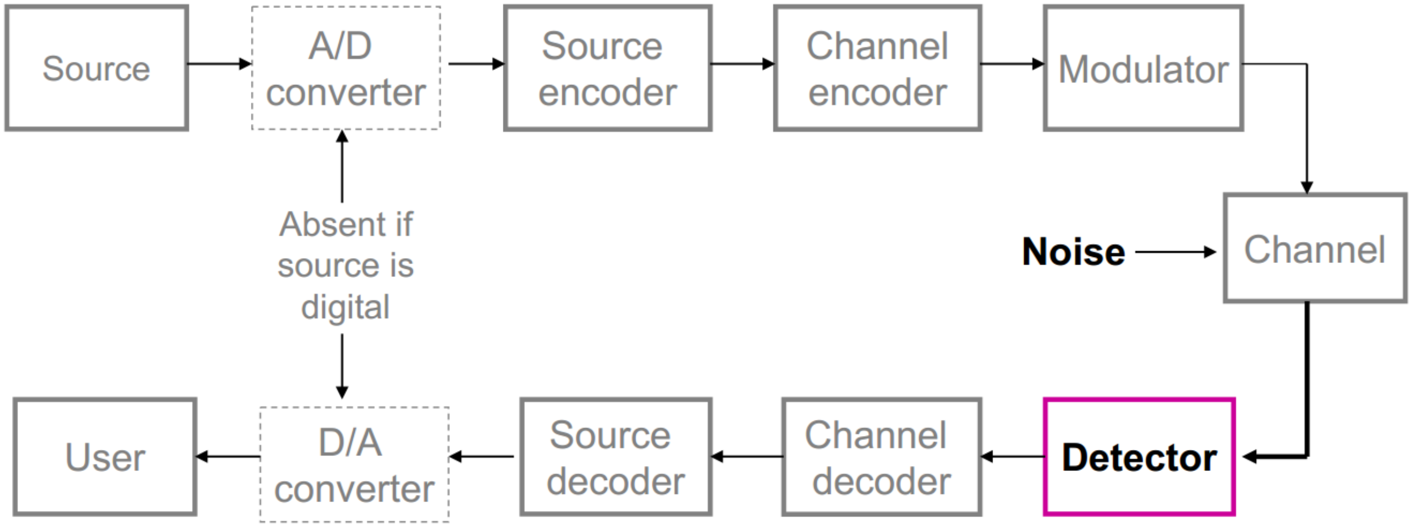
# Outline

- Introduction
- Signal, random variable, random process and spectra
- Analog modulation
- Analog to digital conversion
- Digital transmission through baseband channels
- Signal space representation
- **Optimal receivers**
- Digital modulation techniques
- Channel coding
- Synchronization
- Information theory

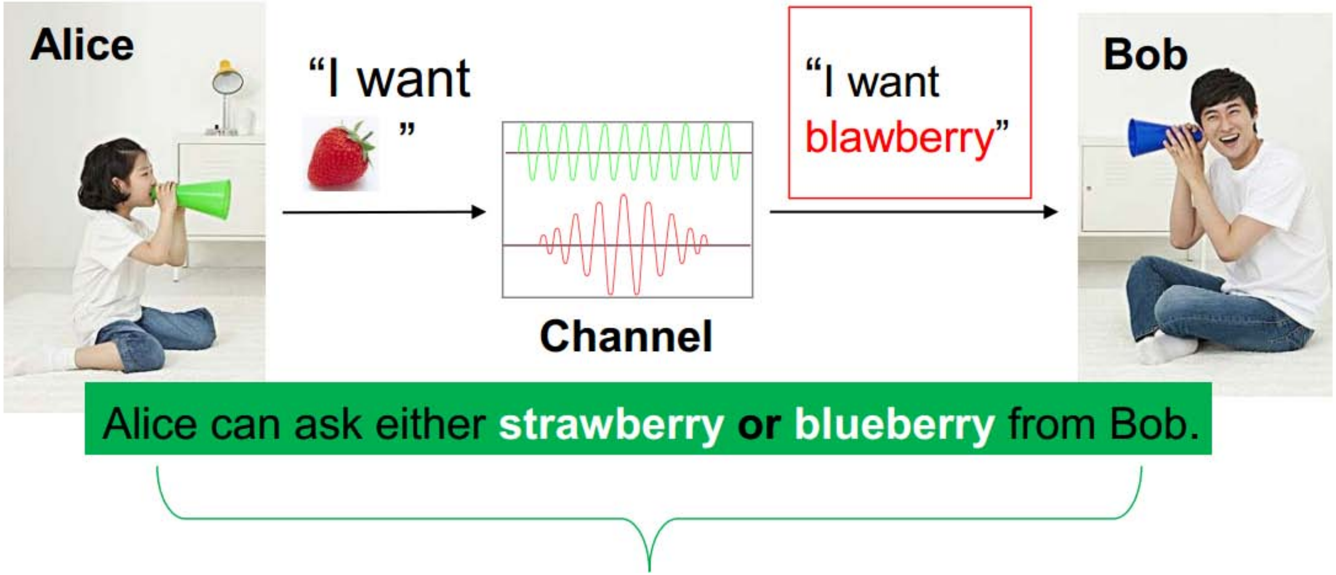
# Optimal receivers




- Detection theory
- Optimal receiver structure
- Matched filter
- Decision regions
- Error probability analysis

**Chapter 8.2-8.4, 8.5.3**

# Optimal receivers




- In digital communications, hypotheses are the possible messages and observations are the output of a channel
- Based on the observed values of the channel output, we are interested in the best decision making rule in the sense of minimizing the probability of error

# Detection theory

- Given M possible hypotheses $H_i$ (signal $m_i$) with probability

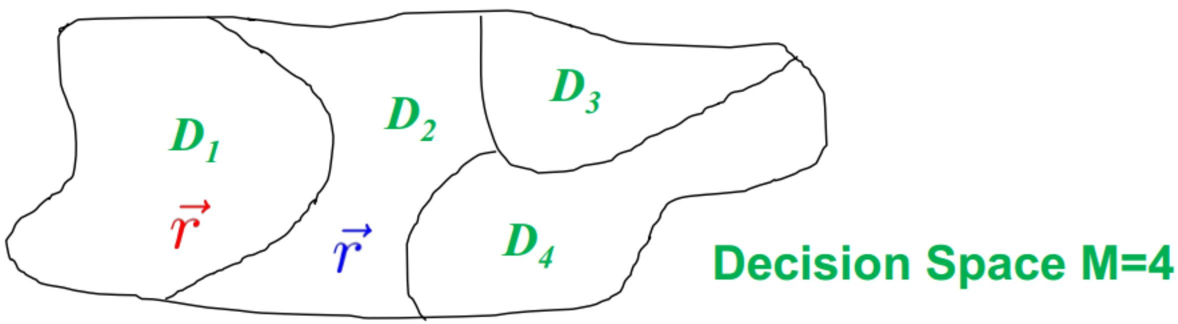
$$P_i = P(m_i) \quad , \quad i = 1, 2, \ldots, M$$

  where $P_i$ represents the priori knowledge concerning the probability of the signal mi (priori probability)
- The observation is some collection of N real values denoted by $\vec{r} = (r_1, r_2, \ldots, r_N)$ with conditional pdf

  $f(\vec{r}|m_i)$ -- conditional pdf of observation $\vec{r}$ given the signal $m_i$
- Our goal is to find the best decision-making rule in the sense of minimizing the probability of error

# Detection theory

- In general, $\vec{r}$ can be regarded as a point in some observation space
- Each hypothesis $H_i$ is associated with a decision region $D_i$:
  If $\vec{r}$ falls into $D_i$, the decision is $H_i$
- Error occurs when a decision is in favor of another when the signal $\vec{r}$ falls outside the decision region $D_i$



# Detection theory

- Consider a decision rule based on the computation of the posterior probabilities defined as

$$P(m_i|\vec{r}) = P(\text{ signal } m_i \text{ was transmitted given } \vec{r} \text{ observed })$$
$$\text{for } i = 1, \ldots, M$$

- A posterior since the decision is made after (or given) the observation
- Different from the a priori where some information about the decision known before the observation
- By Bayes' Rule: $P(m_i|\vec{r}) = \dfrac{P_i f(\vec{r}|m_i)}{f(\vec{r})}$
- Minimizing the probability of detection error given $\vec{r}$ is equivalent to maximize the probability of correct dection
- Maximum a posterior (MAP) decision rule:

Choose $\hat{m} = m_k$ if and only if
$P_k f(\vec{r}|m_k) \geq P_i f(\vec{r}|m_i)$; for all $i \neq k$

# Detection theory

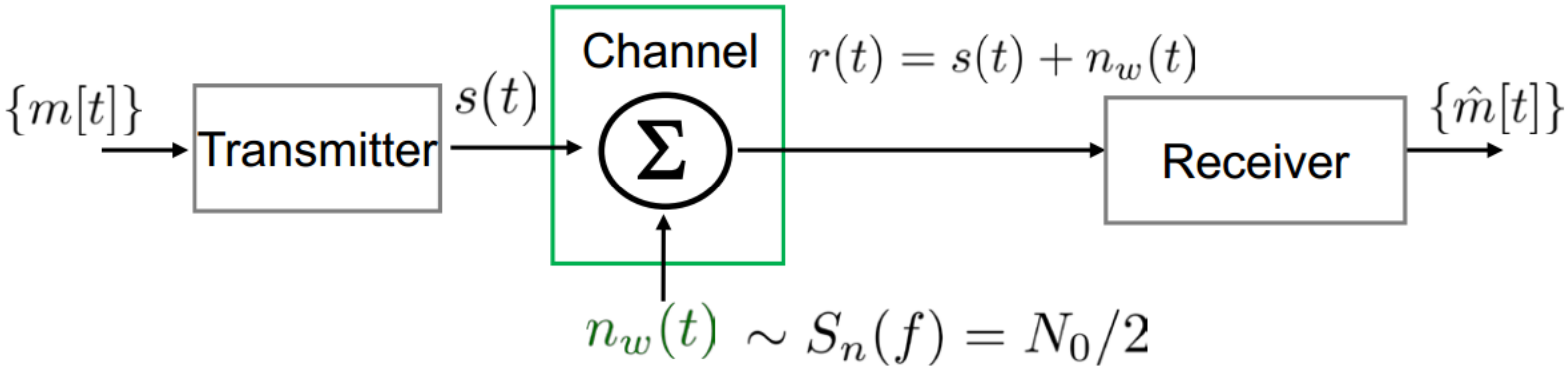
- If $p_1 = p_2 = ... = p_M$., the signals are equiprobable, finding the signal that maximizes $P(m_k|\vec{r})$ is equivalent to finding the signal that maximizes $f(\vec{r}|m_k)$
- The conditional pdf $f(\vec{r}|m_k)$ is usually called the likelihood function. The decision criterion based on the maixmum of $f(\vec{r}|m_k)$ is called the maximum likelihood (ML) dectection
- ML decision rule:

