Update/Key Lecture Materials

Geometric approach!

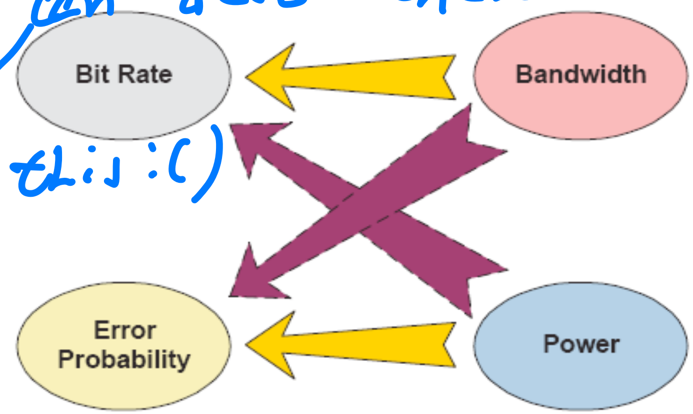
- **Have Considered**
  - Geometrical representation of signals.
  - Signal space, signals as vectors.
  - Signal space dimension, independent vectors.
  - Basis Vectors/functions, Orthogonal/orthonormal signals.
  - Systematic determination of an orthogonal basis set (Gram-Schmidt Orthogonalization or GSO process).

» **We will now consider**
  - M-ary Modulation
  - Optimum signal detection.
  - We will determine the optimum receiver (in the sense of minimizing P_e) for general M-ary signaling in the presence of AWGN.

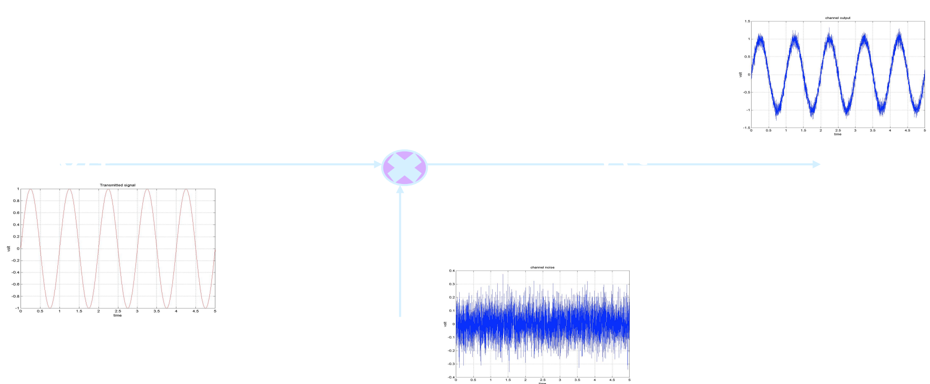
Review of Digital Communication

- Three components:
  - **Transmitter (or digital modulator),**
  - **Receiver (or demodulator),**
  - **Channel.** (no freedom design this :()

can design these!



- Quality (Performance) Measure
  - **Bit Error Rate (BER), Bit Rate.**
- Resource
  - **Bandwidth, Power**
- Channel Example
  - **AWGN (Additive White Gaussian Noise).** $y(t) = x(t) + n(t)$
  - **Channel just introduce noise without further distortion.**



Digital Modulator

- **Modulation symbols**
  - **signal pulses of finite duration** T_s.
  - **For example, binary modulator takes in 1 bit {0,1} and output one modulation symbol from the set** $\{s_0(t), s_1(t)\}$
- **Baud Rate (Symbol Rate)**

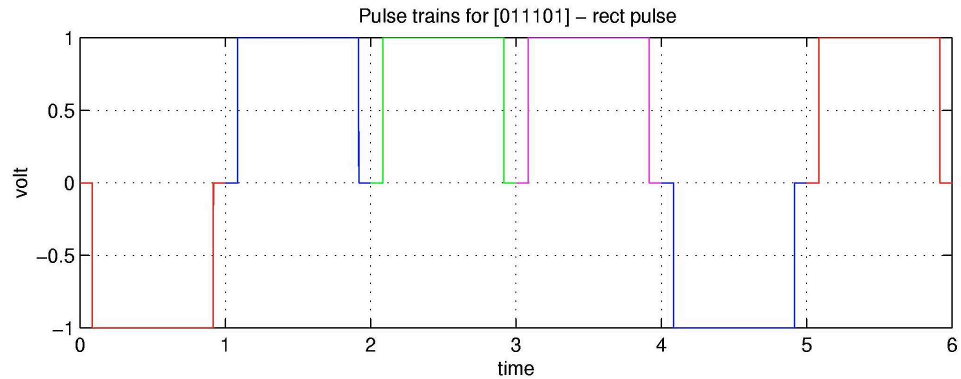
Symbol Rate: $\frac{1}{T_s}$

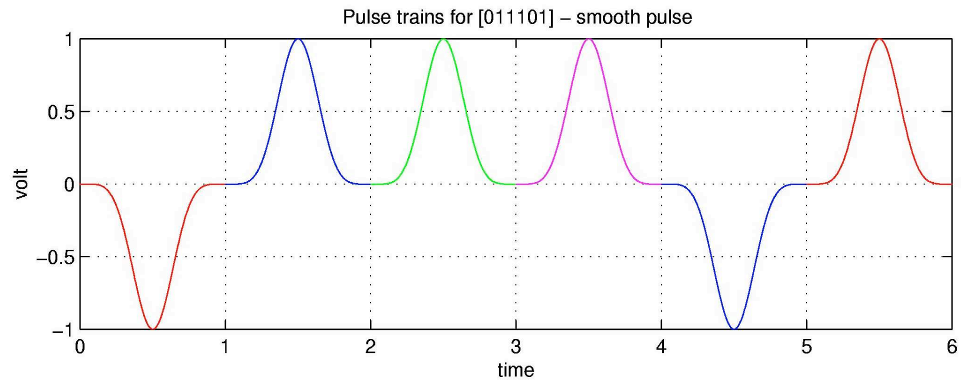
- **Bit Rate (bps) =** $\left(\frac{1}{T_s}\right) \times \left(\frac{bits}{symbol}\right)$
- **Transmission BW (bandpass)**

$$W = \left(\frac{1}{T_s}\right)(1+\alpha)$$

- **M-ary Modulator**
  - **One out of M symbols**

$$\{s_1(t), s_2(t), \ldots, s_M(t)\}$$





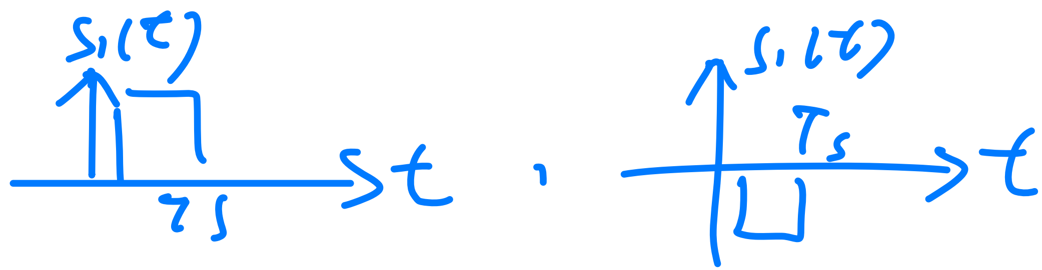
$\rightarrow$ Binary modulator $\rightarrow$ $s(t) = \begin{cases} s_1(t) & b=1 \\ s_2(t) & b=0 \end{cases}$

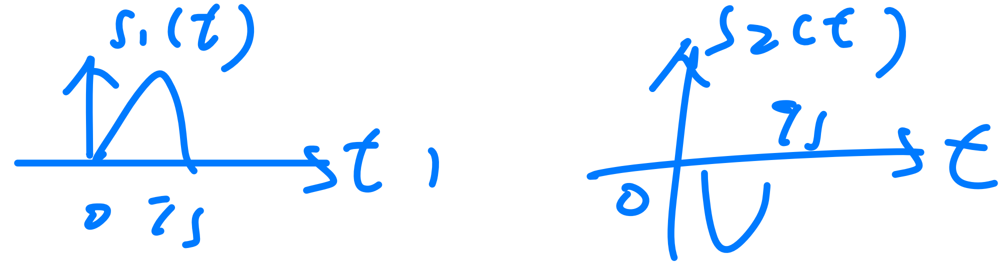
① T_s

② $S = \{ \underset{[1]}{s_1(t)} \quad \underset{[0]}{s_2(t)} \}$

$\Rightarrow$ able to completely specify the behaviour of this modulator

$S^A = \{$



$S^B = \{$  $\}$

① $R_b = \frac{1}{\tau_s}$ b/s - bit rate

② $W_{tx} = (1+\beta)\frac{1}{\tau_s} \rightarrow = \frac{1}{T_s} \rightarrow$ expensive!!!

$\uparrow$ Bandpass; $\beta \downarrow 0$

If reduce the duration by $\frac{1}{2}$, the W_{tx} will $\times 2 \Rightarrow$ Tradeoff

①, ② are tradeoffs!

$E_s = \frac{1}{2}(E_1 + E_2)$

③ $P_t = \frac{E_s}{T_s}$

= 4 possibilities!

in general M-ary modulator

$b_1 b_2 \ldots \rightarrow$ Quanturary mod $\rightarrow s(t)$

①

② $S = \{ \overset{T_s}{s_1(t)}, s_2(t), s_3(t), s_4(t) \ldots \}$

[00] [01] [10] [11]

① $R_b = 2 \times \frac{1}{T_s} = (\log_2 M) \frac{1}{T_s}$

\# of 'passengers'

Think about like minibus, in binary modulation, only carry one passenger in one time

- M is dimension of the signal space S

Expand bit sequence, expand the

$\Rightarrow$ multiple of R_b

Trade-off of sending more bits per symbol

- Error probability p_e ↑↑

Specification of Digital Modulator

Symbol Duration T_s: The symbol duration T_s specifies how often the modulator produces a *modulation symbol*. The *baud rate* or *symbol rate* of a digital modulator is defined as the number of modulation symbols per second, which is given by:

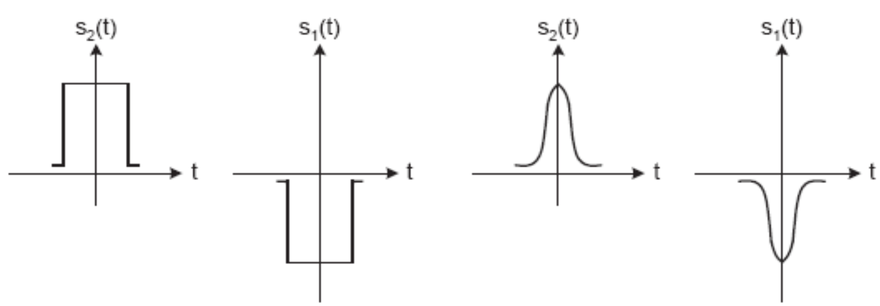
$$R_{baud} = \frac{1}{T_s} \tag{2.7}$$

The baud rate or symbol rate is a very important parameter because it determines the required channel bandwidth of a digital modulator.

Signal Set $\mathcal{S}$: The signal set of a digital modulator is a collection of M time domain signals given by:

$$\mathcal{S}_M = \{s_1(t), \ldots, s_M(t)\} \tag{2.8}$$

for some integer M. Each of the signal element in the signal set $\mathcal{S}$ is labeled with a unique bit pattern which will be explained below.



(a) binary modulator A (b) binary modulator B

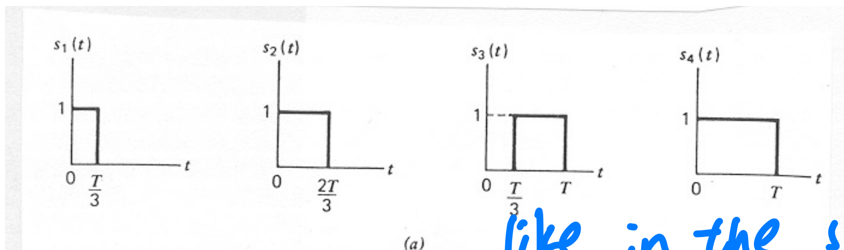
$$W_{tx} = (1+\alpha)\frac{1}{T_s}$$

$$P_{tx} = \frac{1}{MT_s}\|\vec{s}_m\|^2 = \frac{1}{MT_s}\int_0^{T_s} |s_m(t)|^2 dt.$$

Specification of a Digital Modulator

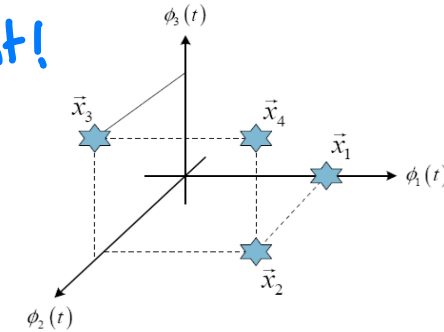
Energy $= \frac{1}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2)$

- How to specify the signal set of a digital modulator?
  - Time Domain



Equivalent!

like in the sky



  - Geometric Domain (Constellation)
    - » **The set of M modulation signals could be represented geometrically as points in the signal space.**
    - » **Hence, M-ary modulator could be represented by M-points in a D-dim signal space. This is called the constellation of the modulator.**
    - » **Average Energy out of modulator**
      - (equivalent to average distance-square of all the constellation points from origin)

$$E = \frac{1}{M}\sum_{i=1}^{M} |\vec{s}_i|^2$$

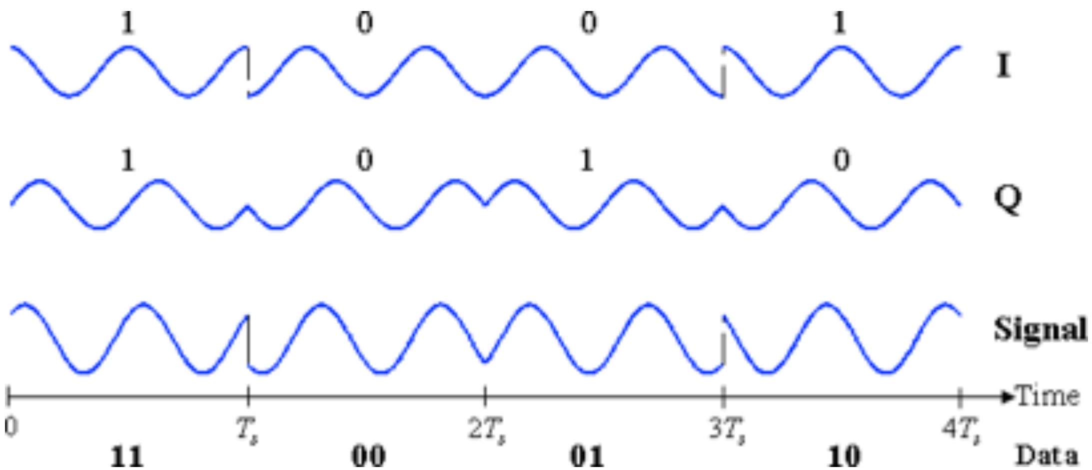
    - » **Average Transmit Power**

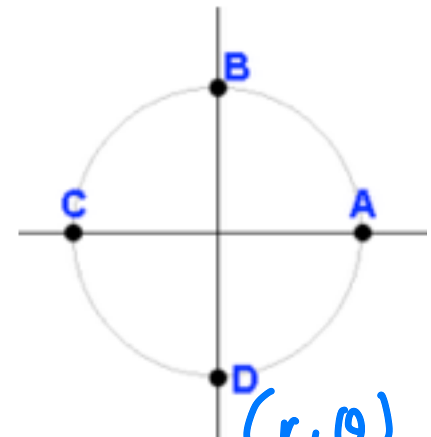
$$P_{tx} = \frac{1}{T_s}E = \frac{1}{MT_s}\sum_{i=1}^{M} \|\vec{s}_i\|^2$$

Examples of M-ary Modulation

- **QPSK Modulation**
  - 2 bits per symbol
  - Bit rate = 2 x symbol rate

Draw 4 diagram





(r, θ)

$\vec{s} = (s_1, s_2) \in \mathbb{R}$

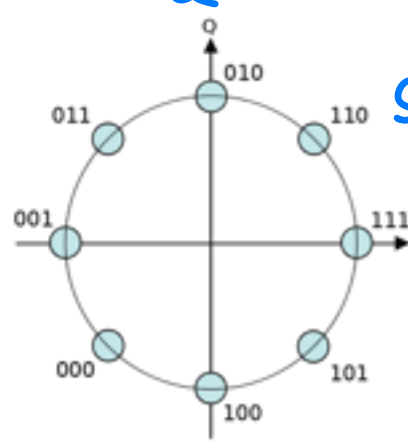
$\Updownarrow$ isomorphism

$s = s_1 + i s_2 \in \mathbb{C}'$

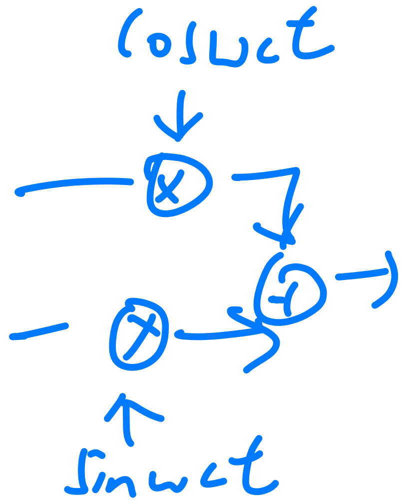
- **8PSK Modulation**
  - 3 bits per symbol
  - Bit rate = 3 x symbol rate

Too troublesome in time domain!

