Point-to-Point Wireless Communication (I): Digital Basics, Modulation, Detection, Pulse Shaping

Shiv Kalyanaraman

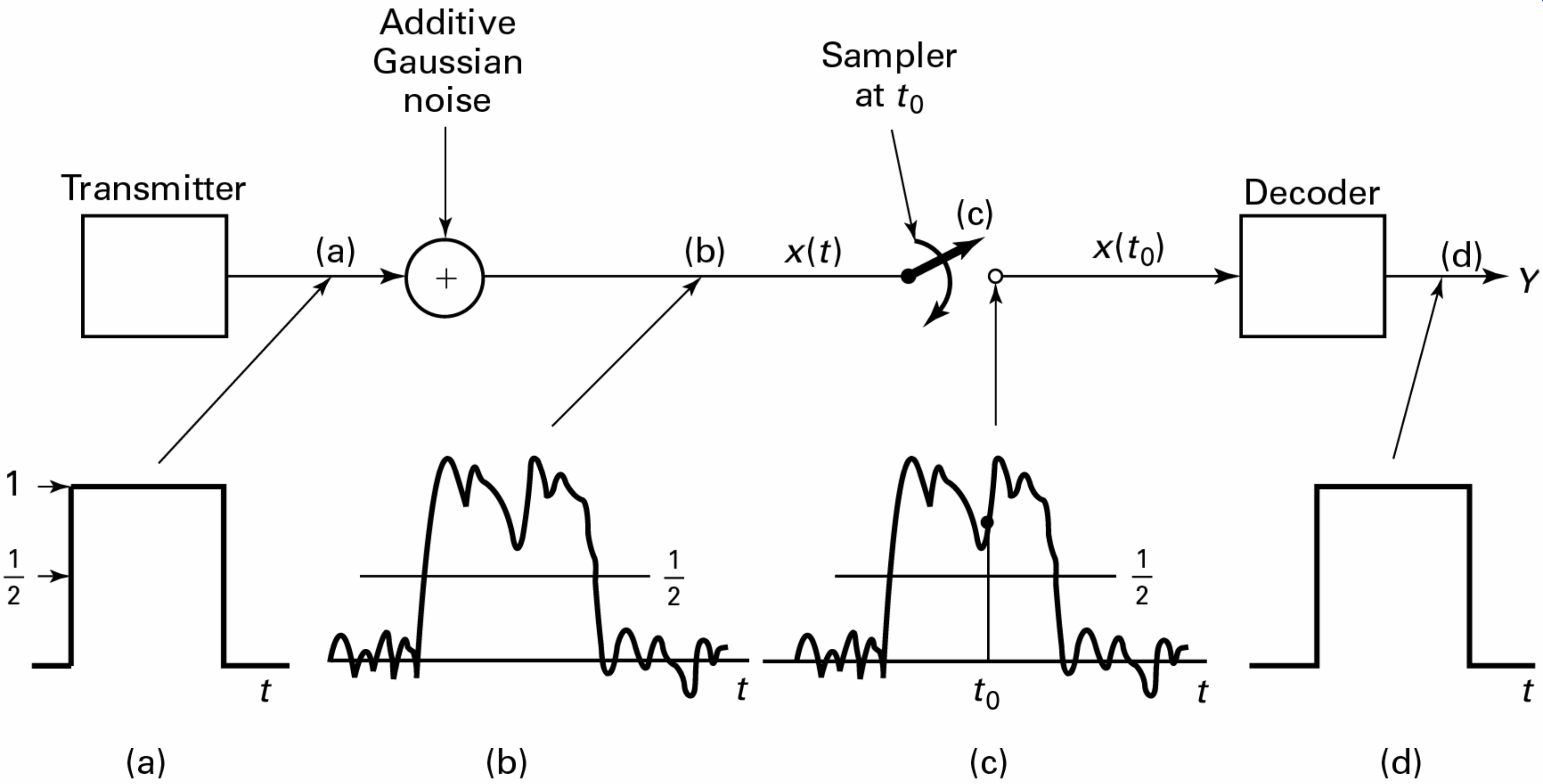
Google: “Shiv RPI”

shivkuma@ecse.rpi.edu

Based upon slides of Sorour Falahati, A. Goldsmith, & textbooks by Bernard Sklar & A. Goldsmith

The Basics

Big Picture: Detection under AWGN

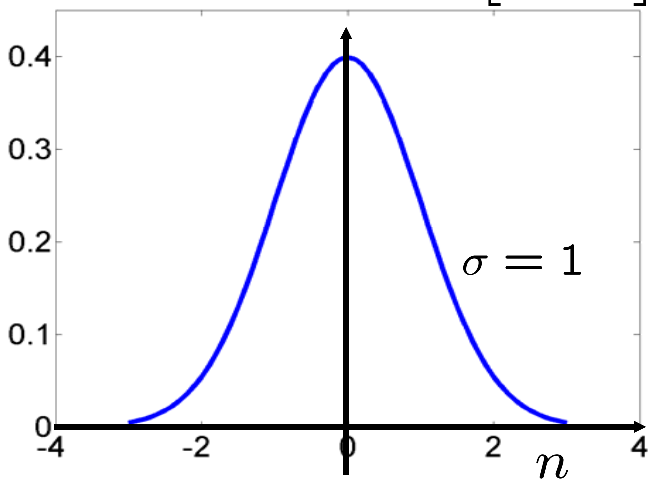


Decoding of a noise-corrupted digital pulse by sampling and hard clipping.

Additive White Gaussian Noise (AWGN)

- Thermal noise is described by a zero-mean **Gaussian** random process, n(t) that ADDS on to the signal => "**additive**"
- Its PSD is flat, hence, it is called ***white** noise*.
  - Autocorrelation is a spike at 0: uncorrelated at any non-zero lag

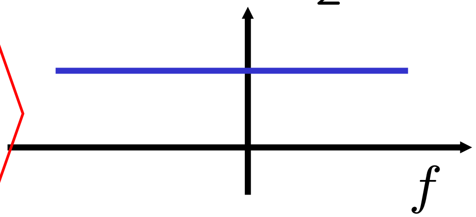
$$p(n) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{n^2}{2\sigma^2}\right]$$



Probability density function (gaussian)

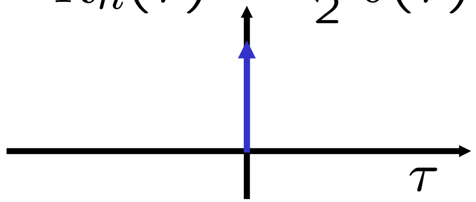
$$G_n(f) = \frac{N_0}{2} \text{ [w/Hz]}$$

Power spectral Density (flat => "white")



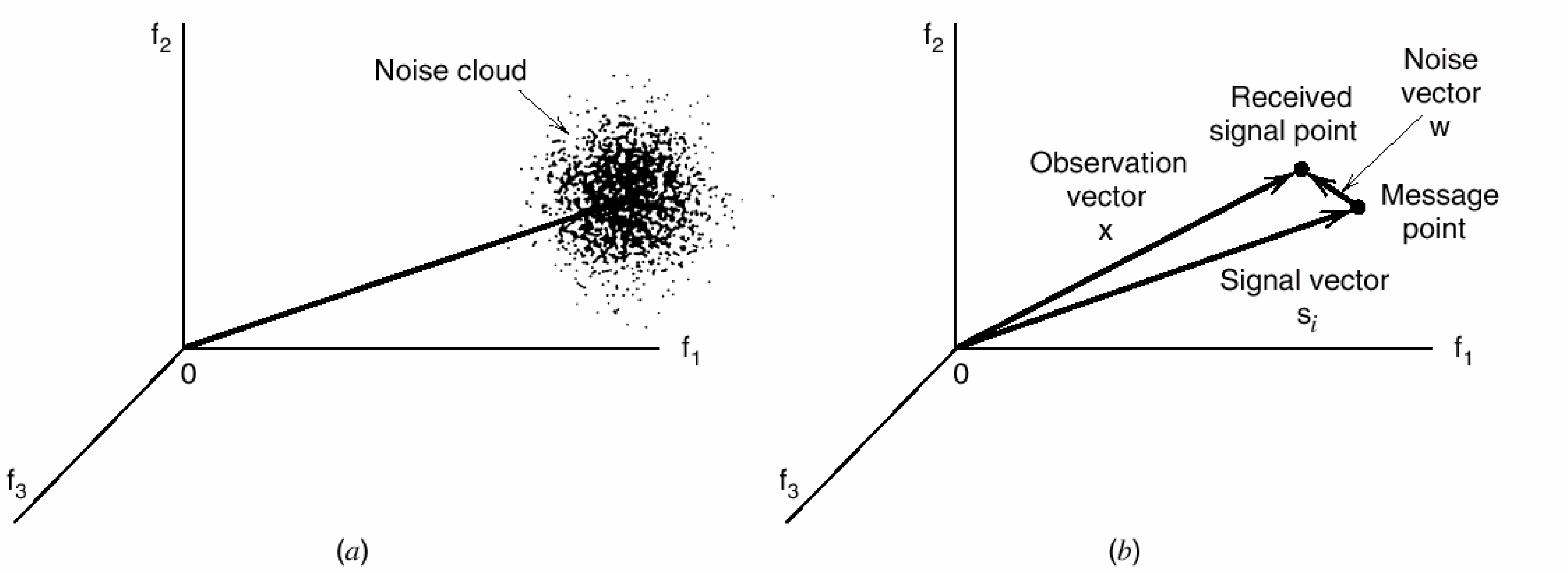
$$R_n(\tau) = \frac{N_0}{2}\delta(\tau)$$

Autocorrelation Function (uncorrelated)

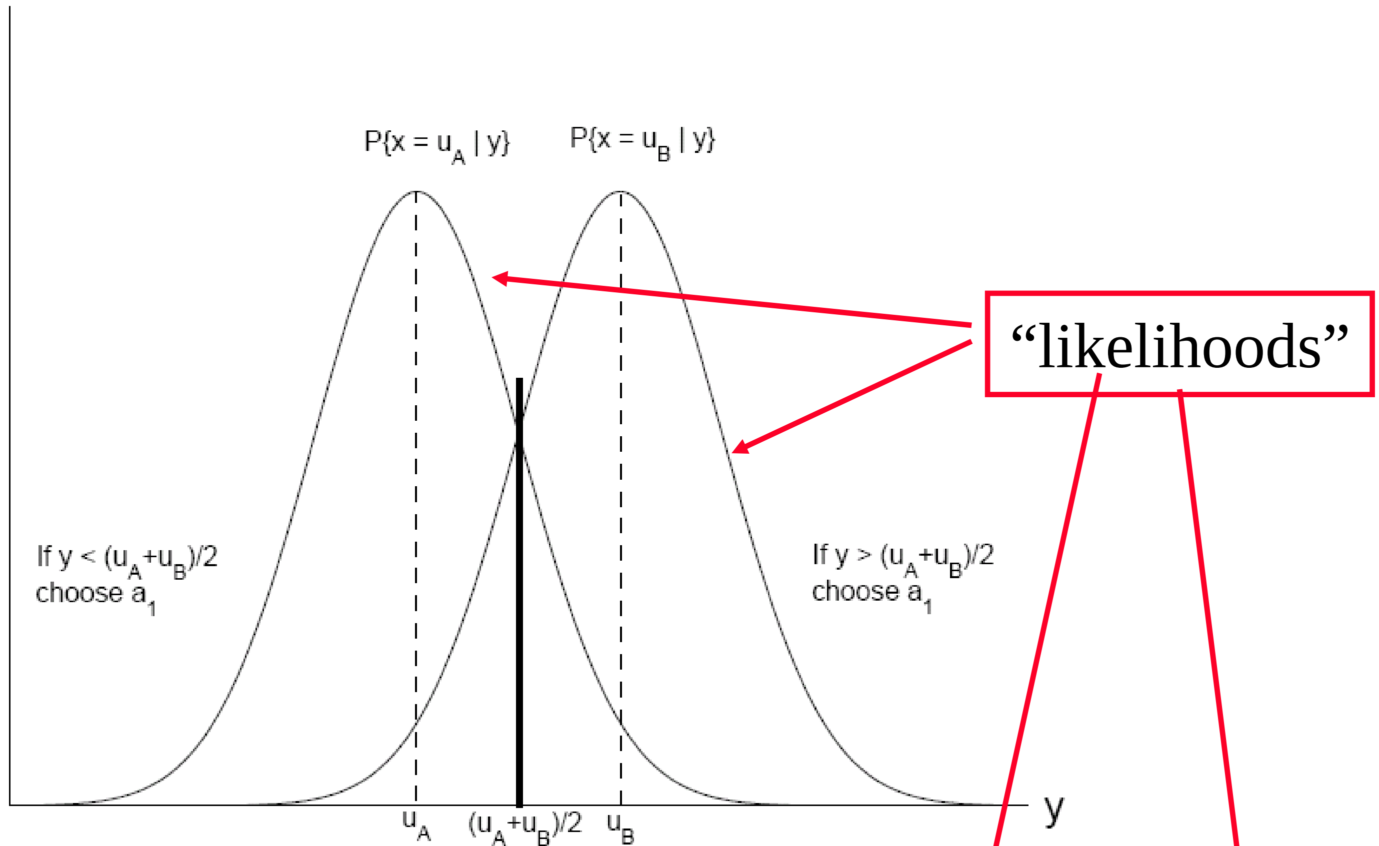


Effect of Noise in Signal Space

- The cloud falls off exponentially (gaussian).
- Vector viewpoint can be used in signal space, with a random noise vector **w**



