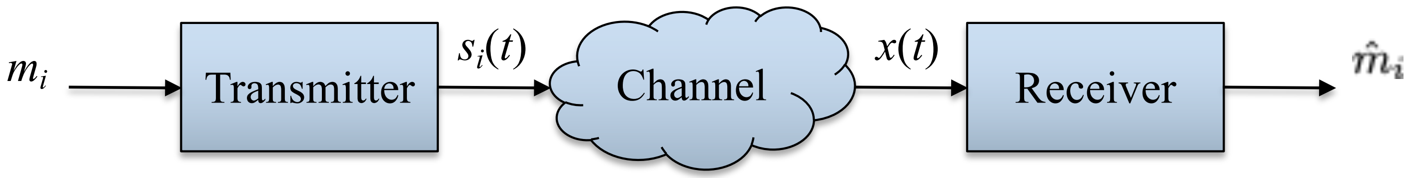
Part 9 Signal-Space Analysis

Introduction



- Statistical model for a genetic digital communication system
  - *Message source*: *A priori* probabilities for information source

$$p_i = P(m_i) \text{ for } i = 1,2,\ldots,M$$

  - *Transmitter*: The transmitter takes the message source output m_i and (en-)codes it into a distinct signal $s_i(t)$ suitable for transmission over the channel. So:

$$p_i = P(m_i) = P(s_i(t)) \text{ for } i = 1,2,\ldots,M$$

Introduction

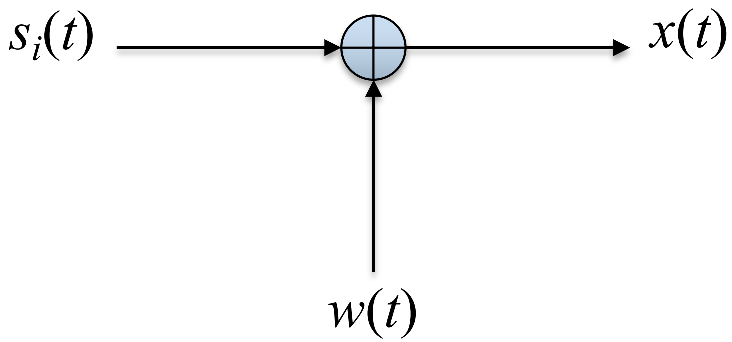
- $s_i(t)$ must be a real-valued *energy signal* (i.e., a signal with finite energy) with duration T.

$$E_i = \int_0^T s_i^2(t)dt < \infty.$$

- *Channel*: The channel is assumed *linear* and with a bandwidth wide enough to pass $s_i(t)$ with no distortion.
  - A zero-mean additive white Gaussian noise (AWGN) is also assumed to facilitate the analysis.

A Mathematical Model

- We can simplify the previous system block diagram to:



- Upon the reception of $x(t)$ with a duration of T, the receiver makes the *best* estimate of m_i. (We haven't defined what "the best" means.)

Criterion for the "Best" Decision

- Best = Minimization of the average probability of symbol error.

$$P_e = \sum_{i=1}^{M} p_i \cdot P(\hat{m} \neq m_i \mid m_i)$$

  - It is *optimum in the minimum-probability-of-error* sense.
  - Based on this criterion, we begin to design the receiver that can give the best decision.

Geometric Representation of Signals

- Signal space concept
  - **Vectorization** of the (discrete or continuous) signals removes the redundancy in signals, and provides a compact representation for them.
  - Determination of the vectorization basis
    - Gram-Schmidt orthogonalization procedure

Gram-Schmidt Orthogonalization Procedure

Given $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$, how to find an orthonormal basis for them?

(step *i*) Let $\vec{u}_1 = \dfrac{\vec{v}_1}{\|\vec{v}_1\|}$.

(step *ii*) $\vec{u}_2^{'} = \vec{v}_2 - (\vec{v}_2 \cdot \vec{u}_1)\vec{u}_1$. Set $\vec{u}_2 = \dfrac{\vec{u}_2^{'}}{\|\vec{u}_2^{'}\|}$.

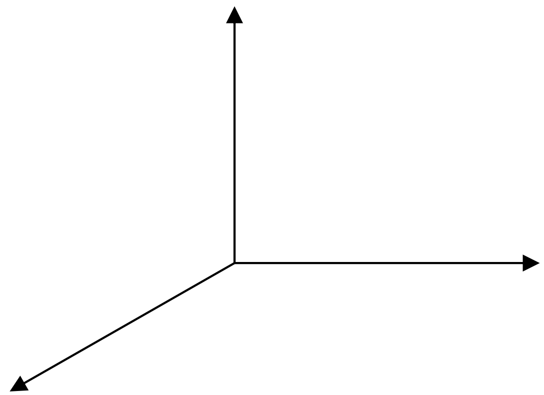
(step *iii*) For $i = 3, 4, \ldots,$

Let $\vec{u}_i^{'} = \vec{v}_i - (\vec{v}_i \cdot \vec{u}_{i-1})\vec{u}_{i-1} - (\vec{v}_i \cdot \vec{u}_{i-2})\vec{u}_{i-2} - \cdots - (\vec{v}_i \cdot \vec{u}_1)\vec{u}_1$.

Set $\vec{u}_i = \dfrac{\vec{u}_i^{'}}{\|\vec{u}_i^{'}\|}$.

(step *iv*) Then $\left(\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_k\right)$ forms an orthonormal basis.

Properties:



(i) vector: $\vec{v} = (v_1, \ldots, v_n)$

(ii) inner product: $\vec{v}_1 \cdot \vec{v}_2 = \sum_{i=1}^{n} v_{i1} v_{i2}$

(iii) orthogonal, if inner product = 0

(iv) norm: $\|\vec{v}\| = \sqrt{v_1^2 + \cdots + v_n^2}$

(v) orthonormal, if inner product = 0, and individual norm = 1

(vi) linearly independent, if none can be represented as a linear combination of others

(vii) triangle inequality: $\|\vec{v}_1 + \vec{v}_2\| \leq \|\vec{v}_1\| + \|\vec{v}_2\|$

(viii) Cauchy Schwartz inequality:

$$|\vec{v}_1 \cdot \vec{v}_2| \leq \|\vec{v}_1\| \cdot \|\vec{v}_2\|$$

with equality holding if $\vec{v}_1 = a\vec{v}_2$

(xi) Norm square:

$$\|\vec{v}_1 + \vec{v}_2\|^2 = \|\vec{v}_1\|^2 + \|\vec{v}_2\|^2 + 2\vec{v}_1 \cdot \vec{v}_2$$

(x) Pythagorean property: If orthogonal,

$$\|\vec{v}_1 + \vec{v}_2\|^2 = \|\vec{v}_1\|^2 + \|\vec{v}_2\|^2$$

(xi) matrix transformation w.r.t. matrix *A*:

$$\vec{v}_1 = A\vec{v}_2$$

(xii) eigenvalues w.r.t. matrix *A*:

$$\text{solution } \lambda \text{ of } \det[A - \lambda I] = 0$$

(xiii) eigenvectors w.r.t. eigenvalue λ :

$$\text{solution } \vec{v} \text{ of } A\vec{v} = \lambda\vec{v}$$

Signal Space Concept for Continuous Functions

Properties for continuous functions

(i) (complex-valued) signal: $z(t)$

(ii) inner product: $\langle z(t), \hat{z}(t)\rangle = \int_a^b z(t)\hat{z}^*(t)dt$

(iii) orthogonal, if inner product = 0

(iv) norm: $\|z(t)\| = \sqrt{\int_a^b |z(t)|^2 dt}$

(v) orthonormal, if inner product = 0, and individual norm = 1

(vi) linearly independent, if none can be represented as a linear combination of others

(vii) triangle inequality: $\|z(t) + \hat{z}(t)\| \le \|z(t)\| + \|\hat{z}(t)\|$

(viii) Cauchy Schwartz inequality:

$$|\langle z(t), \hat{z}(t)\rangle| \le \|z(t)\| \cdot \|\hat{z}(t)\|$$

with equality holding if $z(t) = a \cdot \hat{z}(t)$, where a is a complex number

(xi) norm square:

$$\|z(t) + \hat{z}(t)\|^2 = \|z(t)\|^2 + \|\hat{z}(t)\|^2 + \langle z(t), \hat{z}(t)\rangle + \langle \hat{z}(t), z(t)\rangle$$

(x) Pythagorean property: If orthogonal,

$$\|z(t) + \hat{z}(t)\|^2 = \|z(t)\|^2 + \|\hat{z}(t)\|^2$$

(xi) transformation w.r.t. a function $C(t,\tau)$:

$$\hat{z}(t) = \int_a^b C(t,\tau) z(\tau) d\tau$$ (Recall $v_{1j} = \sum_{i=1}^n a_{ji} v_{2i}$.)

