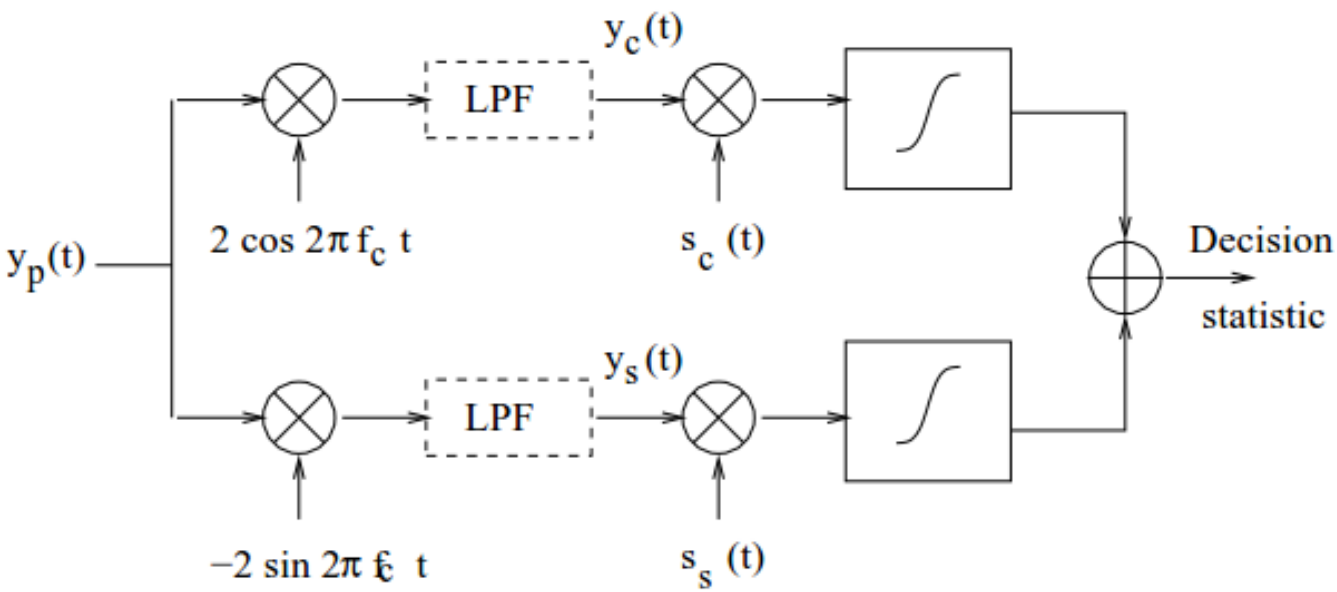
Implementation in Complex Baseband

- For equiprobable prior probabilities, MAP is equivalent to ML

$$\delta_{\mathrm{ML}}(\mathbf{y}) = \arg\min_i ||\mathbf{y} - \mathbf{s}_i||^2$$
$$= \arg\max_i \langle y, s_i \rangle - ||s_i||^2/2$$

- Note the relation between correlation in passband and baseband is given by

$$\langle u_p, v_p \rangle = \frac{1}{2}\mathrm{Re}\langle u, v \rangle = \frac{1}{2}(\langle u_c, v_c \rangle + \langle u_s, v_s \rangle)$$



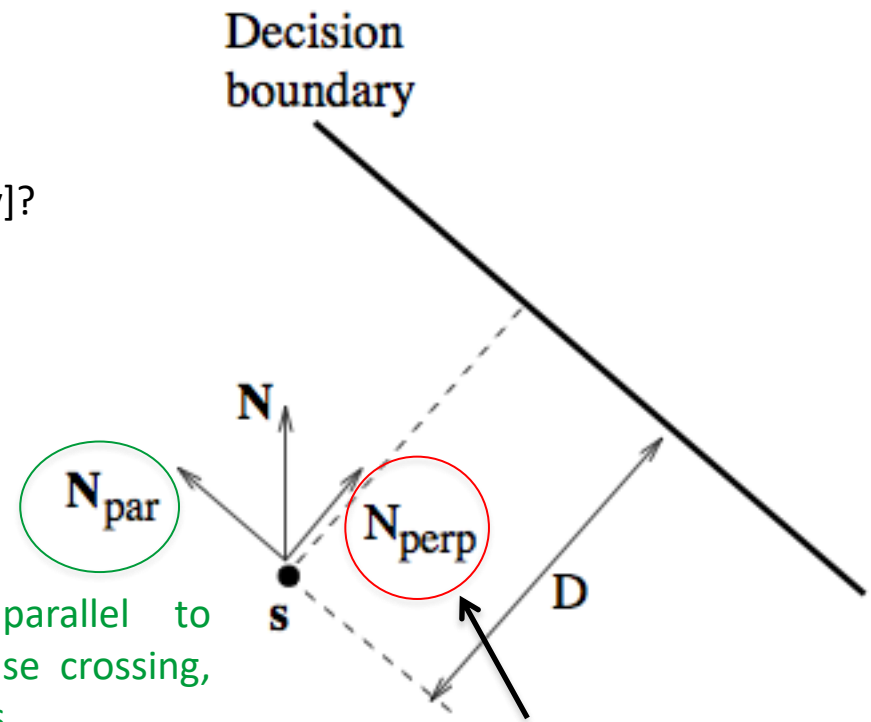
Performance Analysis of ML Reception

Performance Analysis of ML Reception

- Performance analysis of ML under the assumption of equiprobable priors.
- Although performance analysis of MPE or MAP is skipped, it is simple extension of ML case.

The Geometry of Errors

Basic building block:
Send signal **s.**
Noise **N** gets added.
P[crossing a given boundary]?