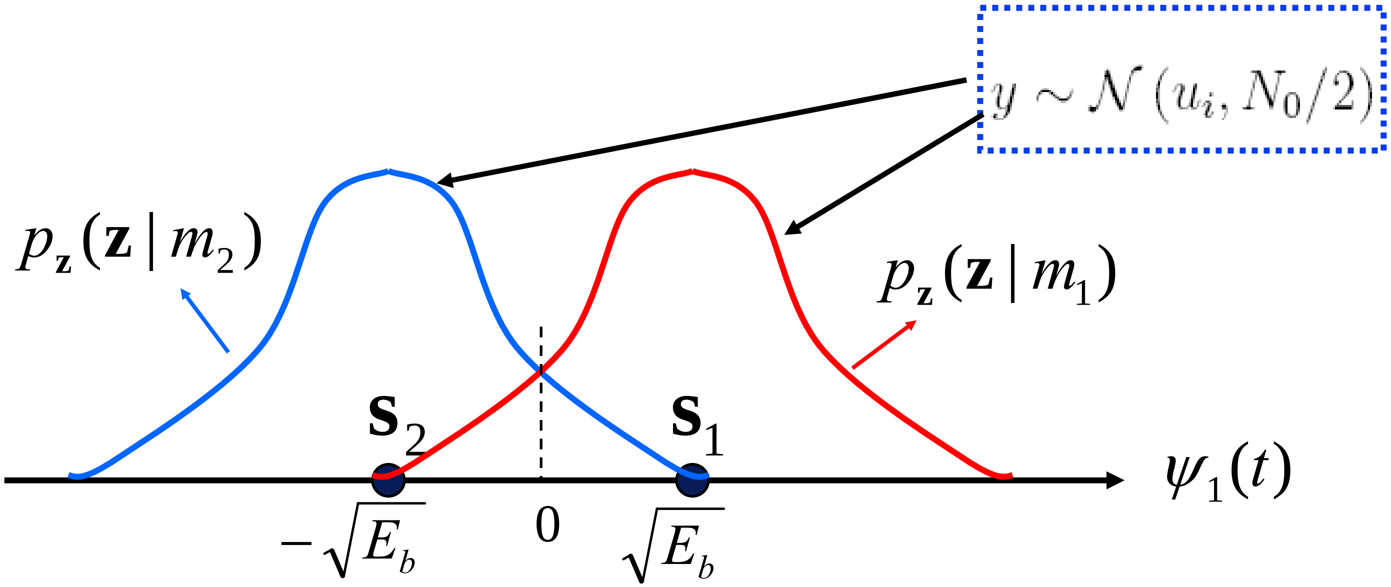
Maximum Likelihood (ML) Detection: *Scalar* Case



Assuming both symbols equally likely: u_A is chosen if $\mathbb{P}\{u = u_A|y\} \geq \mathbb{P}\{u = u_B|y\}$

Log-Likelihood => $|y - u_A| < |y - u_B|.$ **A simple distance criterion!**

AWGN Detection for Binary PAM

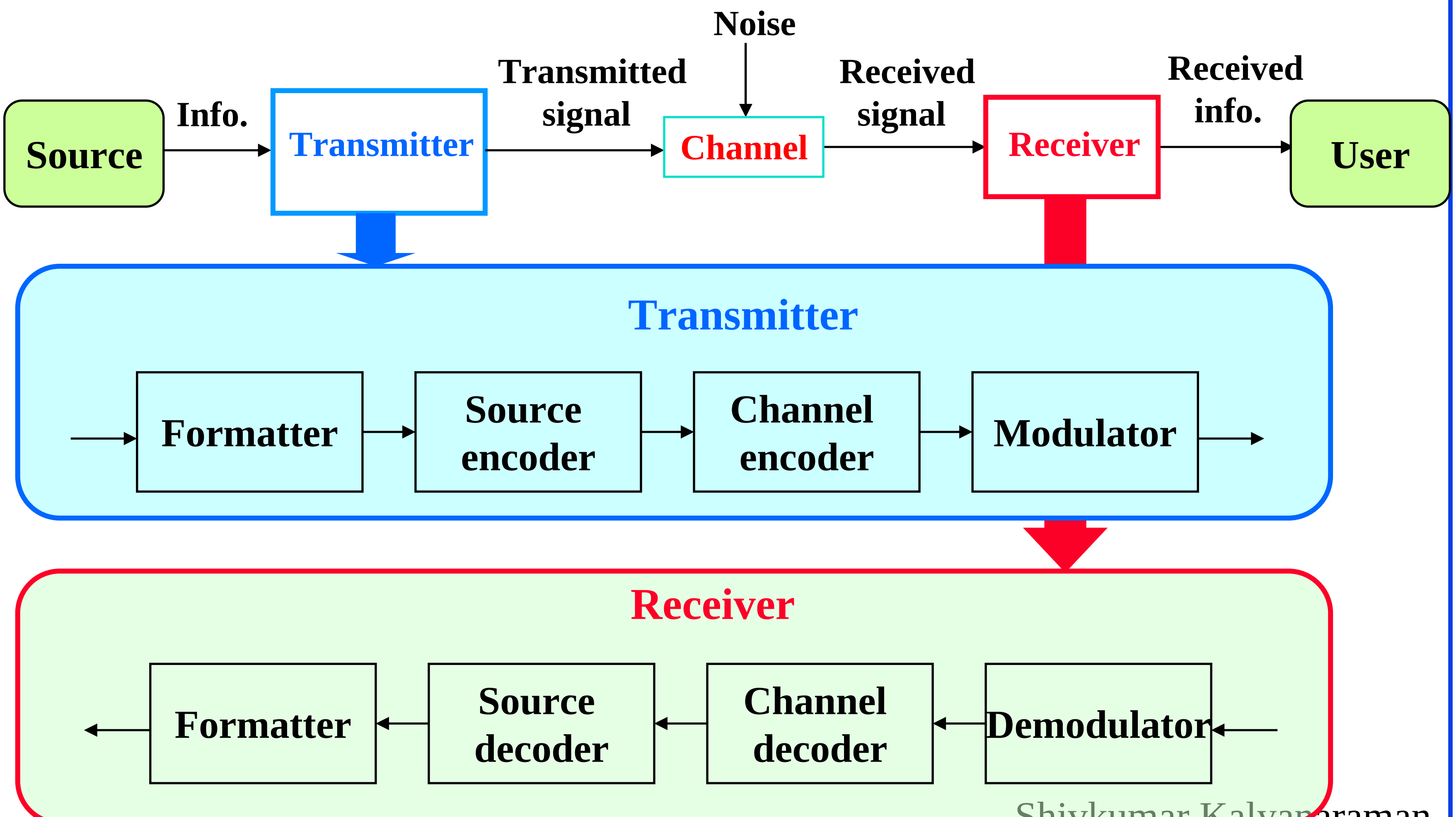


$$P_e(m_1) = P_e(m_2) = Q\left(\frac{\|\mathbf{s}_1 - \mathbf{s}_2\| / 2}{\sqrt{N_0 / 2}}\right)$$

$$P_B = P_E(2) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

Bigger Picture

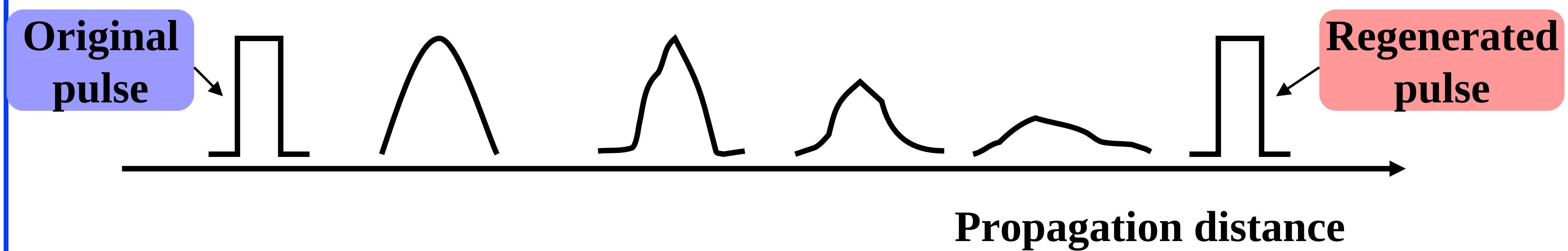
<u>General structure of a communication systems</u>



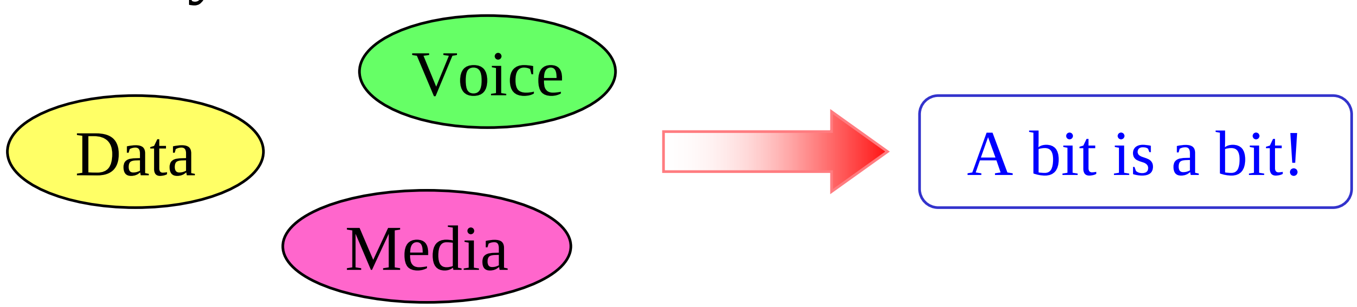
Digital vs Analog Comm: Basics

Digital versus analog

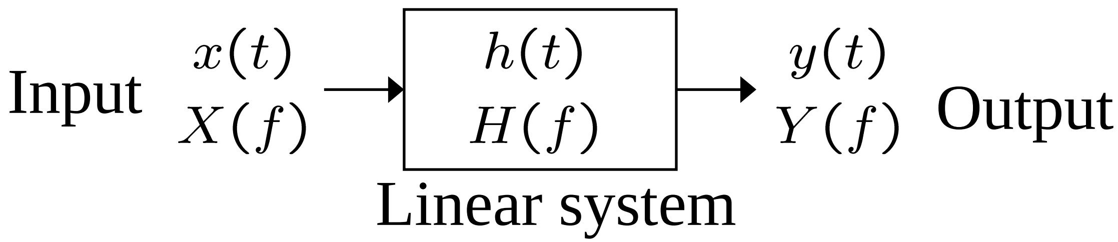
- Advantages of digital communications:
  - Regenerator receiver



  - Different kinds of digital signal are treated identically.