Q



I

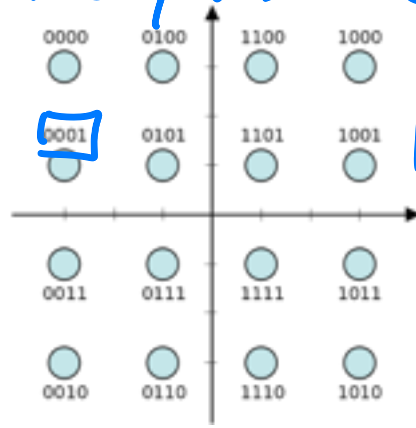

cos ωct
sin ωct

Examples of M-ary Modulation

nobody uses time domain

special structure

- **16-QAM Modulation**
  - 4 bits per symbol
  - Bit rate = 4 x symbol rate $= \frac{1}{T_s}$



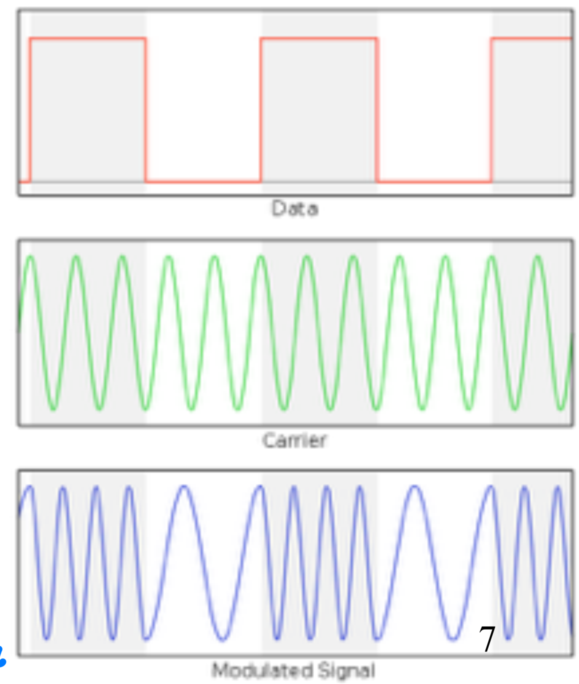
= waveform

1 symbol

→ 4 bits

Digital FM

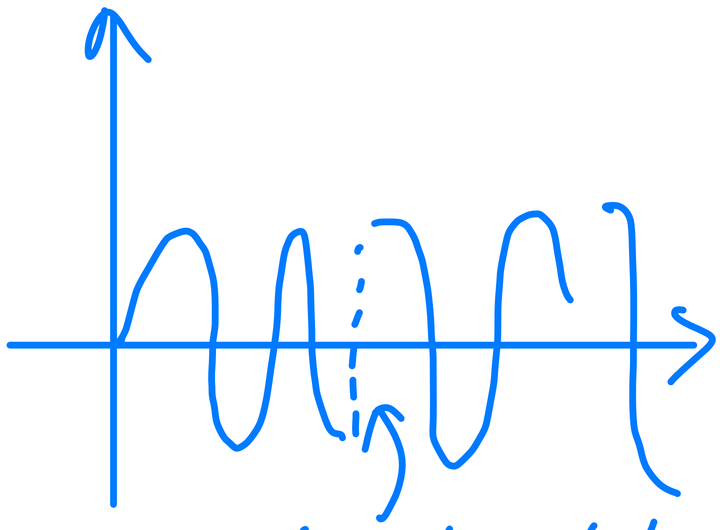
- **Frequency Shift Keying (FSK)**
  - Binary FSK
    - » 1 bit per symbol
    - » Bit rate = 1 x symbol rate
  - 4FSK
    - $\omega_1, \omega_2, \omega_3, \omega_4$
    - » Signal set has 4 tones
    - » 2 bits per symbol
    - » Bit rate = 2 x symbol rate
  - MFSK



$\Delta\omega = \frac{1}{T_s} \Rightarrow \omega_1 + \frac{1}{T_s} = \omega_2$

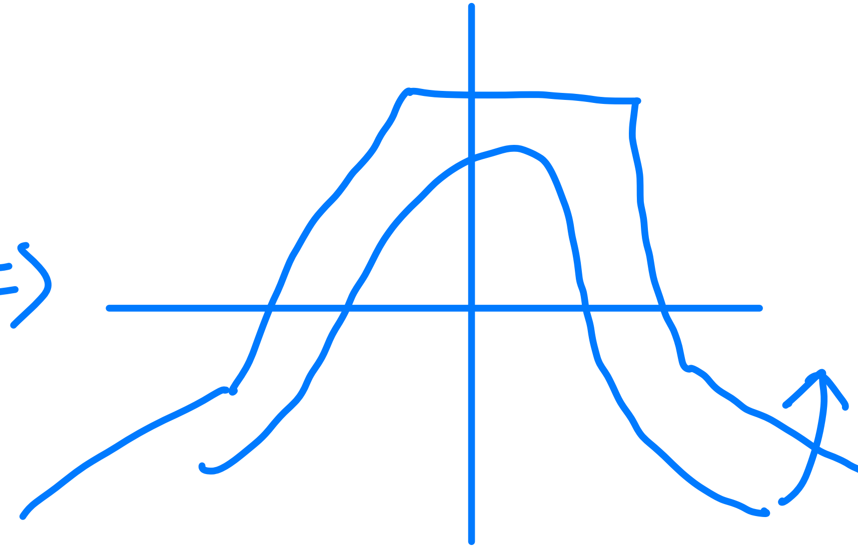
$\Rightarrow$ switch w_1, w_2, w_3, w_4

$\Rightarrow$ phase will be continuous!



discontinuity!

$\Rightarrow$



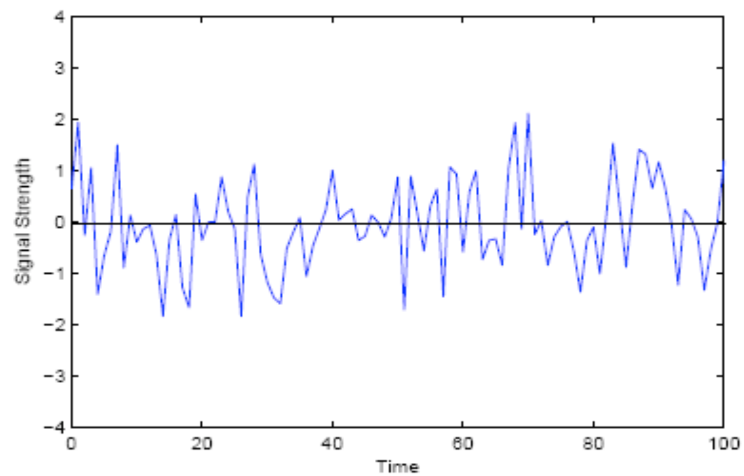
spectral regrowth!!!

犯法的!

Received Signal - Time Domain View

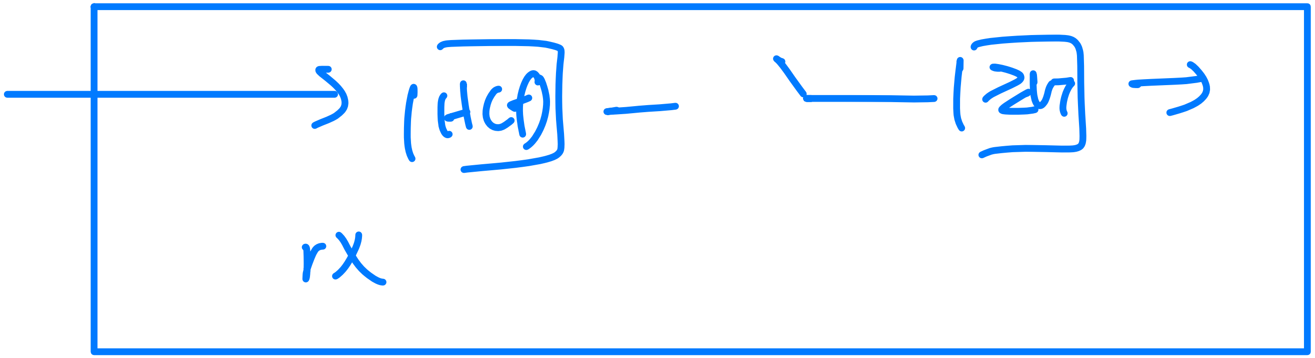
- Received signal after AWGN channel is given by:

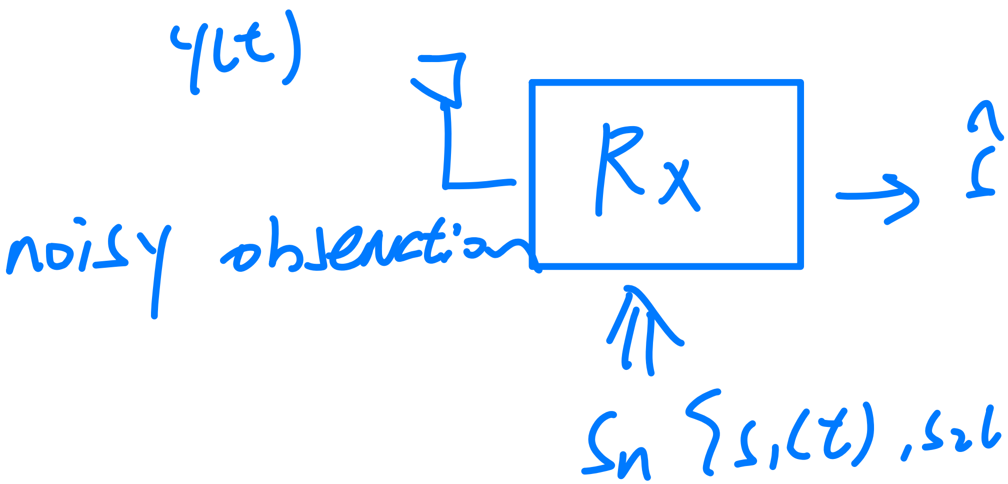
$$y(t) = \overbrace{x(t)}^{\substack{\text{transmitted}\\ \text{signal}}} + \overbrace{n(t)}^{\substack{\text{random}\\ \text{noise}}}$$



- Q) What is the optimal way of detection based on this observed waveform?

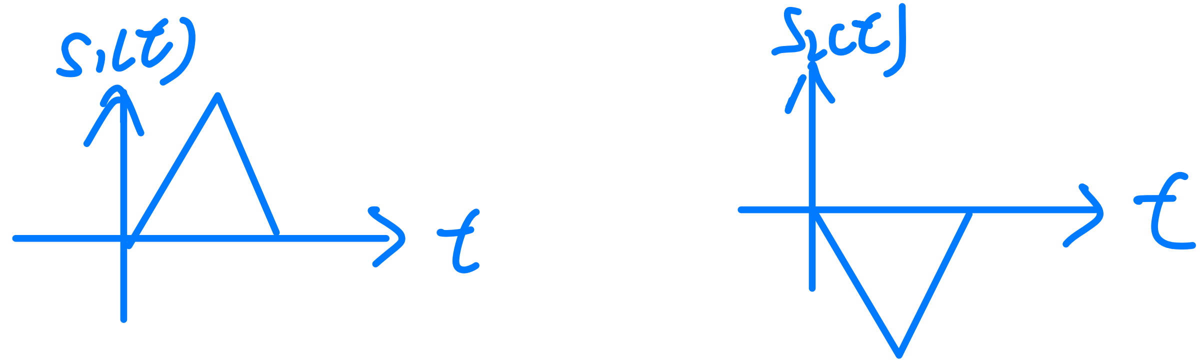
Why this filter is optimal in universal?



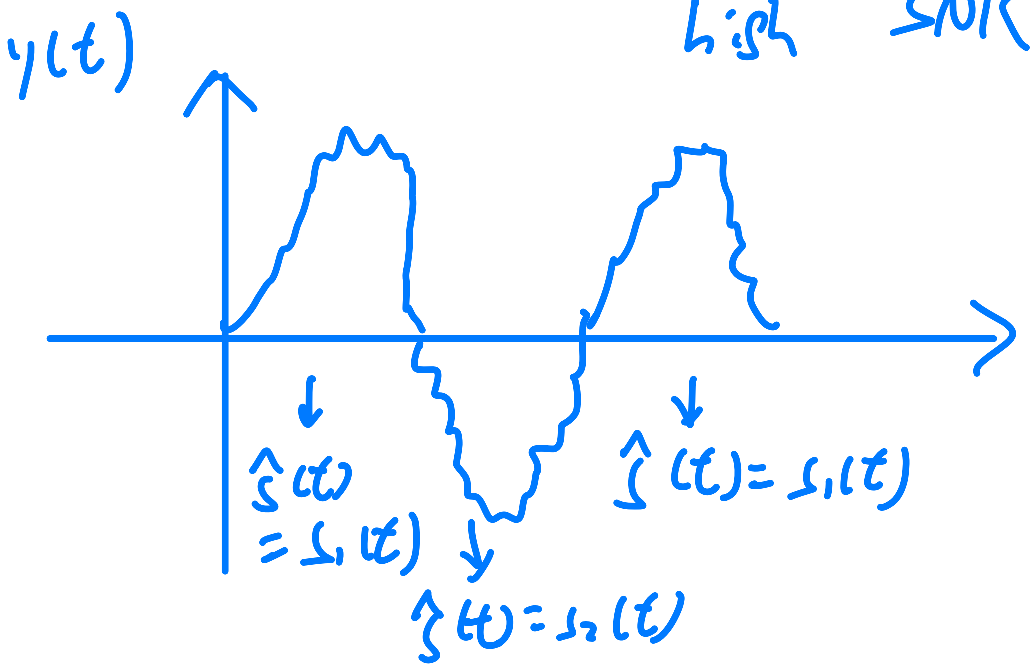


guessing game / hypothesis testing

$S^A = \{$

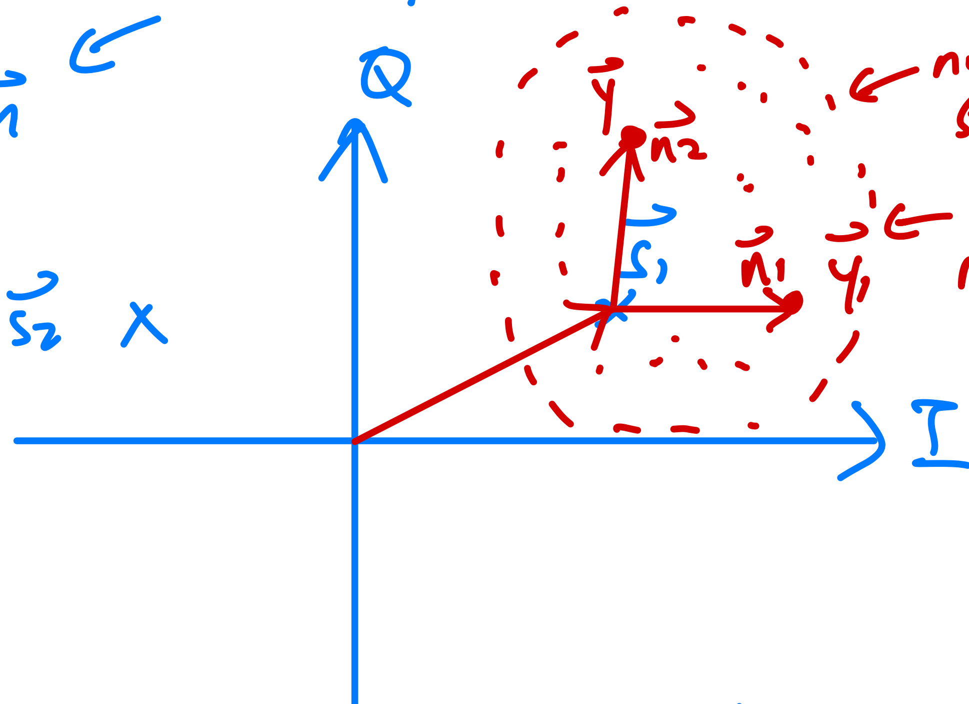


high SNR



isotropic, length = random!