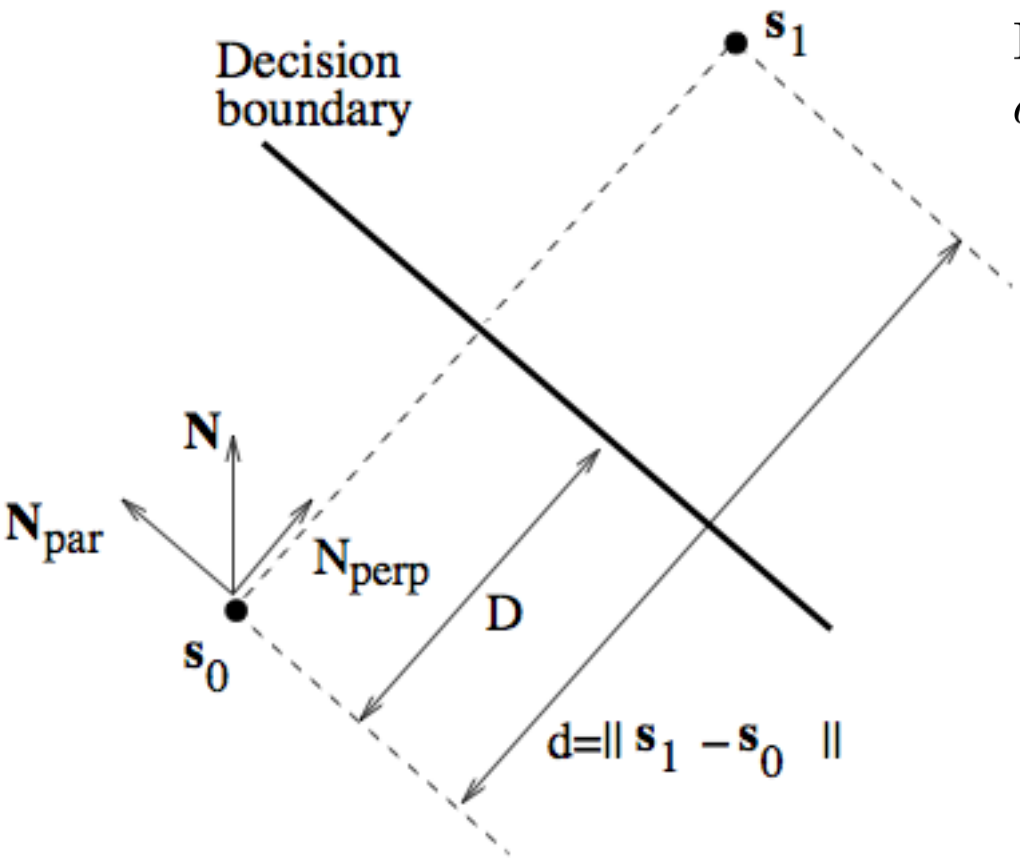
Noise component parallel to boundary cannot cause crossing, no matter how big it is.
In *n*-dimensional space, this component is *(n-1)*-dimensional

Boundary crossing determined by length and sign of noise component perpendicular to boundary. 1-dimensional, regardless of the signal space dimension

$$N_{perp} \sim N(0, \sigma^2)$$

$$\boxed{P[\text{cross a boundary at distance } D] = P[N_{perp} > D] = Q\left(\frac{D}{\sigma}\right)}$$

Geometry for binary signaling



Distance between signals
$d = ||\mathbf{s}_1 - \mathbf{s}_0||$

Distance from boundary
(starting from either signal)
$D = d/2 = ||\mathbf{s}_1 - \mathbf{s}_0||/2$

$$P[\text{cross ML boundary between } \mathbf{s}_0 \text{ and } \mathbf{s}_1] = Q\left(\frac{||\mathbf{s}_1 - \mathbf{s}_0||}{2\sigma}\right) = \boxed{Q\left(\frac{||s_1 - s_0||}{2\sigma}\right)}$$

(starting from either signal)

$$P_{e|0} = P_{e|1} = P_e$$

Important Note

- As the equivalence between CT signal s_i and vector $\mathbf{s}_i$ has been already established, we drop the boldface notation, using y, s_i, and n to denote the received signal, the transmitted signal, and the noise, respectively, in both the settings

Performance with binary signaling: *Algebraic*

- For the binary hypothesis test

$$H_0 : y(t) = n(t)$$
$$H_1 : y(t) = s(t) + n(t)$$

show that

$$P_e = P_{e|1} = P_{e|0} = Q\left(\frac{||s||}{2\sigma}\right) = Q\left(\frac{d}{2\sigma}\right)$$

- Moreover for following binary hypothesis test

$$H_0 : y(t) = s_0(t) + n(t)$$
$$H_1 : y(t) = s_1(t) + n(t)$$

show that

$$P_e = P_{e|1} = P_{e|0} = Q\left(\frac{||s_1 - s_0||}{2\sigma}\right) = Q\left(\frac{d}{2\sigma}\right)$$

Importance of Scale Invariance

- The error probability is given by

$$P_e = P_{e|1} = P_{e|0} = Q\left(\frac{||s_1 - s_0||}{2\sigma}\right) = Q\left(\frac{d}{2\sigma}\right)$$

- Same scaling in the signal and noise does not change the performance
- The performance depends on ratio rather than individually on the signal and noise strengths

Some Standard Measures

- Energy Bit E_b: For binary signaling it is given by

$$E_b = \frac{1}{2}(||s_0||^2 + ||s_1||^2)$$

assuming that 0 and 1 are equally likely

- Scale invariant parameter

$$\eta_p = \frac{d^2}{E_b}$$

Note that this does not change if we scale both s_1 and s_0 by A

- In terms of energy per bit and scale-invariant parameter, the probability of error for binary signaling is given by

$$\begin{aligned} P_e &= Q\left(\frac{d}{2\sigma}\right) \\ &= Q\left(\sqrt{\frac{\eta_p E_b}{2N_0}}\right) \end{aligned}$$

where $\sigma^2 = N_0/2$

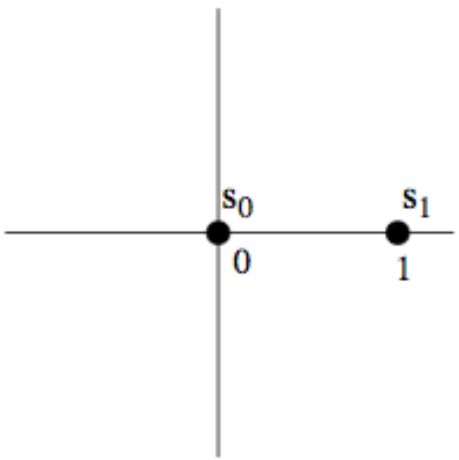
ML Binary Signaling Performance..

- In terms of energy per bit and scale-invariant parameter, the probability of error for binary signaling is given by

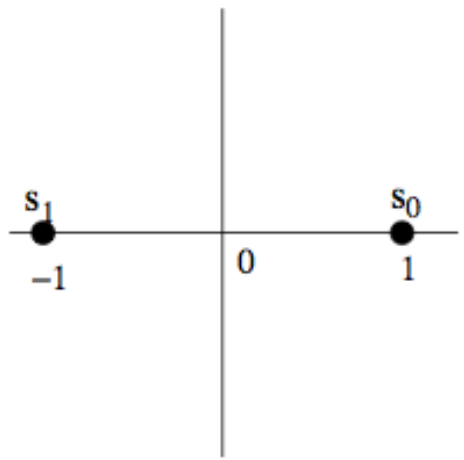
$$P_e = Q\left(\sqrt{\frac{\eta_p E_b}{2N_0}}\right)$$

- Important observations
  - Performance depends on the signal-to-noise ratio
  - For a fixed E_b/N_0, the performance is better for higher for higher value of η_p. The parameter η_p is also called as power efficiency of the constellation

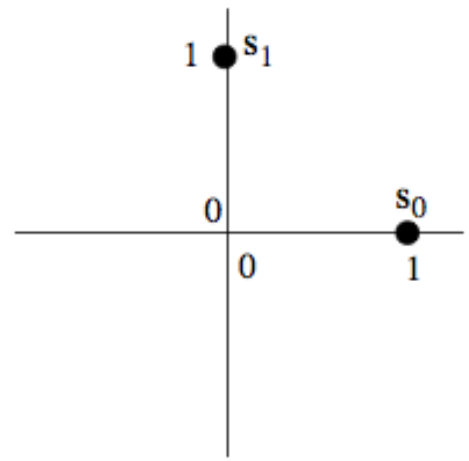
Performance for different binary schemes