(xii.a) eigenvalues and eigenfunctions w.r.t. a function $C(t,\tau)$:

solution λ_k and $\{\phi_k(t)\}_{k=1}^{\infty}$ of $\lambda_k \cdot \phi_k(t) = \int_a^b C(t,\tau)\phi_k(\tau) d\tau$
and $C(t,\tau)$ can be represented as

$$C(t,\tau) = \sum_{k=1}^{\infty} \phi_k(t) \cdot \lambda_k \cdot \phi_k^*(\tau)$$

(xii.b) Give a deterministic function $\{s(t), t \in [0, T)\}$ and a set of orthonormal basis $\{\psi_k(t)\}_{1 \le k < \infty}$ that can span $s(t)$. Then

$$s(t) = \sum_{k=1}^{\infty} a_k \psi_k(t) \; 0 \le t < T,$$

where $a_k = \int_0^T s(t)\psi_k^*(t)dt.$

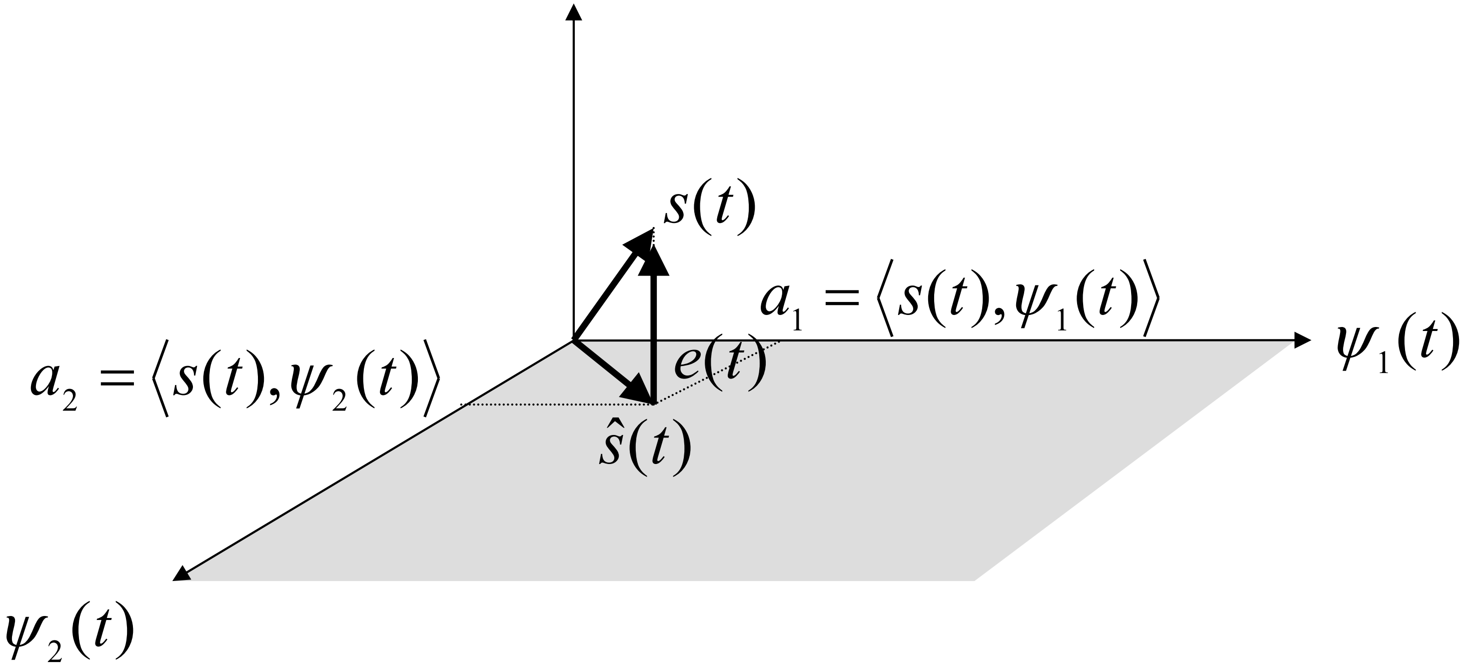
(xii.c) If orthonormal set $\{\psi_k(t)\}_{1 \le k \le K}$ does not span the space, then it is possible that $\hat{s}(t) = \sum_{k=1}^{K} a_k \psi_k(t) \neq s(t)$ for all choices of $\{a_k\}_{1 \le k \le K}$.

☐ *Problem : How to minimize the "energy" of* $e(t) = s(t) - \hat{s}(t)$?

$$
\begin{aligned}
&\int_{-\infty}^{\infty} |e(t)|^2 dt \\
&= \int_{-\infty}^{\infty} \left[s(t) - \sum_{k=1}^{K} a_k \psi_k(t) \right] \left[s(t) - \sum_{k=1}^{K} a_k \psi_k(t) \right]^* dt \\
&= \sum_{k=1}^{K} |a_k|^2 - \sum_{k=1}^{K} a_k \int_{-\infty}^{\infty} \psi_k(t) s^*(t) dt - \sum_{k=1}^{K} a_k^* \int_{-\infty}^{\infty} \psi_k^*(t) s(t) dt + \int_{-\infty}^{\infty} |s(t)|^2 dt \\
&= \sum_{k=1}^{K} \left| a_k - \int_{-\infty}^{\infty} s(t) \psi_k^*(t) dt \right|^2 + \int_{-\infty}^{\infty} |s(t)|^2 dt - \sum_{k=1}^{K} \left| \int_{-\infty}^{\infty} s(t) \psi_k^*(t) dt \right|^2
\end{aligned}
$$

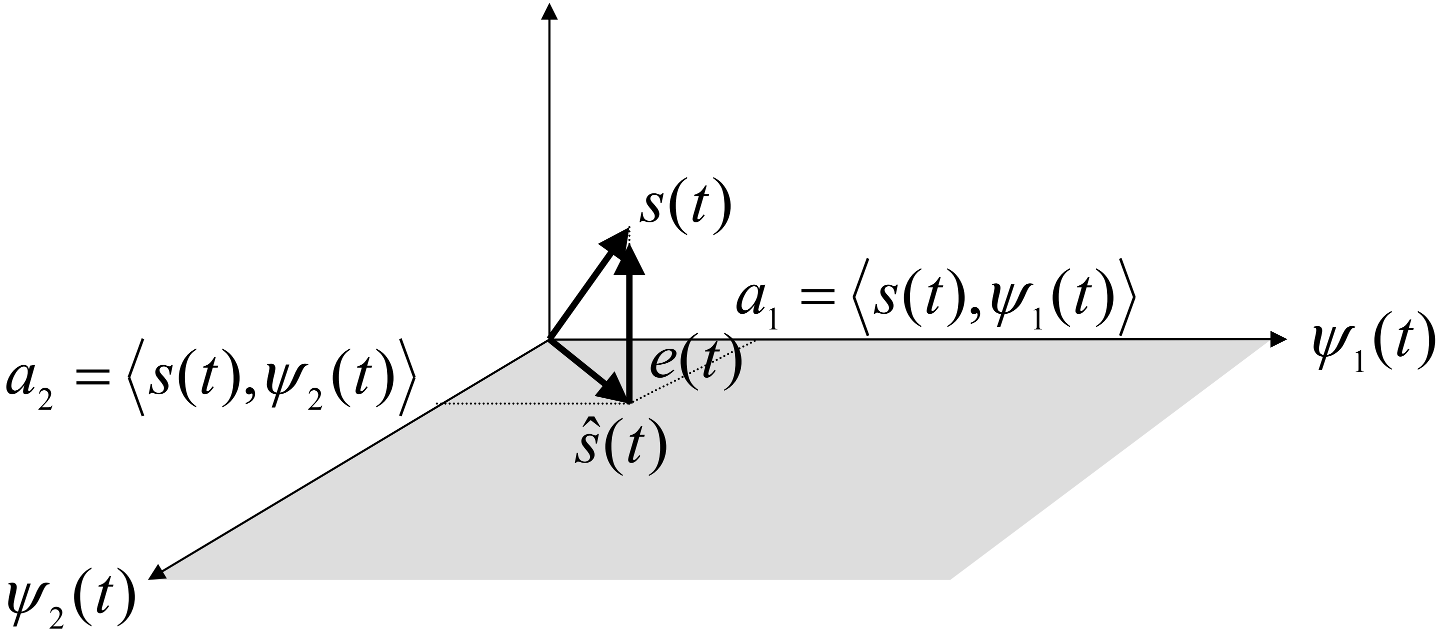
$$
\Rightarrow a_k = \int_{-\infty}^{\infty} s(t) \psi_k^*(t) dt
$$

Q.E.D.



Hence, $\langle e(t), \hat{s}(t) \rangle = \int_{-\infty}^{\infty} e(t)\hat{s}^*(t)dt = 0.$

- Interpretation
  - a_j is the projection of $s(t)$ onto the $\Psi_j(t)$-axis.
  - $|a_j|^2$ is the energy-projection of $s(t)$ onto the $\Psi_j(t)$-axis.