Choose $\hat{m} = m_k$ if and only if

$$f(\vec{r}|m_k) \geq f(\vec{r}|m_i);\ \text{for all } i \neq k$$

# Optimal receiver structure

- Signal model
  - Transmitter transmits a sequence of symbols or messages from a set of M symbols $m_1, m_2, \ldots, m_M$ with priori probabilities
    $$p_1 = P(m_1),\ p_2 = P(m_2),\ p_M = P(m_M)$$
  - The symbols are represented by finite energy waveforms $s_1(t), s_2(t), \ldots, s_M(t)$ defined in intervals [0,T]
  - The signal is assumed to be corrupted by additive



# Optimal receiver structure

- Signal space representation
  - Signal space of $\{s_1(t), s_2(t), \ldots, s_M(t)\}$ is assumed to be of dimension N ($N \le M$)
  - $\phi_k(t)$ for k=1,…,N will denote the orthonormal basis functions
  - Then each transmitted signal waveform can be represented as

$$s_m(t) = \sum_{k=1}^{N} s_{mk}\phi_k(t) \quad \text{where} \quad s_{mk} = \int_0^T s_m(t)\phi_k(t)dt$$

  - Note that the noise $n_w(t)$ can be written as

$$n_w(t) = \underbrace{n_0(t)}_{\text{orthogonal to the space, falls outside the signal space spanned by } \{\phi_k(t), k=1,\ldots N\}} + \underbrace{\sum_{k=1}^{N} n_k\phi_k(t)}_{\text{Projection of } n_w(t) \text{ on the N-dim space}} \quad \text{where} \quad n_k = \int_0^T n_w(t)\phi_k(t)dt$$

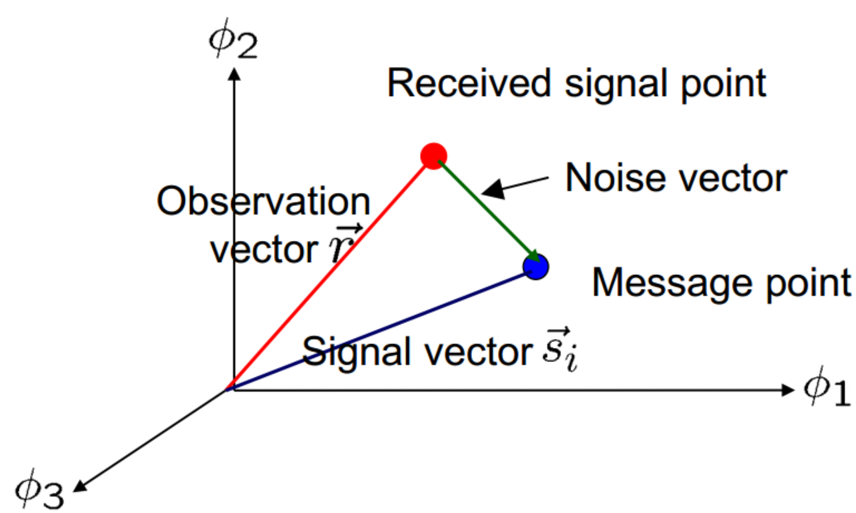
# Optimal receiver structure

- Signal space representation
  - The received signal can thus be represented as

$$
\begin{aligned}
r(t) &= s(t) + n_w(t) \\
&= \sum_{k=1}^{N} s_{mk}\phi_k(t) + \sum_{k=1}^{N} n_k\phi_k(t) + n_0(t) \\
&= \underbrace{\sum_{k=1}^{N} r_k\phi_k(t)}_{\textbf{Projection of r(t) on N-dim signal space}} + n_0(t) \qquad \text{where} \quad r_k = s_{mk} + n_k
\end{aligned}
$$

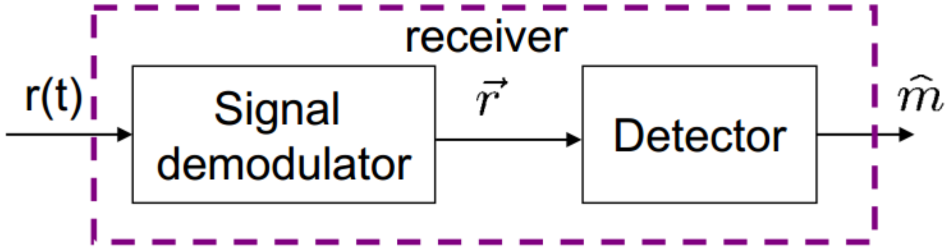
  - In vector form, we have

$$\vec{r} = \vec{s}_i + \vec{n}$$



# Optimal receiver structure

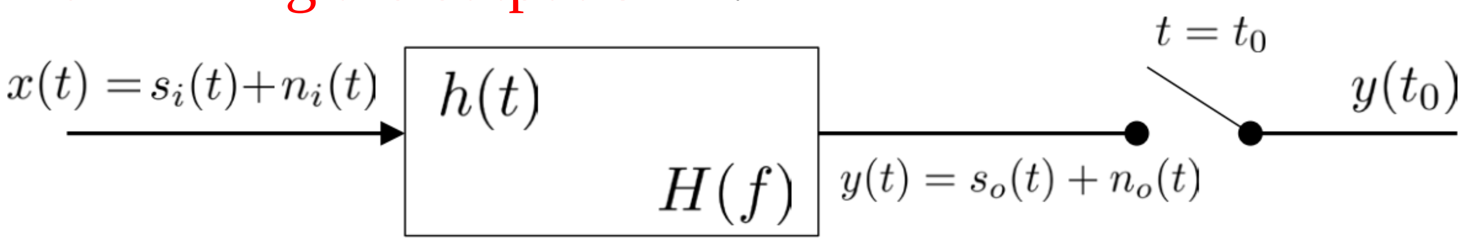
- Receiver structure
  - Signal demodulator: to convert the received wave form r(t) into an N-dim vector $\vec{r} = (r_1, r_2, \ldots, r_N)$
  - Detector: to decide which of the M possible signal waveforms was transmitted based on the observation vector $\vec{r}$




  - Two realizations of the signal demodulator: correlation type and matched-filter type

# Matched filter

- Derivation
  - The matched-filter (MF) is the optimal linear filter for maximizing the output SNR.




  - Input signal component $s_i(t) \leftrightarrow A(f)=\int_{-\infty}^{\infty} s_i(t)e^{-j\omega t}dt$
  - Input noise component $n_i(t)$ with PSD $S_{n_i}(f)=N_0/2$
  - Output signal component

$$s_o(t)=\int_{-\infty}^{\infty} s_i(t-\tau)h(\tau)d\tau$$

$$=\int_{-\infty}^{\infty} A(f)H(f)e^{j\omega t}df$$

  - Sample at $t=t_0$

# Matched filter

- Derivation
  - At the sampling instance $t = t_0$ , $s_o(t_0) = \int_{-\infty}^{\infty} A(f) H(f) e^{j\omega t_0} df$
  - Average power of the output noise is

$$N = E\{n_o^2(t)\} = \frac{N_0}{2} \int_{-\infty}^{\infty} |H(f)|^2 df$$

  - Output SNR

$$\boxed{d = \frac{s_o^2(t_0)}{E\{n_o^2(t)\}} = \frac{\left[\int_{-\infty}^{\infty} A(f) H(f) e^{j\omega t_0} df\right]^2}{\frac{N_0}{2} \int_{-\infty}^{\infty} |H(f)|^2 df}}$$

**Find $H(f)$ that can maximize $d$**

# Matched filter

- Derivation
  - Schwarz's inequality

$$\int_{-\infty}^{\infty}\left|F(x)\right|^2 dx\int_{-\infty}^{\infty}\left|Q(x)\right|^2 dx \geq \left|\int_{-\infty}^{\infty} F^*(x)Q(x)\,dx\right|^2$$

equality holds when $F(x)=CQ(x)$

  - Let $\begin{cases} F^*(x)=A(f)e^{j\omega t_0} \\ Q(f)=H(f) \end{cases}$, then

*E* : signal energy

$$\boxed{d \leq \frac{\int_{-\infty}^{\infty}\left|A(f)\right|^2 df\int_{-\infty}^{\infty}\left|H(f)\right|^2 df}{\frac{N_0}{2}\int_{-\infty}^{\infty}\left|H(f)\right|^2 df}=\frac{\int_{-\infty}^{\infty}\left|A(f)\right|^2 df}{\frac{N_0}{2}}=\frac{2E}{N_0}}$$

# Matched filter

- Derivation
  - When the max output SNR $2E/N_0$ is achieved, we have

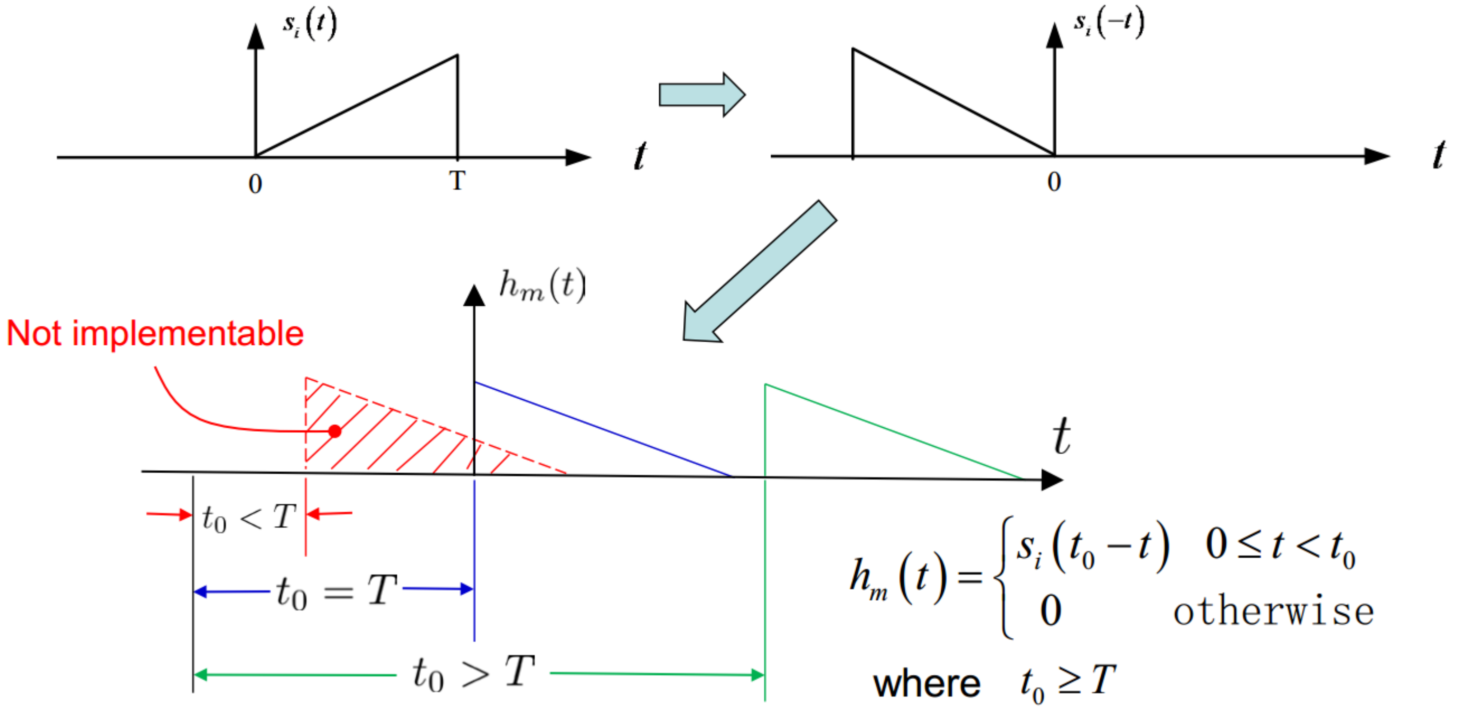
$$\boxed{\begin{aligned} H_m(f) &= A^*(f)e^{-j\omega t_0} \\ &\updownarrow \\ h_m(t) &= s_i^*(t_0 - t) \end{aligned}}$$

$$\begin{aligned} h_m(t) &= \int_{-\infty}^{\infty} H_m(f)e^{j\omega t}df \\ &= \int_{-\infty}^{\infty} A^*(f)e^{-j\omega(t_0 - t)}df \\ &= s_i^*(t_0 - t) \end{aligned}$$