$\vec{y} = \vec{s} + \vec{n}$



noise cloud, spherical!

$\vec{n}$ specify, $\{\vec{s}_j\}$ 有 99% received within this boundary

$\|\vec{n}\|^2 \in (0, \infty)$

isotropic: direction to all, with same probability

Received Waveform - Geometric Domain View

– Expressing in vector form, we have: $\vec{y} = \vec{x} + \vec{n}$

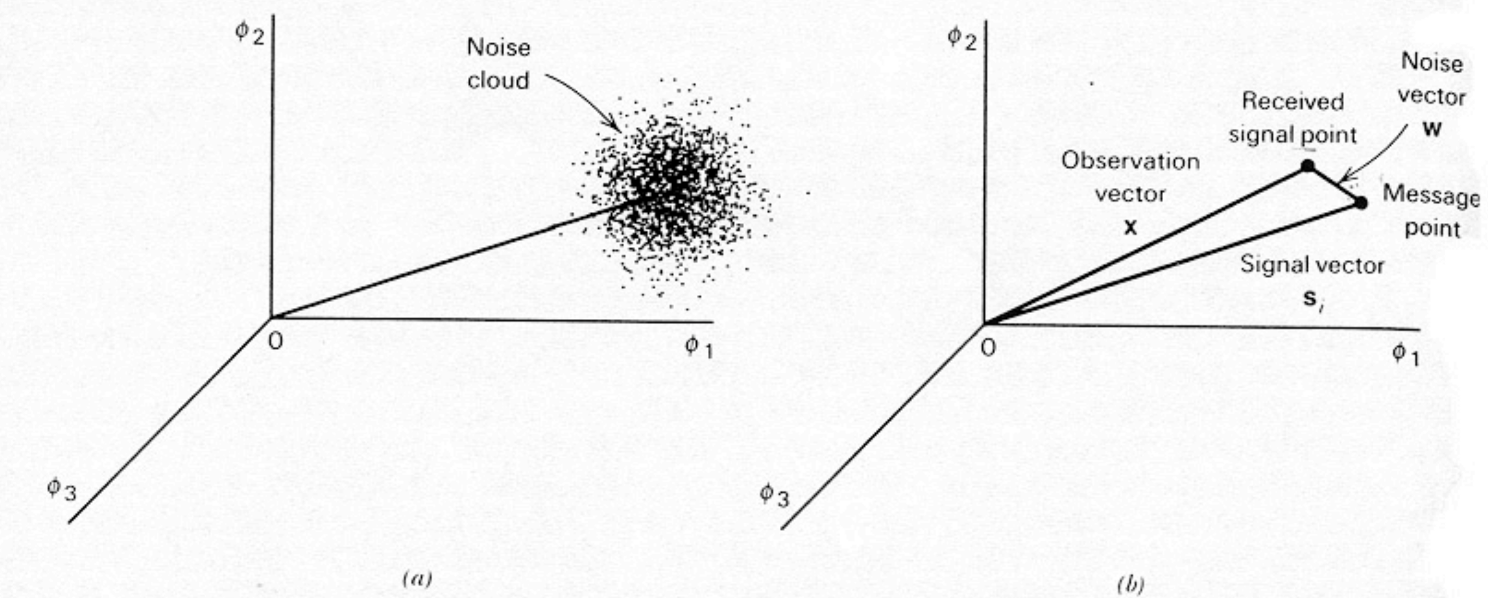


Figure 8.7 Illustrating the effect of (*a*) noise perturbation on (*b*) the location of the received signal point.

Property of the noise vector

- Suppose the M possible transmitted signals (points) are contained in a N-dim signal space (with orthonormal basis function $\{\vec{s}_1, \vec{s}_2, ..., \vec{s}_M\}$). $\{\phi_1(t), ..., \phi_N(t)\}$
  - The transmitted signal x(t) is completely contained in the signal space.

$$\boxed{\vec{x} = (x_1, .., x_N) \text{ where } x_n =< x(t), \phi_n(t) >= \int_0^{T_s} x(t)\phi_n^*(t)dt}$$

  - The channel white noise signal n(t) ***may not*** lie completely on the signal space. Hence, the received signal y(t) may not be precisely represented by the signal space.
    - » However, since the components of y(t) outside the signal space DOES NOT carry any useful observation about the transmitted information, it is irrelevant to the detection process.
    - » Without loss of optimality, we could consider the components of y(t) that lie within the signal space. (i.e. we could do a projection of y(t) onto the N-dim signal space to get all the useful components).

$$\boxed{\vec{y} = (y_1, .., y_N) \text{ where } y_n =< y(t), \phi_n(t) >= \int_0^{T_s} y(t)\phi_n^*(t)dt}$$

Property of the noise vector

projection n to the basis ϕ_n !

weighted sum of jointly gaussian

- Writing y = x + n, the noise vector is given by:

$$\boxed{\vec{n} = (n_1, .., n_N) \text{ where } n_n = < n(t), \phi_n(t) > = \int_0^{T_s} n(t)\phi_n^*(t)dt}$$

- Property I: The noise vector is a Gaussian random vector.
  - This follows directly from $n_n = \int_0^{T_s} n(t)\phi_n^*(t)dt$ because n(t) is a white Gaussian process.

$\vec{n} \sim$ iid. jointly gaussian $E(n_n)$ $E(n_n^2)$

- Property II: Mean of the noise vector is 0.

$$\mathbb{E}[n_n] = \mathbb{E}\left[\int_0^{T_s} n(t)\phi_n^*(t)dt\right] = \int_0^{T_s} \mathbb{E}[n(t)]\phi_n^*(t)dt = 0$$

- Property III: The noise vector is iid with variance $\eta_0/2$

$$\mathbb{E}[n_i n_j^*] = \mathbb{E}\left[\int_0^{T_s}\int_0^{T_s} n(t)n(t')\phi_i^*(t)\phi_j(t')dtdt'\right] = \int_0^{T_s}\int_0^{T_s} \mathbb{E}[n(t)n(t')]\phi_i^*(t)\phi_j(t')dtdt'$$

sifting-prop $t=t'$

$$= \frac{\eta_0}{2}\int_0^{T_s}\int_0^{T_s} \delta(t-t')\phi_i^*(t)\phi_j(t')dtdt' = \frac{\eta_0}{2}\int_0^{T_s} \phi_i^*(t)\phi_j(t)dt = \begin{cases} \eta_0/2 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

$\sum x_i \sum y_i = \sum_i \sum_j x_i y_j$

$S_n(f) = \frac{N_0}{2} \rightsquigarrow R_n(\tau) \triangleq E[n(t)n(t+\tau)] = \frac{N_0}{2}\delta(\tau)$

$$f(\vec{n}) = \text{const} \exp -\left(\frac{-\vec{n}^{T}\vec{n}}{N_0}\right)$$

$$= \text{const} \exp -\left(\frac{-\|\vec{n}\|^2}{N_0}\right)$$

Task of the Demodulator

- **Detection:**
  - Given the received observation (point)$(\vec{y})$, decode or find out what has been transmitted (transmitted point $(\vec{s})$from one of the M possible points

    $\{\vec{s}_1, \vec{s}_2, \ldots, \vec{s}_M\}$

- **Optimal Detection:**
  - Given the observation $(\vec{y})$, what is the ***best guess*** on the information bit carried.
  - Best = minimize BER.

can get universal optimal receiver!

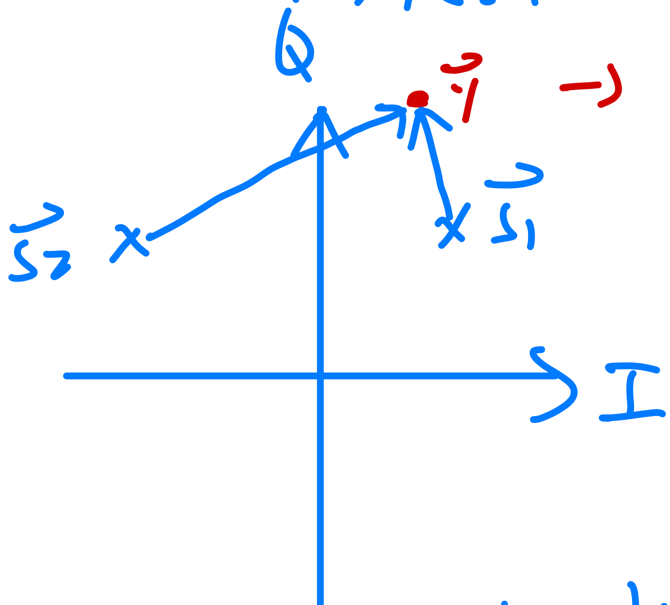
- Q1) Optimal Demodulator:
  - Minimum Distance Detection is optimal when the channel is AWGN
  - What is the associated structure?
- Q2) Error Performance using “Minimum Distance Detection”:
  - Will we commit error when using the best demodulator?
  - What is the mechanism for error?

$\rightarrow \hat{s} \in \{\vec{s}_1, \vec{s}_2\}$

noisy observation

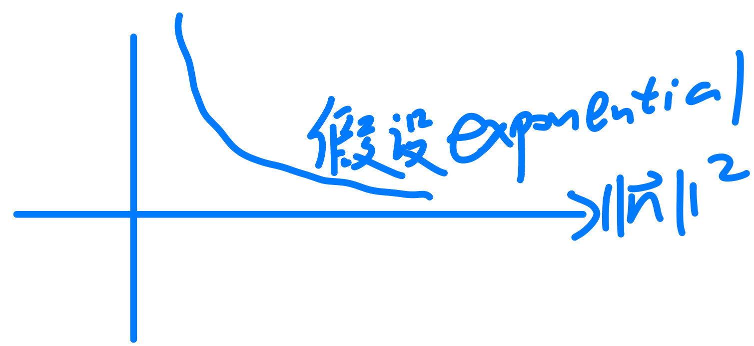
$\vec{y} = \vec{s} + \vec{n}$

$\vec{s} \in S = \{\vec{s}_1, \vec{s}_2\}$



$\rightarrow$ Thinking process:
pick one point that
is closer! (universal optimal)

noise distribution

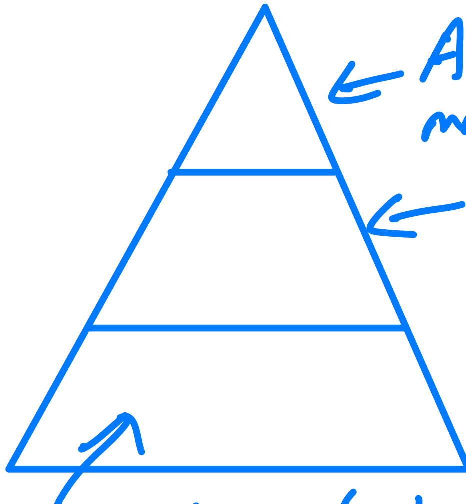


- Minimum distance demodulation

$$\hat{s} = \underset{\vec{s} \in \{\vec{s}_1, \cdots \vec{s}_M\}}{\arg\min} \|\vec{y} - \vec{s}\|^2$$

Not end yet!