On–off keying



Antipodal signaling



Equal energy, orthogonal signaling

$d = 1$

$E_b = (0^2 + 1^2)/2 = 1/2$

$\eta_P = d^2 / E_b = 2$

$$P_{e,ML} = Q\left(\sqrt{\frac{\eta_P E_b}{2N_0}}\right) = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$$

$d = 2$

$E_b = ((-1)^2 + 1^2)/2 = 1$

$\eta_P = d^2 / E_b = 4$

$$P_{e,ML} = Q\left(\sqrt{\frac{\eta_P E_b}{2N_0}}\right) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

$d = \sqrt{2}$

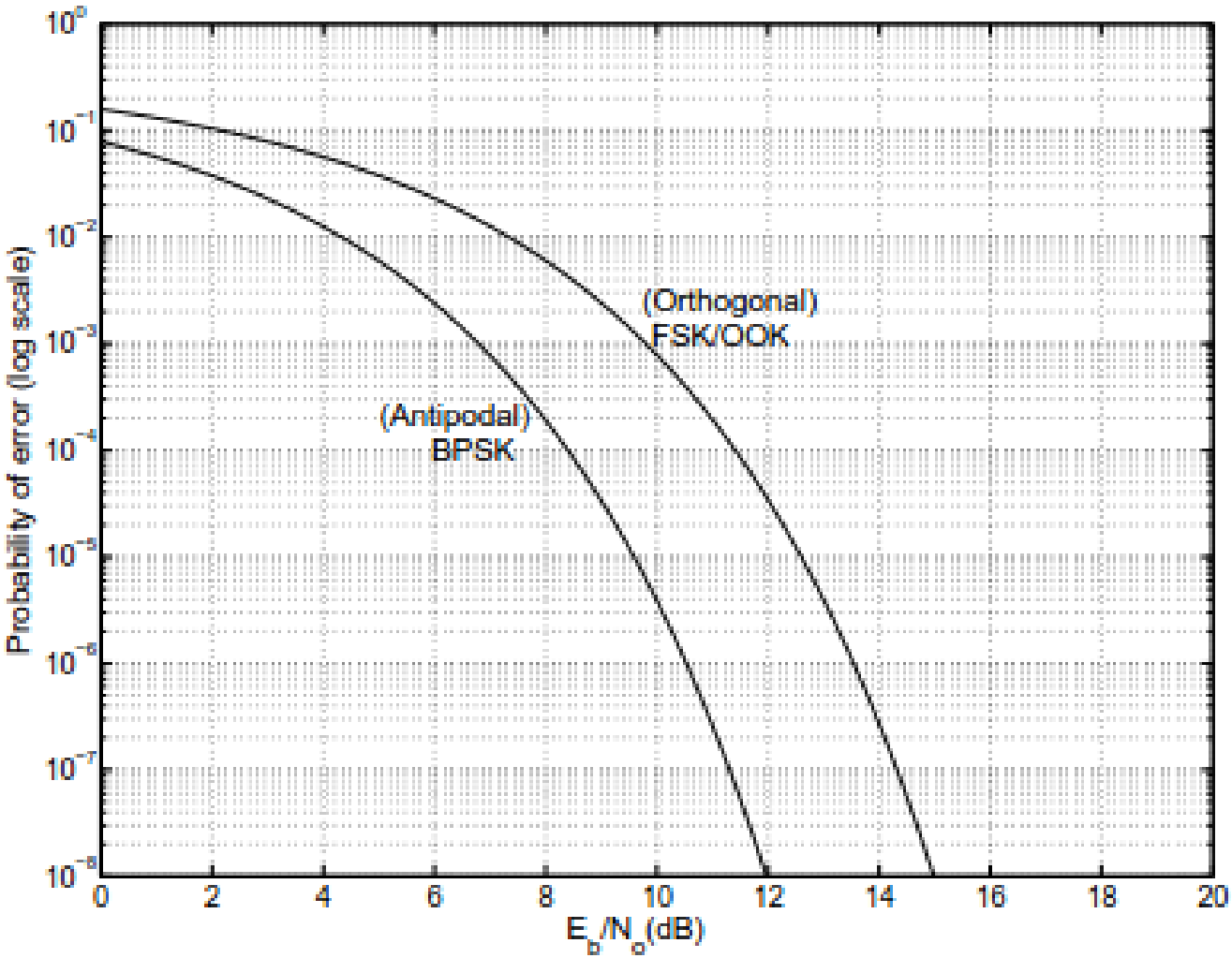
$E_b = (1^2 + 1^2)/2 = 1$

$\eta_P = d^2 / E_b = 2$

$$P_{e,ML} = Q\left(\sqrt{\frac{\eta_P E_b}{2N_0}}\right) = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$$

- OOK and equal energy orthogonal signaling have same power efficiency. Antipodal is 3 dB better.

Performance for different binary schemes..



- OOK and equal energy orthogonal signaling have same power efficiency. Antipodal is 3 dB better.

Performance analysis for larger constellations

M-ary signaling in AWGN

- For the multiple hypothesis testing problem given by

$$H_i : y(t) = s_i(t) + n(t)$$

where $i = 0, 1, \ldots, M-1$, the ML rule is given by

$$\delta_{\text{ML}} = \arg\max_i Z_i$$

with decision statistics

$$Z_i = \langle y, s_i \rangle - \frac{1}{2}||s_i||^2$$

and corresponding decision region is

$$\Gamma_i = \{y : Z_i > Z_j \text{ for all } \ j \neq i\}$$

- The conditional error probability, conditioned on H_i is given by

$$P_{e|i} = P[y \notin \Gamma_i | i \text{ sent}] = P[Z_i < Z_j \text{ for some } j \neq i \mid i \text{ sent}]$$