  - Transfer function: complex conjugate of the input signal spectrum
  - Impulse response: time-reversal and delayed version of the input signal s(t)

# Matched filter

- Properties
  - Choice of $t_0$ versus the causality




$$h_m(t) = \begin{cases} s_i(t_0 - t) & 0 \le t < t_0 \\ 0 & \text{otherwise} \end{cases}$$

where $t_0 \ge T$

# Matched filter

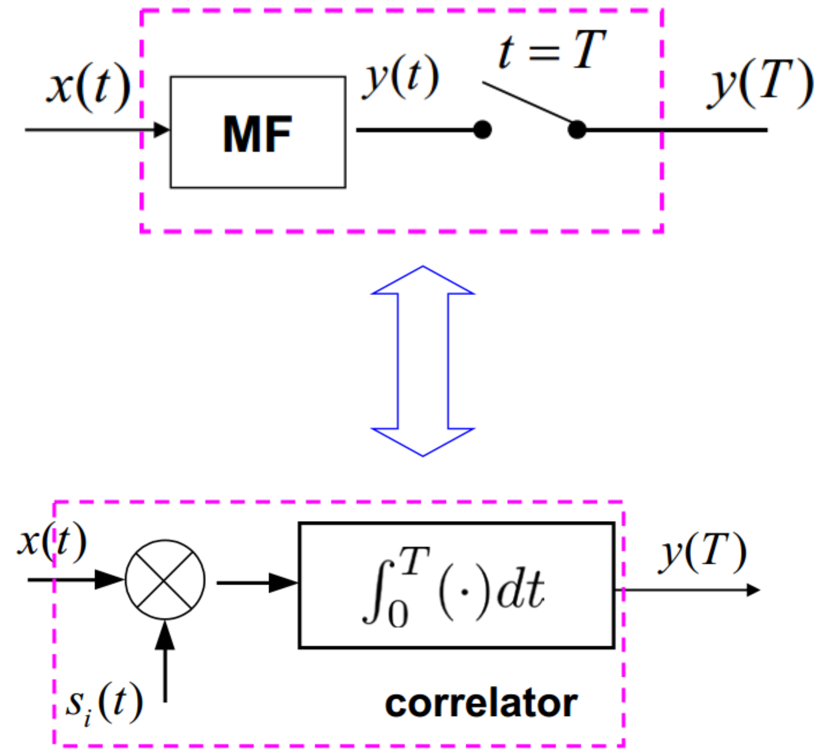
- Properties
  - Equivalent form in correlator
  - Let $s_i(t)$ be within [0,T]

$$y(t) = x(t) * h_m(t) = x(t) * s_i(T-t)$$

$$= \int_0^T x(\tau) s_i(T-t+\tau) d\tau$$

  - Observe at sampling time t=T

$$y(T) = \int_0^T x(\tau) s_i(\tau) d\tau = \int_0^T x(t) s_i(t) dt$$




Correlation integration
(相关积分)

# Matched filter

- Properties
  - Correlation function

$$R_{12}(\tau)=\int_{-\infty}^{\infty}s_1(t)s_2(t+\tau)dt=\int_{-\infty}^{\infty}s_1(t-\tau)s_2(t)dt=R_{21}(-\tau)$$

  - Auto-correlation function

$$R(\tau)=\int_{-\infty}^{\infty}s(t)s(t+\tau)dt$$

1. $R(\tau)=R(-\tau)$
2. $R(0)\geq R(\tau)$
3. $R(0)=\int_{-\infty}^{\infty}s^2(t)dt=E$
4. $R(\tau)\leftrightarrow|A(f)|^2 \qquad R(0)=\int_{-\infty}^{\infty}s^2(t)dt=\int_{-\infty}^{\infty}|A(f)|^2\,df$

# Matched filter

- Properties
  - MF output is the auto-correlation function of input signal

$$s_o(t)=\int_{-\infty}^{\infty}s_i(t-u)h_m(u)du=\int_{-\infty}^{\infty}s_i(t-u)s_i(t_0-u)du$$

$$=\int_{-\infty}^{\infty}s_i(\mu)s_i[\mu+t-t_0]d\mu=R_{s_0}(t-t_0)$$

  - The peak value of $s_0(t)$ happens

$$s_o(t_0)=\int_{-\infty}^{\infty}s_i^2(\mu)du=E$$

  - $s_0(t)$ is symmetric at $t=t_0$

$$A_o(f)=A(f)H_m(f)=|A(f)|^2e^{-j\omega t_0}$$

# Matched filter

- Properties
  - MF output noise
  - The statistical auto-correlation of $n_0(t)$ depends on the auto-correlation of $s_i(t)$

$$R_{n_o}(\tau) = E\{n_o(t)n_o(t+\tau)\} = \frac{N_0}{2}\int_{-\infty}^{\infty} h_m(u)h_m(u+\tau)du$$

$$= \frac{N_0}{2}\int_{-\infty}^{\infty} s_i(t)s_i(t-\tau)dt$$

  - Average power

$$E\{n_o^2(t)\} = R_{n_o}(0) = \frac{N_0}{2}\int_{-\infty}^{\infty} s_i^2(\mu)du$$ **Time domain**

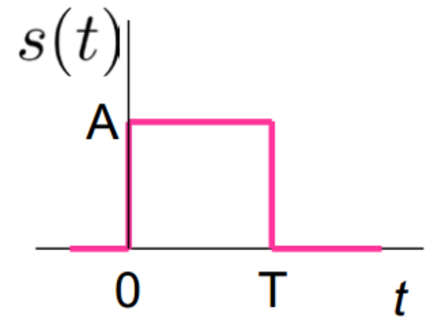
$$= \frac{N_0}{2}\int_{-\infty}^{\infty}|A(f)|^2 df = \frac{N_0}{2}\int_{-\infty}^{\infty}|H_m(f)|^2 df$$ **Frequency domain**

$$= \frac{N_0}{2}E$$

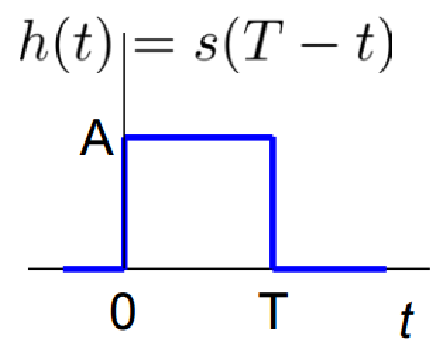
# Matched filter

- Example
  - Consider a rectangular pulse s(t)

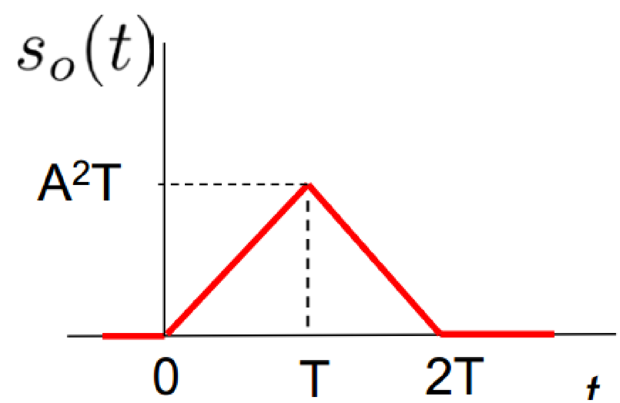
$$E_s = A^2 T$$




  - The impulse response of a filter matched to s(t) is also a rectangular pulse




  - The output of the matched filter $s_0(t)$ is $h(t) * s(t)$




  - The output SNR is

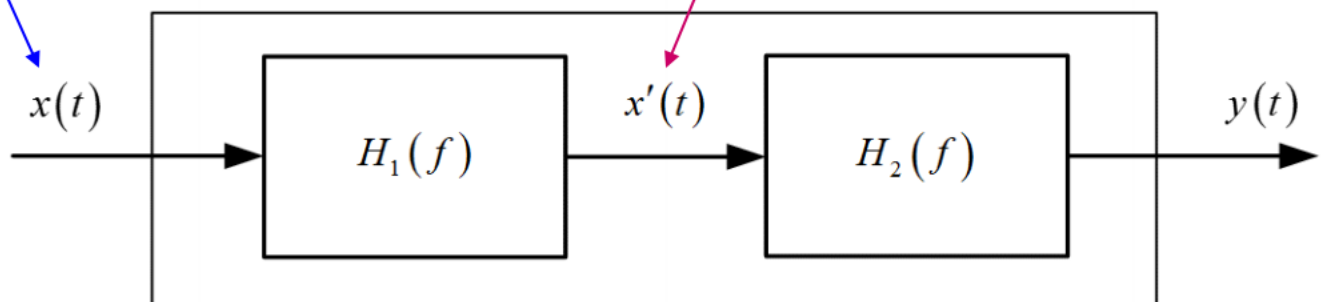
$$(SNR)_o = \frac{2}{N_0}\int_0^T s^2(t)dt = \frac{2A^2T}{N_0}$$

# Matched filter

- Colored noise
  - In case of colored noise, we need to preprocess the combined signal and noise such that the non-white noise becomes white noise- Whitening Process