R&D

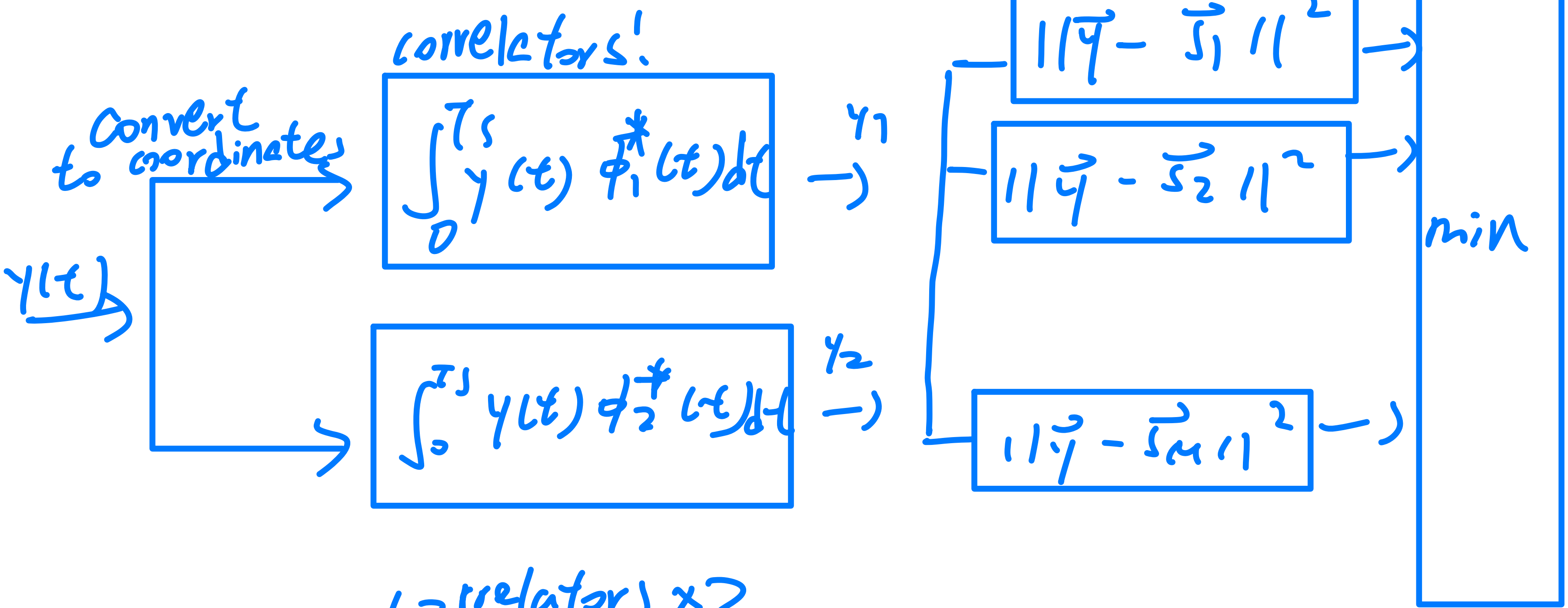


Architect, must be experienced
map algorithm and implementation

System engineer (PhD)
~Do algorithm

Developer (HW/SW) — BEng
M.Phil, M.S.C

$\vec{s}_1, \cdots \vec{s}_M \in S = \text{Span}\{\phi_1(t), \phi_2(t)\}$



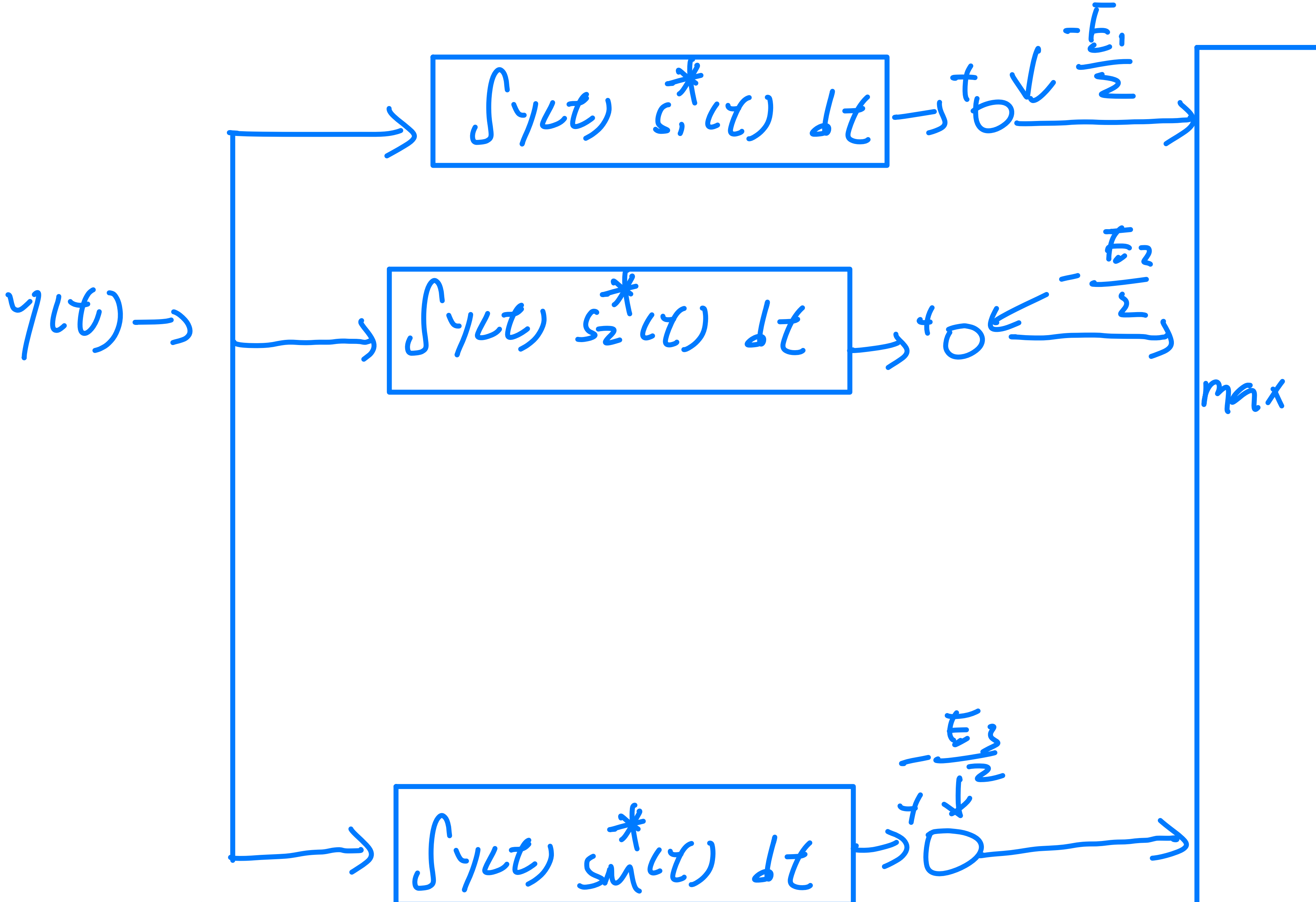
correlators x2 $\rightarrow$ # of dimensions

(I) (II) (III)

$$\hat{s} = \underset{\vec{s} \in \{\vec{s}_1, \cdots \vec{s}_M\}}{\arg\min} \|\vec{y} - \vec{s}\|^2 = \underset{s}{\arg\min}\left[-2\langle \vec{y}, \vec{s}\rangle + \|s\|^2\right]$$

$$= \underset{s}{\arg\max}\left[\langle \vec{y}, \vec{s}\rangle - \tfrac{1}{2}\|s\|^2\right]$$

can ignore this in MPSK!!!



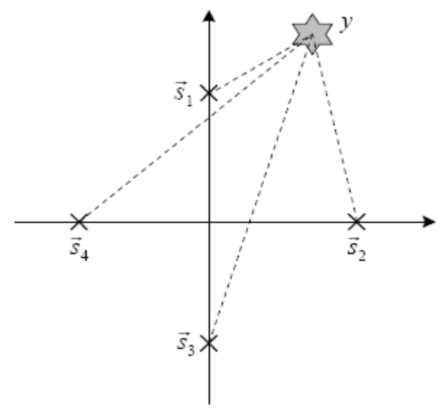
- Same performance as before, but M correlators! not good!

Optimal detector → without any pre-defined structure!

Minimum Distance Detection

- **Optimality**
  - It can be shown (later) that the optimal detector is very simple:

$$\vec{s}^* = \arg \min_{\vec{x} \in \{\vec{s}_1, \ldots, \vec{s}_M\}} \|\vec{y} - \vec{x}\|^2$$



- **Architecture:**

Step 1: Convert the time domain received signal $y(t)$ to the received vector $\mathbf{y}_n$. As mentioned in the previous section, suppose the constellation signal space $\mathcal{X}$ (with D dimension) has basis $\{\phi_1(t), \ldots, \phi_D(t)\}$. The j-th coordinate of the received vector $\mathbf{y}_n$ is given by:

$$\mathbf{y}_n(j) = < y(t), \phi_j(t) > = \int_{(n-1)T_s}^{T_s} y(t)\phi_j^*(t)dt \qquad (2.14)$$

for $j = \{1, 2, \ldots, D\}$.

Step 2: Determine the M distances. Given the received vector $\mathbf{y}_n$ as determined in step 1, we measure M distances between the received vector and the M possible *hypothesis points* $\{\mathbf{s}_1, \ldots, \mathbf{s}_M\}$. That is,

$$\begin{aligned} d(\mathbf{y}_n, \mathbf{s}_m)^2 &= \|\mathbf{y}_n - \mathbf{s}_m\|^2 = < y(t) - s_m(t), y(t) - s_m(t) > \\ &= \int_{(n-1)T_s}^{nT_s} |y(t) - s_m(t)|^2 dt \qquad (2.15) \end{aligned}$$

Step 3: Determine the minimum distance. The detected symbol $\hat{\mathbf{x}}_n$ is given by:

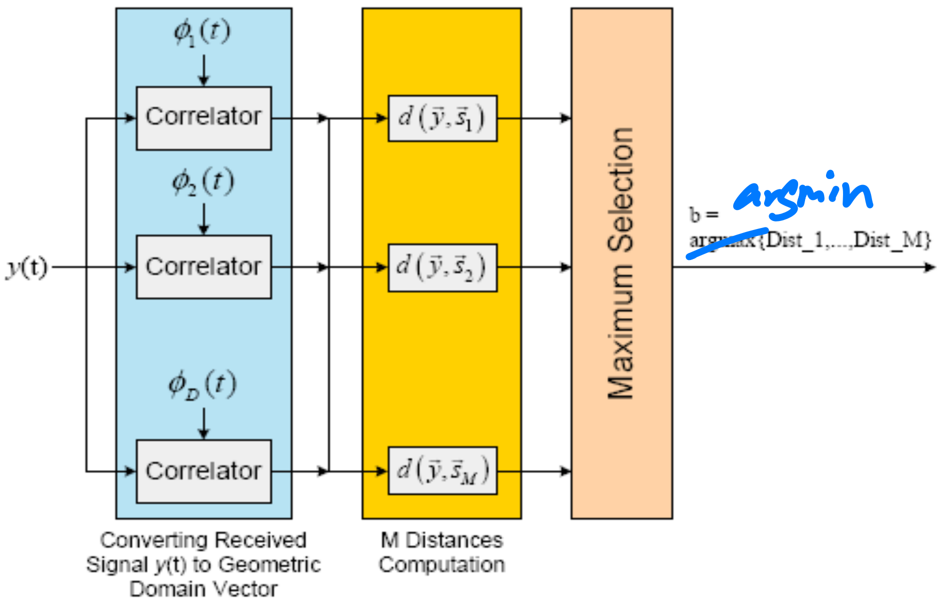
$$\hat{\mathbf{x}}_n = \mathbf{s}_{m^*} \qquad (2.16)$$

where

$$m^* = \arg\min_m d(\mathbf{y}_n, \mathbf{s}_m)^2 \qquad (2.17)$$

what is optimal structure?

Structure of the Demodulator



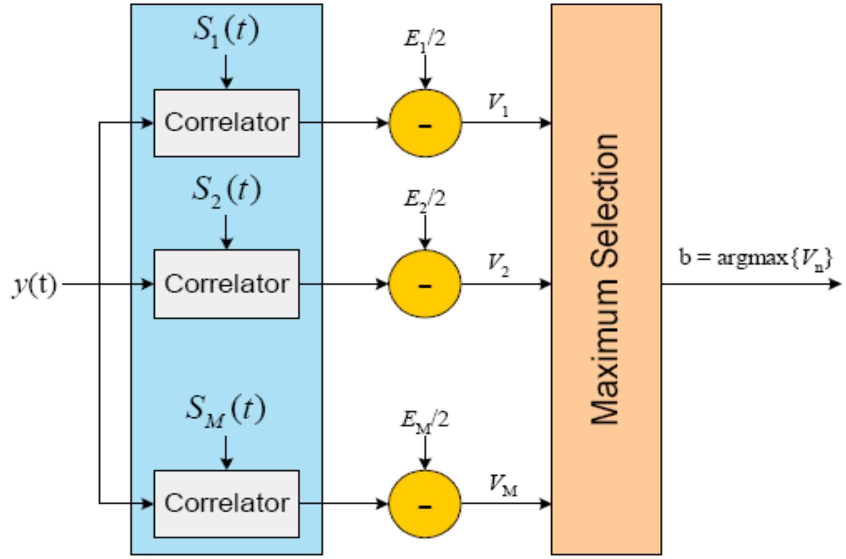


Figure 2.15 A simplified structure of the minimum distance detector.

$$d(y(t), s_m(t))^2 = \|\mathbf{y}\|^2 + \|\mathbf{s}_m\|^2 - 2 < \mathbf{y}, \mathbf{s}_m >$$
$$= \|\mathbf{y}\|^2 + E_m - 2\int_{(n-1)T_s}^{nT_s} y(t) s_m^*(t) dt$$

minimum!

$$m^* = \arg\max_m \int_{(n-1)T_s}^{nT_s} y(t) s_m^*(t) dt - E_m/2$$

Example

In an additive white Gaussian noise channel with a noise power-spectral density of $N_0/2$, two equiprobable messages are transmitted by

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

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

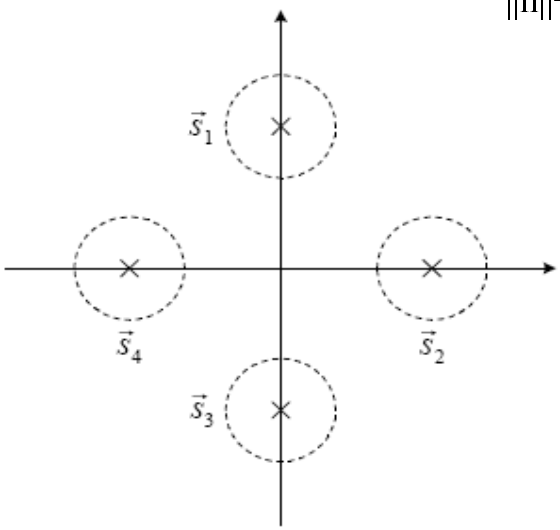
1. Determine the structure of the optimal receiver.

what is optimal demodulation

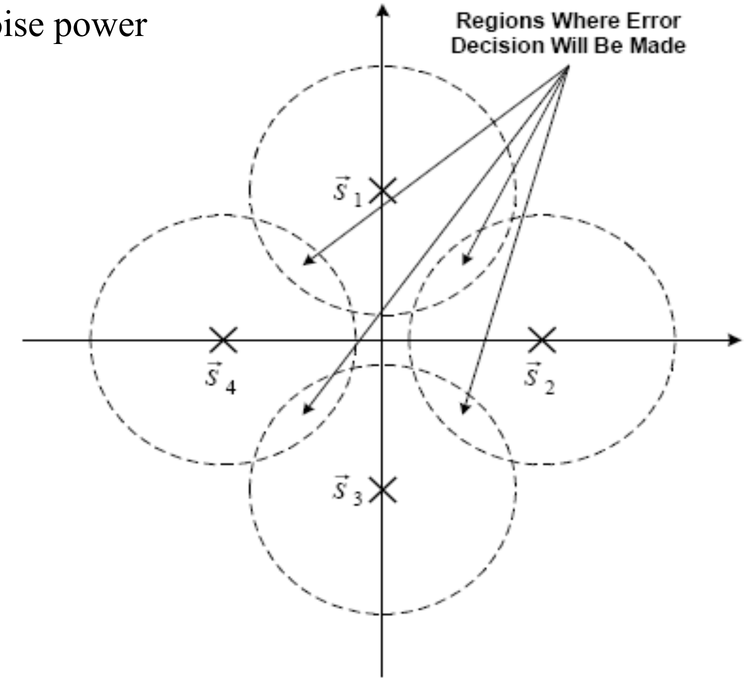
what is Pe?

Error Analysis - Qualitative

$\|n\|^2$ determines the noise power

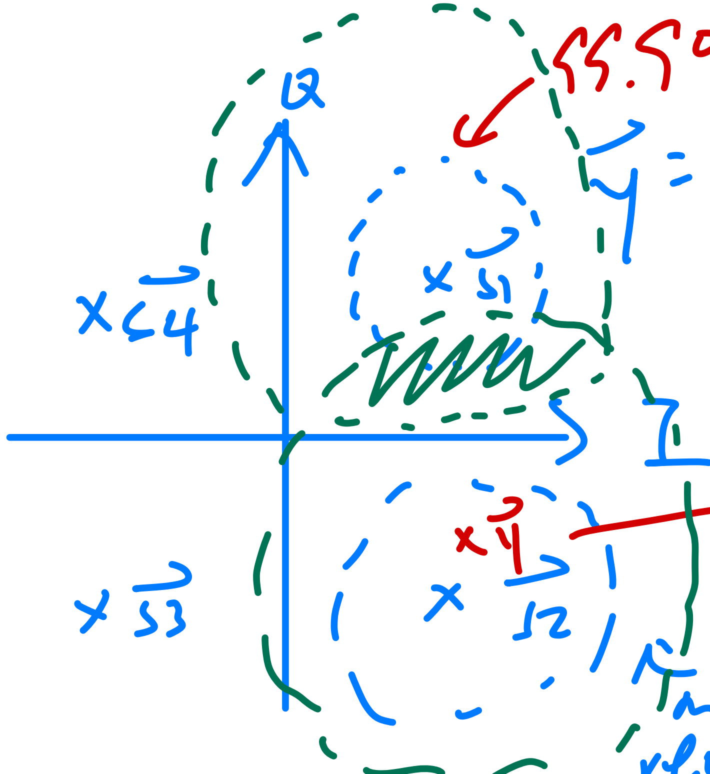


(a) small noise



(b) large noise

$$P_e \approx Q\left(\sqrt{\frac{d_{min}^2}{\eta_0}}\right) \qquad d_{min} = \min_{i,j} \|\vec{s}_i - \vec{s}_j\|$$



95.5% noise cloud!

spherical

isotropic

$\vec{y} = \vec{s} + \vec{n}$

"random" vector

if cloud is big enough

$\Downarrow$

error $\downarrow$

most of time receive this!

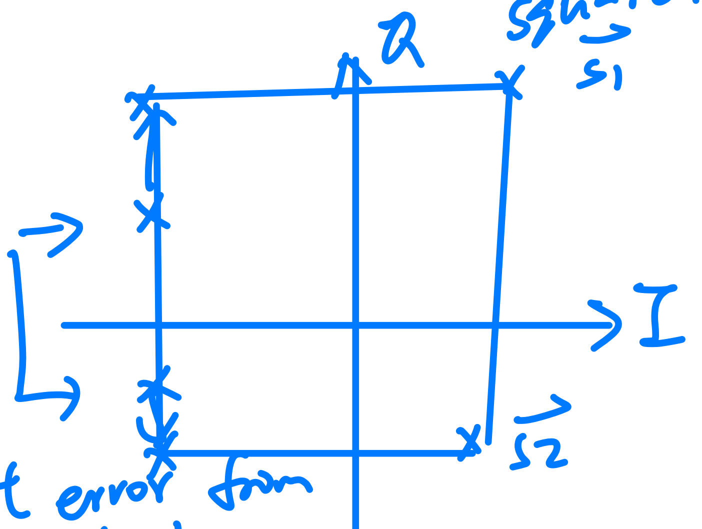
Always use minimum distance demodulation!