Some important parameters

- Energy per symbol is given by

$$E_s = \frac{1}{M}\sum_{i=1}^{M} ||s_i||^2$$

- Energy per bit is given by

$$E_b = \frac{E_s}{\log_2 M}$$

- Bandwidth efficiency

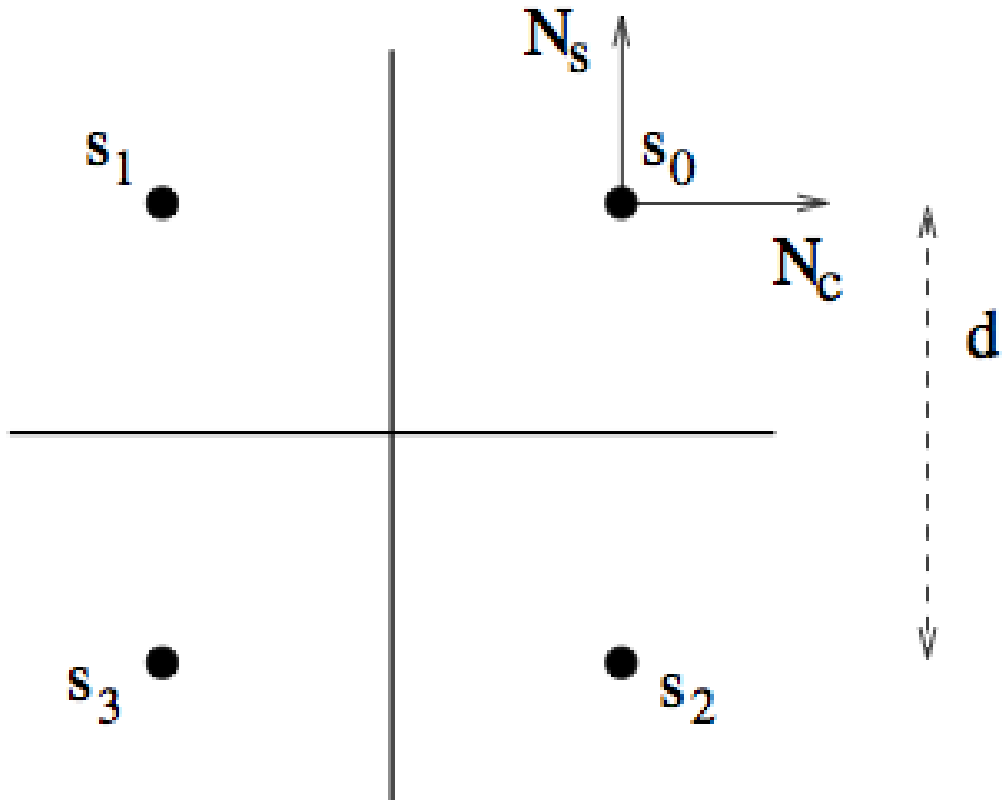
$$\begin{aligned}\eta_B &= \frac{R}{B_{\min}} = \frac{\text{Number of Bits per sec}}{\text{Min. Bandwidth}} \\ &= \frac{\text{Number of Bits per symbol} \times \text{Symbol Rate}}{\text{Number of Complex Dimensions} \times \text{Symbol Rate}} \\ &= \frac{\log_2 M}{N_D}\end{aligned}$$

- Show that performance of M-ary signaling depends only on E_b/N_0 and on constellation shape

Exact Error Probability of QPSK

- Show that the exact error probability of QPSK constellation shown below is given by

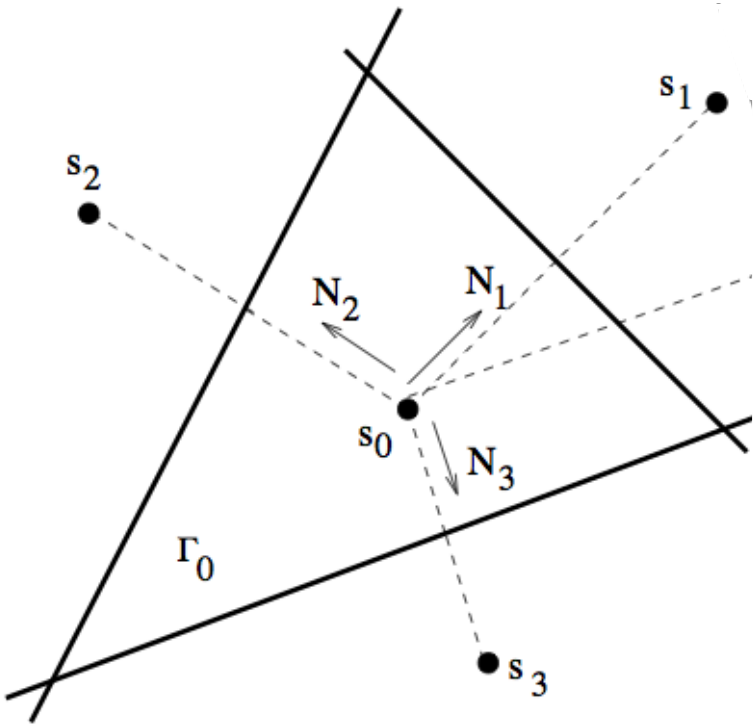
$$P_e = P_{e|1} = 2Q\left(\sqrt{\frac{2E_b}{N_0}}\right) - 2Q^2\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

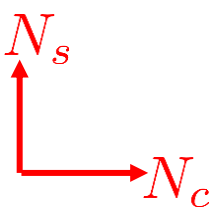


Performance analysis for larger constellations:

Union Bounds and Variants

Exact analysis might be difficult!





- Conditioned on s_0 sent, error occurs if we cross one of the 3 boundaries

$$P_{e|0} = P\left[N_1 > \frac{||s_1 - s_0||}{2} \text{ or } N_2 > \frac{||s_2 - s_0||}{2} \text{ or } N_3 > \frac{||s_3 - s_0||}{2}\right]$$

- Problem: The three noise variances, which cause crossing of boundary, are not independent! $P(A \cap B \cap C) \neq P(A)P(B)P(C)$

- Exact evaluation would require integrating a 3D Gaussian density function!

Approximation using Union Bound

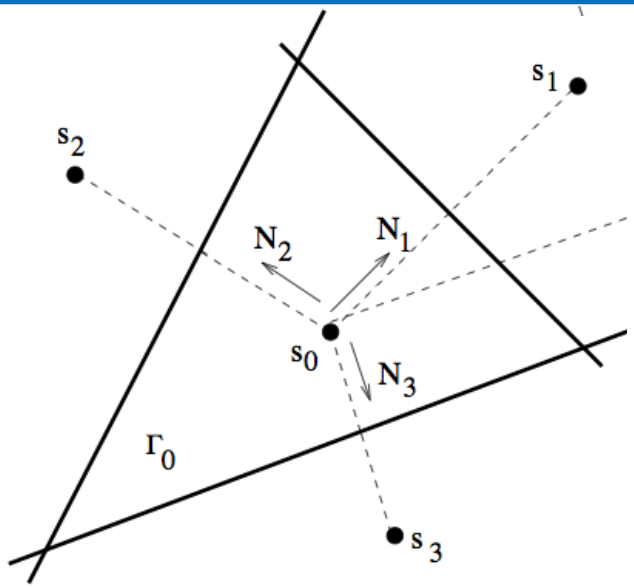
- Union bound is given by

  $$P\left[A_1 \text{ or } A_2 \text{ or } \ldots \text{ or } A_2\right] = P[A_1 \cup A_2 \ldots \cup A_n] \leq P[A_1] + P[A_2] + \ldots + P[A_n]$$

  The bound is satisfied with equalilty only for mutually exclusive events

- This allows us to estimate the union (OR) of events which need not be independent!