$x(t) = s_i(t) + n(t)$ where $n(t)$ is colored noise with PSD $S_n(f)$

$x'(t) = s'(t) + n'(t)$




**Choose $H_1(f)$ so that $n'(t)$ is white, i.e.**

$$S_n'(f) = |H_1(f)|^2 S_n(f) = C$$

# Matched filter

- Colored noise
  - We choose

$$H_1(f):\ |H_1(f)|^2=\frac{C}{S_n(f)}$$

$H_2(f)$ should match with $S'(t)$ $\quad A'(f)=H_1(f)A(f)$

$$H_2(f)=A'^*(f)e^{-j2\pi ft_0}=H_1^*(f)A^*(f)e^{-j2\pi ft_0}$$

  - Therefore, the overall transfer function of the cascaded system:

$$H(f)=H_1(f)\cdot H_2(f)=H_1(f)H_1^*(f)A^*(f)e^{-j2\pi ft_0}$$

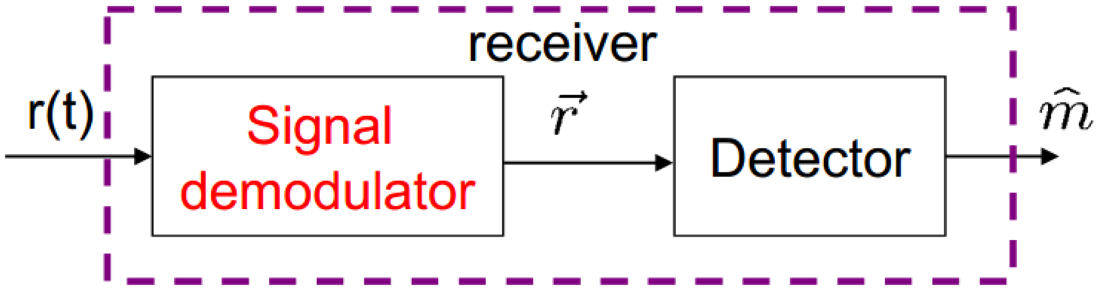
$$=|H_1(f)|^2A^*(f)e^{-j2\pi ft_0}$$

$$\boxed{H(f)=\frac{A^*(f)}{S_n(f)}e^{-j2\pi ft_0}}$$

**MF for colored noise**

# Updates on the receiver

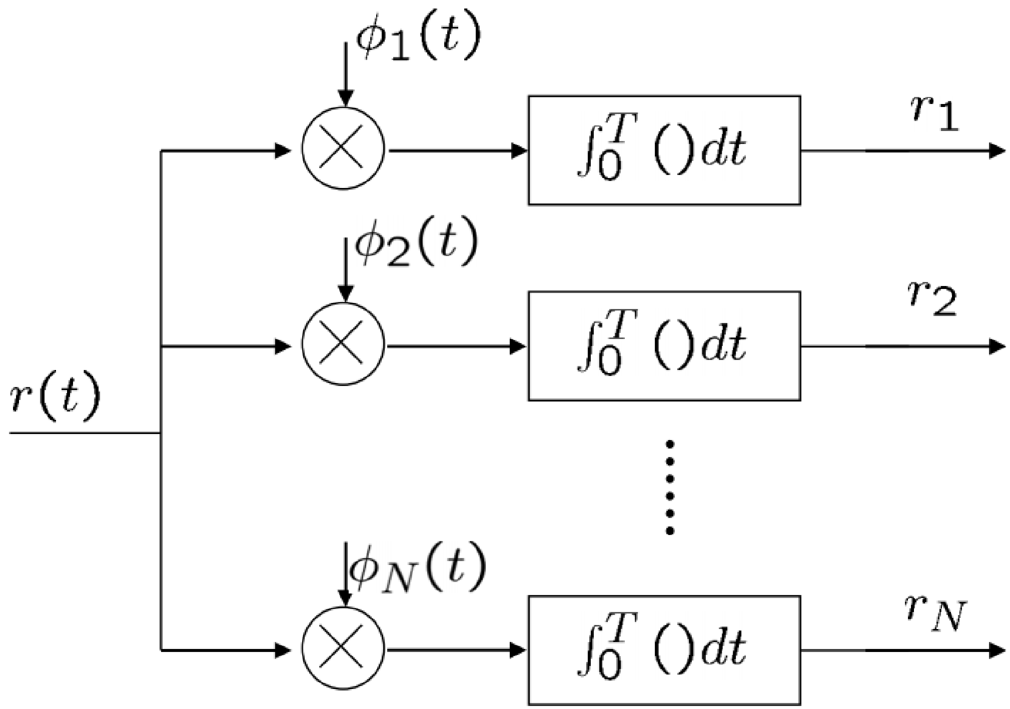
- We have talked about matched filter
- Consider the optimal receiver structure again




  - Two realizations of the signal demodulator: correlation type and matched-filter type

# Updates on the receiver

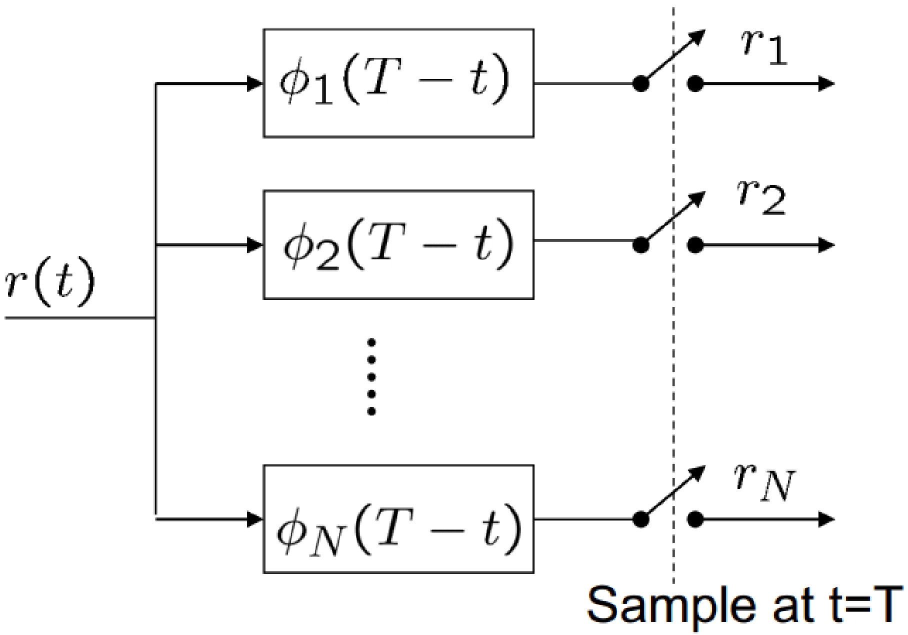
- Correlation type demodulator
  - The received signal r(t) is passed through a parallel bank of N cross correlators which basically compute the projection of r(t) onto the N basis functions $\{\phi_k(t), k = 1, \ldots N\}$



# Updates on the receiver

- Matched filter type demodulator
  - Alternatively, we may apply the received signal r(t) to a bank of N matched filters and sample the output of filters at t=T. The impulse responses of the filters are

$$h_k(t) = \phi_k(T-t), \quad 0 \le t \le T$$



# Updates on the receiver

- For a signal transmitted over an AWGN channel, either a correlation type demodulator or a matched filter type demodulator produces the vector $\vec{r} = (r_1, r_2, \ldots, r_N)$ which contains all the necessary information in r(t)
- The next step is to design a signal detector that makes a decision of the transmitted signal in each signal interval based on the observation of $\vec{r}$ , such that the probability of error is minimized (or correct probability is maximized)
- Decision rules:

likelihood function $f(\vec{r}|m_k)$

MAP decision rule:

choose $\hat{m} = m_k$ if and only if

$$P_k f(\vec{r}|m_k) > P_i f(\vec{r}|m_i); \text{ for all } i \neq k$$

ML decision rule

choose $\hat{m} = m_k$ if and only if

$$f(\vec{r}|m_k) > f(\vec{r}|m_i); \text{ for all } i \neq k$$

# Likelihood function

- Distribution of the noise vector
  - Since $n_w(t)$ is a Gaussian random process, the noise component of output $n_k = \int_0^T n_w(t)\phi_k(t)dt$ is Gaussian r.v.
  - Mean:

$$E[n_k] = \int_0^T E[n_w(t)]\phi_k(t)dt = 0 \quad , \quad k = 1, \ldots, N$$

  - Correlation between $n_j$ and $n_k$

$$\begin{aligned} E[n_j n_k] &= E\left[\int_0^T n_w(t)\phi_j(t)dt \cdot \int_0^T n_w(\tau)\phi_k(\tau)d\tau\right] \\ &= E\left[\int_0^T \int_0^T n_w(t)n_w(\tau)\phi_j(t)\phi_k(\tau)dtd\tau\right] \\ &= \int_0^T \int_0^T E[n_w(t)n_w(\tau)]\phi_j(t)\phi_k(\tau)dtd\tau \\ &= \int_0^T \int_0^T \frac{N_0}{2}\delta(t-\tau)\phi_j(t)\phi_k(\tau)dtd\tau \\ &= \frac{N_0}{2}\int_0^T \phi_j(\tau)\phi_k(\tau)d\tau = \begin{cases} \frac{N_0}{2}, & j = k \\ 0, & j \neq k \end{cases} \end{aligned}$$

**PSD of $n_w(t)$ is**
$S_n(f) = N_0/2$

# Likelihood function

- Distribution of the noise vector
  - Therefore, $n_j$ and $n_k$ ($j \neq k$) are uncorrelated Gaussian r.v.s, and hence independent with zero mean and variance $N_0/2$
  - The joint pdf of $\vec{n} = (n_1, \ldots, n_N)$

$$
\begin{aligned}
p(n_1, \ldots, n_N) &= \prod_{k=1}^{N} p(n_k) = \prod_{k=1}^{N} \frac{1}{\sqrt{\pi N_0}} \exp\left(-n_k^2/N_0\right) \\
&= (\pi N_0)^{-N/2} \exp\left(-\sum_{k=1}^{N} n_k^2/N_0\right)
\end{aligned}
$$

# Likelihood function

- Conditional probability
  - If mk is transmitted, $\vec{r} = \vec{s}_k + \vec{n}$ with $r_j = s_{kj} + n_j$

    and $E[r_j|m_k] = s_{kj} + E[n_j] = s_{kj}$

    $\boxed{\text{Transmitted signal values in each dimension represent the mean values for each received signal}}$

    and $Var[r_j|m_k] = Var[n_j] = N_0/2$
  - Therefore, the conditional pdf of

$$f(\vec{r}|m_k) = \prod_{j=1}^{N} \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{(r_j - s_{kj})^2}{N_0}\right)$$

$$= (\pi N_0)^{-N/2} \exp\left(-\frac{\sum_{j=1}^{N} (r_j - s_{kj})^2}{N_0}\right)$$

# Likelihood function

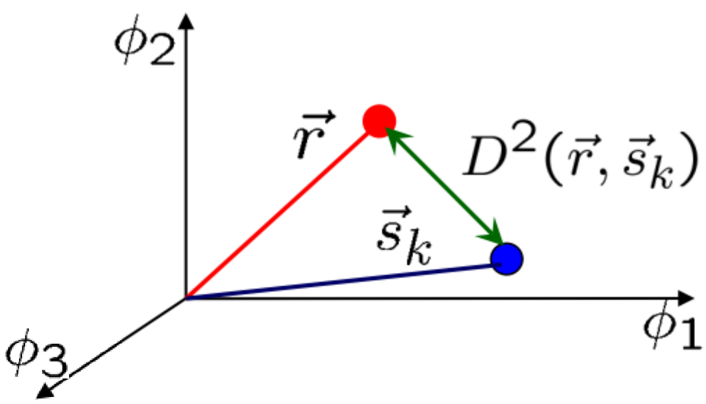
- Log-likelihood function
  - To simplify the computation, we take the natural logarithm of $f(\vec{r}|m_k)$ , which is a monotonic function. Thus,

$$\ln f(\vec{r}|m_k) = -\frac{N}{2}\ln(\pi N_0) - \frac{1}{N_0}\sum_{j=1}^{N}(r_j - s_{kj})^2$$

  - Let

$$D^2(\vec{r},\vec{s}_k) = \sum_{j=1}^{N}(r_j - s_{k,j})^2 \quad = \|\vec{r} - \vec{s}_k\|^2$$

denote the Euclidean distance between $\vec{r}$ and $\vec{s}_k$ in the N-dim signal space. It is also called the distance metric.