- error probability is related to overlapping (ambiguous) area in the noise cloud!

- $\sqrt{2}A_c$ more further to the origin!

↓ overlapping area, ↓ error probability!

⇒ ↑ energy!



most error from those two!

→ bottleneck!

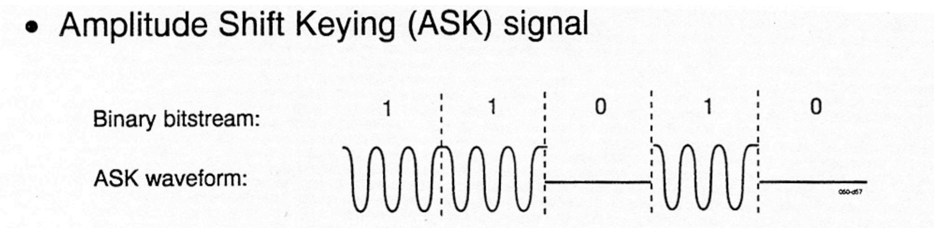
$P_e \sim$ overlapping area of noise clouds $\sim$ worst case separation between

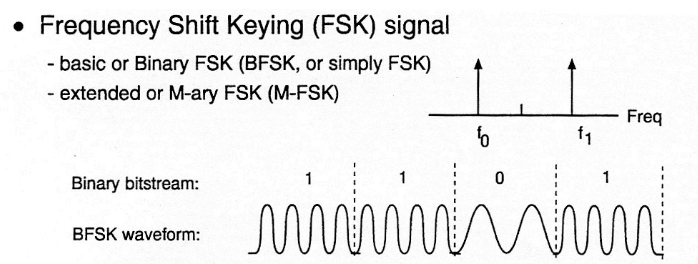
$$d_{min} = \min_{i \neq j} \|\vec{s}_i - \vec{s}_j\|^2$$

- Also spend some energy to improve bottleneck!

Examples - Qualitative Model

- **Useful Insights**
  - Using this qualitative model, we could have very useful design insights
  - We could use this to compare which modulator design is better.
  - We could use this to explain the tradeoff mechanism between "performance" and "resource".
- **Example 1**
  - Which one (ASK or FSK) is a better modulation scheme?



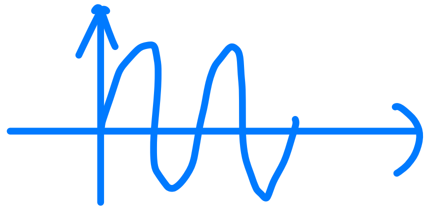
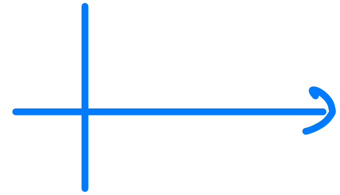


BASK

$s_1(t) = A\cos\omega_0 t$

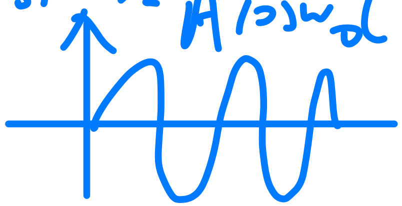
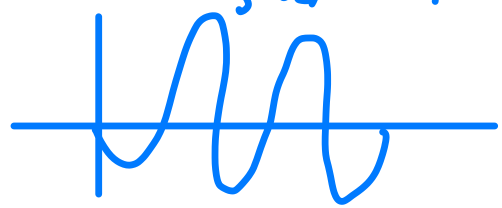
$s_0(t) = 0 \rightarrow \text{Energy} = 0$

$\Rightarrow$ origin

$S = \Big\{$   $\Big\}$

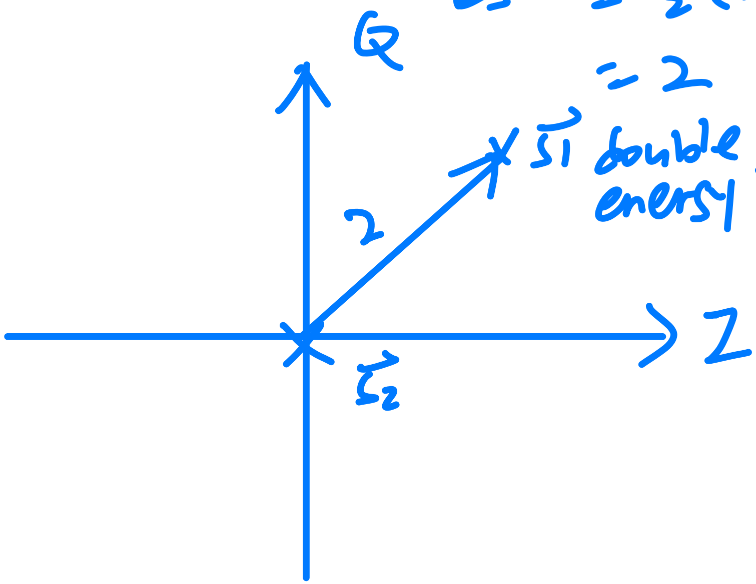
$s_1(t) = A\cos\omega_0 t$

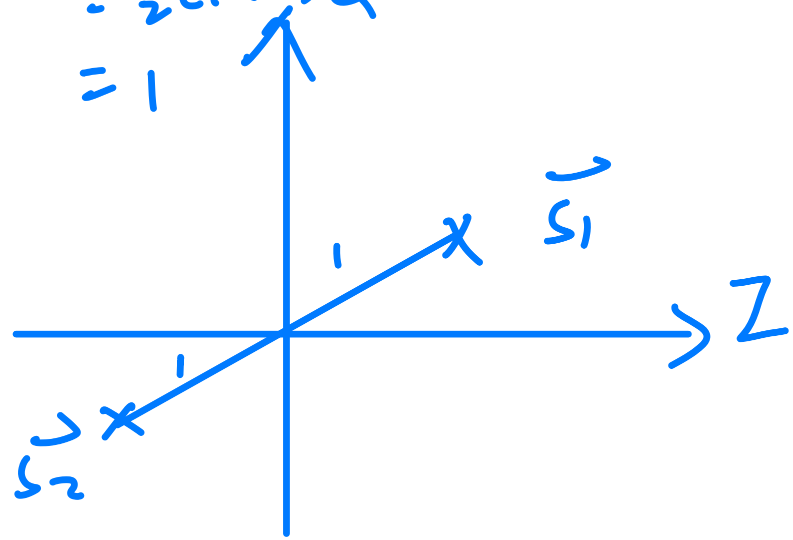
$s_0(t) = A\cos(\omega_0 t + \pi)$

$S^{BPSK} = \Big\{$  

$E_s^{BASK} = \frac{1}{2}(2^2+0^2) = 2$

$E_s^{BPSK} = \frac{1}{2}(1^2+1^2) = 1$





turn off sig,
but other sig
much more energy!

same error probability!

still used by remote control!

↓

easy implementation

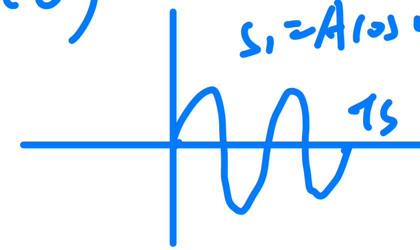
low cost! just diode is okay!

Just measure the energy!

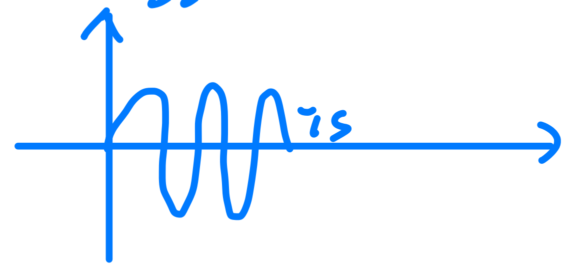
This needs coherent receiver! much harder implementation!

- need lots of circuits to maintain phase synchronization

$S^{BPSK} = \begin{cases} s_1(t) \end{cases}$



,

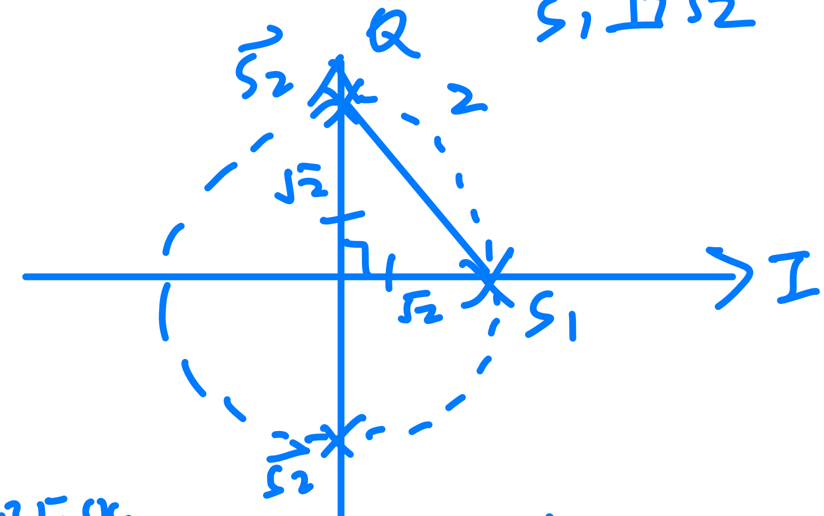


$w_1 = \frac{n_1}{T_s}$ $\qquad$ $w_2 = \frac{n_2}{T_s}$

$\Delta w = \frac{1}{T_s}$

$\Rightarrow$ continuous phrase constraint!

$s_1 \perp s_2$



$E_1 = E_2$

$||s_1||^2 = ||s_2||^2$

how much feature extracted from time domain to determine in geometric domain?

$\vec{s_1}, \vec{s_2}$ are orthogonal!

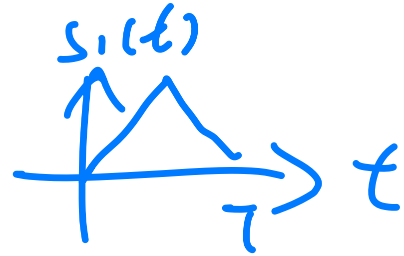
$E_1^{BFSK} = \frac{1}{2}(\sqrt{2}^2 + \sqrt{2}^2)$
$= 2$

Equally bad as BASK!

Time domain will hide some details

So if M is large for M-ary,
constellation diagram is usually preferred!
You can directly visualize the performance

-Much easy to derive intuition!

$S = \{$  ,  $\}$

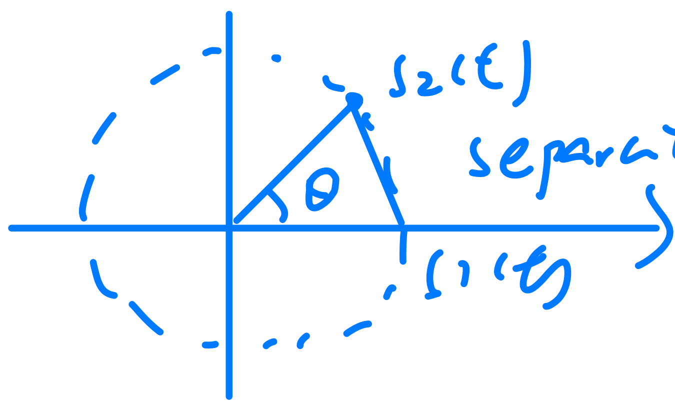
$\bar{E}_1 = \bar{E}_2 \Rightarrow \|\vec{s}_1\| = \|\vec{s}_2\|$

$R_b = \log_2 M \cdot \frac{1}{T_s}$ b/s

Not a good design!

$\uparrow R_b$, can increase bit rate, $\uparrow$ passengers, keep the same bandwidth!

$\theta < 90^\circ$

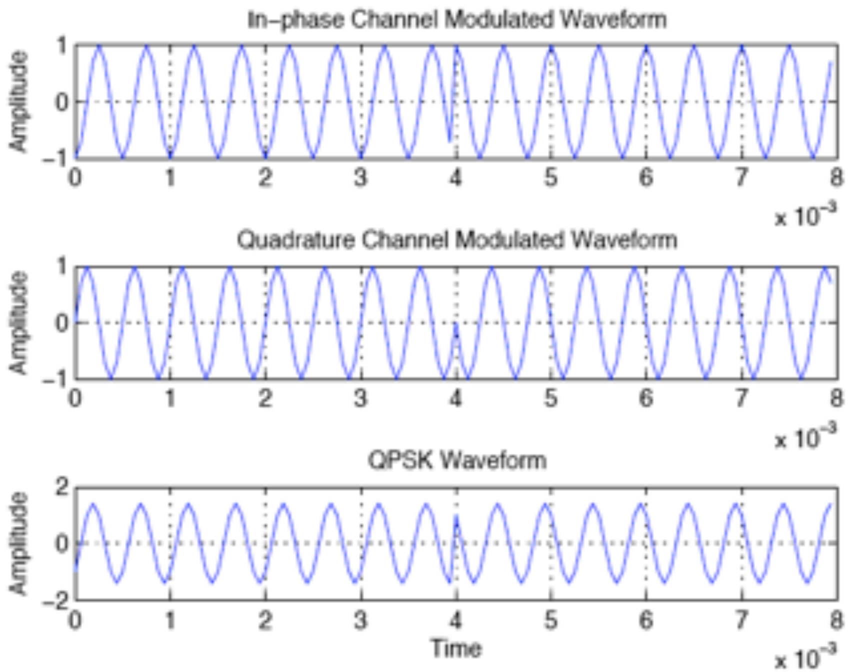


separation not high enough

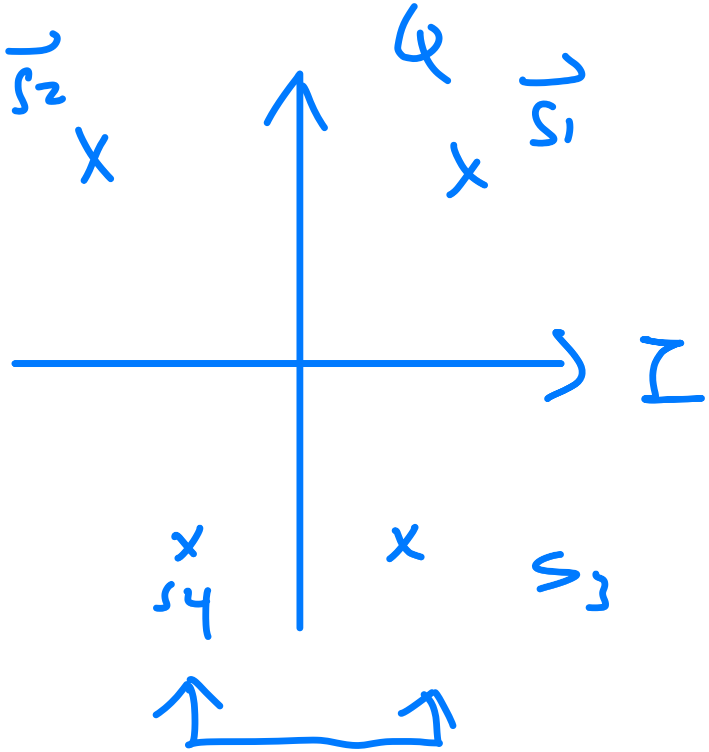
$\Rightarrow$ BPSK is the best design!

Application of Geometric Model

- **Example 2: Is there any better Quartenary modulation than QPSK (on a 2 dimensional signal space)?**



This is the best!