Slide 9-13 also yields:

$$\begin{aligned}\int_{-\infty}^{\infty} |e(t)|^2 dt &= \int_{-\infty}^{\infty} |s(t)|^2 dt - \sum_{k=1}^{K} \left| \int_{-\infty}^{\infty} s(t)\psi_k^*(t) dt \right|^2 \\ &= \int_{-\infty}^{\infty} |s(t)|^2 dt - \sum_{k=1}^{K} \left|a_k\right|^2\end{aligned}$$

and $\int_{-\infty}^{\infty} |\hat{s}(t)|^2 dt = \sum_{k=1}^{K} |a_k|^2.$

Signal Space Concept for Continuous Functions

- For simplicity, we now focus on real-valued functions.
- Completeness
  - If every finite energy signal $s(t)$ satisfies

$$\int_{-\infty}^{\infty} s^2(t)dt = \sum_{k=1}^{K} a_k^2$$

$\{\psi_k(t)\}_{k=1}^{K}$ is a *complete* orthonormal set.

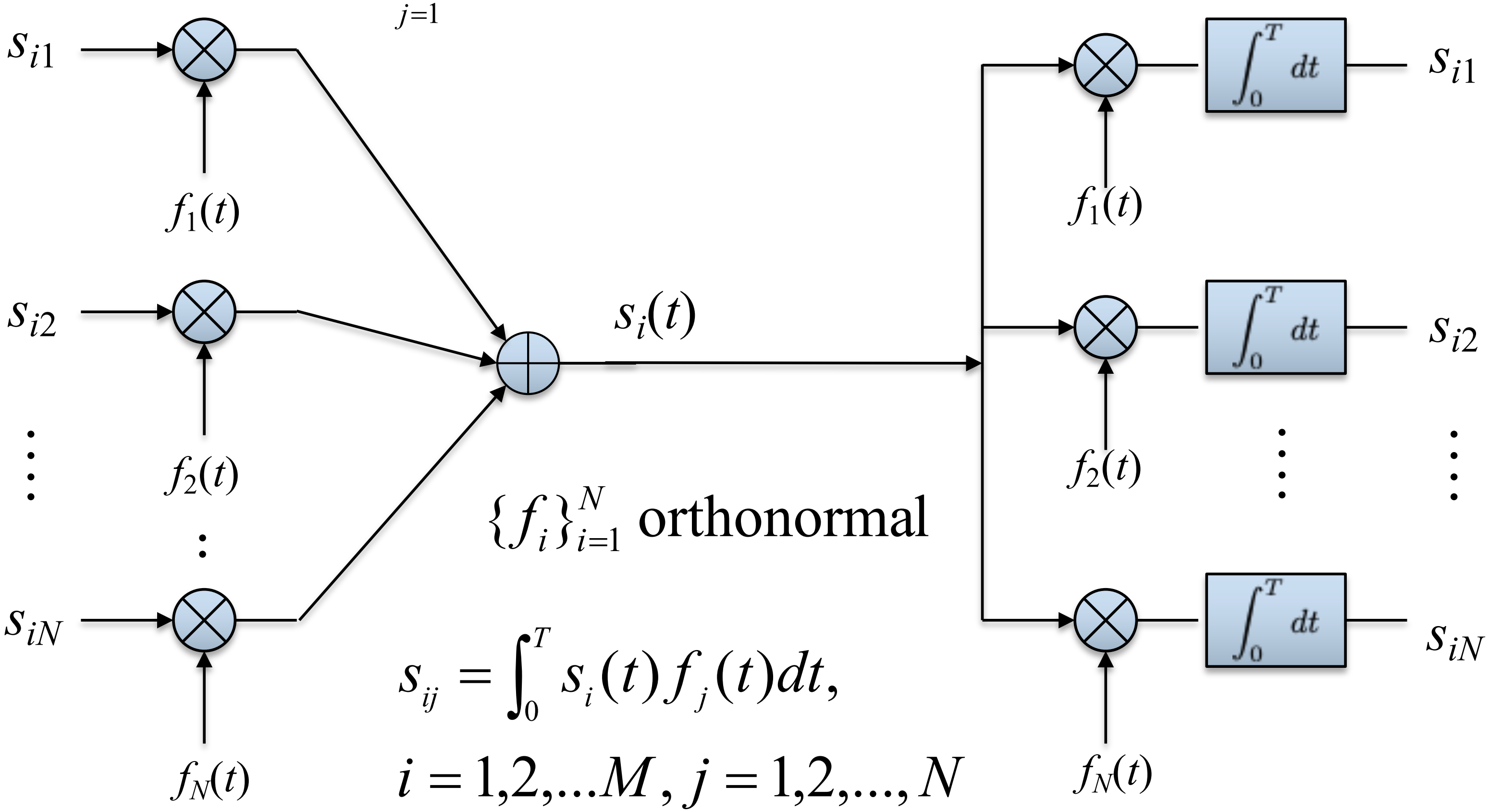
  - Example. Fourier series

$$\left\{\sqrt{\frac{2}{T}}\cos\left(\frac{2\pi kt}{T}\right), \sqrt{\frac{2}{T}}\sin\left(\frac{2\pi kt}{T}\right)\right\}_{0\le k\le\infty}$$

complete for signals defined over $[0, T)$

Geometric Representation of Signals

$$s_i(t) = \sum_{j=1}^{N} s_{ij} f_j(t),\ 0 \le t < T,\ i = 1,2,\ldots,M$$



$$\{f_i\}_{i=1}^{N} \text{ orthonormal}$$

$$s_{ij} = \int_0^T s_i(t) f_j(t)\,dt,$$

$$i = 1,2,\ldots M,\ j = 1,2,\ldots,N$$

Geometric Representation of Signals

- Through the signal space concept, $s_i(t)$ (where $1 \le i \le M$) can be unambiguously represented by an N-dimensional *signal vector* $(s_{i1}, s_{i2}, \ldots, s_{iN})$ over an N-dimensional *signal space*.
- The design of transmitters becomes the selection of M points over the signal space, and the receivers make a guess about which of the M points was transmitted.
- In the N-dimensional signal space,
  - length square of the vector = energy of the signal
  - angle between vectors = energy correlation between signals

$$\cos(\theta_{ik}) = \frac{\langle s_i(t), s_k(t) \rangle}{\| s_i(t) \| \cdot \| s_k(t) \|}$$

$$\| s_i(t) \|^2 = \int_0^T s_i^2(t)dt = \sum_{j=1}^{N} s_{ij}^2$$

Geometric Representation of Signals

- The length square of a vector and the angle between vectors are independent of the basis used (Note that no translation of the origin is allowed).

- From this view,
  - the transmitter may be viewed as a ***synthesizer***, which *synthesizes* the transmitted signal by a bank of N multipliers.
  - the receiver may be viewed as an ***analyzer***, which *correlates* (*product-integrate*) the common input into individual informational signal.

Euclidean Distance

- After vectorization, we can calculate the *Euclidean distance* between two signals, which is the squared root of:

$$\int_0^T (s_i(t) - s_k(t))^2 dt = \| s_i(t) - s_k(t) \|^2 = \sum_{j=1}^{N} (s_{ij} - s_{kj})^2$$

Cauchy-Schwarz Inequality

- Cauchy-Schwarz inequality and angle between signals
  - Cauchy-Schwarz inequality said that

$$\left|\langle s_1(t), s_2(t)\rangle\right|^2 \leq \| s_1(t) \|^2 \cdot \| s_2(t) \|^2 \text{ with equality holding if } s_1(t) = c s_2(t).$$

  - Also, the angle between signals gives that

$$\cos(\theta_{12}) = \frac{\langle s_1(t), s_2(t)\rangle}{\| s_1(t) \| \cdot \| s_2(t) \|}$$

  - Hence, Cauchy-Schwarz inequality can be equivalently stated as:

$$|\cos(\theta_{12})|^2 \leq 1 \text{ with equality holding if } \theta_{12} = 0 \text{ or } \pi$$

Basis

- The (complete) orthonormal basis for a signal space is not unique!
  - So, the synthesizer and the analyzer for the transmission of the same informational messages are not unique!
- One way to determine a set of orthonormal basis is the Gram-Schmidt orthogonalization procedure.

Vectorization of Continuous AWGN Channel

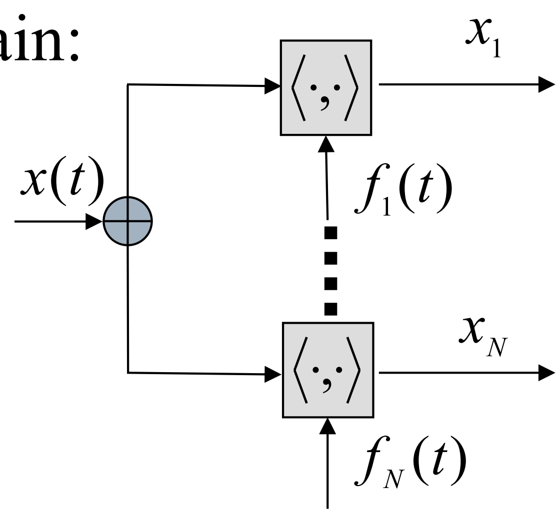
- Influence of the AWGN noise to the signal space concept

$$x(t) = s_i(t) + w(t)$$

where $w(t)$ is zero-mean AWGN with PSD $N_0/2$.