# Likelihood function

- Optimal detector
  - MAP rule:

$$\hat{m} = \arg \max_{\{m_1,\ldots,m_M\}} f(\vec{r}|m_k)P(m_k)$$
$$= \arg \max_{\{m_1,\ldots,m_M\}} \ln\left[f(\vec{r}|m_k)P(m_k)\right]$$
$$= \arg \max_{\{m_1,\ldots,m_M\}} \left\{-\frac{1}{N_0}\|\vec{r}-\vec{s}_k\|^2 + \ln P_k\right\}$$
$$\boxed{= \arg \min_{\{m_1,\ldots,m_M\}} \left\{\|\vec{r}-\vec{s}_k\|^2 - N_0 \ln P_k\right\}}$$

  - ML rule:

$$\boxed{\hat{m} = \arg \min_{\{m_1,\ldots,m_M\}} \|\vec{r}-\vec{s}_k\|^2}$$

ML detector chooses $\hat{m} = m_k$ iff received vector $\vec{r}$ is closer to $\vec{s}_k$ in terms of Euclidean distance than to any other $\vec{s}_i$ for i ≠ k

⇒ Minimum distance detection
(will discuss more in decision region)

# Likelihood function

- Optimal receiver structure
  - With the above expression, we can develop a receiver structure using the following derivation

$$-\sum_{j=1}^{N}(r_j - s_{kj})^2 + N_0 \ln P_k = -\sum_{j=1}^{N} r_j^2 - \sum_{j=1}^{N} s_{kj}^2 + 2\sum_{j=1}^{N} r_j s_{kj} + N_0 \ln P_k$$

$$= -\|\vec{r}\|^2 - \|\vec{s}_k\|^2 + 2\vec{r}\cdot\vec{s}_k + N_0 \ln P_k$$

with

$\|\vec{s}_k\|^2 = \int_0^T s_k^2(t)dt = E_k$ = signal energy

$\vec{r}\cdot\vec{s}_k = \int_0^T s_k(t)r(t)dt$ = correlation between the received signal vector and the transmitted signal vector

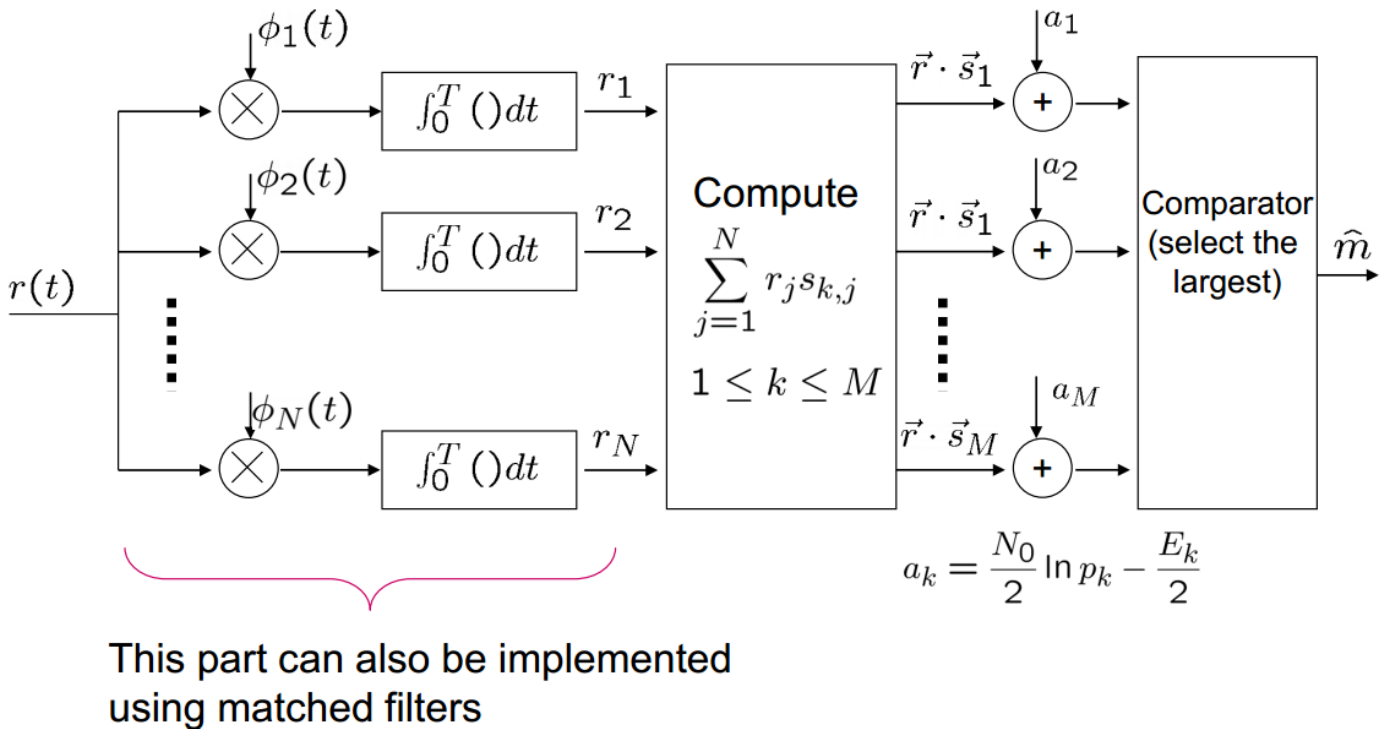
$\|\vec{r}\|^2$ = common to all M decisions and hence can be ignored

# Likelihood function

- Optimal receiver structure
  - Hence, we have

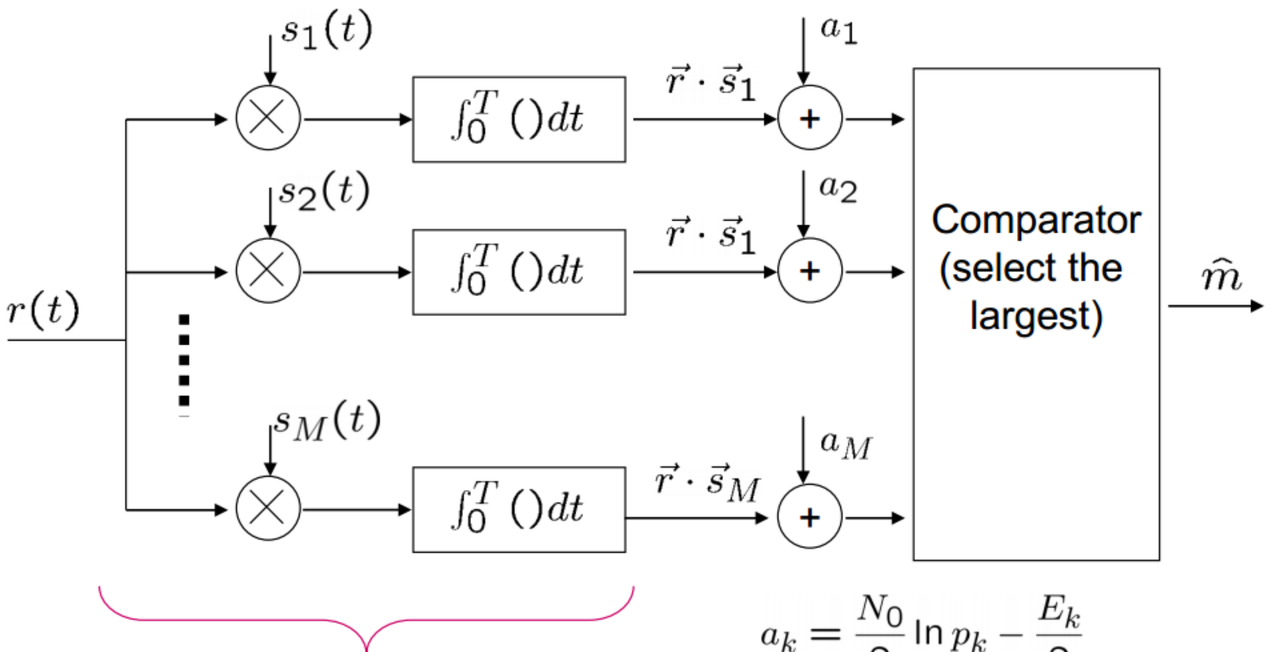
$$\hat{m} = \arg \max_{m_1,\ldots,m_M} \left\{ \vec{r}\cdot\vec{s}_k - \frac{E_k}{2} + \frac{N_0}{2}\ln P_k \right\}$$

  - The diagram of MAP receiver can be (Method 1)



# Likelihood function

- Optimal receiver structure
  - The diagram of MAP receiver can also be (Method 2)




$$a_k = \frac{N_0}{2} \ln p_k - \frac{E_k}{2}$$

This part can also be implemented using matched filters

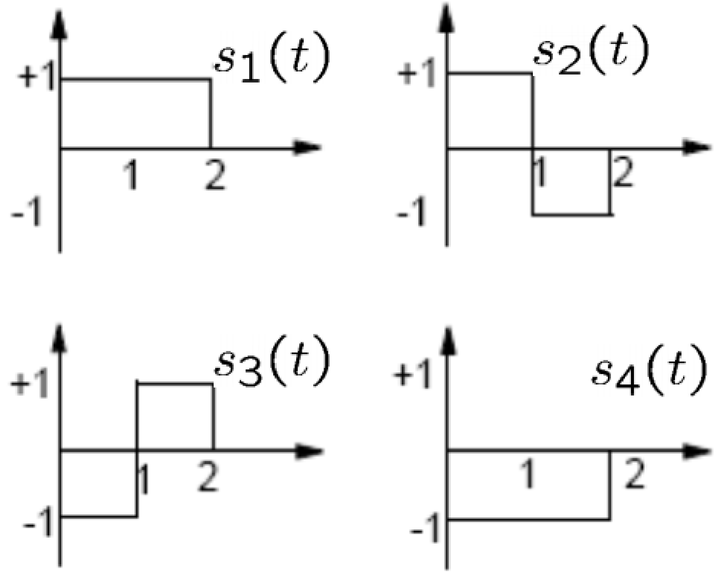
$$\hat{m} = \arg \max_{m_1,\ldots,m_M} \left\{ \vec{r} \cdot \vec{s}_k - \frac{E_k}{2} + \frac{N_0}{2} \ln P_k \right\}$$

- The structures are general and can be simplified in certain cases.