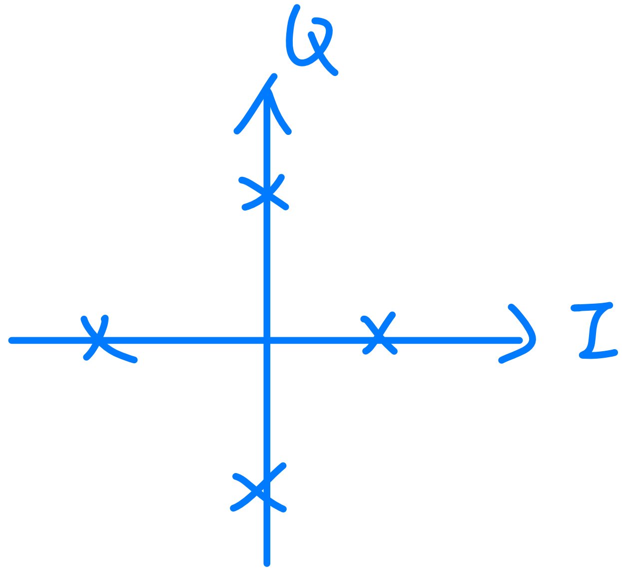


$\vec{E}_s = \text{constant}$

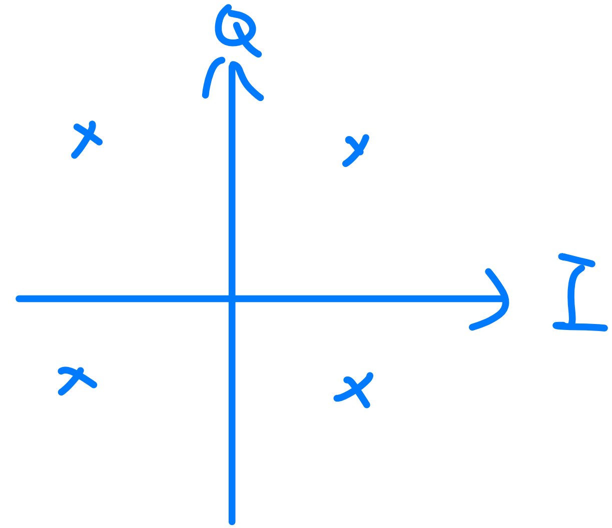
$\frac{1}{4}\left(\|\vec{E}_1\|^2 + \|\vec{E}_2\|^2 + \|\vec{E}_3\|^2 + \|\vec{E}_4\|^2\right)$

can reduce these?
$\downarrow$ energy!

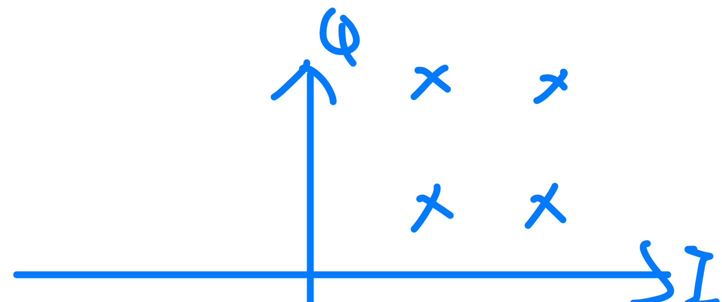
- But can always improve this diagram!



Rotation



Translation



Same energy

Same performance

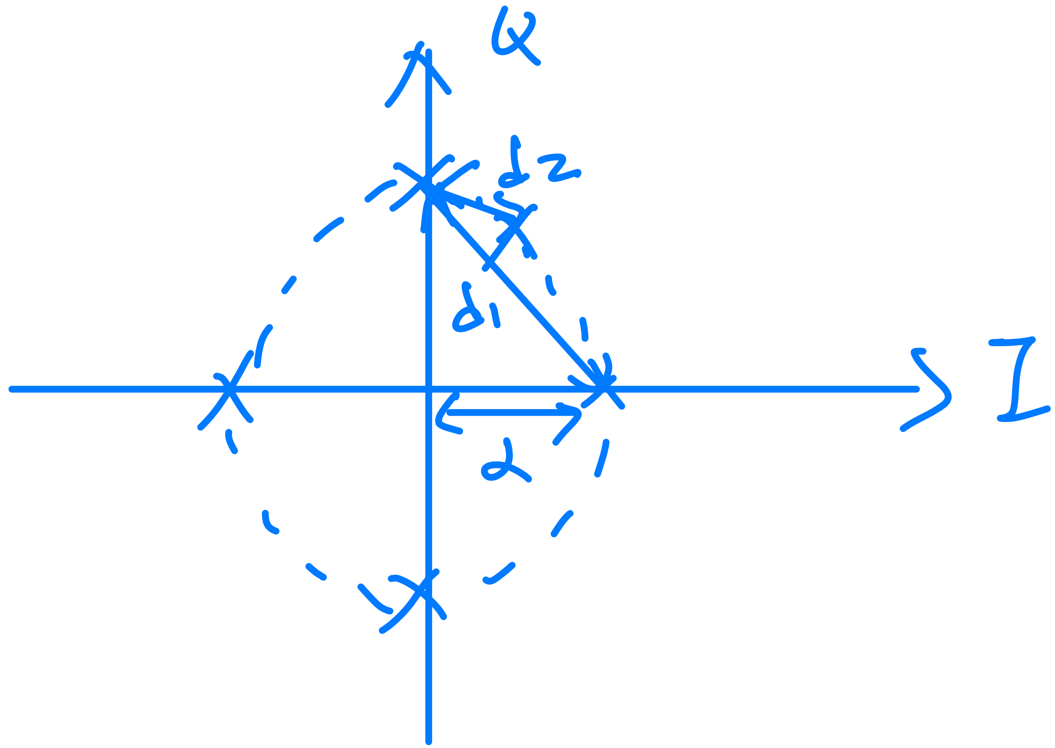
same performance
different
energy!

$R_b = (\log_2 M)\frac{1}{T_s}$

↑ M will be better?

without modifying the bandwidth

- No free lunch :(



Obviously

$d_2 < d_1$

P_e ↑↑

want $d_2 = d_1$, need scale up α,

Transmition power will be higher!!!

High BW efficiency

Bandwidth efficiency:

$\eta = \frac{R_b}{W}$ ← Bit rate
← Bandwidth

$R_b = (\log_2 M) \frac{1}{T_s}$

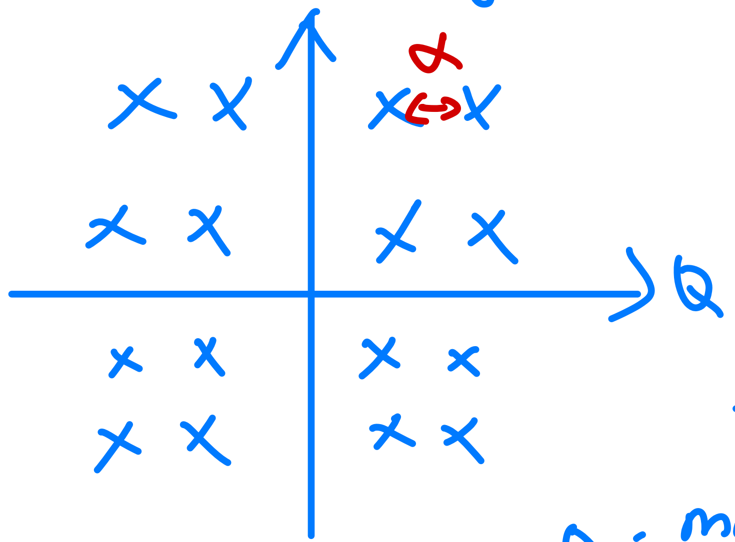
不改 bandwidth, ↑R_b

$M \uparrow$, $\eta \uparrow$

Same bandwidth, can transmit more passengers!

16-QAM

↑ larger d_{min}!



$$P_e^{16PSK} > P_e^{16QAM}$$

$$E_s^{16PSK} = E_s^{16QAM}$$

2-freedom

$$D_1 = \max_{x \in D_2} f(x)$$

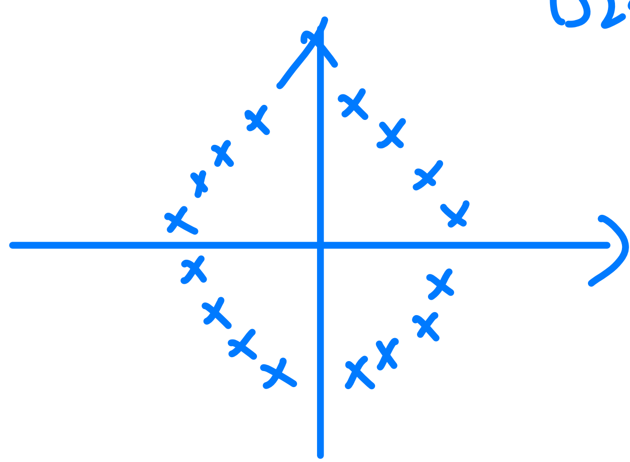
$$D_2 = \max_{x \in D_1} f(x)$$

$$P_2 \geq P_1$$

if optimization domain ↑, then the result will be no worse!

16-PSK

worse!



1-freedom!

r

Applications of the Geometric Model

- **Example 3: Why 16QAM is much more popular than 16PSK?**

smaller P_e, while energy are the same!

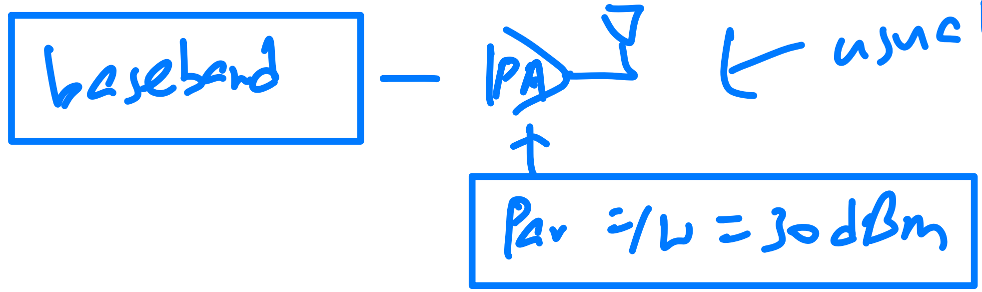
degree of freedom ↑

- **Example 4: Which of the following Binary Modulator is a better design?**

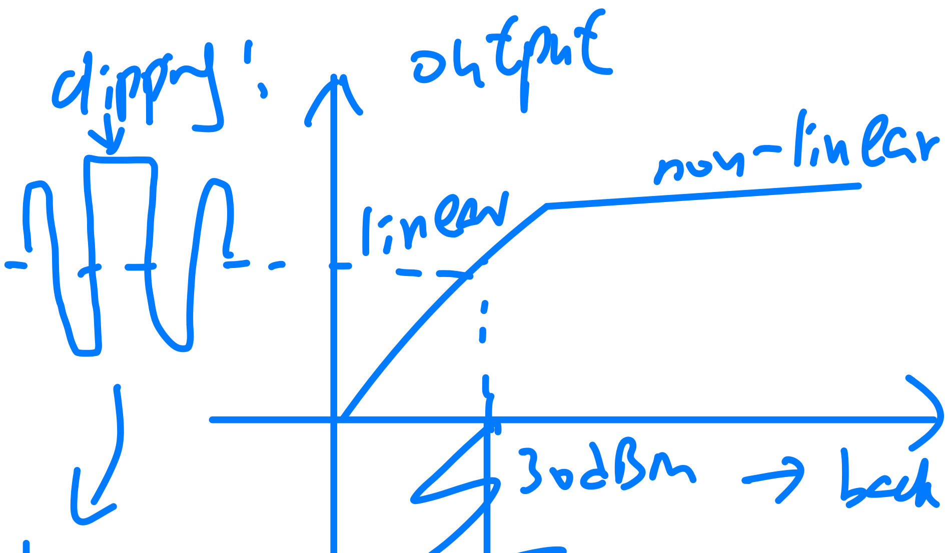
$$\begin{cases} \text{MQAM} \\ \text{MPSK} \end{cases} \rightarrow R_b = \log_2 M \frac{1}{T_s}$$

SNR ↑↑

$$\frac{\text{high BW efficiency}}{\text{Mod}}$$



← usually in wifi!



30dBm → back to 20/10 dBm

↓

some peak (power backoff)

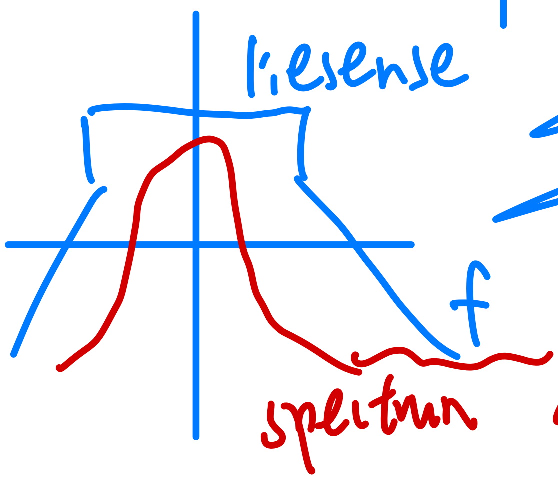
↓

depends on PAR

peak to average ratio!

= 10dB

⇒ backoff at least 10 dB!



clipping → spectral regrowth!

16 QAM , 16 PSK

16 -points on the circle

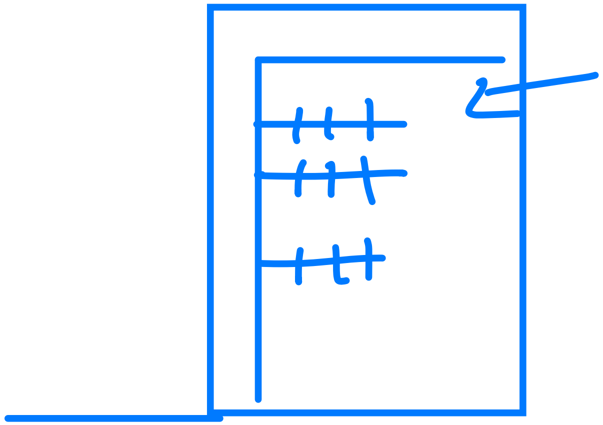
phase term

$PAR^{16PSK} \ll PAR^{16QAM}$

But needs smaller power backup!

2A $PAR^{16PAM} = 3\,dB$, $PAR^{16QAM} = 10\,dB$

cable TV system



coaxial cable

4096-QAM

noise-limited system

which power amplifier needed ?

4096 QAM

very high SNR !

cellular system - not noise limited
(interference limited

bottleneck not SNR,
is SINR

if power ↑, SINR will
be saturated (20dB)

SINR follows a distribution

① high efficiancy mod.

Chapter 6 !

MQAM !
MPSK

② high energy efficiency

MFSK, dimension always 6 ! ! !

total bandwidth

$\Delta\omega = \frac{2\pi}{T_s} \Longrightarrow$ total bandwidth! $BW = \frac{1}{T_s} \cdot M$

mod. $S = \{ s_1(t) \;\cdots\; s_M(t) \}$

$\langle s_i\, s_j \rangle = 0$

mutually orthonormal!

- linearly independent

- need M dimensional space!

- More space to put the points!

compare $\frac{E_b}{N_0}$

MFSK

$$\eta = \frac{R_b}{W} = \frac{O(\log_2 M)}{O(M)}$$

low spectral efficiency

16QAM > 16PSK (3 numbers keep the same, other better!)

$\eta = \frac{R_b}{W} = \log_2 M$ (constant!)

Tradeoffs

Increasing bandwidth can increase bit rate. Recall that bandwidth is directly proportional to the baud rate or symbol rate of a digital communication link. Hence, increasing the bandwidth will directly increase the baud rate. Hence, the bit rate, which is given by $\log_2 M \times$ baud rate, will be increased.

Increasing transmit power can indirectly increase bit rate. On the other hand, one method to increase the bit rate without increasing the bandwidth requirement is to increase the modulation level M. For example, the bit rate can be doubled by using 16QAM constellation instead of QPSK constellation. However, doing so will increase the density of the constellation points unless the transmit power is increased. In that case, more constellation points can be squeezed into the signal space (higher bit rate) and at the same time, maintaining a similar *minimum distance separation between constellation points* (similar error probability) at the expense of higher transmit power.

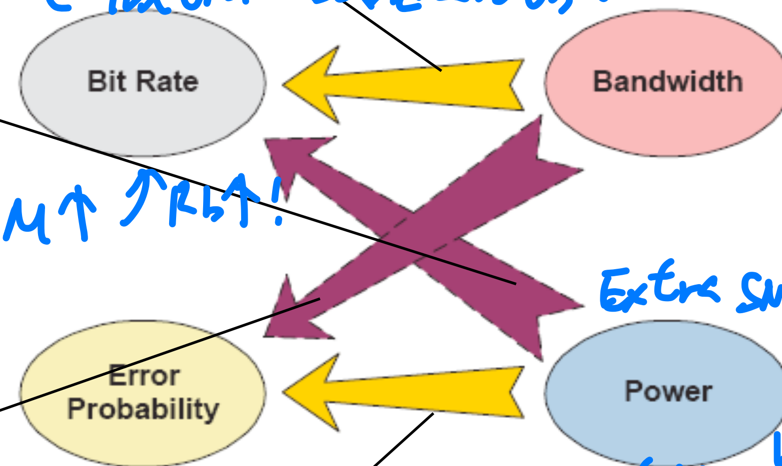
$P = \alpha\left(\frac{E_s}{N_0}\right)$?