- After the correlator at the receiver, we obtain:

$$\langle x(t), f_j(t) \rangle = \langle s_i(t), f_j(t) \rangle + \langle w(t), f_j(t) \rangle$$



  - Equivalently, $x_j = s_{ij} + w_j$.
  - Notably, there is no "information loss" by the signal space representation.

Vectorization of Continuous AWGN Channel

- Statistics of $\{w_j\}$

$$\begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} s_{i1} \\ \vdots \\ s_{iN} \end{bmatrix} + \begin{bmatrix} w_1 \\ \vdots \\ w_N \end{bmatrix}$$

- Since $\{s_{ij}\}$ is deterministic, the distribution of $\boldsymbol{x}$ is a mean-shift of the distribution of $\boldsymbol{w}$.
- Observe that $\boldsymbol{w}$ is Gaussian distributed because $w(t)$ is AWGN. The distribution of $\boldsymbol{w}$ can, therefore, be determined by its *mean vector* and *covariance matrix*.

Vectorization of Continuous AWGN Channel

- Mean

$$E[w_j] = E\left[\int_0^T w(t) f_j(t) dt\right] = \int_0^T E[w(t)] f_j(t) dt = 0$$

- Covariance

$$E[w_i w_j] = E\left[\left(\int_0^T w(s) f_i(s) ds\right)\left(\int_0^T w(t) f_j(t) dt\right)\right]$$

$$= \int_0^T \int_0^T E[w(s) w(t)] f_i(s) f_j(t) ds dt$$

$$= \int_0^T \int_0^T \frac{N_0}{2} \delta(s-t) f_i(s) f_j(t) ds dt$$

$$= \frac{N_0}{2} \int_0^T f_i(t) f_j(t) dt = \frac{N_0}{2} \delta_{ij}$$

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

Vectorization of Continuous AWGN Channel

- As a result, $[w_1, w_2, \ldots, w_N]$ are zero-mean i.i.d. Gaussian distributed with variance $N_0/2$.
- This shows that $\boldsymbol{x}$ is independent Gaussian distributed with common variance $N_0/2$ and mean vector $\boldsymbol{s}_i = [s_{i1}, s_{i2}, \ldots, s_{iN}]$. Equivalently,

$$f(\boldsymbol{x} \mid \boldsymbol{s}_i) = \prod_{j=1}^{N} \frac{1}{\sqrt{\pi N_0}} \exp\left[-\frac{1}{N_0}(x_j - s_{ij})^2\right]$$

Vectorization of Continuous AWGN Channel

- Remaining term in noise
  - It is possible that

$$w'(t) = w(t) - \sum_{i=1}^{N} w_i \cdot f_i(t) \neq 0$$

  - However, it can be shown that (as an error term) $w'(t)$ is orthogonal to $s_i(t)$ for $1 \leq i \leq M$. Hence, $w'(t)$ will not affect the decision error rate on message i.

$$\langle w'(t), s_i(t) \rangle = 0 \text{ with probability } 1.$$

Likelihood Functions under Equal Prior

- An equivalent signal-space channel model

$$m = m_i \text{ for } 1 \le i \le M \to \boldsymbol{s} = c(m) \to \boldsymbol{x} = \boldsymbol{s} + \boldsymbol{w} \\ \to \hat{m} = d(\boldsymbol{x}) \in \{m_1, \ldots, m_M\}$$

- The best decision function $d(\,)$ that minimizes the decision error is:

$$d(\boldsymbol{x}) = m_i, \text{ if } P\{m_i \mid \boldsymbol{x}\} \ge P\{m_k \mid \boldsymbol{x}\} \text{ for all } 1 \le k \le M \\ = \arg \max_{m \in \{m_1, \ldots, m_M\}} P\{m \mid \boldsymbol{x}\}$$

  - This is the *maximum a posteriori probability* (MAP) decision rule.

Likelihood Functions under Equal Prior

- With equal prior probabilities,

$$\begin{aligned}
d(\boldsymbol{x}) &= \arg \max_{m\in\{m_1,\ldots,m_M\}} P\{m|\boldsymbol{x}\} \\
&= \arg\max \{P\{m_1|\boldsymbol{x}\}, P\{m_2|\boldsymbol{x}\}, \ldots, P\{m_M|\boldsymbol{x}\}\} \\
&= \arg\max \left\{ \frac{f(\boldsymbol{x}|m_1)P(m_1)}{f(\boldsymbol{x})}, \frac{f(\boldsymbol{x}|m_2)P(m_2)}{f(\boldsymbol{x})}, \ldots, \frac{f(\boldsymbol{x}|m_M)P(m_M)}{f(\boldsymbol{x})} \right\} \\
&= \arg\max \left\{ \frac{f(\boldsymbol{x}|m_1)\frac{1}{M}}{f(\boldsymbol{x})}, \frac{f(\boldsymbol{x}|m_2)\frac{1}{M}}{f(\boldsymbol{x})}, \ldots, \frac{f(\boldsymbol{x}|m_M)\frac{1}{M}}{f(\boldsymbol{x})} \right\} \\
&= \arg\max\{f(\boldsymbol{x}|m_1), f(\boldsymbol{x}|m_2), \ldots, f(\boldsymbol{x}|m_M)\}
\end{aligned}$$

$f(\boldsymbol{x}|m_i)$ is named the *likelihood function* given m_i is transmitted. Hence, the above rule is named the *maximum-likelihood* (ML) decision rule.

Likelihood Functions under Equal Prior

- MAP rule = ML rule, if equal prior probability is assumed.
- In practice, it is more *convenient* to work on the *log-likelihood functions*, defined by

$$d(\boldsymbol{x}) = \arg\max\{f(\boldsymbol{x}\,|\,m_1), f(\boldsymbol{x}\,|\,m_2),\ldots, f(\boldsymbol{x}\,|\,m_M)\}$$
$$= \arg\max\{\log f(\boldsymbol{x}\,|\,m_1), \log f(\boldsymbol{x}\,|\,m_2),\ldots, \log f(\boldsymbol{x}\,|\,m_M)\}$$

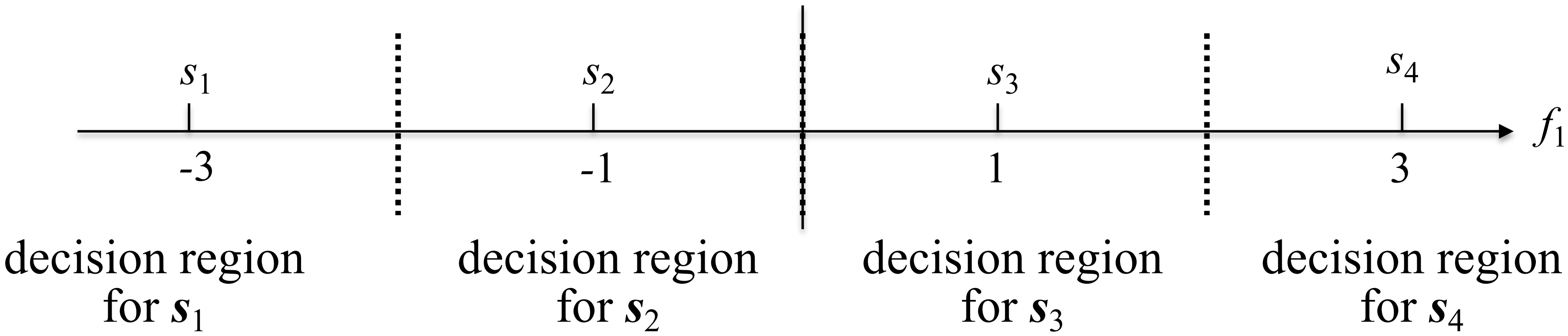
- Why *log-likelihood functions* are more convenient? The decision function becomes "sum of (squared) Euclidean distances" in AWGN channel.

$$
\begin{aligned}
d(\boldsymbol{x}) &= \arg\max_{1\le i\le M} \log f(\boldsymbol{x} \mid m_i) = \arg\max_{1\le i\le M} \log f(\boldsymbol{x} \mid \boldsymbol{s}_i) \\
&= \arg\max_{1\le i\le M} \log \prod_{j=1}^{N} \frac{1}{\sqrt{\pi N_0}} \exp\left[-\frac{1}{N_0}(x_j - s_{ij})^2\right] \\
&= \arg\max_{1\le i\le M} \sum_{j=1}^{N} \left(-\frac{1}{2}\log \pi N_0 - \frac{1}{N_0}(x_j - s_{ij})^2\right) \\
&= \arg\min_{1\le i\le M} \sum_{j=1}^{N} (x_j - s_{ij})^2 \\
&= \arg\min_{1\le i\le M} \| \boldsymbol{x} - \boldsymbol{s}_i \|^2 \quad \left(= \arg\min_{1\le i\le M} \| \boldsymbol{x} - \boldsymbol{s}_i \| \right)
\end{aligned}
$$

Upon the reception of received signal point $\boldsymbol{x}$, find the signal point $\boldsymbol{s}_i$ that is closest in Euclidean distance to $\boldsymbol{x}$.