# Likelihood function

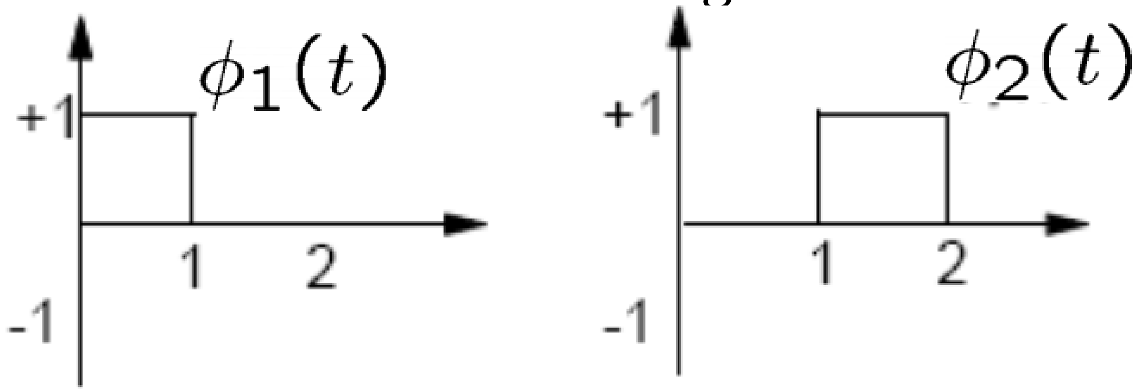
- Optimal receiver structure
  - Both receivers perform identically
  - Choice depends on circumstances
  - For instance, if N<M and $\{\phi_j(t)\}$ are easier to generate than $\{s_k(t)\}$ , then the choice is obvious

  - Consider for example the following signal set



# Likelihood function

- Optimal receiver structure
  - Suppose that we use the following basis functions



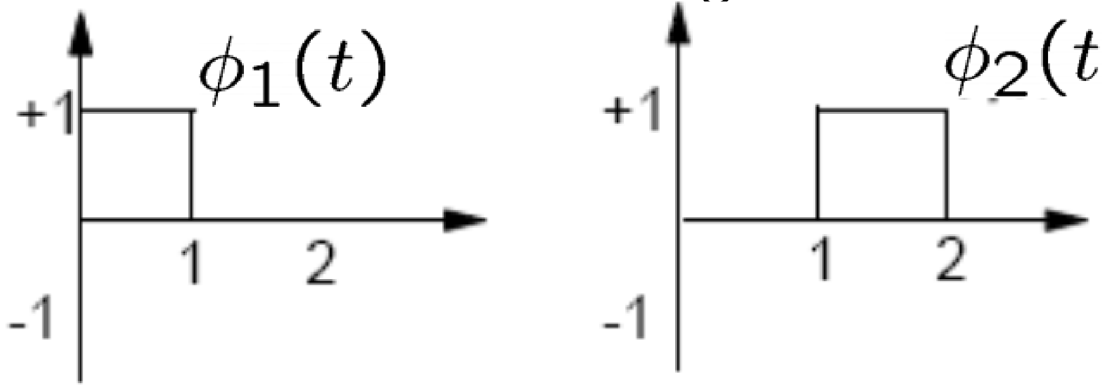

$$s_1(t) = 1\,\phi_1(t) + 1\cdot\phi_2(t) \qquad s_2(t) = 1\,\phi_1(t) - 1\cdot\phi_2(t)$$

$$s_3(t) = -1\,\phi_1(t) + 1\cdot\phi_2(t) \qquad s_4(t) = -1\,\phi_1(t) - 1\,\phi_2(t)$$

  - Since the energy is the same for all four signals, we can drop the energy term, and hence $a_k = \frac{N_0}{2} \ln p_k$

# Likelihood function

- Optimal receiver structure
  - Suppose that we use the following basis functions



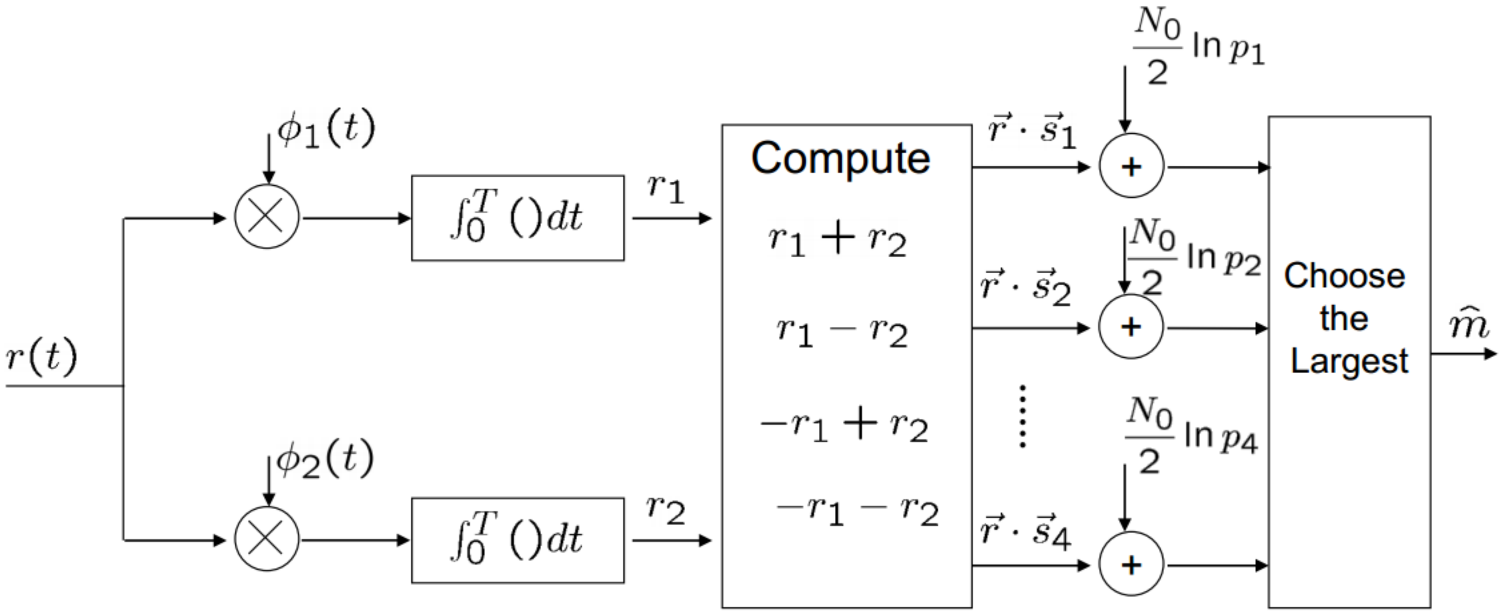

$$s_1(t) = 1\,\phi_1(t) + 1\cdot\phi_2(t) \qquad s_2(t) = 1\,\phi_1(t) - 1\cdot\phi_2(t)$$

$$s_3(t) = -1\,\phi_1(t) + 1\cdot\phi_2(t) \qquad s_4(t) = -1\,\phi_1(t) - 1\,\phi_2(t)$$

  - Since the energy is the same for all four signals, we can drop the energy term, and hence $a_k = \frac{N_0}{2}\ln p_k$

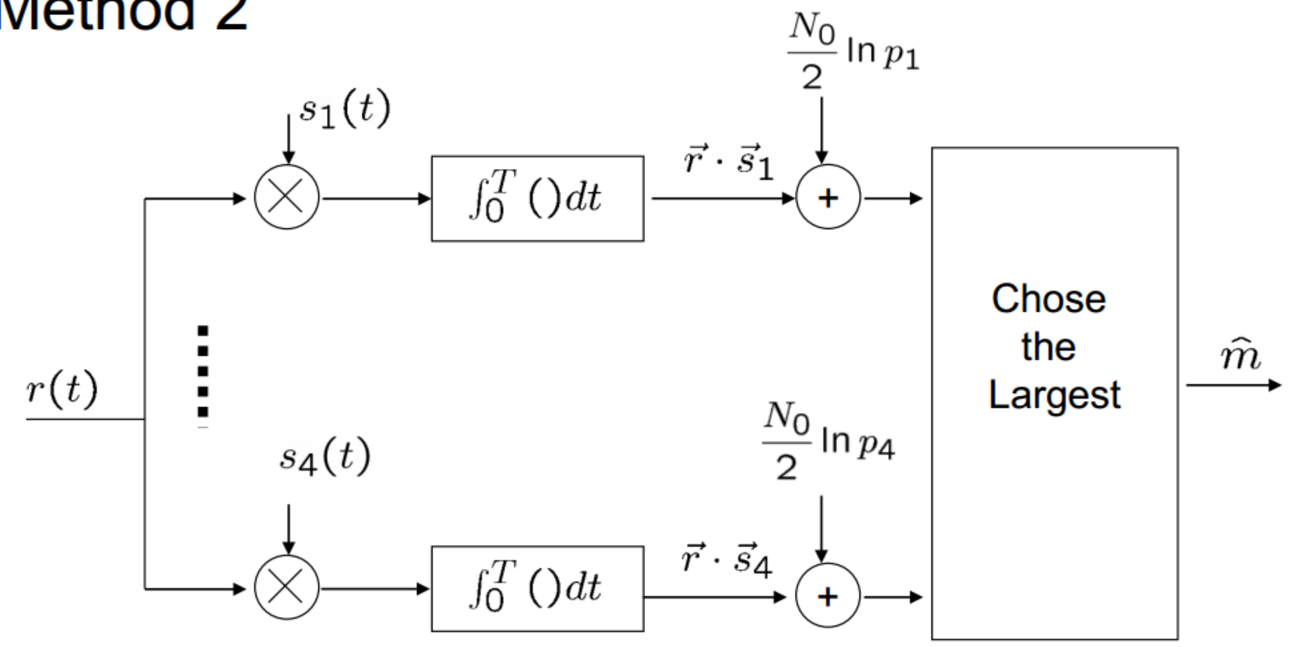
# Likelihood function

- Optimal receiver structure
  - We have (Method 1)



# Likelihood function

- Optimal receiver structure
  - Method 2



# Likelihood function

- Optimal receiver structure
  - Exercise: In an additive white Gaussian noise channel with a noise power-spectral density of N0/2, two equiprobable messages are transmitted by

$$s_1(t) = \begin{cases} \dfrac{At}{T} & 0 \le t \le \mathrm{T} \\ 0 & \text{otherwise} \end{cases}$$

$$s_2(t) = \begin{cases} A - \dfrac{At}{T} & 0 \le t \le \mathrm{T} \\ 0 & \text{otherwise} \end{cases}$$

  - Determine the optimal receiver structure.