Signal transmission through linear systems



- Deterministic signals: $$\boxed{Y(f) = X(f)H(f)}$$
- Random signals: $$\boxed{G_Y(f) = G_X(f)|H(f)|^2}$$

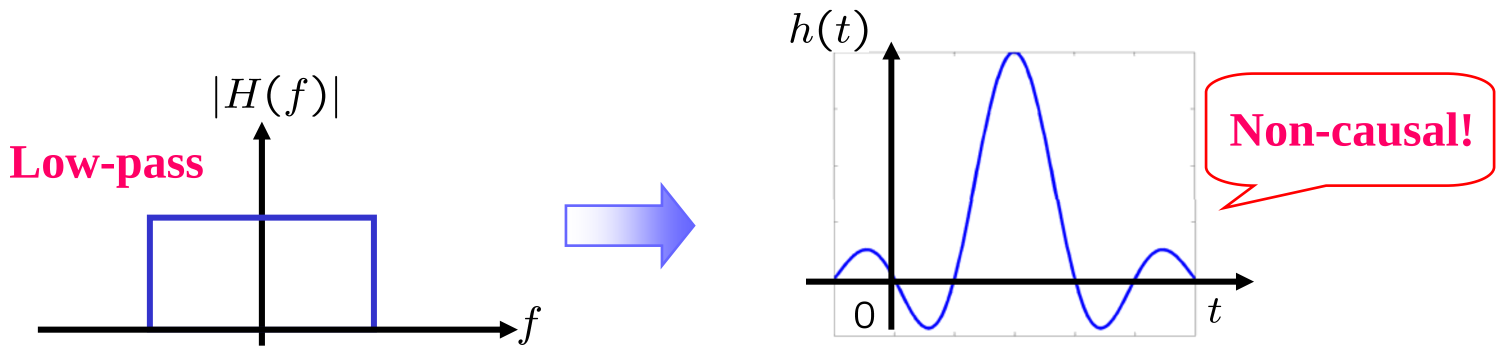
- Ideal distortion less transmission:

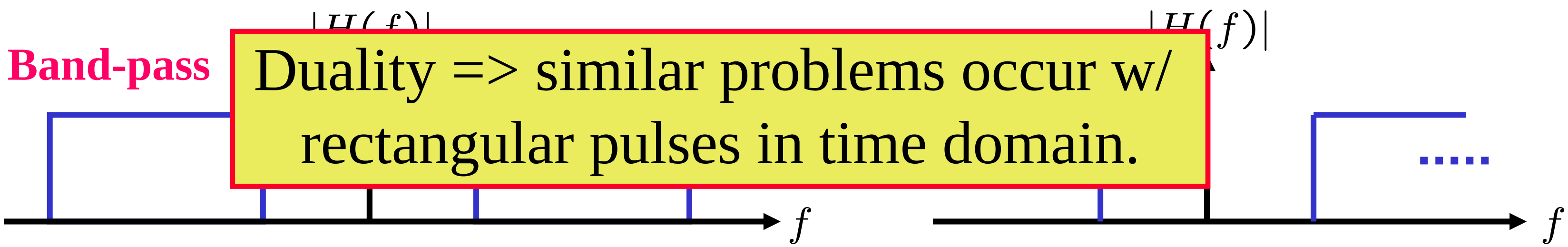
  All the frequency components of the signal not only arrive with an identical time delay, but also are amplified or attenuated equally.

$$\boxed{y(t) = Kx(t - t_0) \text{ or } H(f) = Ke^{-j2\pi f t_0}}$$

Signal transmission (cont'd)

- Ideal filters:





Duality => similar problems occur w/ rectangular pulses in time domain.

- Realizable filters:

RC filters

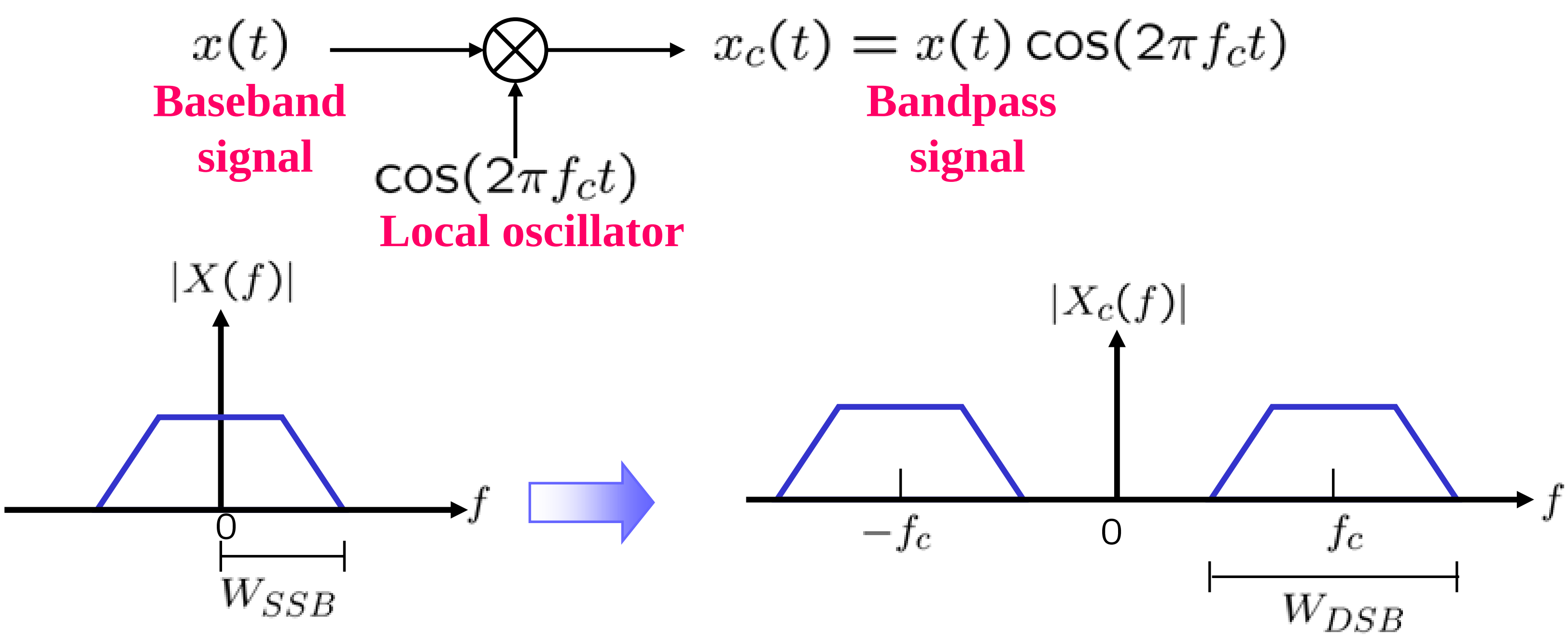
$$H(f) = \frac{1}{1+j2\pi f \mathcal{RC}}$$

Butterworth filter

$$|H_n(f)| = \frac{1}{\sqrt{1+(f/f_u)^{2n}}}$$

Bandwidth of signal

- Baseband versus bandpass:

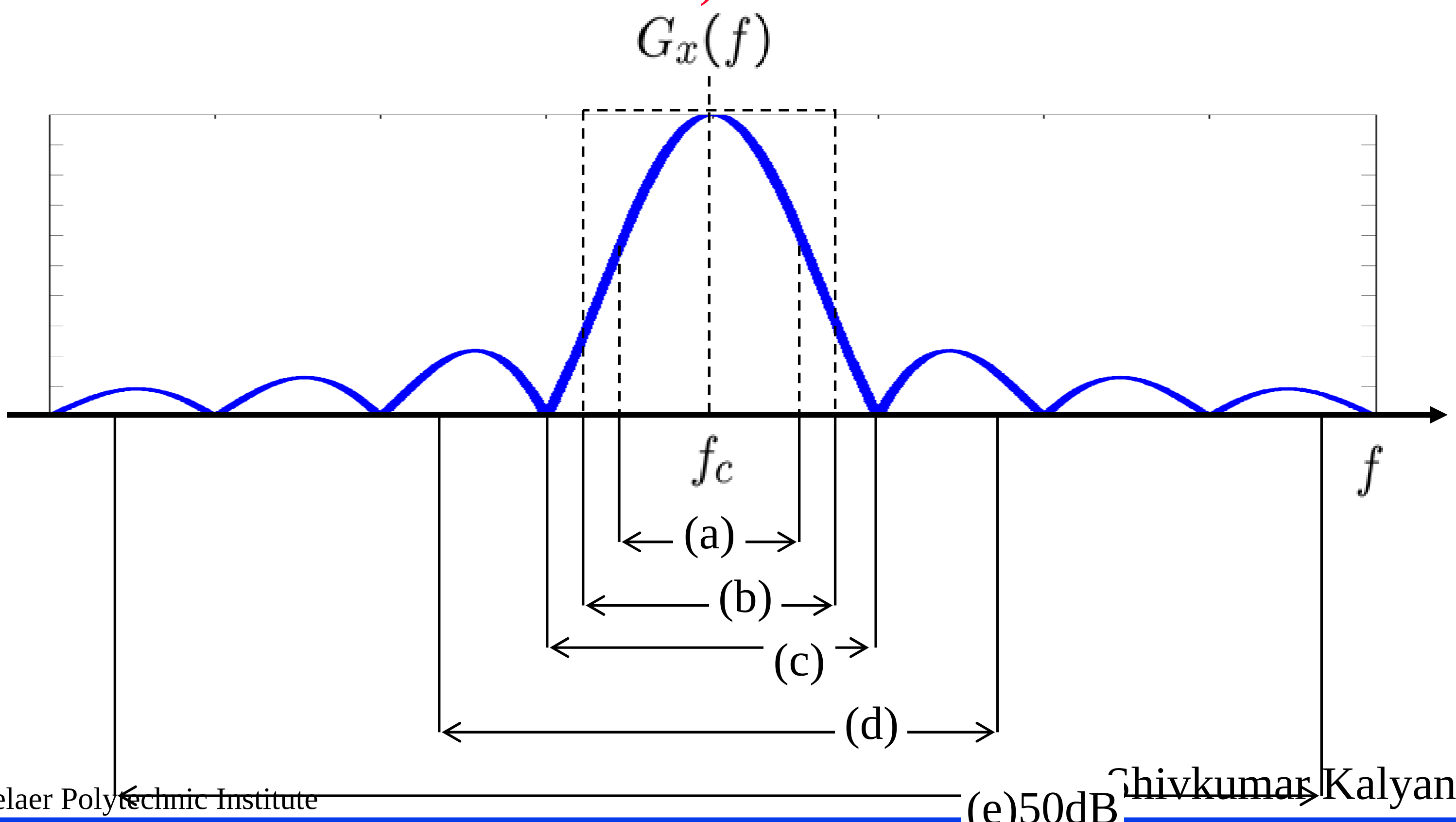


- Bandwidth dilemma:
  - Bandlimited signals are not realizable!
  - Realizable signals have infinite bandwidth!

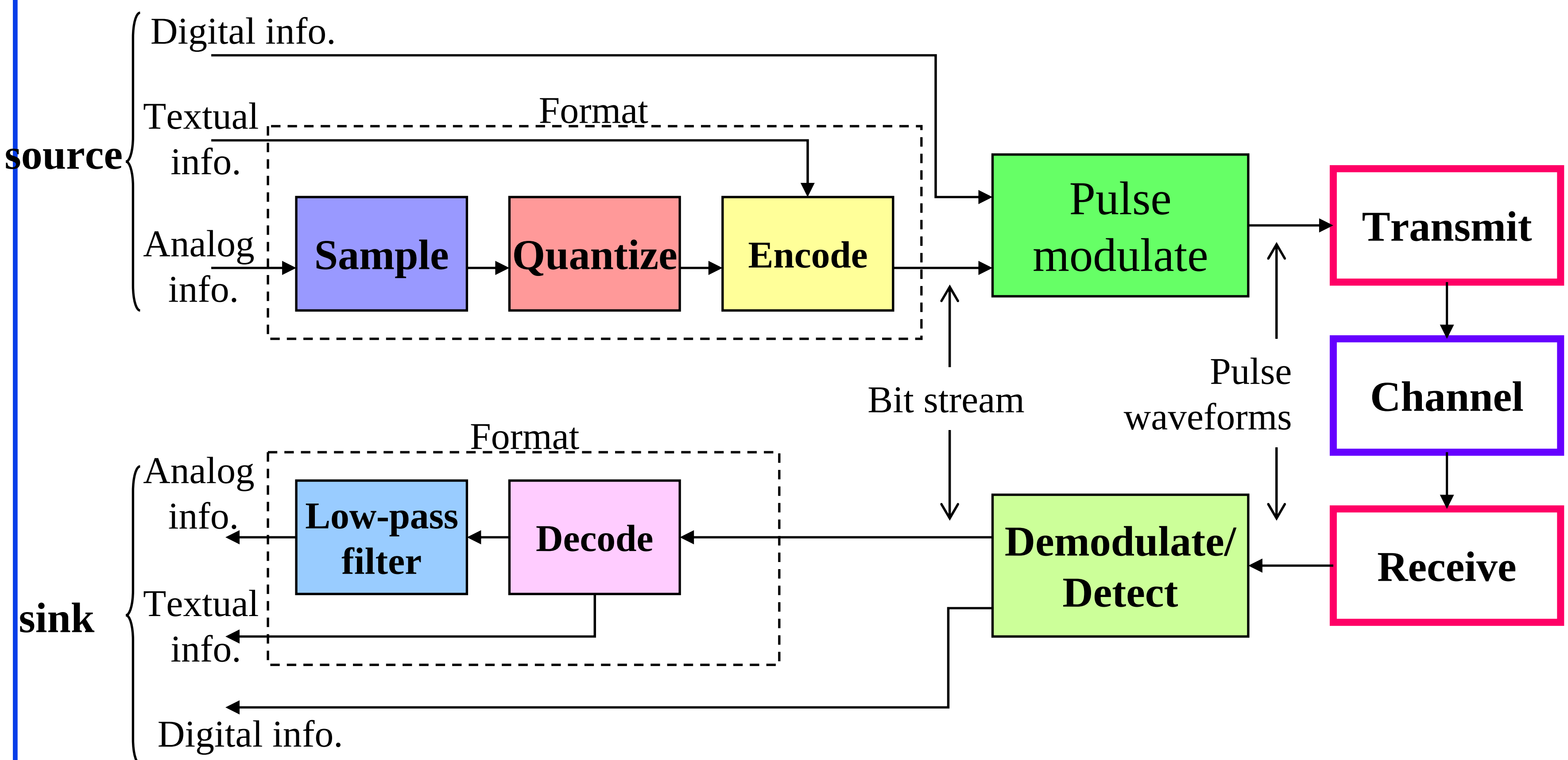
Bandwidth of signal: Approximations

- Different definition of bandwidth:

a) Half-power bandwidth
b) Noise equivalent bandwidth
c) Null-to-null bandwidth
d) Fractional power containment bandwidth
e) Bounded power spectral density
f) Absolute bandwidth