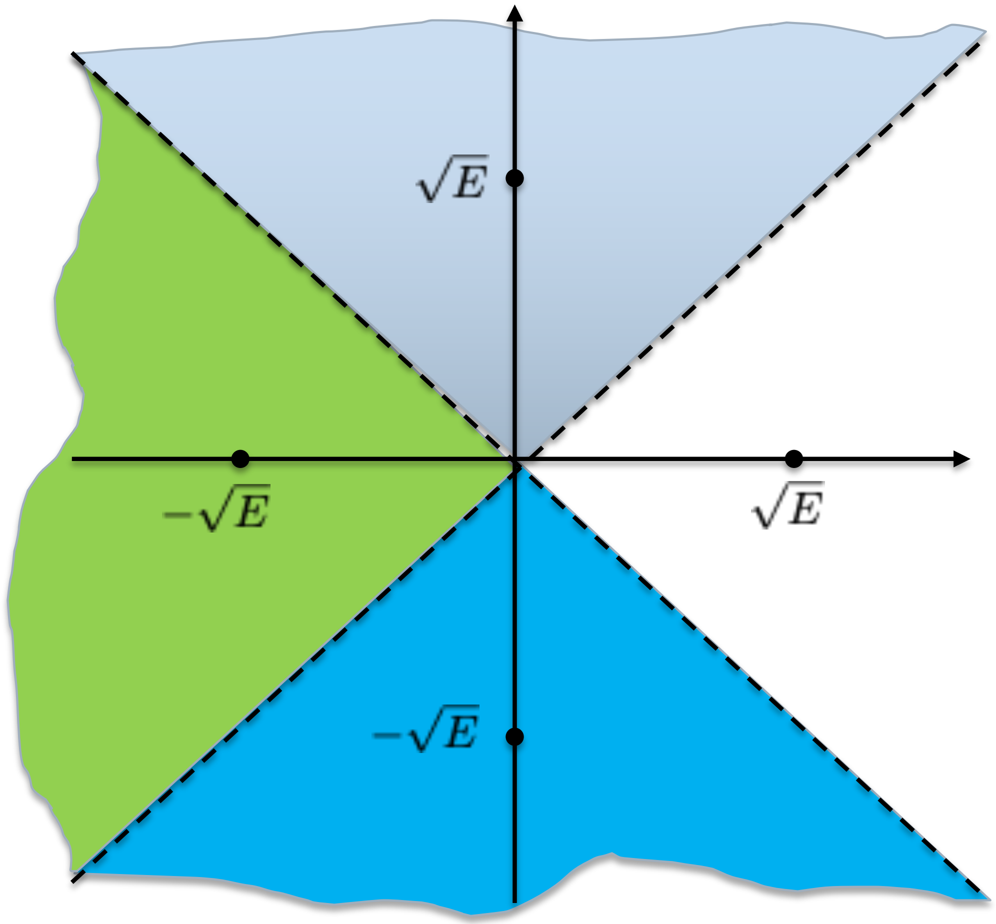
Coherent Detection: Maximum Likelihood Decoding

- Signal constellation
  - The set of M signal points in the signal space
    - Example. Signal constellation for 2B1Q code



Coherent Detection: Maximum Likelihood Decoding

- Decision regions for $N = 2$ and $M = 4$



Coherent Detection: Maximum Likelihood Decoding

- Usually, $\boldsymbol{s}_1$, $\boldsymbol{s}_2$, …, $\boldsymbol{s}_M$ are named the message points.
- The received signal point $\boldsymbol{x}$ wanders about the transmitted message point in a Gaussian-distributed random fashion.
- Constant-energy signal constellation
  - In this case, the ML decision rule can be reduced to an inner-product.

$$
\begin{aligned}
d(\boldsymbol{x}) &= \arg\min_{1\le i\le M} \|\boldsymbol{x}-\boldsymbol{s}_i\|^2 \\
&= \arg\min_{1\le i\le M}\left(\|\boldsymbol{x}\|^2 - 2\langle \boldsymbol{x},\boldsymbol{s}_i\rangle + \|\boldsymbol{s}_i\|^2\right) \\
&= \arg\min_{1\le i\le M}\left(-2\langle \boldsymbol{x},\boldsymbol{s}_i\rangle + E_i\right) \\
&= \arg\max_{1\le i\le M}\langle \boldsymbol{x},\boldsymbol{s}_i\rangle, \text{ if } E_i \text{ is constant.}
\end{aligned}
$$