Increasing BW can indirectly reduce error probability: Take MFSK as an example, increasing M increases BW as well as bit rate. Since Es/N0 = log2(M) Eb/No, Es/N0 (or transmit power) can be increased by log2(M) for a given Eb/No (bit-energy to noise density ratio). On the other hand, the dimension of the space is M and hence, the overall effect is the points are more spacious (distance separation between any two points are increased at the same Eb/No). This results in lower error probability (at the expense of higher BW).

Increasing transmit power can reduce error probability. Recall that P_e is related to the minimum distance between any two constellation points, d_{min}. Increasing the transmit power is equivalent to pulling all the constellation points further away from the origin. In this way, the minimum separation between any two constellation points will be increased and therefore, the error probability will be decreased without decreasing the bit rate.

:) reduce error probability with same energy (→ save energy!)

:(Extra bandwidth!



$M\uparrow \; \uparrow R_b \uparrow$!

Extra SNR

maintain same Peak

R_b for MFSK : $(\log_2 M) \times \frac{1}{T_s}$

$\eta = \frac{R_b}{W} = O(\log_2 M / M)$

if $M\uparrow \Rightarrow \eta\downarrow\downarrow \to 0$

- Both modulation scheme have tradeoffs !

- No universal optimal design !

Modulation for High Spectral Efficiency

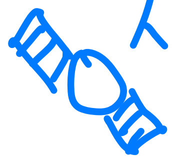
- **M-ary PSK and M-ary QAM**
  - **As we increase M, the dimension of the signal space is always 2**
  - **Transmission bandwidth remains constant for all M**
  - **Bit rate = log2(M) x 1/Ts (Increasing M could increase bit rate without consuming higher BW)** ⇒ need higher energy!
  - **However, the constellation points are more condensed as we increase M.**
    - » **To maintain the same error probability, we need to increase transmission power (to move the points further away from the origin)**
  - **Hence, this family can achieve high spectral efficiency (Bit rate / BW) at the expense of high SNR.**

Modulation for High Power Efficiency

- **M-FSK (Orthogonal Modulation)**
  - This family of modulation is orthogonal (signals within the signal set are mutually orthogonal).
  - As a result, for M-FSK, we need M-dimensional signal space in order to contain all the M signals. These extra dimensions come from the extra BW (Tx-BW = M x (1/Ts))
  - As we squeeze in more points, the signal space gets "bigger" as well.
  - For the same Eb/No, the distance between constellation points are increased!! -> the error probability will drop.
  - Hence, MFSK is good for saving power (at the expense of BW expansion).

Resemble FM

Bandwidth expansion ⇒ Reduction of Pe



人造衛星

MFSK

$\boxed{\text{LOS}}$

$f_c \sim 60$GHz

- Plenty of BW

for low f_c, much expensive,

high f_c, more favourable physical situations!

Resource

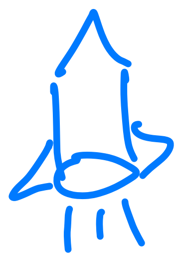
BW $

save $$$

Power $$$

if MFSK

wc



need heavy solar pannel

↳ want to have lower power!

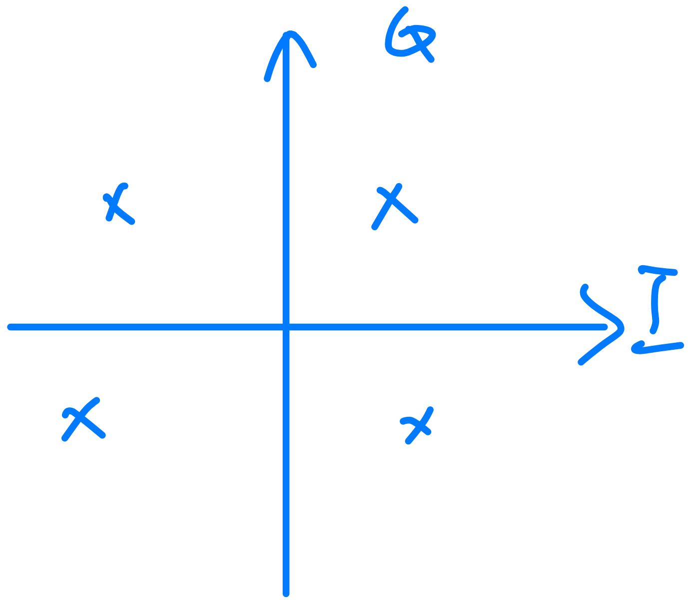
↓↓ solar pannel used

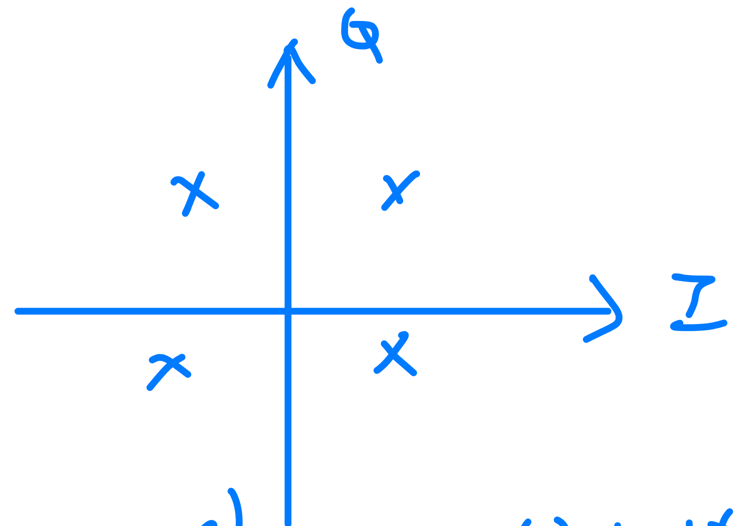
if lower f_c

BW \$\$ $\searrow$ \$\$\$

Power $\nearrow$ \$

Bandwidth is much more expensive!!!





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

2D　　not 2D

(1-1 correspondence)

Projection

$\vec{y} \approx \vec{x}$

但不影响 detection!

- Do projection for $y(t)$

Case Study

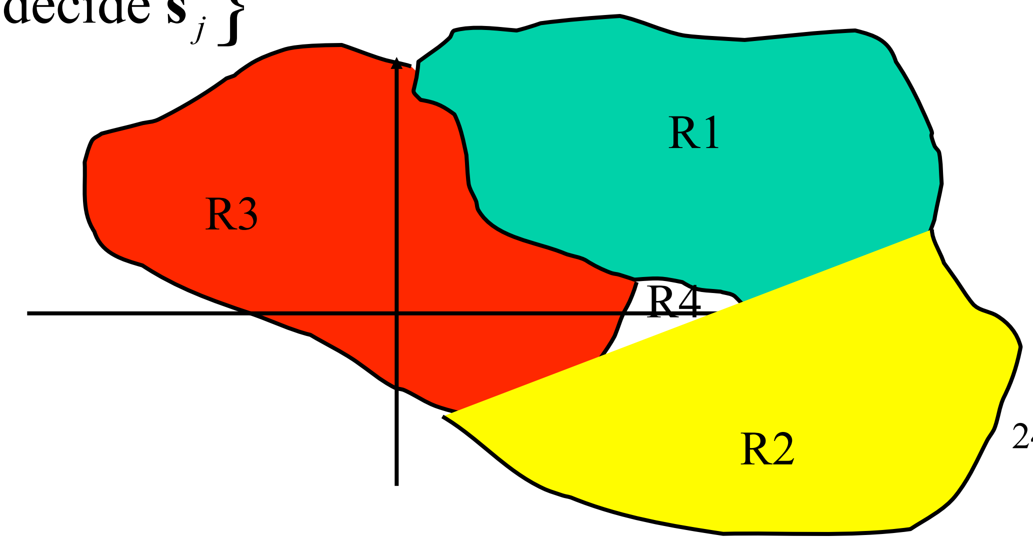
- Case 1) If you are the project manager of a space programme to launch a satellite. You have to decide which modulation family to use for communication from the satellite to an earth station. What would be your choice?

- Case 2) If you are the project manager to set up a wireless video link from Lantau island to HKUST. Which modulation family would you pick?

Proof of Optimal Detection

- We try to establish mathematical proof that "minimum distance detection" is optimal for AWGN.
- Decision Region:
  - Any demodulation scheme can be represented by a decision region on the observation space **y**.
    - » For example, given M possible transmitted symbols $\{\mathbf{s}_1,\ldots,\mathbf{s}_M\}$, the observation space is partitioned into M disjoint regions $\{\mathbf{R}_1,\ldots,\mathbf{R}_M\}$.

$$R_j = \{\mathbf{y} : \text{ decide } \mathbf{s}_j\}$$



$\vec{y} \in \mathbb{R}^N \longrightarrow \boxed{\text{Demod}} \longrightarrow \hat{s} \in \{\vec{s}_1, \dots, \vec{s}_M\}$

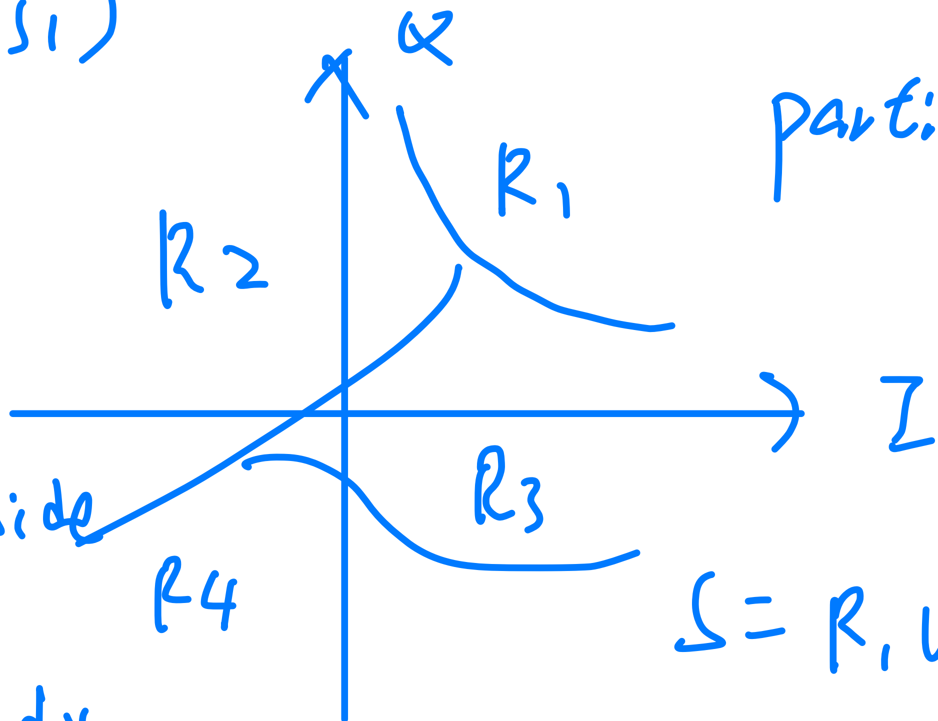
$P(\text{error} \mid s_1)$

$= \int_{R_1^c} f$

integrate all the region outside s_1!

$= \int p(\vec{y} \mid \vec{s}_1)\, dy$

$R_1^c = R_2 \cup R_3 \cup R_4 \cdots \cup R_m$



partition into 4-regions

$S = R_1 \cup R_2 \cup R_3 \cup R_4$

$\phi = R_i \cap R_j \quad \forall i \neq j,$

if $n(t) \sim$ colored

$$f(y|\vec{s}) \sim \exp -\left[(\vec{y}-\vec{s})^T Q^{-1} (\vec{y}-\vec{s})\right]$$

$\downarrow$

minimum distance detection will not hold

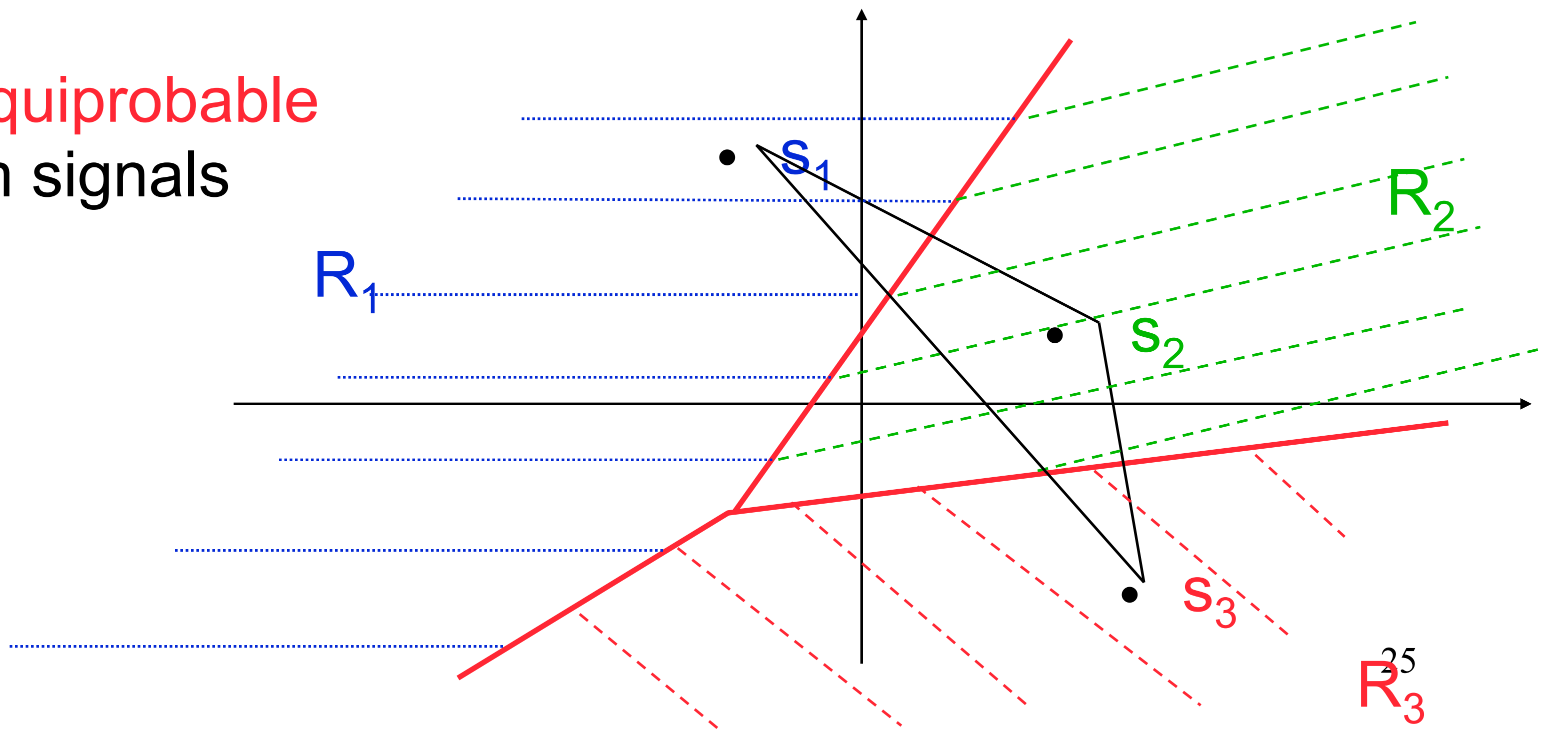
next time why P_e depends on d_{min}?

the minimum distance.

Example - Decision Region of the Min-Distance Detector

In general, boundaries of decision regions are perpendicular bisectors of the lines joining the original transmitted signals. In particular, the pairwise decision regions of two signal points are the half-spaces containing the signal points divided by the bisector.