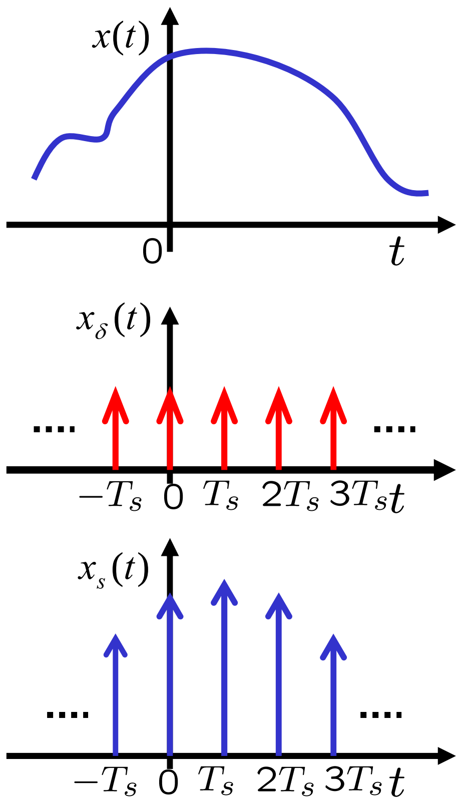
Formatting and transmission of baseband signal



Sampling of Analog Signals

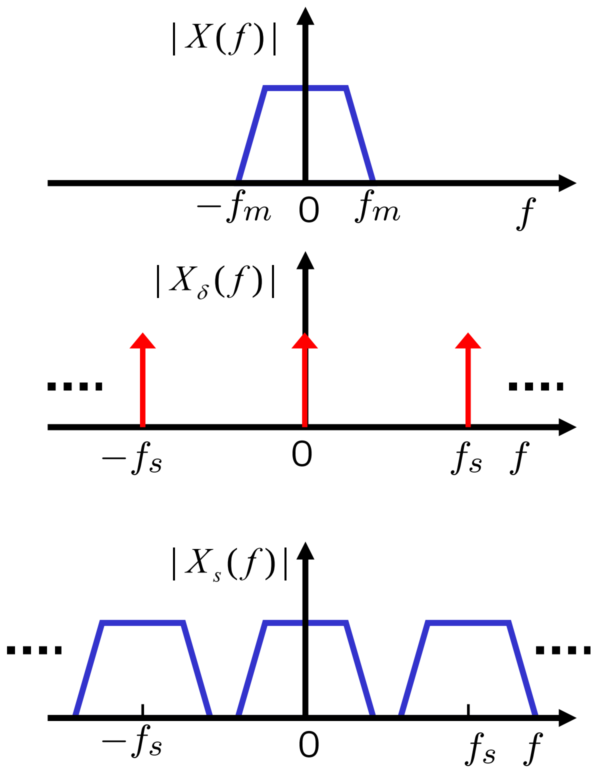
Time domain

$$x_s(t) = x_\delta(t) \times x(t)$$

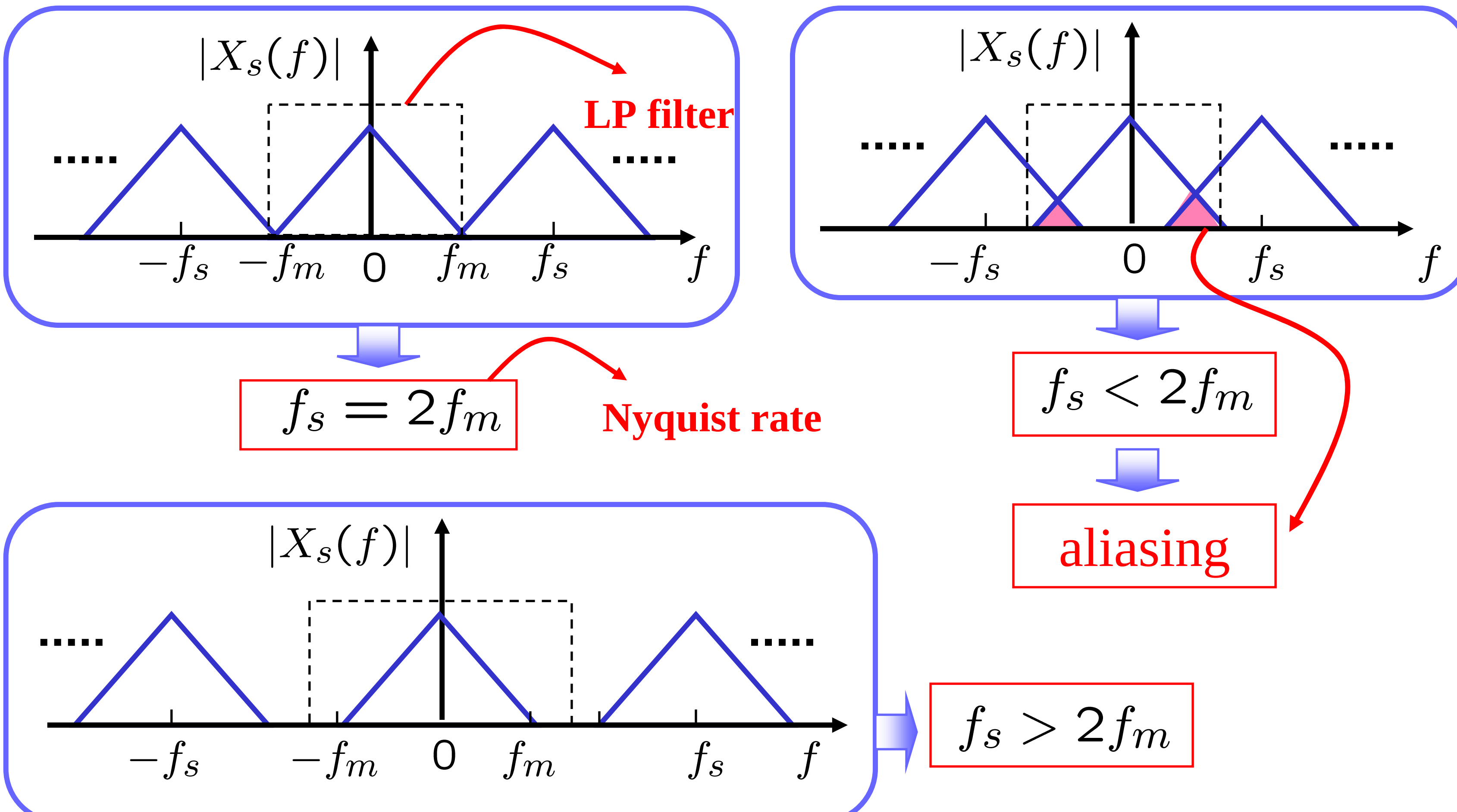


Frequency domain

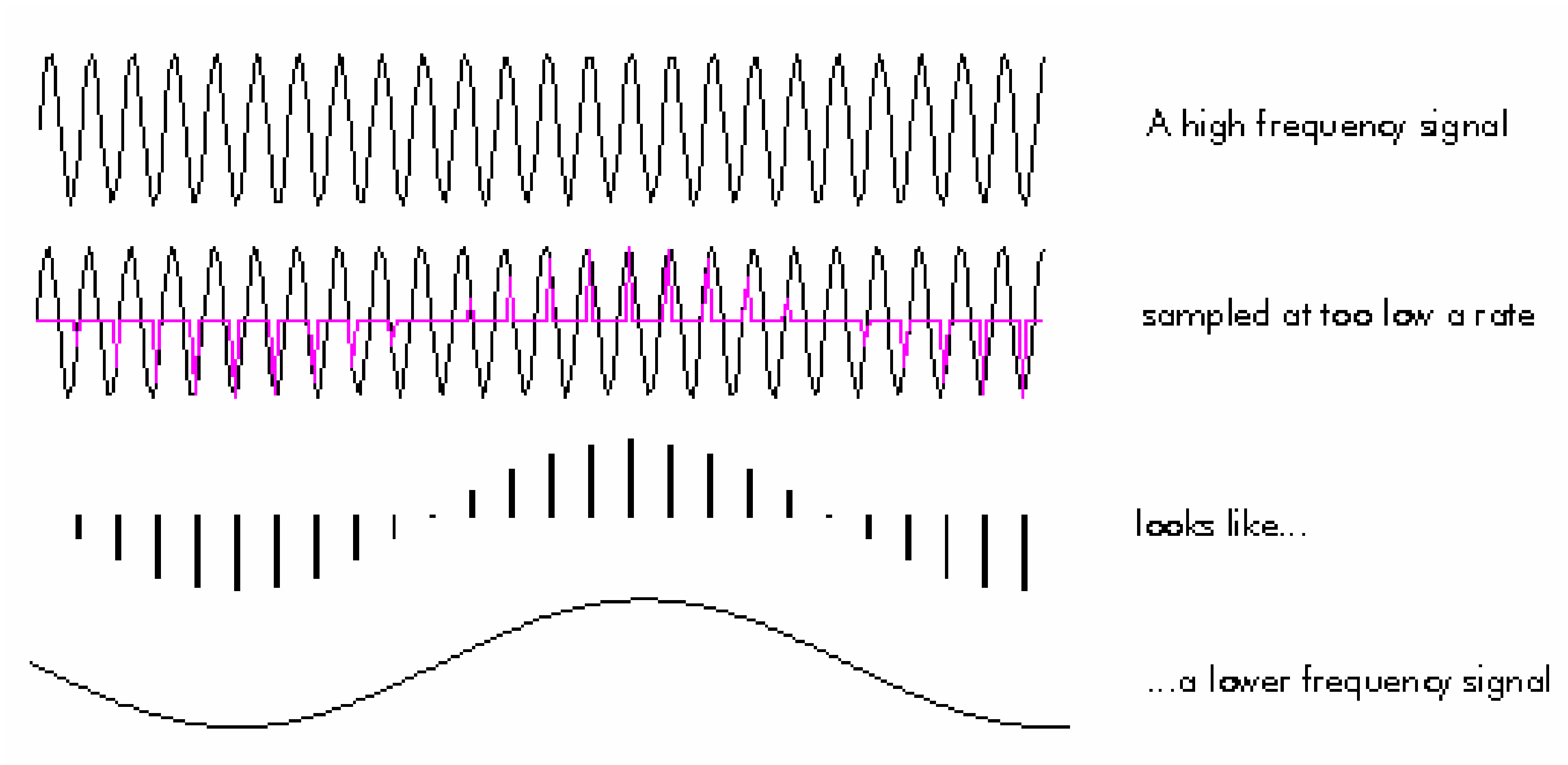
$$X_s(f) = X_\delta(f) * X(f)$$



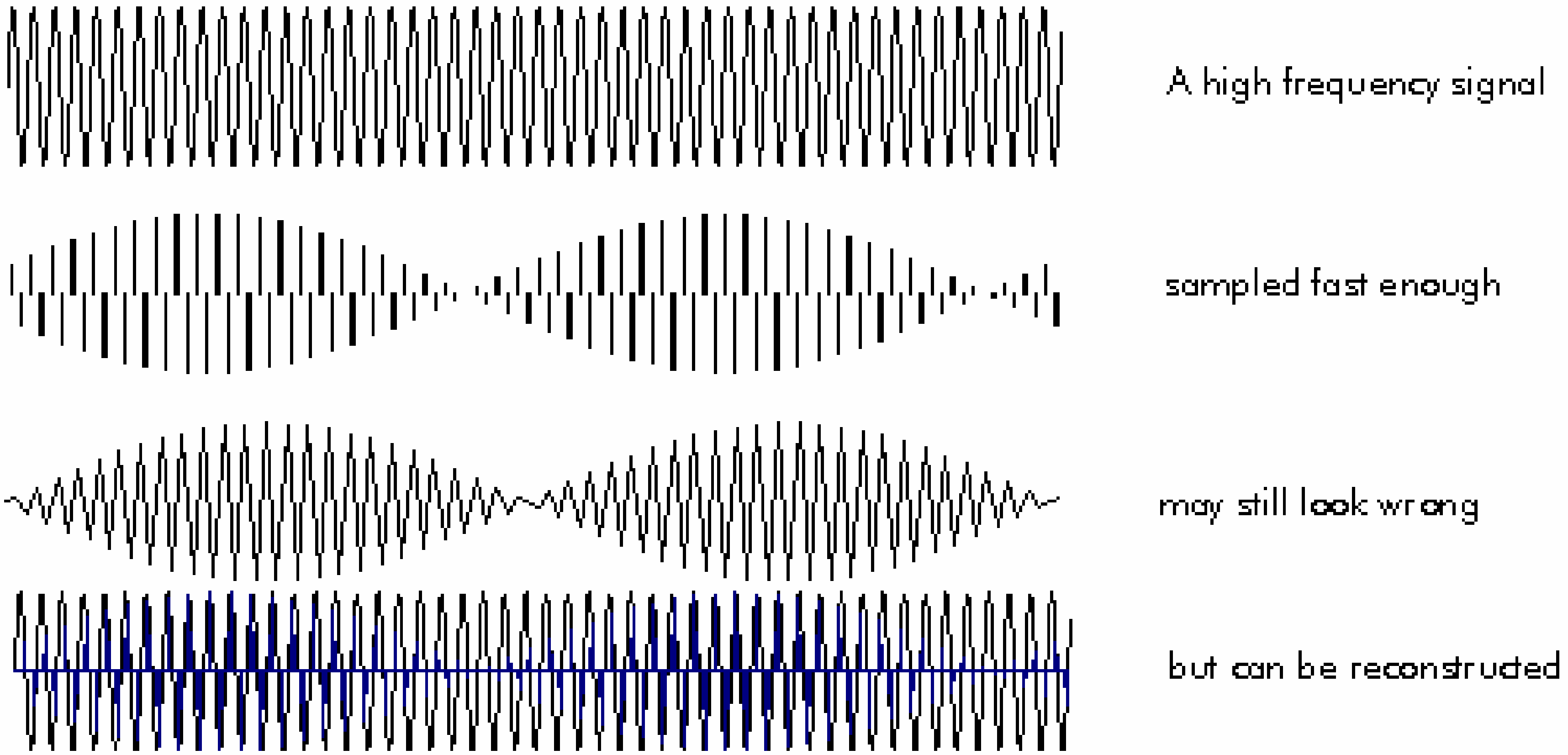
Aliasing effect & Nyquist Rate



Undersampling &Aliasing in Time Domain

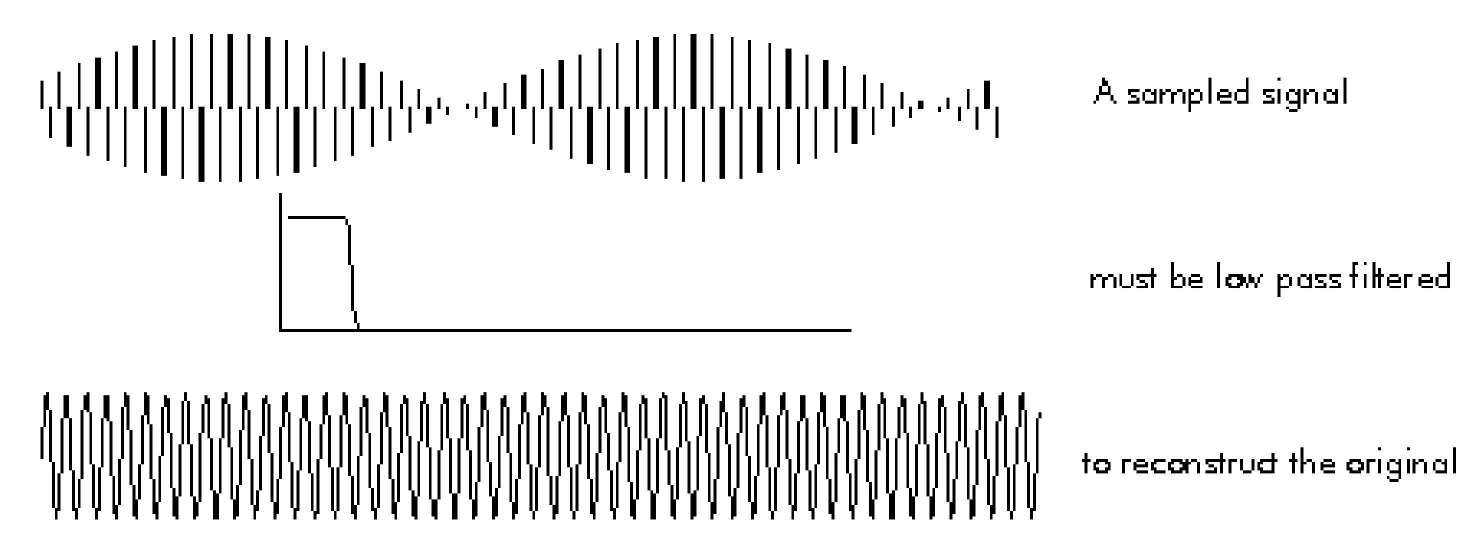


Nyquist Sampling & Reconstruction: Time Domain



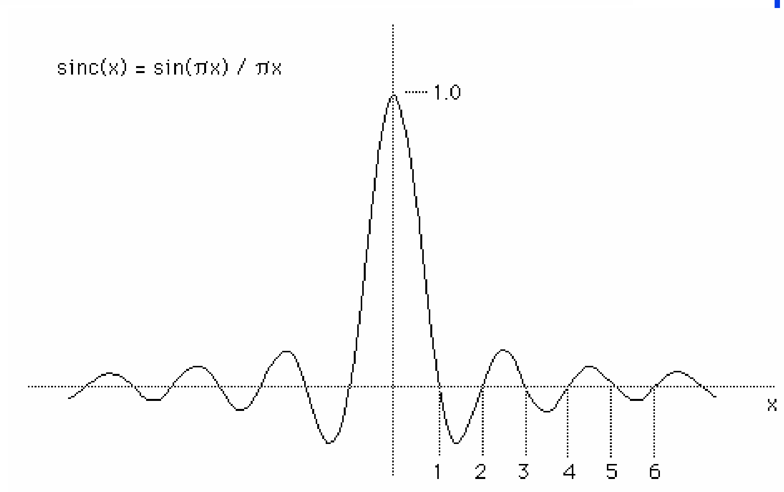
Note: correct reconstruction does not draw straight lines between samples.
Key: use sinc() pulses for reconstruction/interpolation

Nyquist Reconstruction: Frequency Domain



The impulse response of the reconstruction filter has a classic 'sin(x)/x shape.

The stimulus fed to this filter is the series of discrete impulses which are the samples.



Sampling theorem

Analog signal → **Sampling process** → Pulse amplitude modulated (**PAM**) signal

- Sampling theorem: A bandlimited signal with no spectral components beyond f_m, can be uniquely determined by values sampled at uniform intervals of $\boxed{T_s \leq \frac{1}{2f_m}}$