# Decision regions

- Graphical interpretation
  - Signal space can be divided into M disjoint decision regions $R_1, R_2, \ldots, R_M$.

If $\vec{r} \in R_k$ $\Rightarrow$ decide $m_k$ was transmitted

  - Select the decision regions so that $P_e$ is minimized
  - Recall that the optimal receiver sets $\hat{m} = m_k$ iff $\|\vec{r} - \vec{s}_k\|^2 - N_0 \ln P_k$ is minimized
  - For simplicity, if one assumes $p_k = 1/M$ for all k, then the optimal receiver sets $\hat{m} = m_k$ iff $\|\vec{r} - \vec{s}_k\|^2$ is minimized

# Decision regions

- Graphical interpretation
  - Geometrically, we take projection of r(t) in the signal space (i.e., $\vec{r}$ ). Then, decision is made in favor of signal that is the closest to $\vec{r}$ in the sense of minimum Euclidean distance.
  - Specifically, the observations with $\|\vec{r}-\vec{s}_k\|^2 < \|\vec{r}-\vec{s}_i\|^2$ for all $i \neq k$ should be assigned to decision region $R_k$
  - Consider for example the binary data transmission over AWGN channel with PSD $S_n(f)=N0/2$ using

$$s_1(t) = -s_2(t) = \sqrt{E}\phi(t)$$

  - Assume $P(m_1) \neq P(m_2)$. Determine the optimal receiver and the optimal decision regions.

# Decision regions

- Graphical interpretation
  - For the above example, the optimal decision making

$$\|\vec{r}-\vec{s}_1\|^2 - N_0 \ln P(m_1) \underset{\text{Choose } m_2}{\overset{\text{Choose } m_1}{\lessgtr}} \|\vec{r}-\vec{s}_2\|^2 - N_0 \ln P(m_2)$$

  - Let $d_1 = \|\vec{r}-\vec{s}_1\|$ and $d_2 = \|\vec{r}-\vec{s}_2\|$
  - Equivalently,

$$d_1^2 - d_2^2 \underset{\text{Choose } m_2}{\overset{\text{Choose } m_1}{\lessgtr}} N_0 \ln \frac{P(m_1)}{P(m_2)}$$

Constant c

  - Therefore

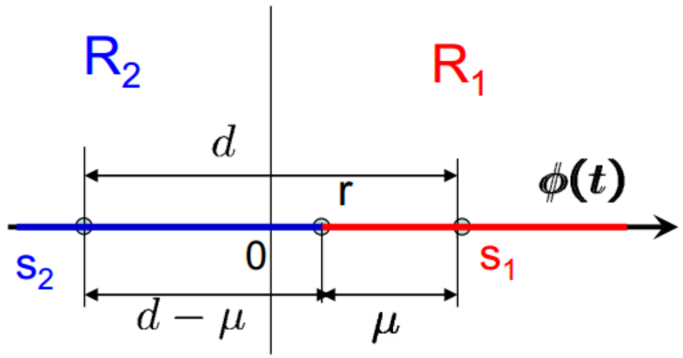
R$_1$: $d_1^2 - d_2^2 < c$ **and** R$_2$: $d_1^2 - d_2^2 > c$

# Decision regions

- Graphical interpretation
  - Now consider the example with $\vec{r}$ on the decision boundary

$$\begin{cases} d = d_1 + d_2 \\ d_1^2 = \mu^2 \\ d_2^2 = (d-\mu)^2 \end{cases} \Rightarrow \quad d_1^2 - d_2^2 = 2d\mu - d^2 \equiv c$$

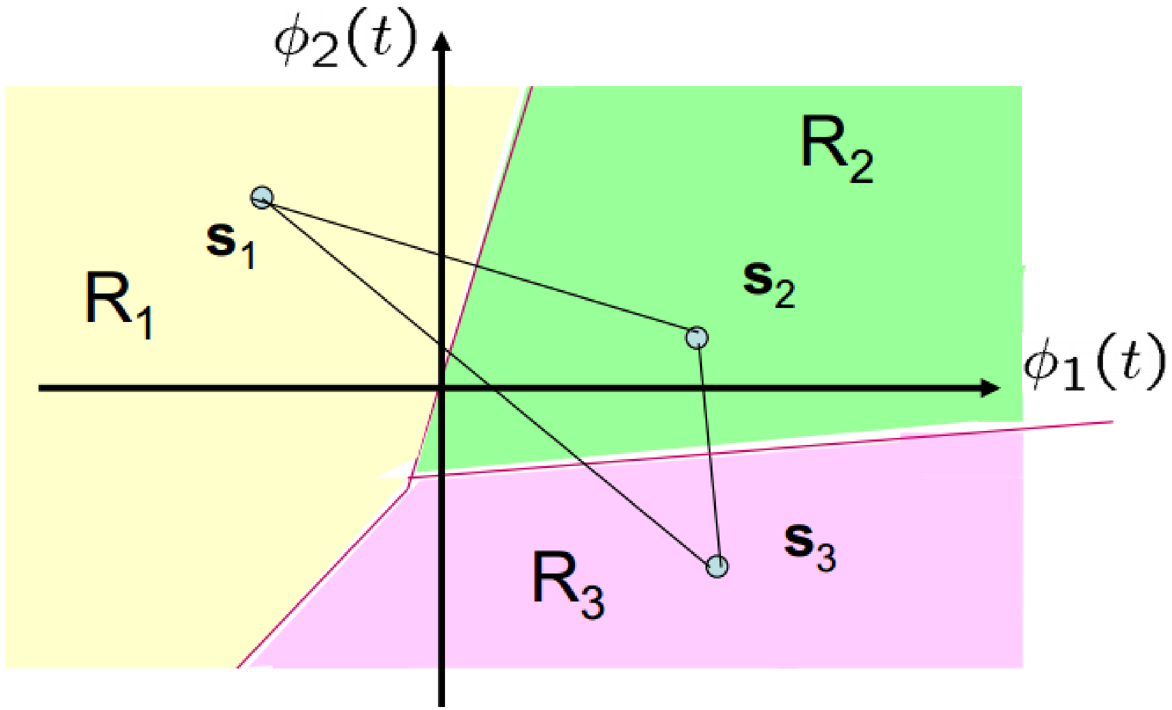
$$\mu = \frac{c + d^2}{2d} = \frac{d}{2} + \frac{N_0}{2d} \ln \frac{P(m_1)}{P(m_2)}$$




$$\mu \begin{cases} = d/2 & \text{if } P(m_1) = P(m_2) \\ > d/2 & \text{if } P(m_1) > P(m_2) \\ < d/2 & \text{if } P(m_1) < P(m_2) \end{cases}$$

# Decision regions

- Graphical interpretation
  - In general, boundaries of decision regions are perpendicular bisectors of the lines joining the original transmitted signals
  - Example: three equiprobable 2-dim signals



# Decision regions

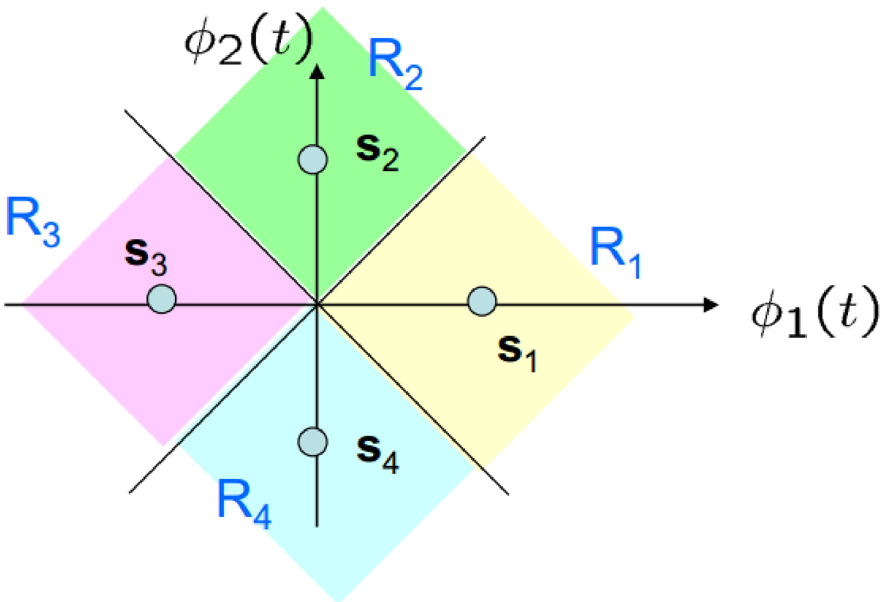
- Graphical interpretation
  - Now consider for example the decision regions for QPSK
  - Assume all signals are equally likely and all 4 signals could be written as the linear combination of two basis functions
  - Constellations of 4 signals

$\mathbf{s}_1=(1,0)$

$\mathbf{s}_2=(0,1)$

$\mathbf{s}_3=(-1,0)$

$\mathbf{s}_4=(0,-1)$



# Decision regions

- Graphical interpretation
  - Exercise: Three equally probable messages $m_1$, $m_2$, and $m_3$ are to be transmitted over an AWGN channel with noise power-spectral density N0/2. The messages are

$$s_1(t)=\begin{cases}1 & 0\le t\le T\\ 0 & otherwise\end{cases}\qquad s_2(t)=-s_3(t)=\begin{cases}1 & 0\le t\le \frac{T}{2}\\ -1 & \frac{T}{2}\le t\le T\\ 0 & otherwise\end{cases}$$

  - What is the dimensionality of the signal space?
  - Find an appropriate basis for the signal space (Hint: you donot need to perform Gram-Schmidt procedure.)
  - Draw the signal constellation for this problem.
  - Sketch the optimal decisions $R_1$, $R_2$, and $R_3$.