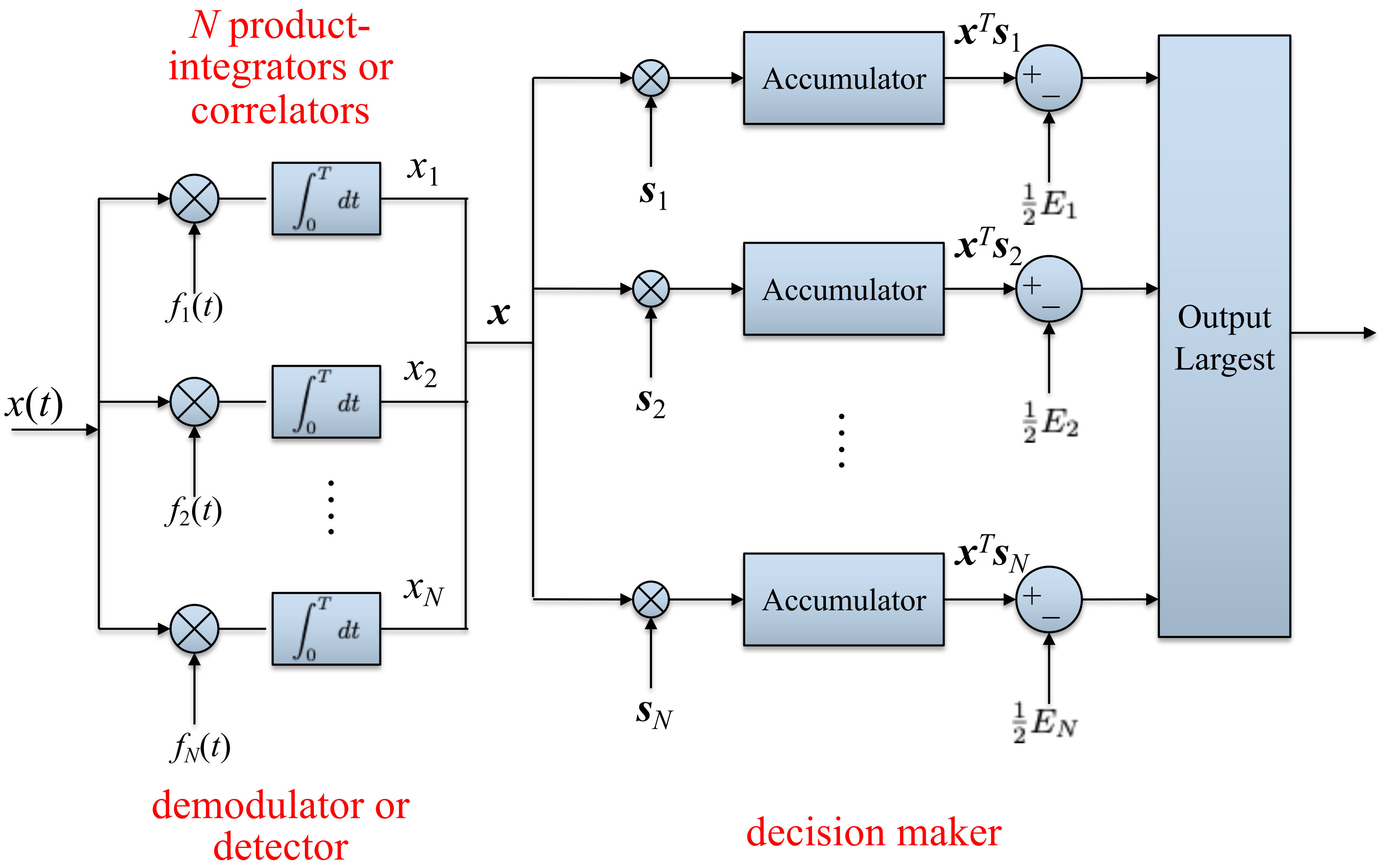
Correlation Receiver

- If signals do not have equal energy, we can use

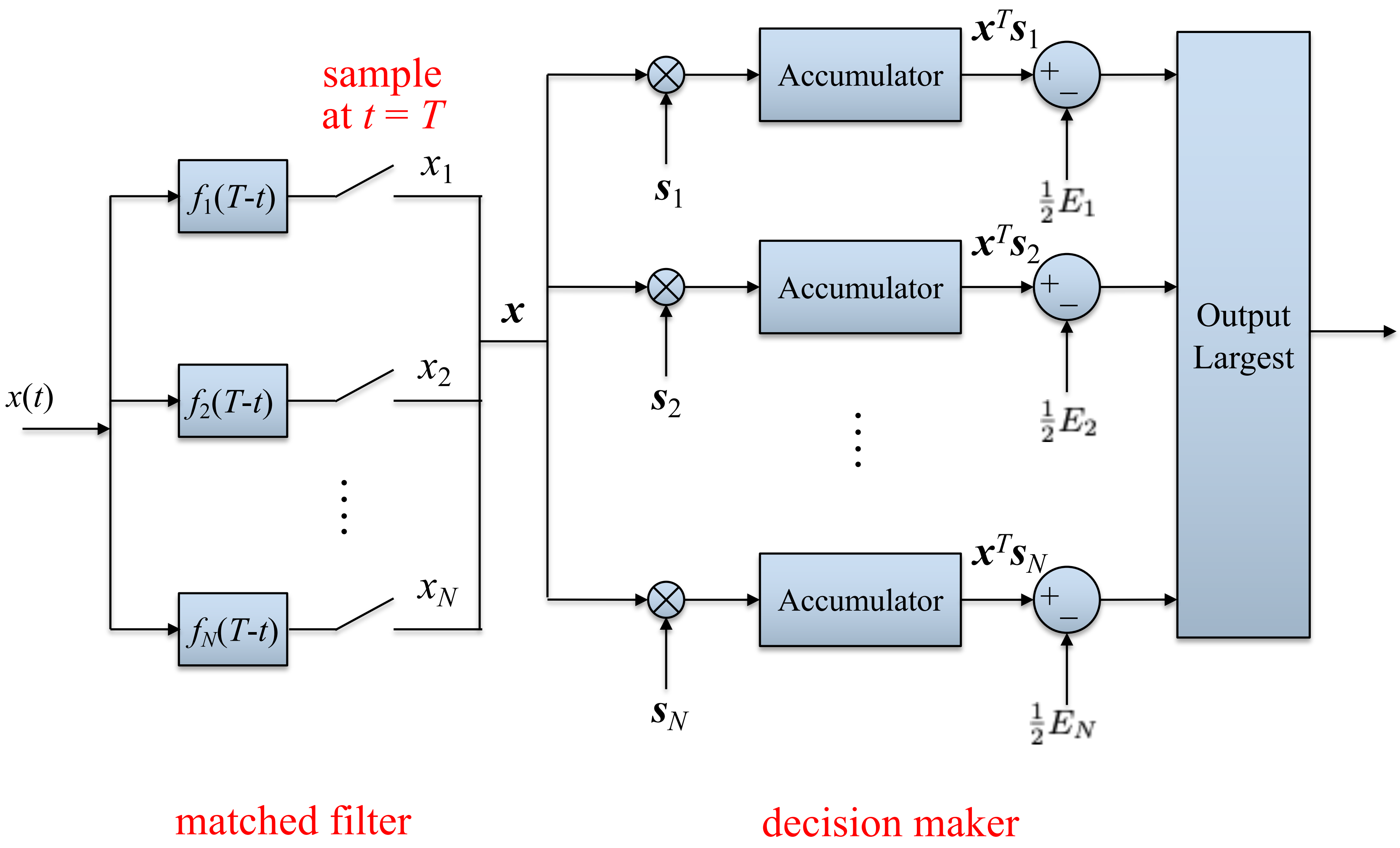
$$d(\boldsymbol{x}) = \arg\max_{1\le i\le M}\left(\langle \boldsymbol{x}, \boldsymbol{s}_i\rangle - \frac{1}{2}E_i\right).$$

  to implement the ML rule.

  - The receiver is *coherent* because the receiver requires to be in perfect synchronization with the transmitter (more specifically, the integration must begin at exactly the right time instance).



Correlation receiver



Correlation receiver

Equivalence
of Correlation and Matched Filter Receivers

- The correlator and matched filter can be made equivalent.
- Specifically,

$$x_i = \int_0^T x(\tau) f_i(t) dt = \int_{-\infty}^{\infty} x(\tau) h_i(T - \tau) d\tau$$

if $h_i(t) = f_i(T - t)$ (and implicitly $f_i(t)$ is zero outside $0 \leq t \leq T$).

Probability of Symbol Error

- Average probability of symbol error

$$P_e = 1 - P_c = 1 - \sum_{i=1}^{M} P(m_i) P(d(\boldsymbol{x}) = m_i \mid m_i \text{ transmitted})$$

$$= 1 - \frac{1}{M} \sum_{i=1}^{M} P(d(\boldsymbol{x}) = m_i \mid m_i \text{ transmitted})$$

$$= 1 - \frac{1}{M} \sum_{i=1}^{M} \Pr\left\{ \| \boldsymbol{x} - \boldsymbol{s}_i \|^2 \le \min_{1 \le j \le M, j \ne i} \| \boldsymbol{x} - \boldsymbol{s}_j \|^2 \,\middle|\, m_i \text{ transmitted} \right\}$$

$$= 1 - \frac{1}{M} \sum_{i=1}^{M} \int_{Z_i} f(\boldsymbol{x} \mid \boldsymbol{s}_i) d\boldsymbol{x}$$

where $Z_i = \left\{ \boldsymbol{x} \in \Re^N : \| \boldsymbol{x} - \boldsymbol{s}_i \|^2 \le \min_{1 \le j \le M, j \ne i} \| \boldsymbol{x} - \boldsymbol{s}_j \|^2 \right\}$

Invariance of Probability of Symbol Error

- Probability of symbol error is invariant with respect to basis change (i.e., rotation and translation of the signal space).
- Specifically, the symbol error rate (SER) only depends on the *relative "Euclidean distances"* between the message points.

$$P_e = 1 - \frac{1}{M} \sum_{i=1}^{M} \Pr\left\{ \| \boldsymbol{x} - \boldsymbol{s}_i \|^2 \le \min_{1 \le j \le M, j \ne i} \| \boldsymbol{x} - \boldsymbol{s}_j \|^2 \middle| m_i \text{ transmitted} \right\}$$

Invariance of Probability of Symbol Error

- Specifically, if **Q** is a reversible transform (matrix), such as rotation, then

$$\left\{\boldsymbol{x} \in \Re^N : \| \boldsymbol{x} - \boldsymbol{s}_i \|^2 \le \min_{1 \le j \le M, j \ne i} \| \boldsymbol{x} - \boldsymbol{s}_j \|^2 \right\}$$
$$= \left\{\boldsymbol{x} \in \Re^N : \| \mathbf{Q}\boldsymbol{x} - \mathbf{Q}\boldsymbol{s}_i \|^2 \le \min_{1 \le j \le M, j \ne i} \| \mathbf{Q}\boldsymbol{x} - \mathbf{Q}\boldsymbol{s}_j \|^2 \right\}$$

- The invariance in SER for translation can be likewise proved.
  - Is the transmission power invariant for rotation and translation?

Minimum Energy Signals

- Since SER is invariant to rotation and translation, we may rotate and translate the signal constellation to minimize the transmission power without affecting SER.

$$E_g = \sum_{i=1}^{M} p_i \|\boldsymbol{s}_i\|^2$$

Find $\boldsymbol{a}$ and $\mathbf{Q}$ such that $E_g(\boldsymbol{a},\mathbf{Q}) = \sum_{i=1}^{M} p_i \|\mathbf{Q}(\boldsymbol{s}_i - \boldsymbol{a})\|^2$ is minimized.

- Since **Q** does not change the norm (i.e., transmission power), we only need to determine the right $\boldsymbol{a}$.

Minimum Energy Signals

- Determine the optimal $\boldsymbol{a}$.

$$E_g(\boldsymbol{a}) = \sum_{i=1}^{M} p_i \|\boldsymbol{s}_i - \boldsymbol{a}\|^2$$

$$= \sum_{i=1}^{M} p_i \left(\|\boldsymbol{s}_i\|^2 - 2\boldsymbol{a}^T \boldsymbol{s}_i + \|\boldsymbol{a}\|^2 \right)$$

$$= \sum_{i=1}^{M} p_i \|\boldsymbol{s}_i\|^2 - 2\boldsymbol{a}^T \left(\sum_{i=1}^{M} p_i \boldsymbol{s}_i \right) + \|\boldsymbol{a}\|^2$$

$$\Rightarrow \boldsymbol{a}_{\text{optimal}} = \sum_{i=1}^{M} p_i \boldsymbol{s}_i \ \text{ and } \ E_g(\boldsymbol{a}_{\text{optimal}}) = \sum_{i=1}^{M} p_i \|\boldsymbol{s}_i\|^2 - \left\| \sum_{i=1}^{M} p_i \boldsymbol{s}_i \right\|^2$$