Applying the union bound



- Conditioned on s_0 sent, error occurs if we cross one of the 3 boundaries

$$P_{e|0} = P\left[N_1 > \frac{||s_1 - s_0||}{2} \text{ or } N_2 > \frac{||s_2 - s_0||}{2} \text{ or } N_3 > \frac{||s_3 - s_0||}{2}\right]$$

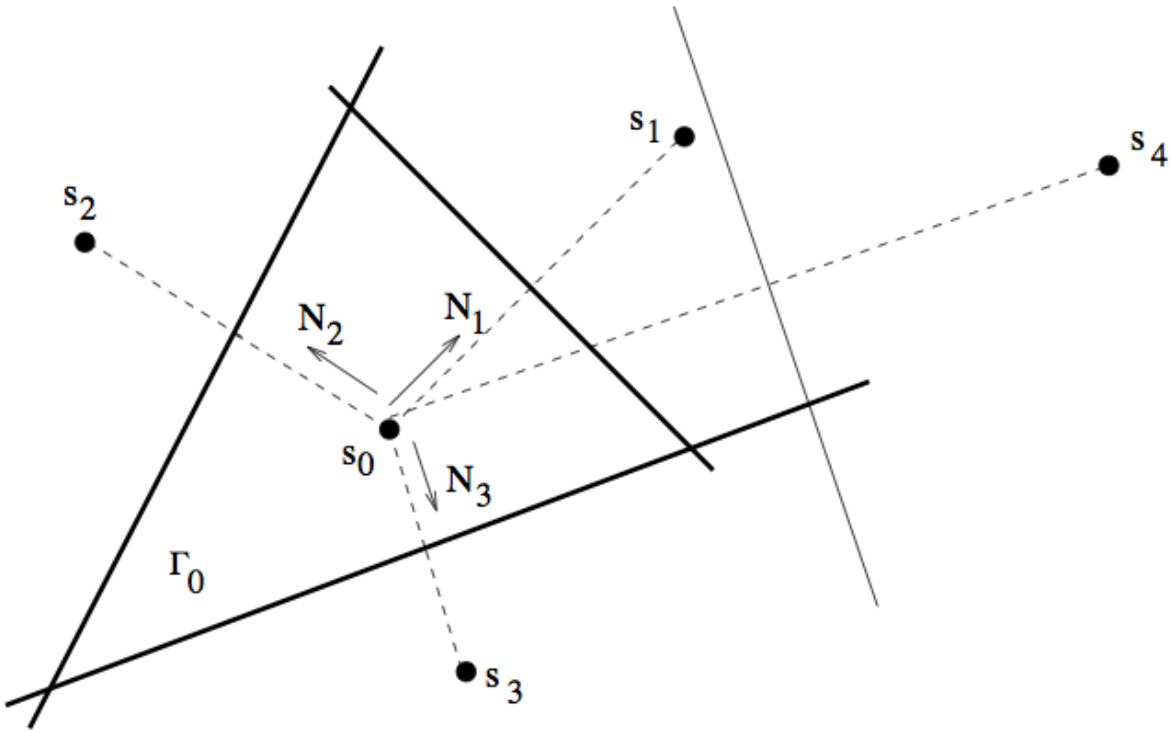
$$\leq P\left[N_1 > \frac{||s_1 - s_0||}{2}\right] + P\left[N_2 > \frac{||s_2 - s_0||}{2}\right] + P\left[N_3 > \frac{||s_3 - s_0||}{2}\right]$$

$$= Q\left(\frac{||s_1 - s_0||}{2\sigma}\right) + Q\left(\frac{||s_2 - s_0||}{2\sigma}\right) + Q\left(\frac{||s_3 - s_0||}{2\sigma}\right)$$

Simple evaluation using Q-functions

Exact Analysis for higher dimensionals difficult!

Can’t draw pictures like this in a 20-dimensional signal space!



General Union Bound for M-ary Signal

- Conditional error probability is a union of events involving pairwise comparison

$$P_{e|i} = P\left(\cup_{j \neq i}\{Z_i < Z_j\} \mid i \text{ sent}\right)$$

  Error occurs if the decision statistic for the right hypothesis loses to any other decision statistic.

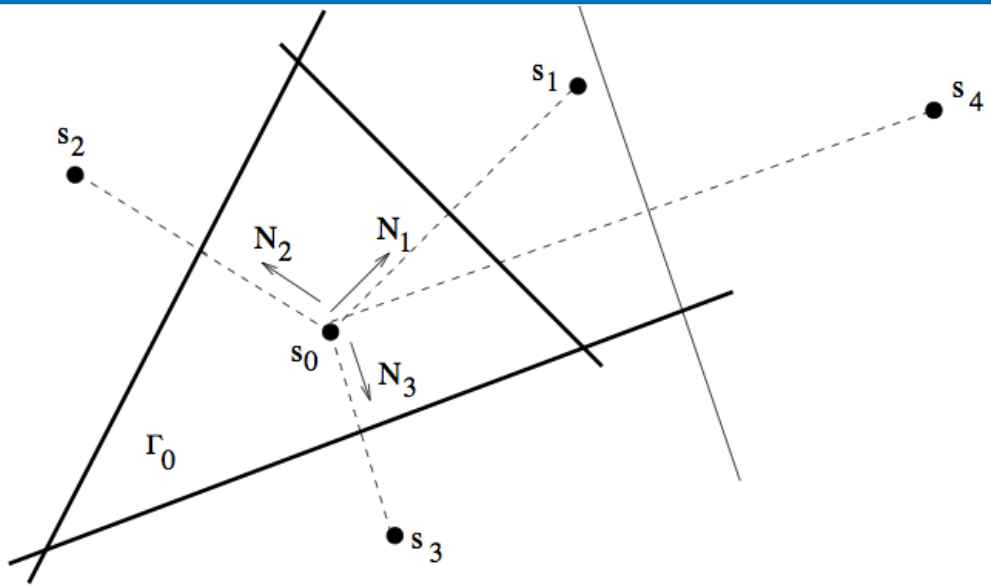
- Note that union bound is sum of pairwise error probabilities

$$P_{e|i} \leq \sum_{j \neq i} P[Z_i < Z_j \mid i \text{ sent}]$$

  where the error probability for binary hypothesis test: H_i vs H_j

$$P[Z_i < Z_j \mid i \text{ sent}] = Q\left(\frac{||s_j - s_i||}{2\sigma}\right)$$

Applying the general union bound



- Conditioned on s_0 sent, error occurs if we cross one of the 4 boundaries

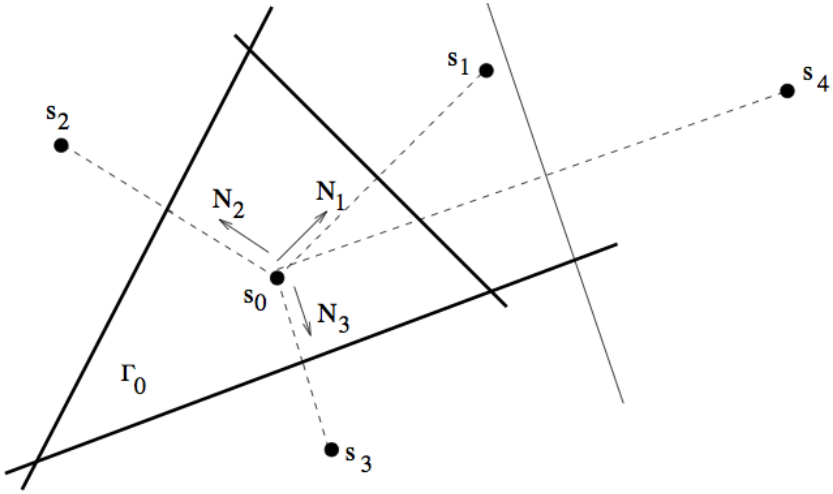
$$P_{e|0} \leq \boxed{Q\left(\frac{||s_1 - s_0||}{2\sigma}\right) + Q\left(\frac{||s_2 - s_0||}{2\sigma}\right) + Q\left(\frac{||s_3 - s_0||}{2\sigma}\right)} + \boxed{Q\left(\frac{||s_4 - s_0||}{2\sigma}\right)}$$

Intelligent union bound

Intelligent Union Bound

- Denote by $N_{\mathrm{NL}}(i)$ the indices of the set of neighbors of signal s_i that characterizes the ML decision region Γ_i.