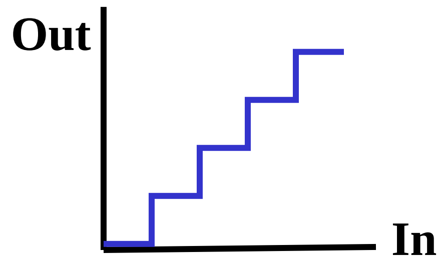
- The sampling rate, $f_s = \frac{1}{T_s} = 2f_m$

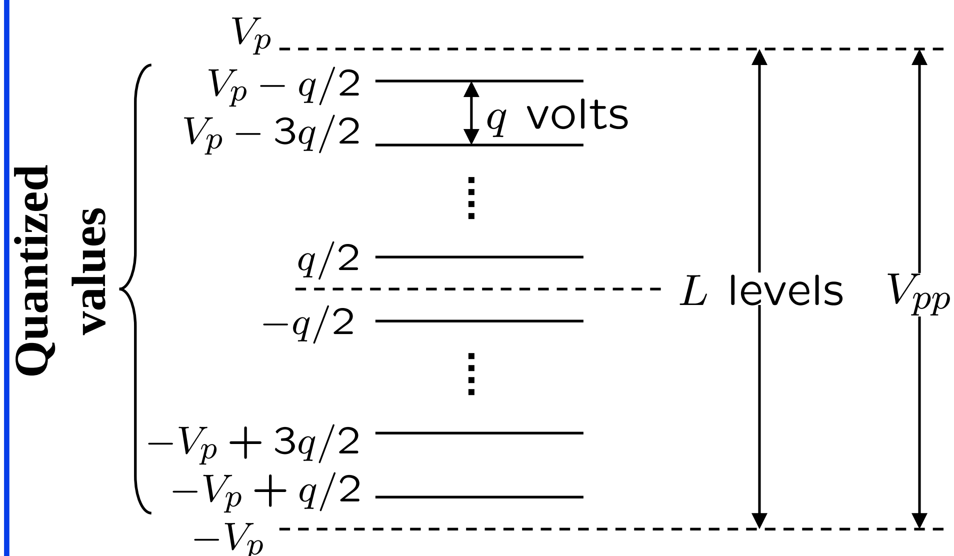
  is called Nyquist rate.

- In practice need to sample faster than this because the receiving filter will not be sharp.

Quantization

- Amplitude quantizing: Mapping samples of a continuous amplitude waveform to a finite set of amplitudes.





- Average quantization noise power

$$\sigma^2 = \frac{q^2}{12}$$

- Signal peak power

$$V_p^2 = \frac{L^2 q^2}{4}$$

- Signal power to average quantization noise power

$$\boxed{\left(\frac{S}{N}\right)_q = \frac{V_p^2}{\sigma^2} = 3L^2}$$

Encoding (PCM)

- A uniform linear quantizer is called Pulse Code Modulation (PCM).
- Pulse code modulation (PCM): Encoding the quantized signals into a digital word (PCM word or codeword).
  - Each quantized sample is digitally encoded into an l bits codeword where L in the number of quantization levels and

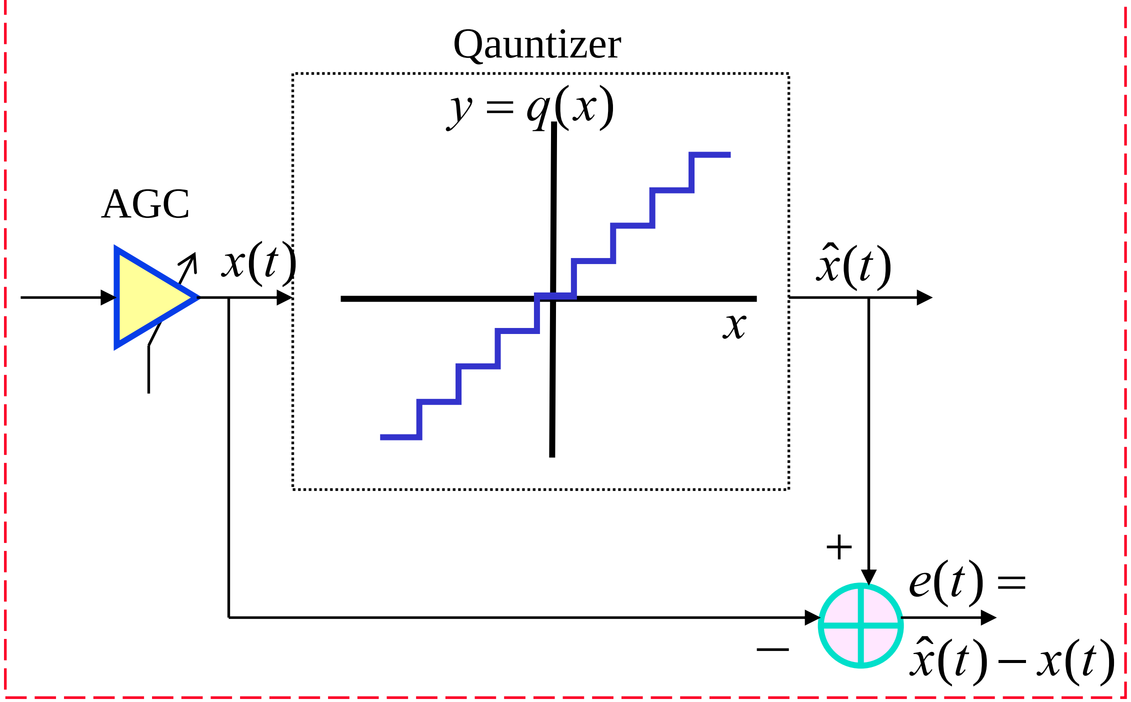
$$l = \log_2 L$$

Quantization error

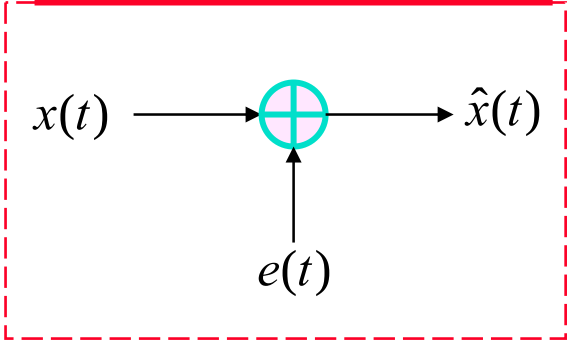
- Quantizing error: The difference between the input and output of a quantizer