Union Bound on Probability of Error

- Union bound $\boxed{P(A \bigcup B) \le P(A) + P(B)}$

$$P_e = 1 - \frac{1}{M}\sum_{i=1}^{M} \Pr\left\{ \| \boldsymbol{x} - \boldsymbol{s}_i \|^2 \le \min_{1 \le j \le M, j \ne i} \| \boldsymbol{x} - \boldsymbol{s}_j \|^2 \middle| m_i \text{ transmitted} \right\}$$

$$= \left(\frac{1}{M}\sum_{i=1}^{M} 1 \right) - \frac{1}{M}\sum_{i=1}^{M} \Pr\left\{ \begin{array}{l} \left(\| \boldsymbol{x} - \boldsymbol{s}_i \|^2 \le \| \boldsymbol{x} - \boldsymbol{s}_1 \|^2 \right) \wedge \cdots \\ \underbrace{\cdots}_{j \ne i} \wedge \left(\| \boldsymbol{x} - \boldsymbol{s}_i \|^2 \le \| \boldsymbol{x} - \boldsymbol{s}_M \|^2 \right) \end{array} \middle| m_i \text{ transmitted} \right\}$$

$$= \frac{1}{M}\sum_{i=1}^{M} \Pr\left\{ \begin{array}{l} \left(\| \boldsymbol{x} - \boldsymbol{s}_i \|^2 > \| \boldsymbol{x} - \boldsymbol{s}_1 \|^2 \right) \vee \cdots \\ \underbrace{\cdots}_{j \ne i} \vee \left(\| \boldsymbol{x} - \boldsymbol{s}_i \|^2 > \| \boldsymbol{x} - \boldsymbol{s}_M \|^2 \right) \end{array} \middle| m_i \text{ transmitted} \right\}$$

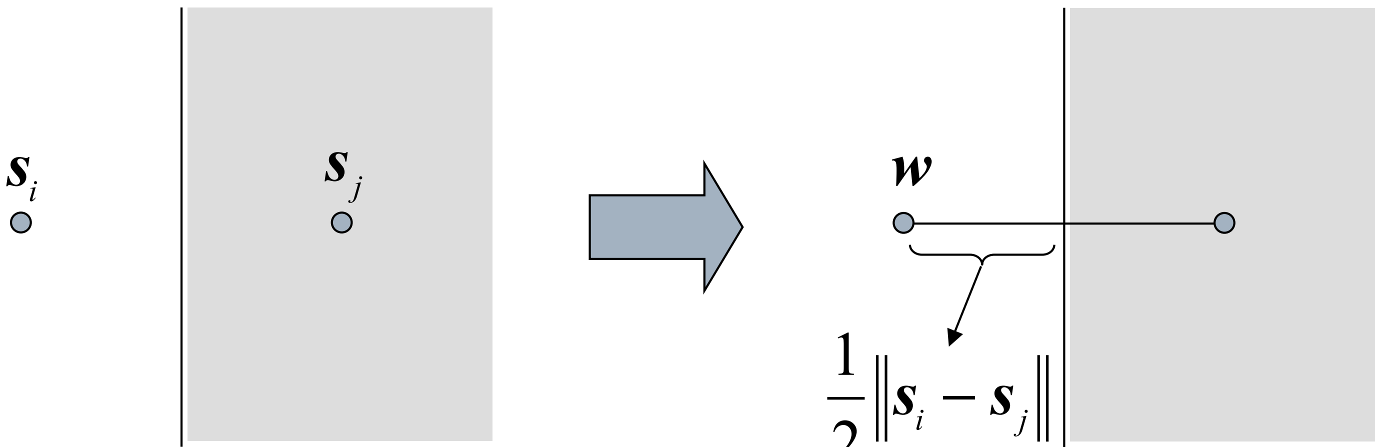
$$\le \frac{1}{M}\sum_{i=1}^{M} \sum_{j=1, j \ne i}^{M} \Pr\left\{ \| \boldsymbol{x} - \boldsymbol{s}_i \|^2 > \| \boldsymbol{x} - \boldsymbol{s}_j \|^2 \middle| m_i \text{ transmitted} \right\}$$

Union Bound on Probability of Error

$$P_e \le \frac{1}{M}\sum_{i=1}^{M}\sum_{j=1, j\ne i}^{M} P_2(\boldsymbol{s}_i, \boldsymbol{s}_j)$$

where $P_2(\boldsymbol{s}_i, \boldsymbol{s}_j) = \Pr\left\{\|\boldsymbol{x}-\boldsymbol{s}_i\|^2 > \|\boldsymbol{x}-\boldsymbol{s}_j\|^2 \,\middle|\, m_i \text{ transmitted}\right\}$.

Notably, given m_i transmitted, $\boldsymbol{x}$ is Gaussian distributed with mean $\boldsymbol{s}_i$.



Union Bound on Probability of Error

Since $\boldsymbol{x} = \boldsymbol{s}_i + \boldsymbol{w}$ when $\boldsymbol{s}_i$ was transmitted, we have

$$\begin{aligned}
&\Pr\left\{\|\boldsymbol{x}-\boldsymbol{s}_i\|^2 > \|\boldsymbol{x}-\boldsymbol{s}_j\|^2 \middle| \boldsymbol{s}_i \text{ transmitted}\right\} \\
&= \Pr\left\{\|(\boldsymbol{s}_i+\boldsymbol{w})-\boldsymbol{s}_i\|^2 > \|(\boldsymbol{s}_i+\boldsymbol{w})-\boldsymbol{s}_j\|^2 \middle| \boldsymbol{s}_i \text{ transmitted}\right\} \\
&= \Pr\left\{\|\boldsymbol{w}\|^2 > \|\boldsymbol{w}+(\boldsymbol{s}_i-\boldsymbol{s}_j)\|^2 \middle| \boldsymbol{s}_i \text{ transmitted}\right\} \\
&= \Pr\left\{\|\boldsymbol{w}\|^2 > \|\boldsymbol{w}\|^2 + \|\boldsymbol{s}_i-\boldsymbol{s}_j\|^2 + 2(\boldsymbol{s}_i-\boldsymbol{s}_j)^T\boldsymbol{w} \middle| \boldsymbol{s}_i \text{ transmitted}\right\} \\
&= \Pr\left\{(\boldsymbol{s}_i-\boldsymbol{s}_j)^T\boldsymbol{w} < -\frac{1}{2}\|\boldsymbol{s}_i-\boldsymbol{s}_j\|^2 \middle| \boldsymbol{s}_i \text{ transmitted}\right\} \\
&= \Pr\left\{n < -\frac{1}{2}\|\boldsymbol{s}_i-\boldsymbol{s}_j\|^2 \middle| \boldsymbol{s}_i \text{ transmitted}\right\}
\end{aligned}$$

where $n \triangleq (\boldsymbol{s}_i-\boldsymbol{s}_j)^T\boldsymbol{w}$.

Union Bound on Probability of Error

Observe that $\boldsymbol{w}$ is zero-mean Gaussian distributed
with covariance matrix $E[\boldsymbol{w}\boldsymbol{w}^T] = \frac{N_0}{2}\mathbb{I}$, where $\mathbb{I}$ is the identity matrix.
Hence, $n \triangleq (\boldsymbol{s}_i - \boldsymbol{s}_j)^T \boldsymbol{w}$ is Gaussian distributed with

$$E[n] = E\big[(\boldsymbol{s}_i - \boldsymbol{s}_j)^T \boldsymbol{w}\big] = (\boldsymbol{s}_i - \boldsymbol{s}_j)^T E[\boldsymbol{w}] = (\boldsymbol{s}_i - \boldsymbol{s}_j)^T \mathbf{0} = 0$$

$$\begin{aligned} E[n^2] &= E\big[(\boldsymbol{s}_i - \boldsymbol{s}_j)^T \boldsymbol{w} \cdot ((\boldsymbol{s}_i - \boldsymbol{s}_j)^T \boldsymbol{w})^T\big] \\ &= E\big[(\boldsymbol{s}_i - \boldsymbol{s}_j)^T \boldsymbol{w}\boldsymbol{w}^T (\boldsymbol{s}_i - \boldsymbol{s}_j)\big] \\ &= (\boldsymbol{s}_i - \boldsymbol{s}_j)^T E\big[\boldsymbol{w}\boldsymbol{w}^T\big](\boldsymbol{s}_i - \boldsymbol{s}_j) \\ &= \frac{N_0}{2}(\boldsymbol{s}_i - \boldsymbol{s}_j)^T \mathbb{I} (\boldsymbol{s}_i - \boldsymbol{s}_j) \\ &= \frac{N_0}{2}\|\boldsymbol{s}_i - \boldsymbol{s}_j\|^2. \end{aligned}$$

This implies that $w \triangleq n/\|\boldsymbol{s}_i - \boldsymbol{s}_j\|$ is Gaussian distributed
with mean zero and variance $N_0/2$.

Union Bound on Probability of Error

As a result,

$$\begin{aligned}
&\Pr\left\{ n < -\frac{1}{2}\|\boldsymbol{s}_i - \boldsymbol{s}_j\|^2 \,\middle|\, \boldsymbol{s}_i \text{ transmitted} \right\} \\
&= \Pr\left\{ \|\boldsymbol{s}_i - \boldsymbol{s}_j\| w < -\frac{1}{2}\|\boldsymbol{s}_i - \boldsymbol{s}_j\|^2 \,\middle|\, \boldsymbol{s}_i \text{ transmitted} \right\} \\
&= \Pr\left\{ w < -\frac{1}{2}\|\boldsymbol{s}_i - \boldsymbol{s}_j\| \,\middle|\, \boldsymbol{s}_i \text{ transmitted} \right\} \\
&= \Pr\left\{ w > \frac{1}{2}\|\boldsymbol{s}_i - \boldsymbol{s}_j\| \,\middle|\, \boldsymbol{s}_i \text{ transmitted} \right\},
\end{aligned}$$

where the last equality is valid because the probability density function of a zero-mean Gaussian random variable is symmetric with respect to $w = 0$ (hence, $\Pr[w > a] = \Pr[w < -a]$ for any $a > 0$).