- For example below, $N_{\mathrm{NL}}(i) = \{1, 2, 3\}$

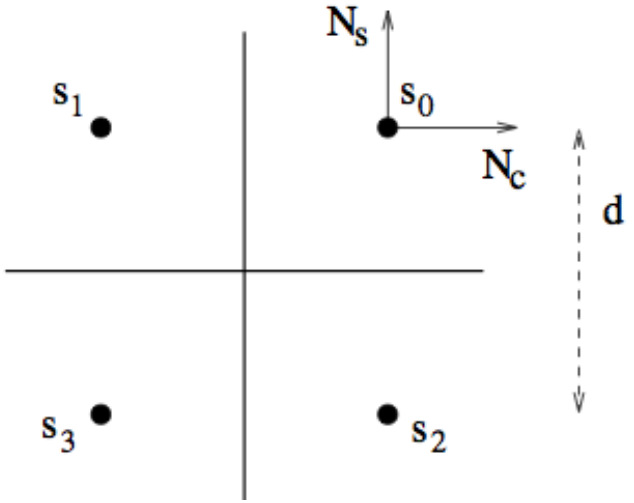


- Intelligent union bound in general is given by

$$P_{e|i} \leq \sum_{j \in N_{\mathrm{ML}}(i)} Q\left(\frac{||s_j - s_i}{2\sigma}\right)$$

Example: QPSK

- Find union bound and intelligent union bound for QPSK



- Union bound for QPSK is given by

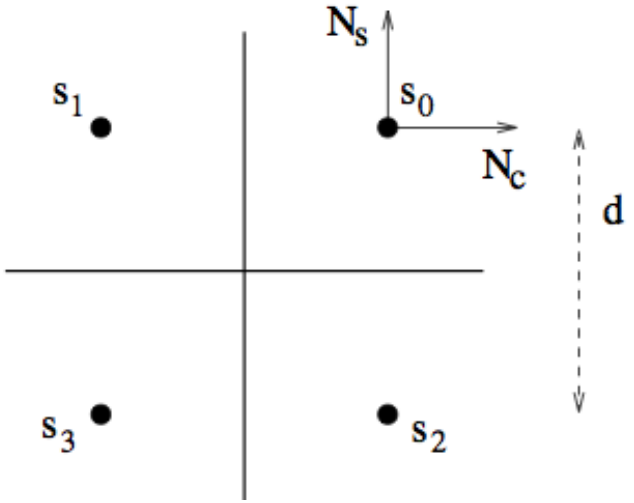
$$P_e = P_{e|0} \leq Q\left(\frac{d_{01}}{2\sigma}\right) + Q\left(\frac{d_{02}}{2\sigma}\right) + Q\left(\frac{d_{03}}{2\sigma}\right) = 2Q\left(\frac{d}{2\sigma}\right) + \boxed{Q\left(\frac{\sqrt{2}d}{2\sigma}\right)}$$

- Using $\frac{d^2}{E_b} = \frac{d^2 \log_2 M}{E_s} = 4$ with $E_s = 2(\frac{d}{2})^2 = \frac{d^2}{2}$, we get

$$P_e \leq 2Q\left(\sqrt{\frac{2E_b}{N_0}}\right) + \boxed{Q\left(\sqrt{\frac{4E_b}{N_0}}\right)}$$

Example: QPSK ..

- Find union bound and intelligent union bound for QPSK



- Intelligent union bound for QPSK is given by

$$P_e = P_{e|0} \leq Q\left(\frac{d_{01}}{2\sigma}\right) + Q\left(\frac{d_{02}}{2\sigma}\right) = 2Q\left(\frac{d}{2\sigma}\right)$$

- Using $\frac{d^2}{E_b} = \frac{d^2 \log_2 M}{E_s} = 4$ with $E_s = 2(\frac{d}{2})^2 = \frac{d^2}{2}$, we get

$$P_e \leq 2Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

Nearest Neighbor Approximation

- Rationale: $Q(x)$ decays rapidly so that the terms in the union bound corresponding to the smallest arguments for the Q dominate at high SNR.

$$P_{e|i} \approx N_{d_{\min}}(i) Q\left(\frac{d_{\min}}{2\sigma}\right)$$

- By averaging over i, we obtain that

$$P_e \approx \bar{N}_{d_{\min}} Q\left(\frac{d_{\min}}{2\sigma}\right) = \bar{N}_{d_{\min}} Q\left(\sqrt{\frac{\eta_p E_b}{2N_0}}\right)$$

where $\bar{N}_{d_{\min}}$ denotes the average number of nearest neighbors for a signal point and is given by

$$\bar{N}_{d_{\min}} = \frac{1}{M} \sum_{i=1}^{M} N_{d_{\min}}(i)$$

is the average number of nearest neighbors for the signal points in the constellation.

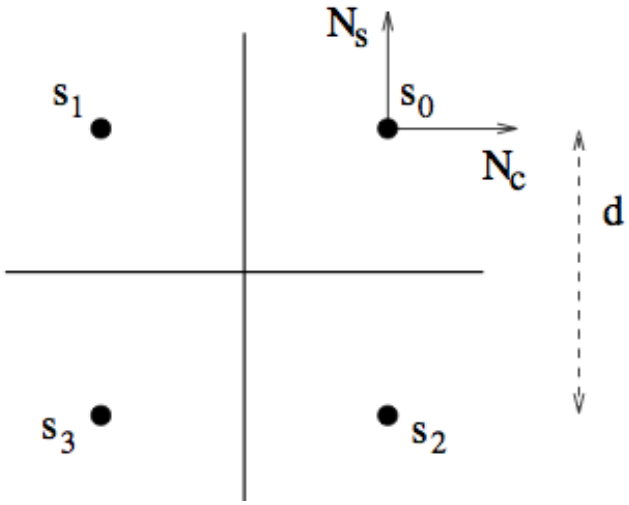
Example: QPSK ...