# Error probability analysis

- Probability of error
  - Suppose $m_k$ is transmitted and $\vec{r}$ is received
  - Correct decision is made when $\vec{r} \in R_k$ with probability

$$P(C|m_k) = P(\vec{r} \in R_k | m_k \text{ is sent})$$

  - Averaging over all possible transmitted symbols, we obtain the average probability of making correct decision

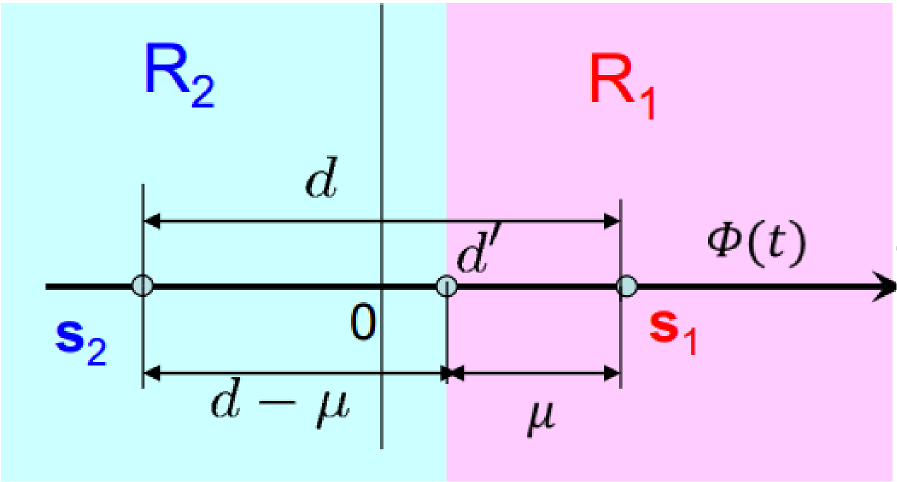
$$P(C) = \sum_{k=1}^{M} P(\vec{r} \in R_k | m_k \text{ is sent}) P(m_k)$$

  - Average probability of error

$$\boxed{P_e = 1 - P(C) = 1 - \sum_{k=1}^{M} P(\vec{r} \in R_k | m_k \text{ is sent}) P(m_k)}$$

# Error probability analysis

- Probability of error
  - Consider for example the binary data transmission




$$\mu = \frac{d}{2} + \frac{N_0}{2d} \ln \frac{P(m_1)}{P(m_2)}$$

  - Given m1 is transmitted, then

$$P(C|s_1) = P(r \in R_1|s_1) = P(s_1 + n > d') = P(n > -\mu)$$

  - Sin n is Gaussian with variance N0/2

$$P(C|s_1) = 1 - Q\left(\frac{\mu}{\sqrt{N_0/2}}\right)$$

# Error probability analysis

- Probability of error
  - Similarly, we have

$$P(C|s_2) = P(s_2 + n < d') = P(n < d - u) = 1 - Q\left(\frac{d-\mu}{\sqrt{N_0/2}}\right)$$

  - Thus,

$$P(C) = P(m_1)\left\{1 - Q\left[\frac{\mu}{\sqrt{N_0/2}}\right]\right\} + P(m_2)\left\{1 - Q\left[\frac{d-\mu}{\sqrt{N_0/2}}\right]\right\}$$

$$= 1 - P(m_1)Q\left[\frac{\mu}{\sqrt{N_0/2}}\right] - P(m_2)Q\left[\frac{d-\mu}{\sqrt{N_0/2}}\right]$$

$$\Rightarrow \boxed{P_e = P(m_1)Q\left[\frac{\mu}{\sqrt{N_0/2}}\right] + P(m_2)Q\left[\frac{d-\mu}{\sqrt{N_0/2}}\right]}$$

where $d = 2\sqrt{E}$

$$\mu = \frac{N_0}{4\sqrt{E}}\log\left[\frac{P(m_1)}{P(m_2)}\right] + \sqrt{E}$$

# Error probability analysis

- Probability of error
  - Note that when $P(m_1) = P(m_2)$

$$\mu = \sqrt{E} = \frac{d}{2}$$

$$\boxed{P_e = Q\left[\frac{d/2}{\sqrt{N_0/2}}\right] = Q\left[\sqrt{\frac{d^2}{2N_0}}\right] = Q\left[\sqrt{\frac{2E}{N_0}}\right]}$$

  - This shows us that:

1. When optimal receiver is used, $P_e$ does not depend on the specific waveform used.
2. $P_e$ depends only on their geometrical representation in signal space
3. In particular, $P_e$ depends on signal waveforms only through their energies (distance)

# Error probability analysis

- Graphical interpretation
  - Exercise: Three equally probable messages $m_1$, $m_2$, and $m_3$ are to be transmitted over an AWGN channel with noise power-spectral density N0/2. The messages are

$$s_1(t) = \begin{cases} 1 & 0 \le t \le T \\ 0 & otherwise \end{cases} \qquad s_2(t) = -s_3(t) = \begin{cases} 1 & 0 \le t \le \frac{T}{2} \\ -1 & \frac{T}{2} \le t \le T \\ 0 & otherwise \end{cases}$$

  - Which of the three messages is more vulnerable to errors and why? In other words, which of the probability of error $p(Error \mid m_i \quad transmitted)$ is larger?

# Error probability analysis

- General expression
  - Average probability of symbol error

$$P_e = 1 - P(C) = 1 - \sum_{k=1}^{M} P(\vec{r} \in R_k | m_k \text{ is sent}) P(m_k)$$

  - Since

$$P(\vec{r} \in R_k | m_k \text{ is sent}) = \underbrace{\int_{R_k}}_{\text{N-dim integration}} \underbrace{f(\vec{r} | m_k)}_{\text{Likelihood function}} d\vec{r}$$

  - Thus, we can rewrite Pe in terms of likelihood functions, assuming that symbols are equally likely to be sent

$$P_e = 1 - \frac{1}{M} \sum_{k=1}^{M} \int_{R_k} f(\vec{r} | m_k) d\vec{r}$$

  - Multi-dimension integrals are quite difficult to evaluate. To overcome the difficulty, we resort to the use of bounds. Then, we can obtain a simple and yet useful bound of $P_e$, called **union bound**.

# Error probability analysis

- General expression
  - Let $A_{kj}$ denote the event that $\vec{r}$ is closer to $\vec{s}_j$ than to $\vec{s}_k$ in the signal space when $m_k$ ( $\vec{s}_k$ ) is sent
  - Conditional probability of symbol error when $m_k$ is sent

$$P(error|m_k) = P(\vec{r} \notin R_k|m_k) = P\left(\bigcup_{j \neq k} A_{kj}\right)$$

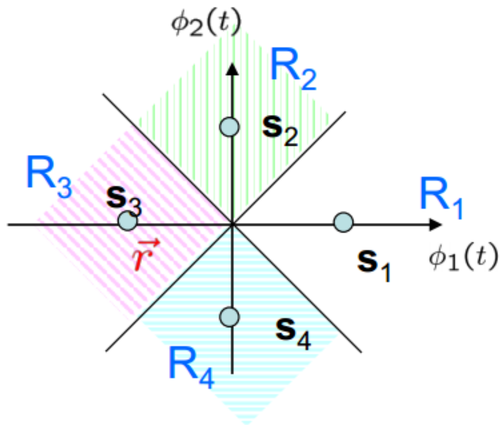
  - Note that

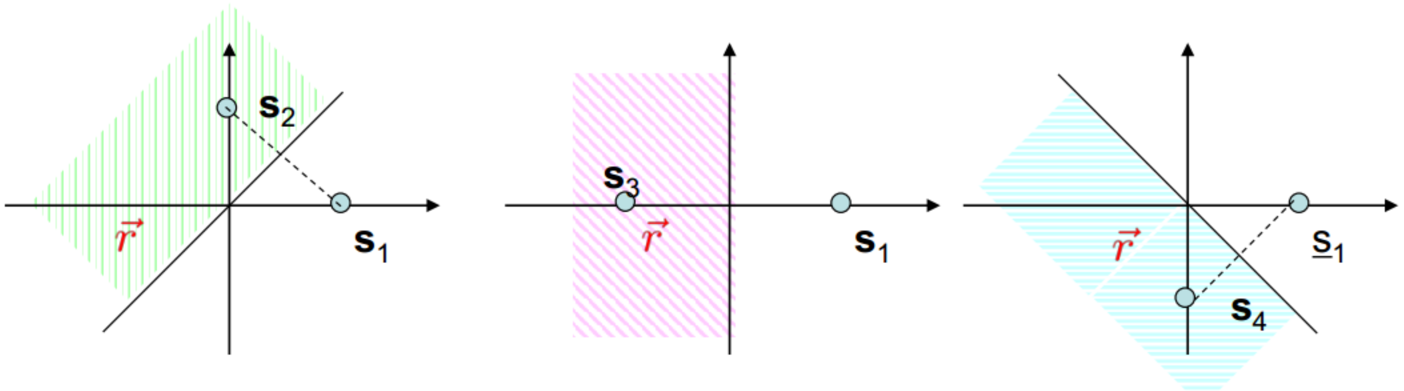
$$\boxed{P\left(\bigcup_{j \neq k} A_{kj}\right) \leq \sum_{\substack{j=1 \\ j \neq k}}^{M} P\left(A_{kj}\right)}$$

# Error probability analysis

- General expression
  - Consider for example

$A_{12} \cup A_{13} \cup A_{14}$




$A_{12}$ $A_{13}$ $A_{14}$

# Error probability analysis

- General expression
  - Define the pair-wise error probability as

$$P(\vec{s}_k \to \vec{s}_j) = P(A_{kj})$$

  - It is equivalent to the probability of deciding in favor of $\vec{s}_j$ when $\vec{s}_k$ was sent in a simplified binary system that involves the use of two equally likely messages $\vec{s}_k$ and $\vec{s}_j$
  - Then

$$P\left(\vec{s}_k \to \vec{s}_j\right) = P\left(n > d_{kj}/2\right) = Q\left(\sqrt{\frac{d_{kj}^2}{2N_0}}\right)$$

where $d_{kj} = \|\vec{s}_k - \vec{s}_j\|$ is the Euclidean distance between $\vec{s}_k$ and $\vec{s}_j$

  - Therefore the conditional error probability

$$P(error|m_k) \le \sum_{\substack{j=1 \\ j \ne k}}^{M} P(\vec{s}_k \to \vec{s}_j) = \sum_{\substack{j=1 \\ j \ne k}}^{M} Q\left(\sqrt{\frac{d_{kj}^2}{2N_0}}\right)$$

# Error probability analysis

- General expression
  - Finally, with M equally likely messages, the average probability of symbol error is upperbounded by

$$\boxed{\begin{aligned} P_e &= \frac{1}{M}\sum_{k=1}^{M} P(error|m_k) \\ &\leq \frac{1}{M}\sum_{k=1}^{M}\sum_{\substack{j=1\\ j\neq k}}^{M} Q\left(\sqrt{\frac{d_{kj}^2}{2N_0}}\right) \end{aligned}}$$

The most general formulation of union bound

  - Let $d_{\min}$ denote the minimum distance, i.e., $d_{\min} = \min_{\substack{k,j\\ k\neq j}} d_{k,j}$
  - Since Q-function is a monotone decreasing function

$$\sum_{\substack{j=1\\ j\neq k}}^{M} Q\left(\sqrt{\frac{d_{kj}^2}{2N_0}}\right) \leq (M-1)Q\left(\sqrt{\frac{d_{\min}^2}{2N_0}}\right)$$

# Error probability analysis

- General expression
  - Consequently, we may simplify the union bound as

$$\boxed{P_e \le (M-1)Q\left(\sqrt{\frac{d_{\min}^2}{2N_0}}\right)}$$

Simplified form of union bound

  - Think about: What is the design criterion of a good signal set?