$$\Longrightarrow e(t) = \hat{x}(t) - x(t)$$

Process of quantizing noise



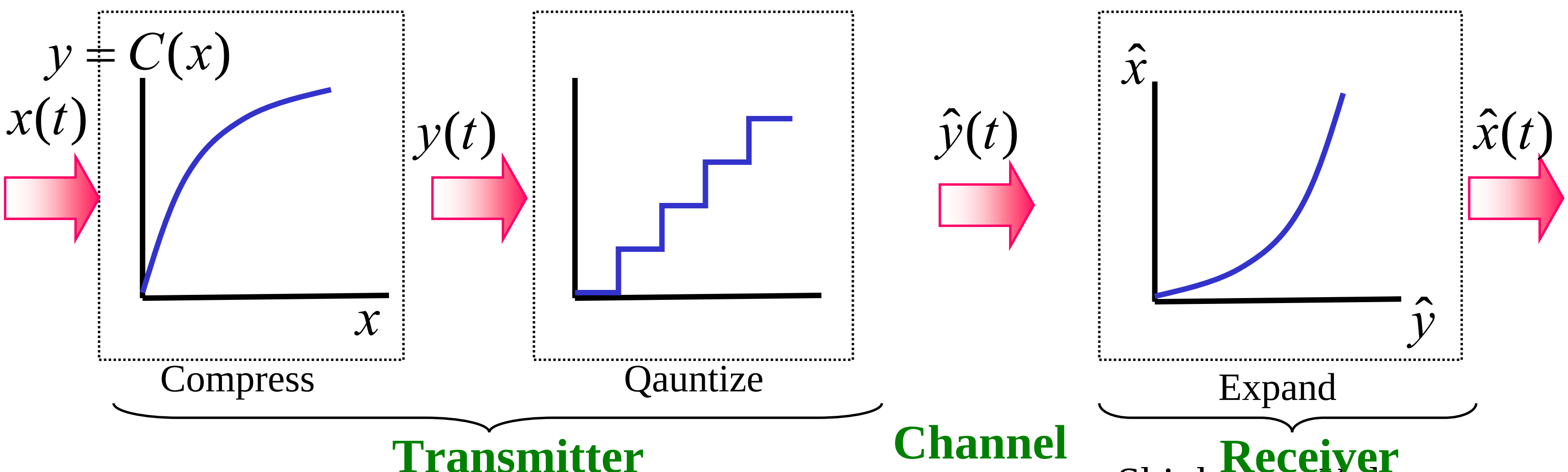
Model of quantizing noise



Non-uniform quantization

- It is done by uniformly quantizing the "compressed" signal.
- At the receiver, an inverse compression characteristic, called "expansion" is employed to avoid signal distortion.

compression+expansion ⇨ companding

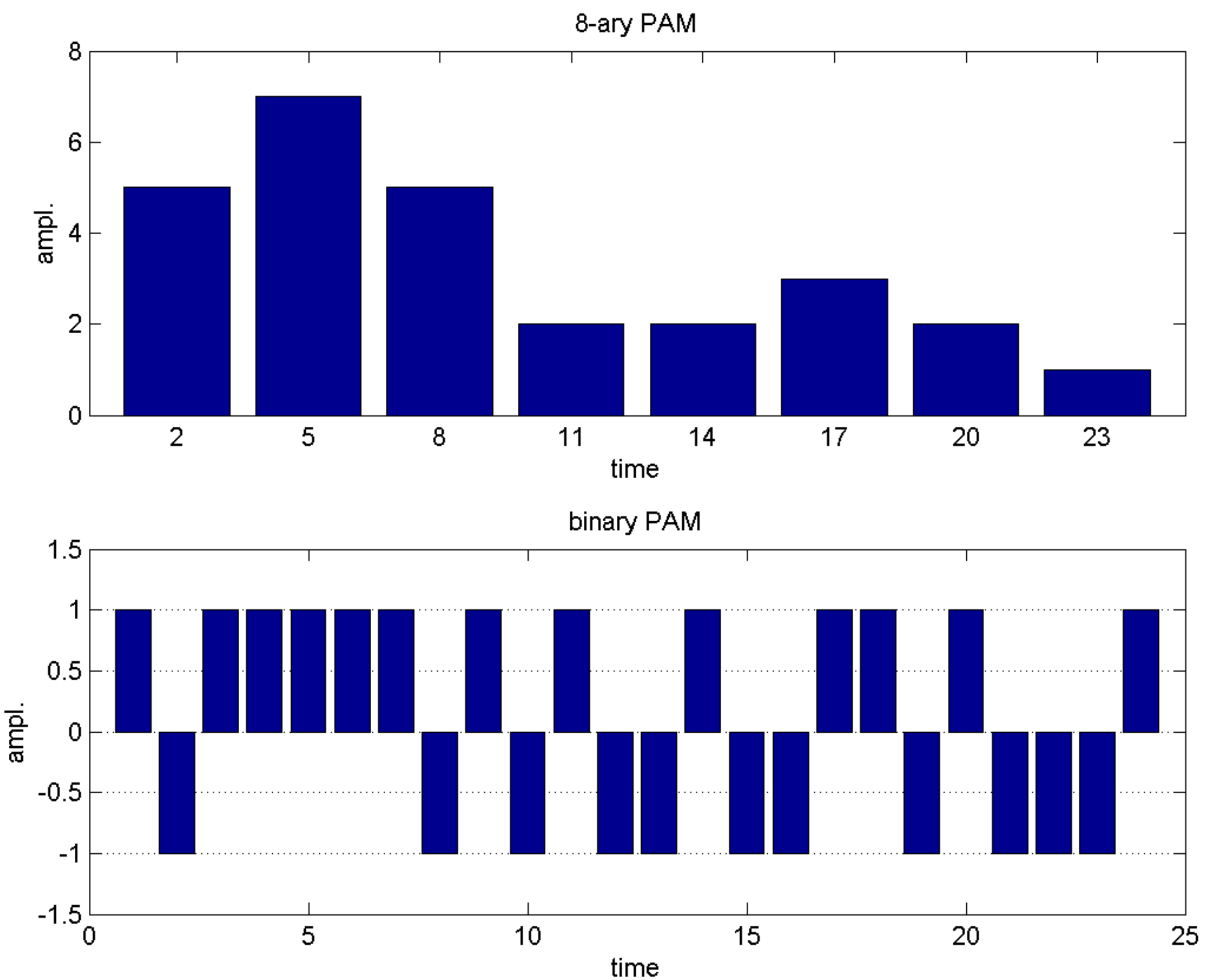


Baseband transmission

- To transmit information thru physical channels, PCM sequences (codewords) are transformed to pulses (waveforms).
  - Each waveform carries a symbol from a set of size M.
  - Each transmit symbol represents $\boxed{k = \log_2 M}$ bits of the PCM words.
  - PCM waveforms (line codes) are used for binary symbols (M=2).

- M-ary pulse modulation are used for non-binary symbols (M>2). Eg: M-ary PAM.
  - For a given data rate, M-ary PAM ($M>2$) requires less bandwidth than binary PCM.
  - For a given average pulse power, binary PCM is easier to detect than M-ary PAM ($M>2$).

PAM example: Binary vs 8-ary

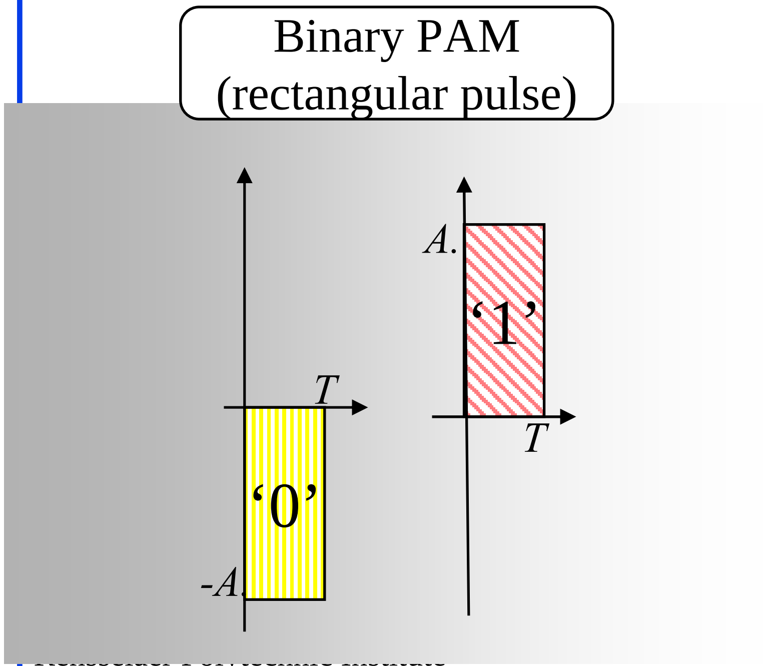


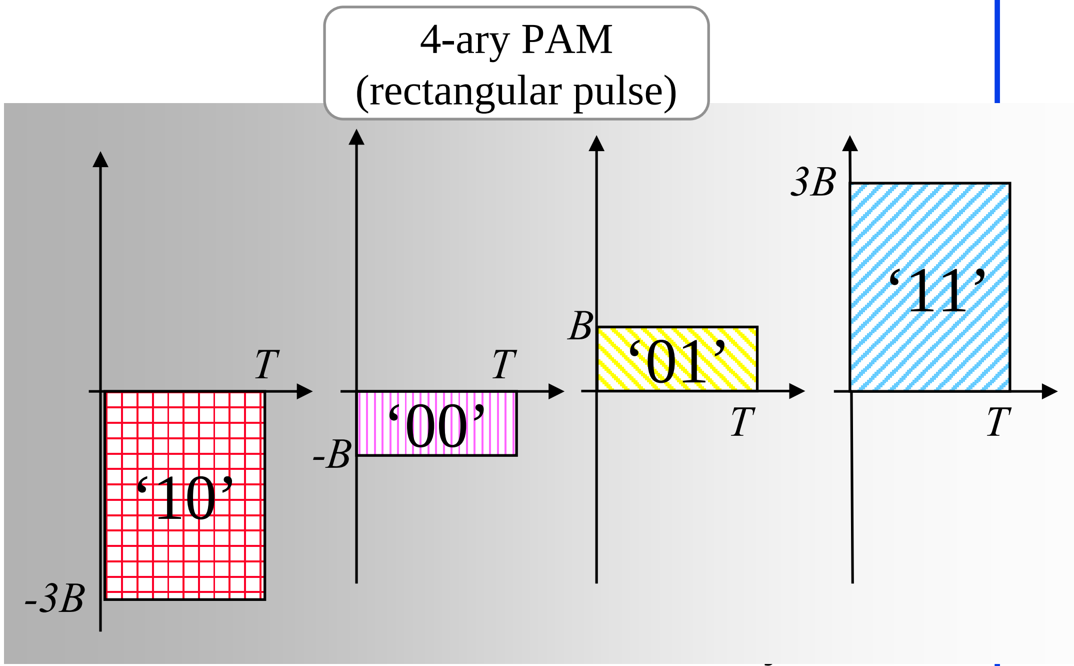
Example of M-ary PAM

Assuming real time Tx and <u>equal energy *per Tx data bit*</u> for binary-PAM and 4-ary PAM:

- 4-ary: $T=2T_b$ and Binary: $T=T_b$

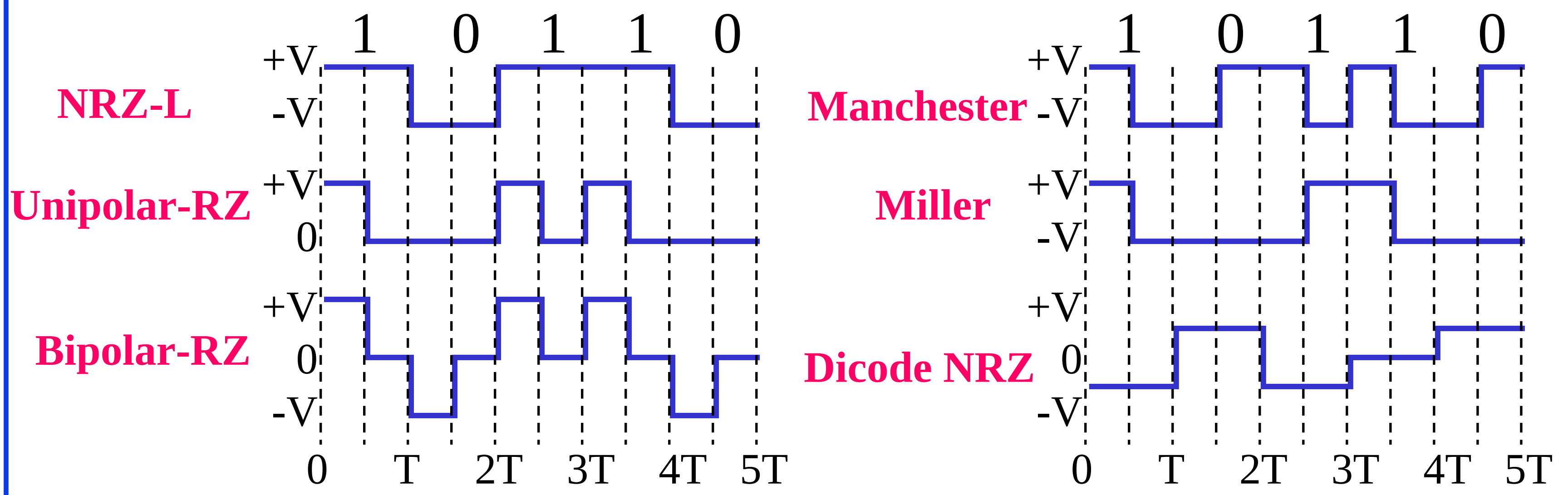
- Energy per symbol in binary-PAM: $A^2 = 10B^2$





Other PCM waveforms: Examples

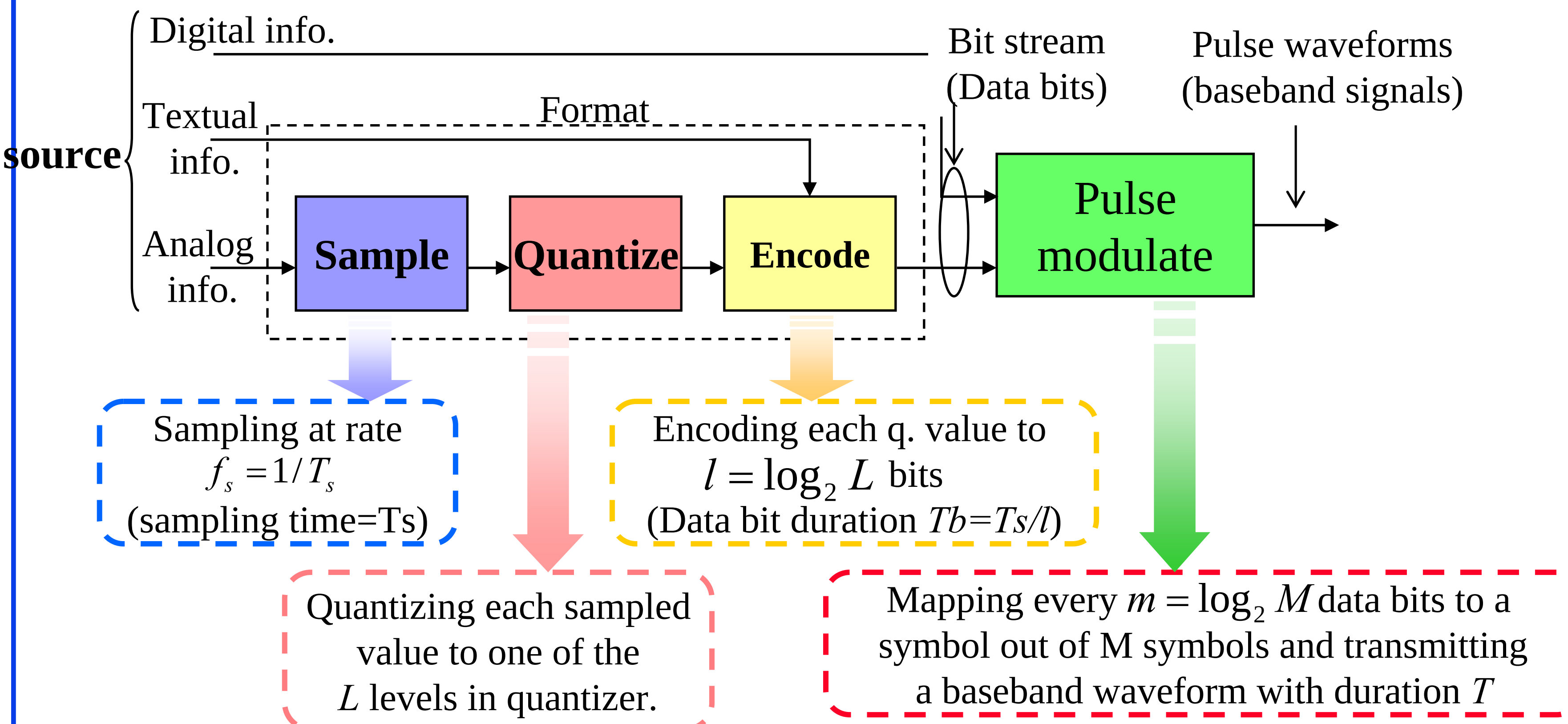
- PCM waveforms category:
  - Nonreturn-to-zero (NRZ)
  - Return-to-zero (RZ)
  - Phase encoded
  - Multilevel binary



PCM waveforms: Selection Criteria

- Criteria for comparing and selecting PCM waveforms:
  - Spectral characteristics (power spectral density and bandwidth efficiency)
  - Bit synchronization capability
  - Error detection capability
  - Interference and noise immunity
  - Implementation cost and complexity

Summary: Baseband Formatting and transmission

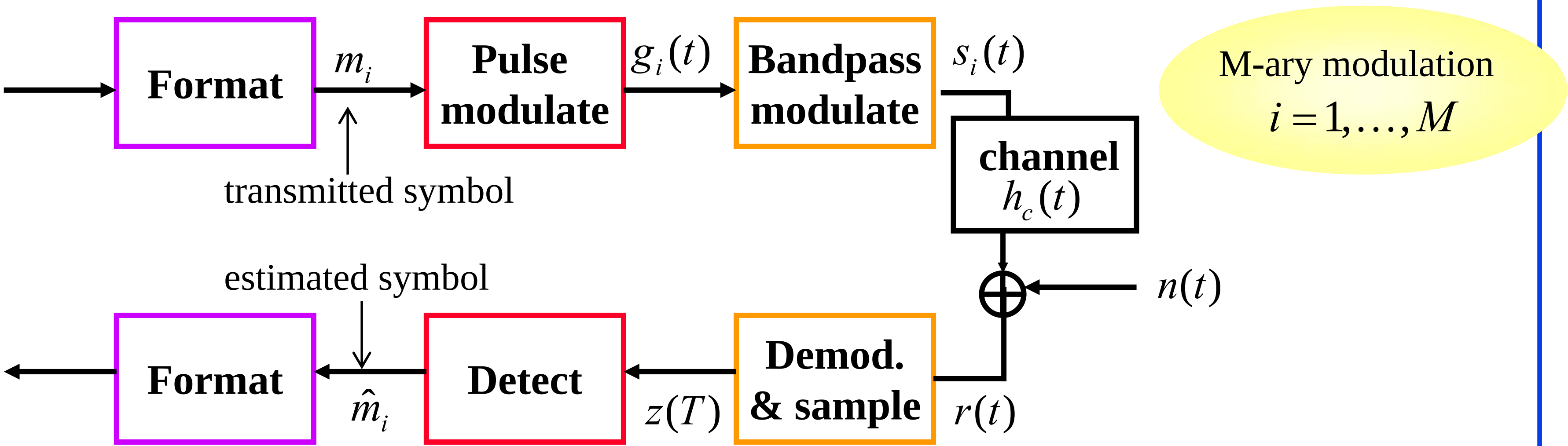


- Information (data- or bit-) rate: $R_b = 1/T_b$ [bits/sec]
- Symbol rate : $R = 1/T$ [symbols/sec]