- For QPSK

$$N_{d_{\min}}(i) = 2 = \bar{N}_{d_{\min}}$$

so that

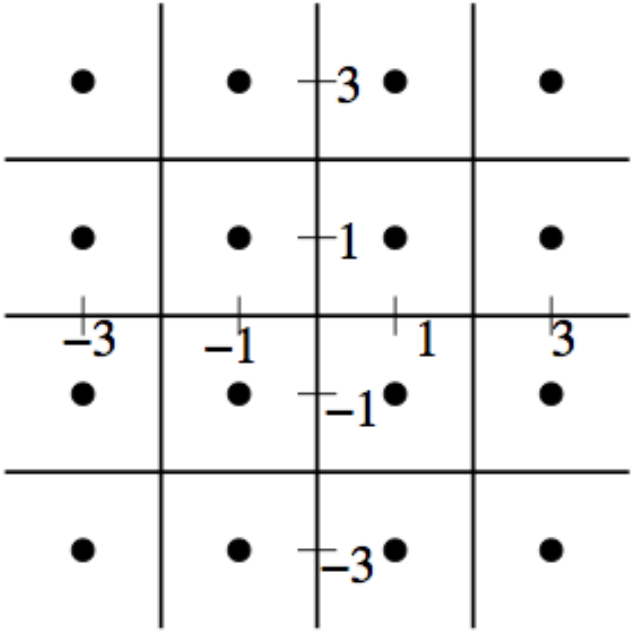
$$P_e \approx 2Q\left(\frac{d_{\min}}{2\sigma}\right)$$



- Using $\frac{d^2}{E_b} = \frac{d^2 \log_2 M}{E_s} = 4$ with $E_s = 2(\frac{d}{2})^2 = \frac{d^2}{2}$, we get

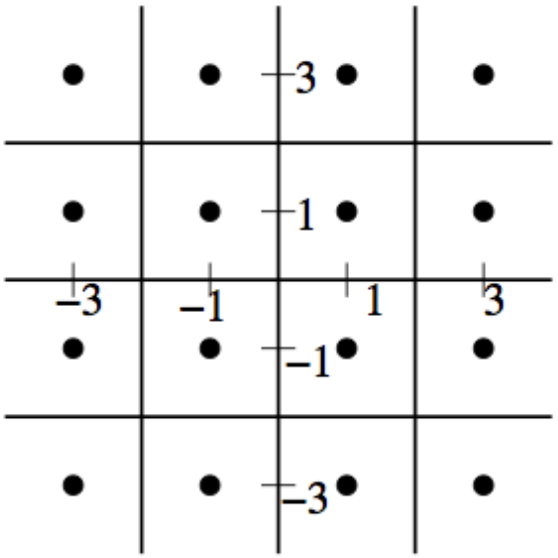
$$P_e \leq 2Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

Performance Analysis of 16 QAM



- Intelligent union bound coincides with nearest neighbors approximation since nearest neighbors define decision boundaries
- Need to compute power efficiency and average number of nearest neighbors
- Since we are computing scale invariant quantities, choose any convenient scaling

16QAM performance…

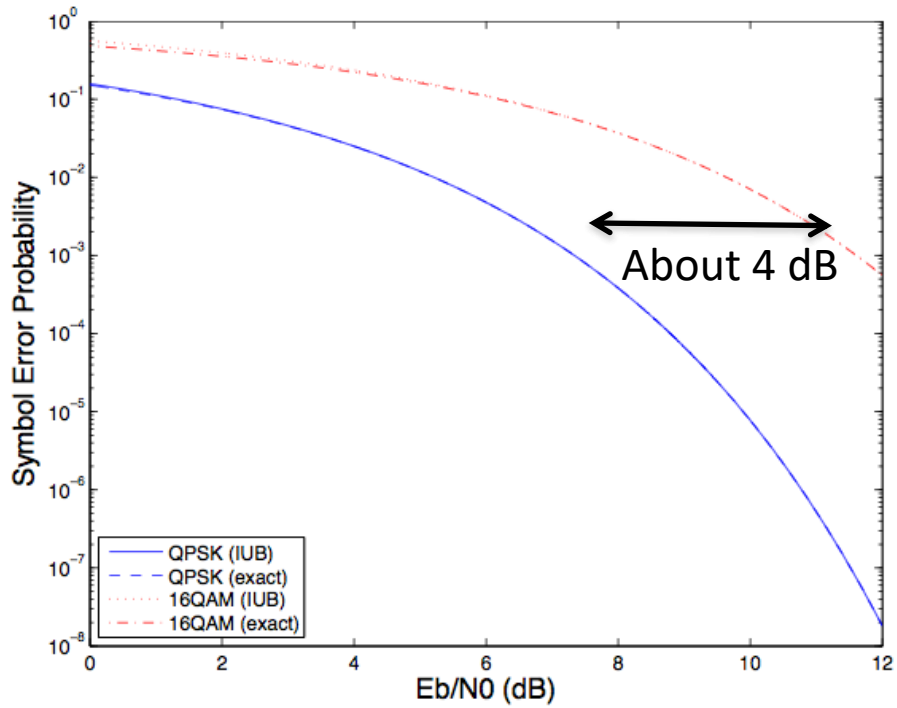


- Using nearest neighbor approximation, prove that the probability of error for 16 QAM is given by

$$P_e \approx 3Q\left(\sqrt{\frac{4E_b}{5N_0}}\right)$$

QPSK vs 16QAM

- Find bandwidth efficiency of 4 QPSK and 16 QAM
- Find power efficiency of 4 QPSK and 16 QAM
- Which out of the two will perform better for same E_b/N_0?



Note: Intelligent union bound is very close to the exact error prob in both cases

Performance Analysis of M-ary orthogonal modulation

- Using union bound, show that the probability of error for M-ary orthogonal modulation can be approximated by