Union Bound on Probability of Error

- Hence,
  - when $N = 1$,

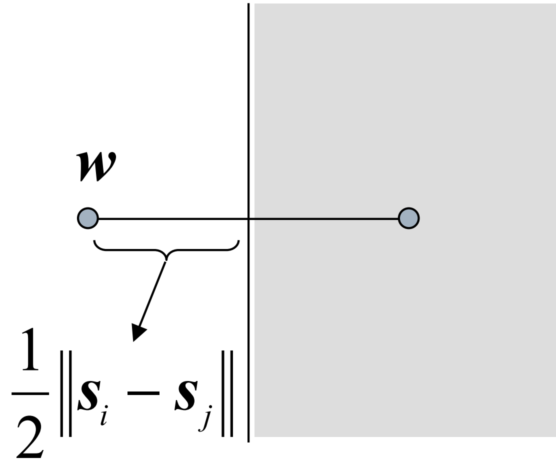
$$P_2(\boldsymbol{s}_i, \boldsymbol{s}_j) = \Pr\left\{ \| \boldsymbol{x} - \boldsymbol{s}_i \|^2 > \| \boldsymbol{x} - \boldsymbol{s}_j \|^2 \,\middle|\, m_i \text{ transmitted} \right\}$$

$$= \Pr\left\{ w > \frac{1}{2} | s_i - s_j | \right\}$$

$$= \int_{d_{ij}/2}^{\infty} \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{v^2}{N_0} \right) dv, \text{ where } d_{ij} = | s_i - s_j |$$

$$= \frac{1}{2} \operatorname{erfc}\left(\frac{d_{ij}}{2\sqrt{N_0}} \right), \text{ where } \operatorname{erfc}(u) = \frac{2}{\sqrt{\pi}} \int_u^{\infty} \exp(-z^2) dz.$$

Union Bound on Probability of Error



- For $N = 2$,

$$P_2(\boldsymbol{s}_i, \boldsymbol{s}_j) = \Pr\left\{ \|\boldsymbol{x} - \boldsymbol{s}_i\|^2 > \|\boldsymbol{x} - \boldsymbol{s}_j\|^2 \;\middle|\; m_i \text{ transmitted} \right\}$$

$$= \Pr\left\{ w_1 > \frac{1}{2} d_{ij} \text{ and } w_2 = \text{don't care} \right\}, \text{ where } d_{ij} = \|\boldsymbol{s}_i - \boldsymbol{s}_j\|$$

$$= \int_{d_{ij}/2}^{\infty} \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{v^2}{N_0} \right) dv$$

$$= \frac{1}{2} \operatorname{erfc}\left(\frac{d_{ij}}{2\sqrt{N_0}} \right), \text{ where } \operatorname{erfc}(u) = \frac{2}{\sqrt{\pi}} \int_u^{\infty} \exp(-z^2)\, dz.$$

- The same formula is valid for any N.

Union Bound on Probability of Error

- Consequently, the *union bound* for symbol error rate is:

$$P_e \le \frac{1}{M}\sum_{i=1}^{M}\sum_{j=1, j\ne i}^{M} P_2(\boldsymbol{s}_i, \boldsymbol{s}_j) = \frac{1}{M}\sum_{i=1}^{M}\sum_{j=1, j\ne i}^{M} \frac{1}{2}\operatorname{erfc}\left(\frac{d_{ij}}{2\sqrt{N_0}}\right)$$

- The above bound can be further simplified when additional condition is given.
  - For example, if the signal constellation is *circularly symmetric* in the sense that "$\{d_{i1}, d_{i2}, \ldots, d_{iM}\}$ is a permutation of $\{d_{k1}, d_{k2}, \ldots, d_{kM}\}$ for $i \ne k$," then

$$P_e \le \sum_{j=1, j\ne i}^{M} \frac{1}{2}\operatorname{erfc}\left(\frac{d_{ij}}{2\sqrt{N_0}}\right)$$

Union Bound on Probability of Error

- Another simplification of union bound
  - Define the minimum distance of a signal constellation as:

$$d_{\min} = \min_{1\le i\le M, 1\le j\le M, i\ne j} d_{ij}$$

Then, by the strict decreasing property of erfc function,

$$\text{erfc}\left(\frac{d_{ij}}{2\sqrt{N_0}}\right) \le \text{erfc}\left(\frac{d_{\min}}{2\sqrt{N_0}}\right)$$

$$\Rightarrow P_e \le \frac{1}{M}\sum_{i=1}^{M}\sum_{j=1, j\ne i}^{M}\frac{1}{2}\text{erfc}\left(\frac{d_{ij}}{2\sqrt{N_0}}\right) \le \frac{1}{M}\sum_{i=1}^{M}\sum_{j=1, j\ne i}^{M}\frac{1}{2}\text{erfc}\left(\frac{d_{\min}}{2\sqrt{N_0}}\right) = \frac{M-1}{2}\text{erfc}\left(\frac{d_{\min}}{2\sqrt{N_0}}\right)$$

Union Bound on Probability of Error

- We may use the bound for erfc function to realize the relation between SER and $d_{\min}$.

$$\operatorname{erfc}(u) \le \frac{\exp(-u^2)}{\sqrt{\pi}} \text{ for } u > 0.608131$$

$$\Rightarrow P_e \le \frac{M-1}{2}\operatorname{erfc}\left(\frac{d_{\min}}{2\sqrt{N_0}}\right) \le \frac{M-1}{2\sqrt{\pi}}\exp\left(-\frac{d_{\min}^2}{4N_0}\right), \text{ if } d_{\min}^2 > 1.47929 N_0.$$

  - Conclusion: *SER* decreases *exponentially* as the *squared minimum distance* grows.

Relation between BER and SER

- The information bits are transmitted in group of $\log_2 M$ bits to form an M-ary symbol.
- This gives the result that a large *symbol error rate* (SER) may not cause a large *bit error rate* (BER).
  - For example, a symbol error (for large M) may be due to only 1 bit error.
  - Optimistically, if every symbol error is due to a single bit error, then (assuming that n symbols are transmitted)

$$\text{BER} = \frac{n \cdot \text{SER}}{n \cdot \log_2(M)} = \frac{\text{SER}}{\log_2(M)}. \quad \left(\text{In general, BER} \geq \frac{\text{SER}}{\log_2(M)}. \right)$$

Relation between BER and SER

- Pessimistically, if every symbol error causes $\log_2 M$ bit errors, then (assuming that n symbols are transmitted)

$$\text{BER} = \frac{n \cdot \log_2 M \cdot \text{SER}}{n \cdot \log_2 M} = \text{SER}. \quad (\text{In general, BER} \le \text{SER.})$$

- Summary:

$$\boxed{\frac{\text{SER}}{\log_2 M} \le \text{BER} \le \text{SER}}$$

Relation between BER and SER

- If the statistics for "number of bit error patterns that causes one symbol error" is known, we can determine the exact relation between BER and SER.

$$\text{BER} = \frac{n \cdot \text{SER} \cdot \sum_{j=1}^{M-1} \#(\boldsymbol{b}_j) \cdot P(\boldsymbol{b}_j)}{n \cdot \log_2 M}$$

where $\#(\boldsymbol{b}_j) = \text{number of 1's in } \boldsymbol{b}_j$,

and $\boldsymbol{b}_j$ represents a binary permutation of $\log_2 M$ bit pattern.

Here, a 1's in $\boldsymbol{b}_j$ means a bit error occurs in the corresponding position; hence, the all-zero pattern is excluded because it represents no symbol error.

Relation between BER and SER

- **Example**. If all bit error patterns (including no error pattern) are equally likely, then

$$
\begin{aligned}
\text{BER} &= \frac{n \cdot \text{SER} \cdot \sum_{j=1}^{M-1} \#(\boldsymbol{b}_j) \cdot P(\boldsymbol{b}_j)}{n \cdot \log_2(M)} = \frac{\text{SER} \cdot \sum_{j=1}^{M-1} \#(\boldsymbol{b}_j) \cdot \frac{1}{M}}{\log_2(M)} \\
&= \frac{\text{SER}}{M \log_2(M)} \sum_{u=1}^{\log_2(M)} u \binom{\log_2(M)}{u} \qquad \left(\text{Note } \sum_{u=1}^{k} u \binom{k}{u} = k 2^{k-1}.\right) \\
&= \frac{\text{SER}}{M \log_2(M)} \log_2(M) \frac{M}{2} \\
&= \frac{1}{2} \text{SER}
\end{aligned}
$$

Summary

- Geometric Representation of Signals
- Gram-Schmidt Orthogonalization Procedure
- Signal Space Concept
- Coherent Detection: Maximum Likelihood Decoding
- Equivalence of Correlation and Matched Filter Receivers
- Union Bound on Probability of Error
- Relation between BER and SER