Three equiprobable 2-dim signals



Proof of Optimal Detection

- Given s_1 is transmitted, the error probability is given by:

$$P_e(\mathbf{s}_1) = \int_{R_1^c} p(\mathbf{y} \mid \mathbf{s}_1) d\mathbf{y} = 1 - \int_{R_1} p(\mathbf{y} \mid \mathbf{s}_1) d\mathbf{y}$$

- Hence, the average error probability is:

$$P_e = \sum_{j=1}^{M} p(\mathbf{s}_j) \int_{\mathcal{R}_j^c} p(\mathbf{y}|\mathbf{s}_j) d\mathbf{y} = 1 - \sum_{j=1}^{M} \int_{\mathcal{R}_j} p(\mathbf{s}_j) p(\mathbf{y}|\mathbf{s}_j) d\mathbf{y}$$

weighted overall probability!

we have freedom to decide this!

Since each y must belong to one decision region only, we should assign y to R_j (for minimizing the error probability) if $p(\mathbf{y} \mid \mathbf{s}_j) p(\mathbf{s}_j)$ is maximized.

Optimal detection:

$$\mathbf{s}^* = \arg \max_{\mathbf{s} \in \{\mathbf{s}_1, \ldots, \mathbf{s}_N\}} p(\mathbf{s}) p(\mathbf{y}|\mathbf{s})$$

Likelihood function

Maximal Likelihood (ML) Detection:

if p(s) are equally likely

assign y = Rj

If the M symbols are equal-likely, the optimal detection becomes

$$\mathbf{s}^* = \arg \max_{\mathbf{s} \in \{\mathbf{s}_1, \ldots, \mathbf{s}_N\}} p(\mathbf{y}|\mathbf{s})$$

Maximal Likelihood Detection

- Since $\mathbf{y} = \mathbf{s} + \mathbf{n}$ and the noise vector is iid Gaussian random vector, we have

$$p(\mathbf{y}|\mathbf{s}) \sim \exp - \left(\frac{1}{2\sigma_n^2} \|\mathbf{y} - \mathbf{s}\|^2 \right)$$

likelihood $= p(\vec{y} | \vec{n})$

$\vec{y} = \vec{s} + \vec{n}$

suppress the randomless!

- Hence, the ML solution is:

$$b = \arg\min_{\mathbf{s} \in \{\mathbf{s}_1, \dots, \mathbf{s}_M\}} \|\mathbf{y} - \mathbf{s}\|^2 = \arg\max_{\mathbf{s} \in \{\mathbf{s}_1, \dots, \mathbf{s}_M\}} \left[2\mathbf{y} \bullet \mathbf{s} - \|\mathbf{s}\|^2 \right] = \arg\max_{\mathbf{s} \in \{\mathbf{s}_1, \dots, \mathbf{s}_M\}} \left[2\int_0^{T_s} y(t)s(t)dt - \int_0^{T_s} s^2(t)dt \right]$$

$f(\vec{y} | \vec{s}) = f(\vec{n})$

$\boxed{\text{Minimum Distance Detection}}$

$\downarrow$

optimal

$E(n_n) = 0$

$E(n_n^2) = \frac{\eta_0}{2}$

Proof: Error Probability Analysis

- Recall that

$$P_e \leq Q\left(\sqrt{\frac{d_{min}^2}{2\eta_0}}\right) \text{ where } d_{min} = \min_{i,j,i\neq j} \|\vec{s}_i - \vec{s}_j\|$$

- We are going to prove this result.
  - Error occurs when the estimated transmitted point is not the original transmitted point. $P_e = \sum_{m=1}^{M} P(\mathbf{s}=\mathbf{s}_m)P(\Psi_m|\mathbf{s}_m) \quad \Psi_m = Event\left\{\|\mathbf{y}-\mathbf{s}_m\|^2 \text{ is not min}\right\}$

    If $P(s_1)=P(s_2)=..=P(s_M)=1/\text{M}$ and $P(\Psi_1|\mathbf{s}_1)=...=P(\Psi_M|\mathbf{s}_M)$

    then $P_e = P(\Psi_1|\mathbf{s}_1)$
  - Without loss of generality, let $s_1(t)$ be the transmitted signal. The error event is given by: $\Psi_1 = \bigcup_{n\neq 1}\left\{\|\mathbf{y}-\mathbf{s}_n\|^2 \leq \|\mathbf{y}-\mathbf{s}_1\|^2\right\}$
  - Error probability is given by: (equality holds for M=2)

$$P_e = P(\Psi_1|s_1) = \Pr\left\{\bigcup_{n\neq 1}\left[\|y-s_1\|^2 \geq \|y-s_n\|^2\right]\right\} \leq \sum_{n\neq 1}\Pr\left[\|y-s_1\|^2 \geq \|y-s_n\|^2\right]$$

$$P_e(\psi_1|s_1) = Pr\left(\|\vec{y}-\vec{s}_1\| \text{ is not min} \,\middle|\, s_1\right)$$

$$= Pr\left(\bigcup_{n\neq 1}^{M} \left\{\|\vec{y}-\vec{s}_n\|^2 < \|\vec{y}-\vec{s}_1\|^2\right\} \middle| s_1\right)$$

$$\le \sum_{n\neq 1}^{M} Pr\left(\|\vec{y}-\vec{s}_n\| < \|\underset{\text{random}}{\vec{y}}-\vec{s}_1\| \,\middle|\, s_1\right)$$

event A_n.

$\Sigma - \Sigma$ mutually exclusive

$$Pr(A_1 \cup A_2) = Pr(A_1) + Pr(A_2) - Pr(A_1 \cap A_2)$$

$$Pr(A_1 \cup A_2) \le P(A_1) + P(A_2)$$

upper bound!

union bound!

$Pr(R.V. \le \text{const})$

Conditioning suppress the randomness!

$$\|\vec{y}-\vec{s}_n\|^2 < \|\vec{y}-\vec{s}_1\|^2$$

$$\cancel{\|y\|^2} - 2\langle \vec{y}, \vec{s}_n\rangle + \|\vec{s}_n\|^2 < \cancel{\|\vec{y}\|^2} - 2\langle \vec{y}, \vec{s}_1\rangle + \|s_1\|^2$$

$$-2\langle \vec{s}_1+\vec{n}, \vec{s}_n\rangle + \|\vec{s}_n\|^2 \leq -2\langle \vec{s}_1+\vec{n}, \vec{s}_1\rangle + \|s_1\|^2$$

$$-2\langle \vec{s}_1, \vec{s}_n\rangle - 2\langle \vec{n}, \vec{s}_n\rangle + \|\vec{s}_n\|^2 \leq -2\langle \vec{s}_1, \vec{s}_1\rangle - 2\langle \vec{n}, \vec{s}_1\rangle + \|s_1\|^2$$

$$\|\vec{s}_n - \vec{s}_1\|^2 \leq 2\langle \vec{n}, \vec{s}_n - \vec{s}_1\rangle$$

Proof: Error Analysis

- The event in the probability term can be simplified as:

$$\|\mathbf{y}-\mathbf{s}_1\|^2 \geq \|\mathbf{y}-\mathbf{s}_j\|^2 \Leftrightarrow (\mathbf{s}_j-\mathbf{s}_1)\bullet\mathbf{n} \geq \frac{1}{2}\|\mathbf{s}_j-\mathbf{s}_1\|^2 \Leftrightarrow \int_0^{T_s}\left[s_j(t)-s_1(t)\right]n(t)dt \geq \frac{1}{2}\int_0^{T_s}\left[s_j(t)-s_1(t)\right]^2 dt$$

- Let $Z_j = \int_0^{T_s}\left[s_j(t)-s_1(t)\right]n(t)dt$

- Since n(t) is a white Gaussian random process, Zj is a Gaussian random variable with zero mean and variance $\sigma_Z^2 = \frac{N_0}{2}\int_0^{T_s}\left[s_j(t)-s_1(t)\right]^2 dt$

- The error probability is given by:

1st s1 condition!

$$P_j = \Pr\left[Z_j \geq \frac{1}{2}\int_0^{T_s}\left[s_j(t)-s_i(t)\right]^2 dt\right] = Q\left(\sqrt{\frac{\|\mathbf{s}_j-\mathbf{s}_1\|^2}{2\eta_0}}\right)$$

mean

$$P_e \leq \sum_{j\neq 1} P_j \leq \sum_{j\neq 1} Q\left(\sqrt{\frac{\|\mathbf{s}_j-\mathbf{s}_1\|^2}{2\eta_0}}\right) \approx Q\left(\sqrt{\frac{\min_{i,j,i\neq j}\|\mathbf{s}_i-\mathbf{s}_j\|^2}{2\eta_0}}\right)$$

$$E(z_j|s_1) = E\left(\int_0^{T_s} (s_i(t) - s_1(t))\, n(t)\, dt \,\Big|\, s_1\right)$$

$$= \int_0^{T_s} (s_i(t) - s_1(t))\, E(n(t))\, dt$$

$$= 0$$

$$Var(z_j|s_1) = E(z_j^2 \,|\, s_1)$$

$$= E\left(\int_0^{T_s}\int_0^{T_s} (\Delta s(t))\, \Delta s(t')\, n(t)\, n(t')\, dt\, dt' \,\Big|\, s_1\right)$$

$$= \int_0^{T_s}\int_0^{T_s} \Delta s(t)\, \Delta s(t')\, E[n(t)\, n(t')]\, dt\, dt'$$

$$= \int_0^{T_s}\int_0^{T_s} \Delta s(t)\, \Delta s(t')\, \frac{N_0}{2}\, \delta(t - t')\, dt\, dt'$$

$$= \int_0^{T_s} \Delta s(t)\, \Delta s(t)\, \frac{N_0}{2}\, dt$$

$$= \frac{N_0}{2} \int_0^{T_s} (\Delta s(t))^2\, dt$$

$$S_n(f) = \frac{N_0}{2}$$

$$\downarrow$$

$$R_n(\tau) \triangleq E(n(t)\, n(t+\tau))$$

$$= \frac{N_0}{2}\, \delta(\tau)$$

$$P_e = P_r(\psi_1 \mid s_1) \le \sum_{j \ne 1}^{M} Q\left(\sqrt{\frac{\|s_j - s_1\|^2}{2N_0}}\right)$$

union bound

$$Q(x) \approx k\exp\left(-\frac{x^2}{2}\right)$$

if $x \geq 3$

$$f(x) = e^{-\frac{x^2}{2}} + e^{-x^2} + e^{-2x^2}$$

$$\approx e^{-\frac{x^2}{2}}$$

$$= 1 \times \sum_{j \ne 1}^{M} Q\sqrt{\frac{\min_{j \ne 1} \|s_j - s_1\|^2}{2N_0}}$$

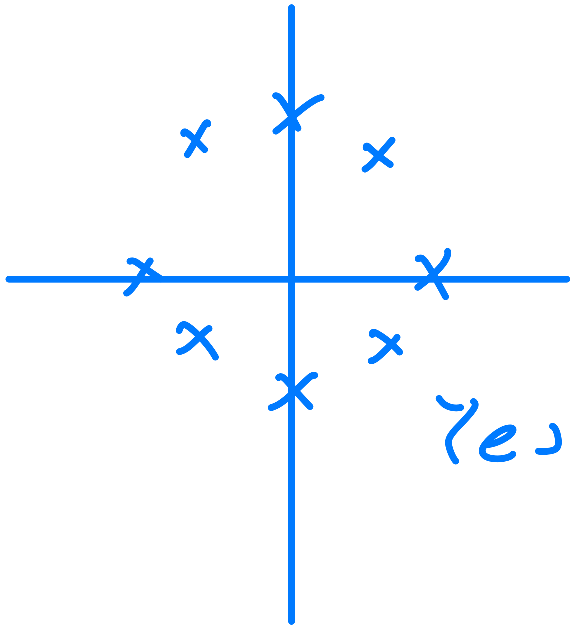
for high SNR

$$P_e = \sum_{m=1}^{M} \left(\frac{1}{M}\right) \Pr(\text{err} \mid s_m)$$

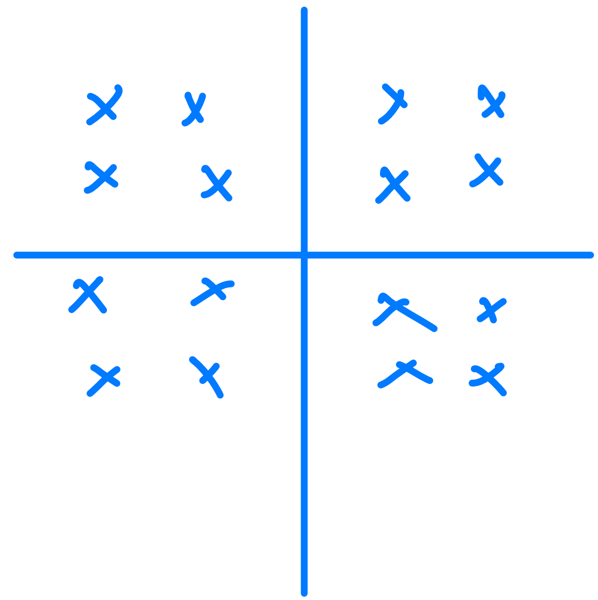
$$\leq \frac{1}{M} \sum_{m=1}^{M} Q\sqrt{\frac{\min_{j \neq m} \|s_j - s_m\|^2}{2N_0}}$$

$$\leq \frac{1}{M} Q\left(\sqrt{\frac{\underbrace{\min_{i \neq j} \|s_j - s_i\|^2}_{d_{min}^2}}{2N_0}}\right)$$

$$p(\psi_1|s_1) = p(\psi_2|s_2) = \cdots = p(\psi_m|s_m) \, ?$$



Yes!



No!

Example

Three equally probable messages m1, m2, and m3 are to be transmitted over an AWGN channel with noise power-spectral density $N_0 / 2$. The messages are

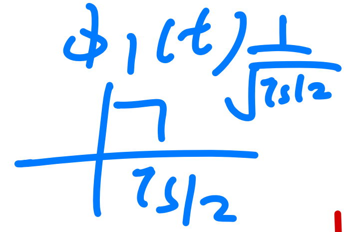
$\left(\sqrt{\frac{T_s}{2}}, \sqrt{\frac{T_s}{2}}\right)$

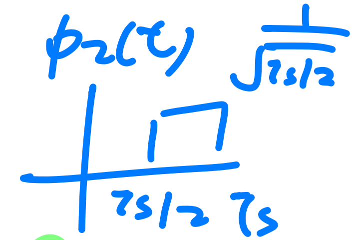
$$s_1(t) = \begin{cases} 1 & 0 \le t \le T \\ 0 & otherwise \end{cases}$$

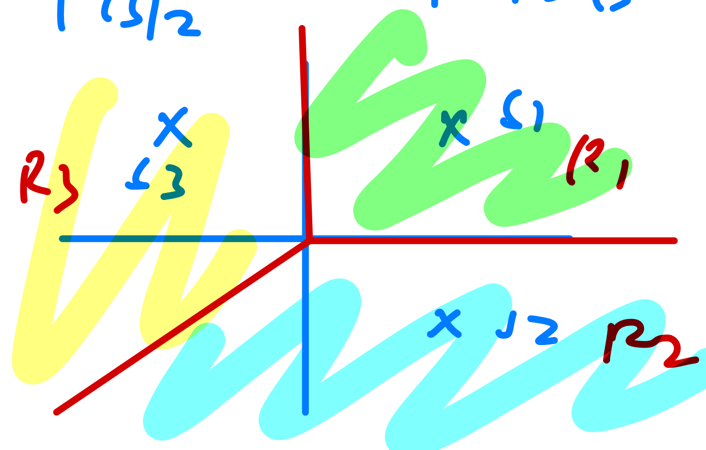
$\left(\sqrt{\frac{T_s}{2}}, -\sqrt{\frac{T_s}{2}}\right)$

$\left(-\sqrt{\frac{T_s}{2}}, \sqrt{\frac{T_s}{2}}\right)$

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







1. What is the dimensionality of the signal space ? 2
2. Find an appropriate basis for the signal space (Hint: You can find the basis without using the Gram-Schmidt procedure).
3. Draw the signal constellation for this problem. minimum distance!
4. Sketch the optimal decision regions R1, R2, and R3.
5. Which of the three messages is more vulnerable to errors and why ? In other words, which of $p(Error \mid m_i \quad transmitted), \quad i = 1, \quad 2, \quad 3$ is larger ?

P(ψ|s1) is maximum! i=1 multiplicity =2, 2 equal ⊗ function