$$P_e = (M-1)Q\left(\sqrt{\frac{E_b \log_2 M}{N_0}}\right)$$

Included in course

- Exact probability of correct reception for M-ary orthogonal signaling

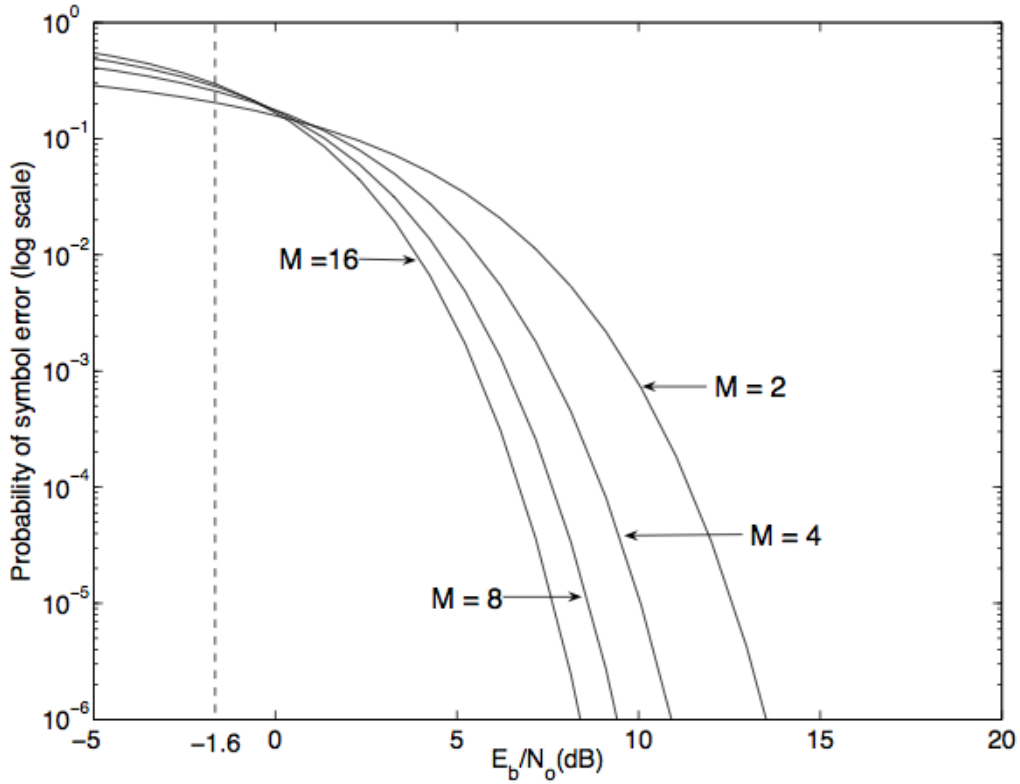
Not included in course

$$P_c = P_{c|i} = \int_{-\infty}^{\infty} [\Phi(x)]^{M-1} \frac{1}{\sqrt{2\pi}} exp(-\frac{(x-m)^2}{2})dx$$

where $\Phi(x)$ is the CDF of the standard Gaussian random variable while

$$m = \sqrt{\frac{2E_s}{N_0}} = \sqrt{\frac{2E_b \log_2 M}{N_0}}$$

Symbol Error Probability for M-ary Orthogonal

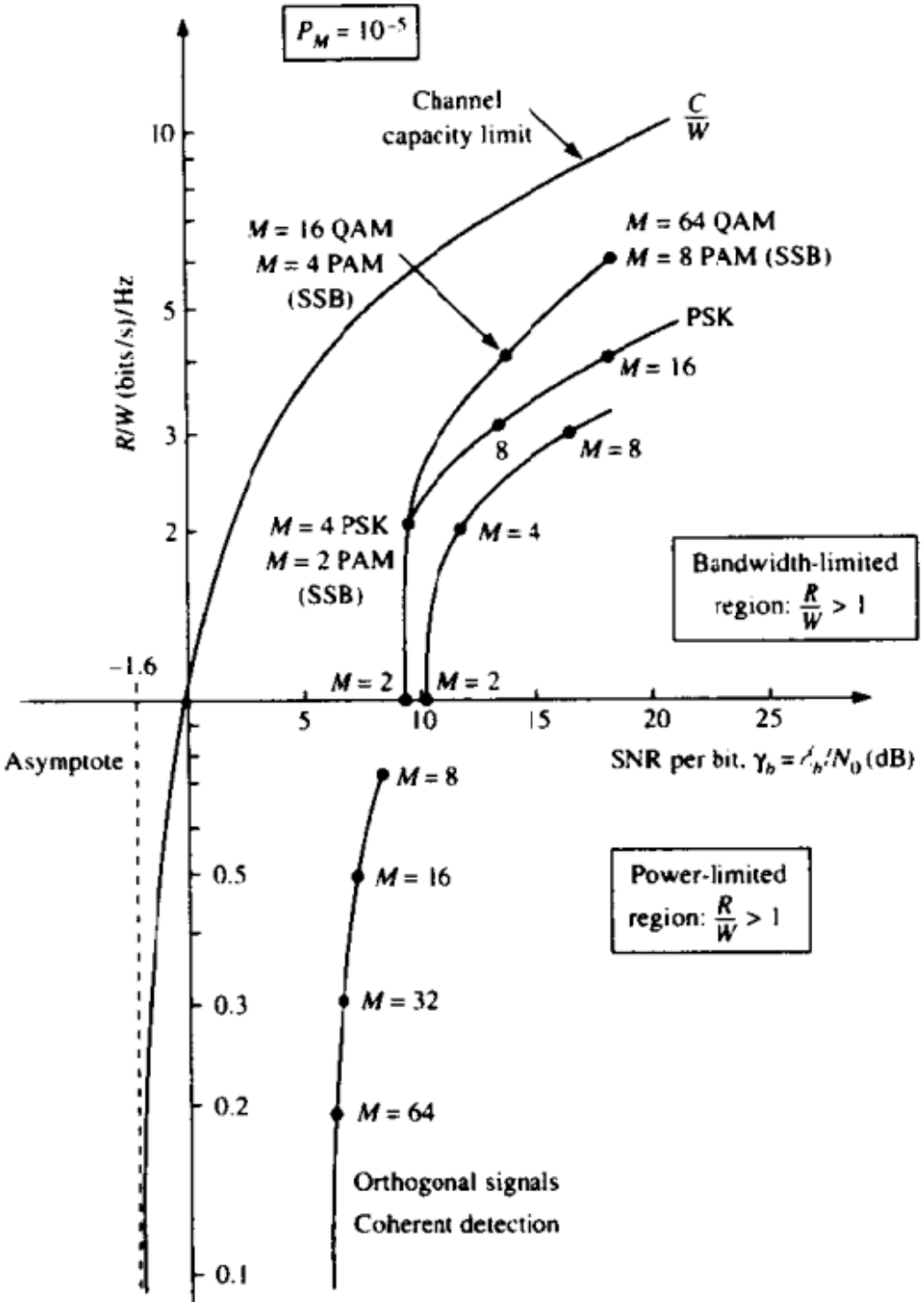


- Performance improves as M increases!
- Asymptotic performance

$$\lim_{M\to\infty} P_e = \begin{cases} 0, & E_b/N_0 > \ln 2 \\ 1, & E_b/N_0 < \ln 2 \end{cases}$$

Power-Bandwidth Tradeoff for M-ary orthogonal!

- Power efficiency $= \log_2 M$
- Bandwidth effificiency is $\eta_B = \frac{\text{No. of Bits per symbol}}{\text{Number of Complex Dimensions}} = \frac{\log_2 M}{M}$
- Trade-off between power and bandwidth efficiency!


$P_M = 10^{-5}$
Channel capacity limit
$\frac{C}{W}$
M = 16 QAM
M = 4 PAM (SSB)
M = 64 QAM
M = 8 PAM (SSB)
PSK
M = 16
8
M = 8
M = 4 PSK
M = 2 PAM (SSB)
M = 4
Bandwidth-limited region: $\frac{R}{W} > 1$
M = 2
M = 2
R/W (bits/s)/Hz
10
5
3
2
−1.6
5
10
15
20
25
SNR per bit, $\gamma_b = \mathcal{E}_b/N_0$ (dB)
Asymptote
M = 8
M = 16
M = 32
M = 64
0.5
0.3
0.2
0.1
Power-limited region: $\frac{R}{W} > 1$
Orthogonal signals
Coherent detection