$$\boxed{R_b = mR}$$

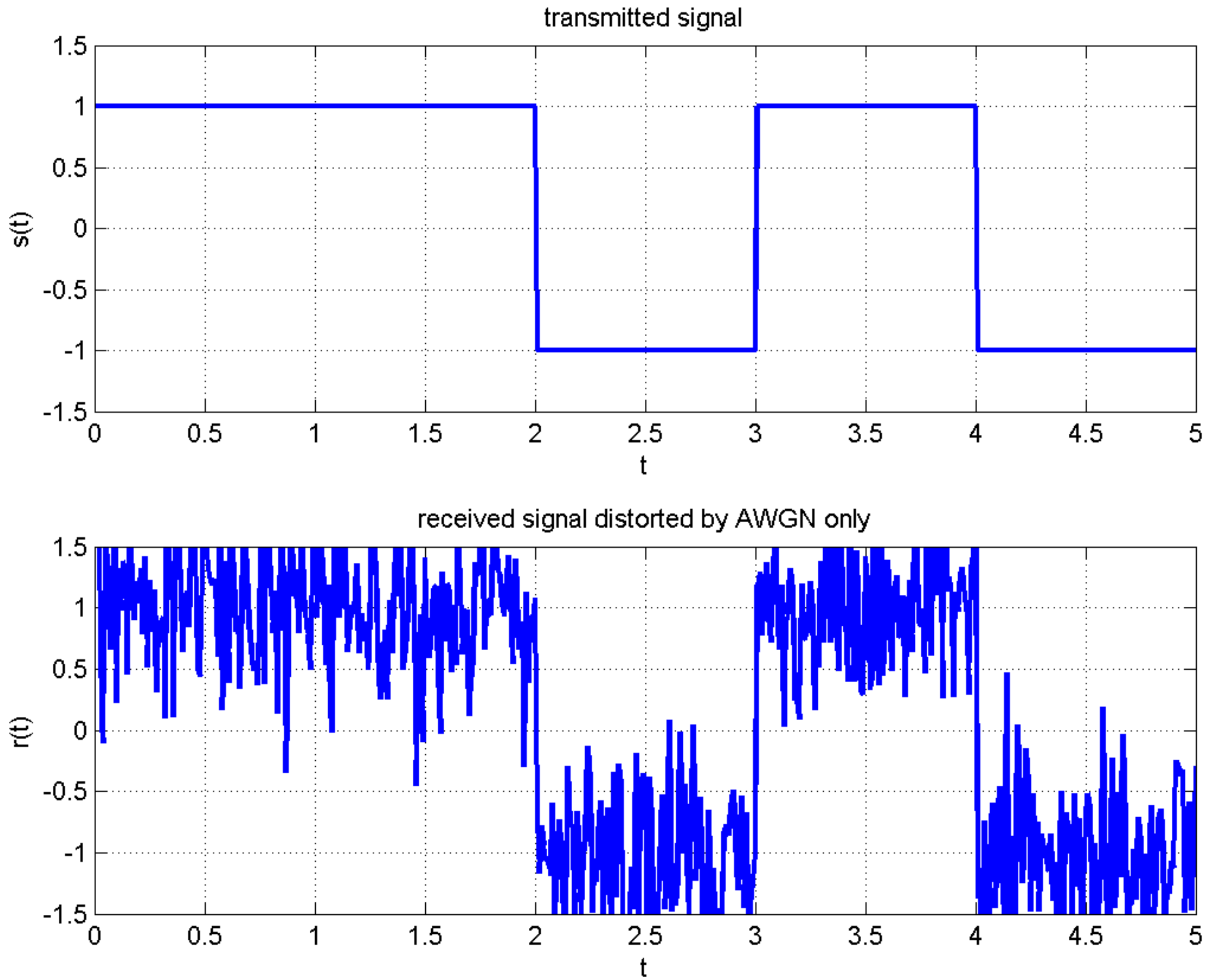
Receiver Structure & Matched Filter

Demodulation and detection

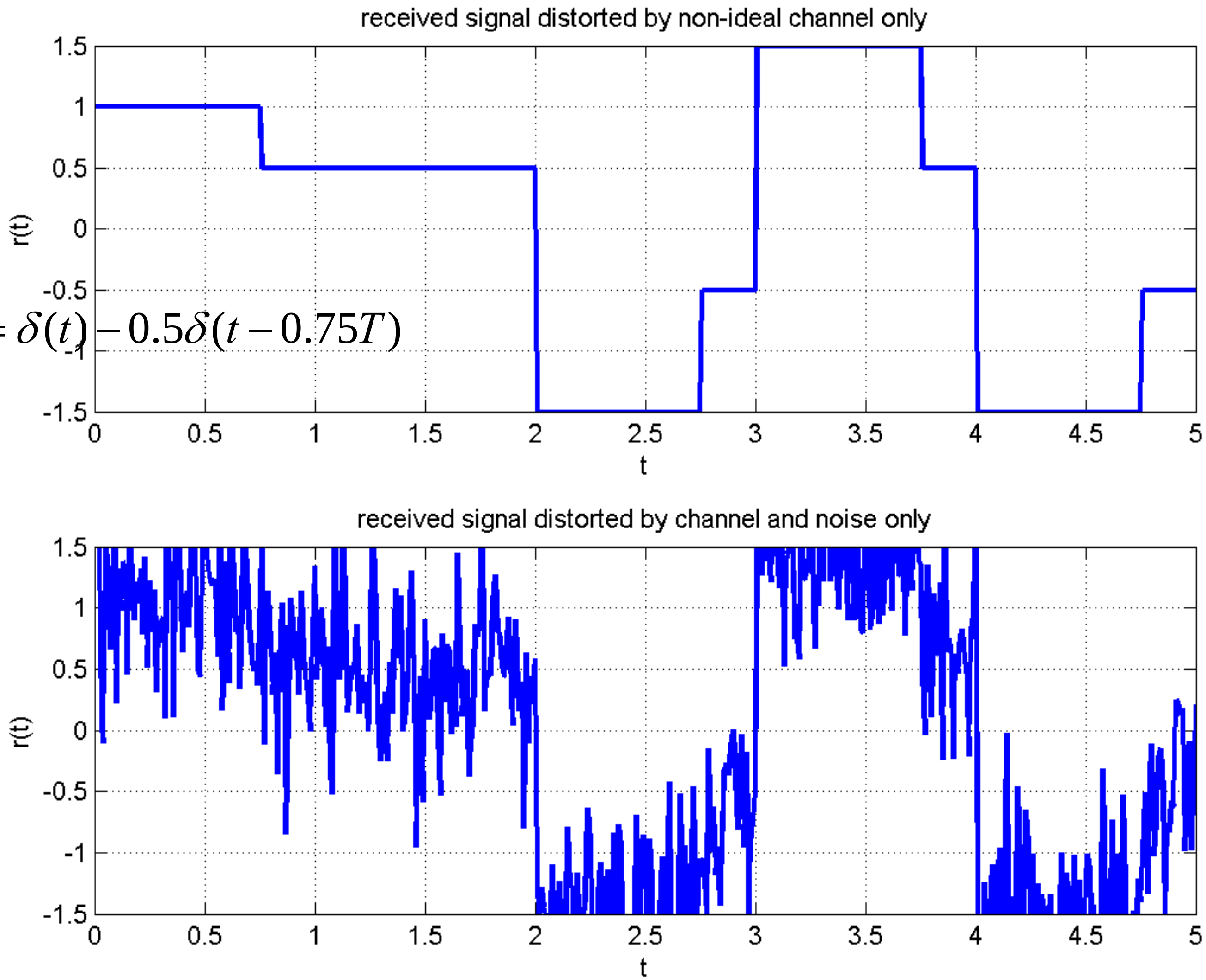


- Major sources of errors:
  - Thermal noise (AWGN)
    - disturbs the signal in an additive fashion (Additive)
    - has flat spectral density for all frequencies of interest (White)
    - is modeled by Gaussian random process (Gaussian Noise)
  - Inter-Symbol Interference (ISI)
    - Due to the filtering effect of transmitter, channel and receiver, symbols are "smeared".

Impact of AWGN



Impact of AWGN & Channel Distortion

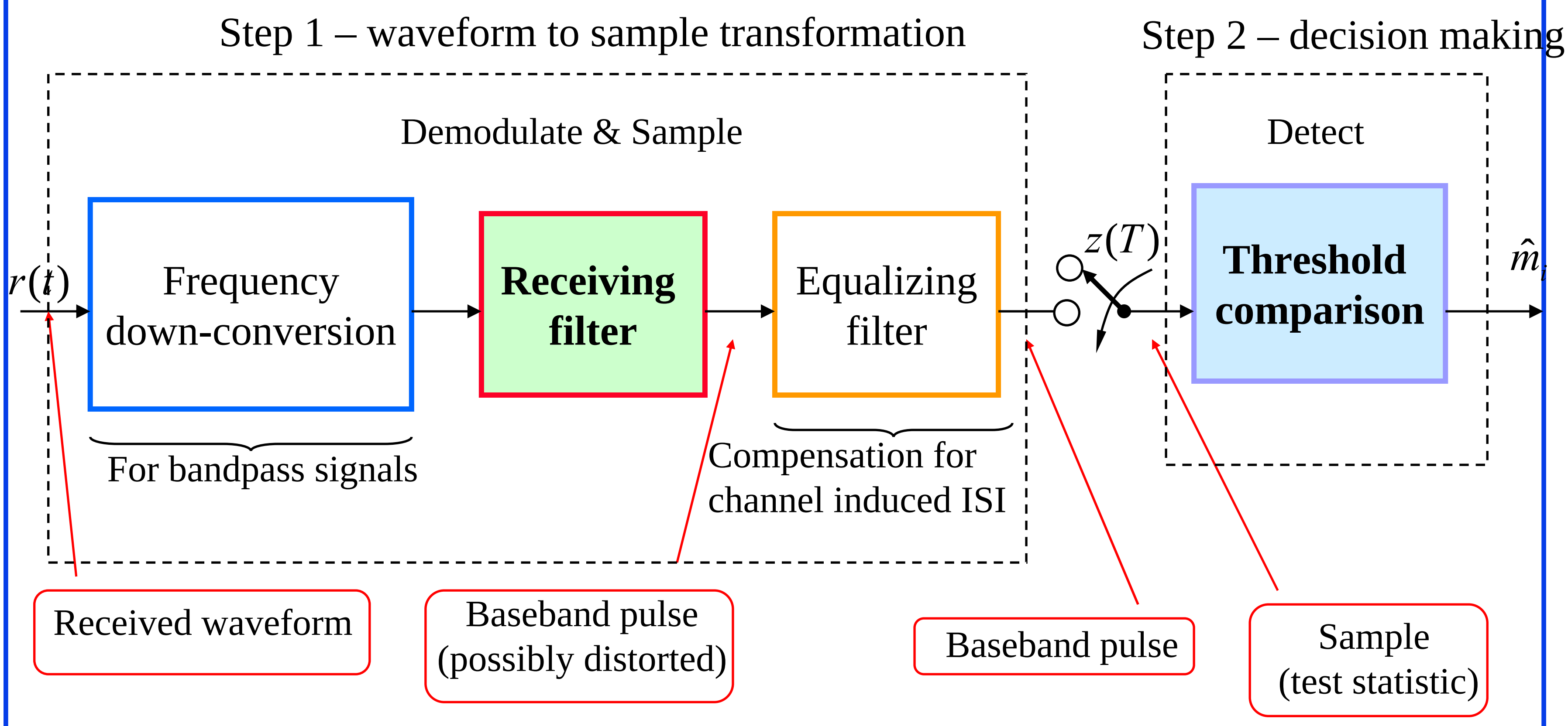


$$h_c(t) = \delta(t) - 0.5\delta(t - 0.75T)$$

Receiver job

- Demodulation and sampling:
  - Waveform recovery and preparing the received signal for detection:
    - Improving the signal power to the noise power (SNR) using matched filter (project to signal space)
    - Reducing ISI using equalizer (remove channel distortion)
    - Sampling the recovered waveform
- Detection:
  - Estimate the transmitted symbol based on the received sample

Receiver structure



Baseband vs Bandpass

- Bandpass model of detection process is equivalent to baseband model because:
  - The received bandpass waveform is first transformed to a baseband waveform.
- <u>Equivalence theorem:</u>
  - Performing bandpass linear signal processing followed by heterodying the signal to the baseband, …
  - … yields the same results as …
  - … heterodying the bandpass signal to the baseband , followed by a baseband linear signal processing.

Steps in designing the receiver

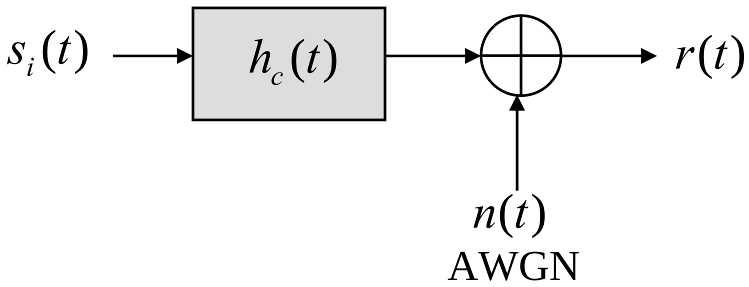
- Find optimum solution for receiver design with the following goals:
  1. Maximize SNR: *matched filter*
  2. Minimize ISI: *equalizer*

- Steps in design:
  - Model the received signal
  - Find separate solutions for each of the goals.

- First, we focus on designing a receiver which maximizes the SNR: *matched filter*

Receiver filter to maximize the SNR

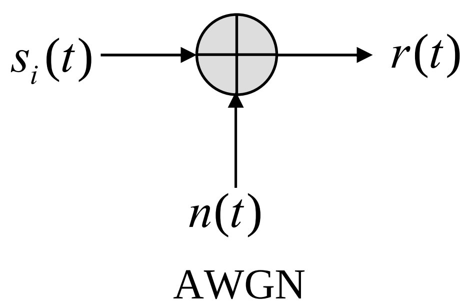
- Model the received signal



$$\boxed{r(t) = s_i(t) * h_c(t) + n(t)}$$

- Simplify the model (ideal channel assumption):
  - Received signal in AWGN

Ideal channels
$h_c(t) = \delta(t)$



$$\boxed{r(t) = s_i(t) + n(t)}$$

Matched Filter Receiver

- Problem:
  - Design the receiver filter $h(t)$ such that the SNR is maximized <u>at the sampling time</u> when $s_i(t), i = 1,\ldots,M$ is transmitted.

Correlator Receiver

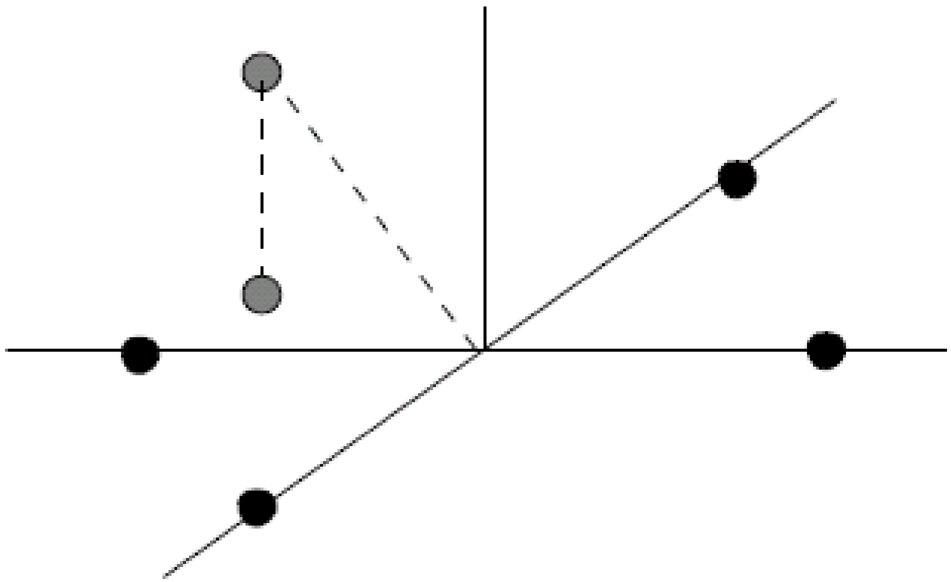
- The matched filter output at the sampling time, can be realized as the correlator output.
  - Matched filtering, i.e. convolution with $s_i^*(T-\tau)$ simplifies to integration w/ $s_i^*(\tau)$, i.e. correlation or inner product!

$$
\begin{aligned}
z(T) &= h_{opt}(T) * r(T) \\
&= \int_0^T r(\tau) s_i^*(\tau)\, d\tau = \langle r(t), s(t) \rangle
\end{aligned}
$$

Recall: correlation operation is the projection of the received signal onto the signal space!

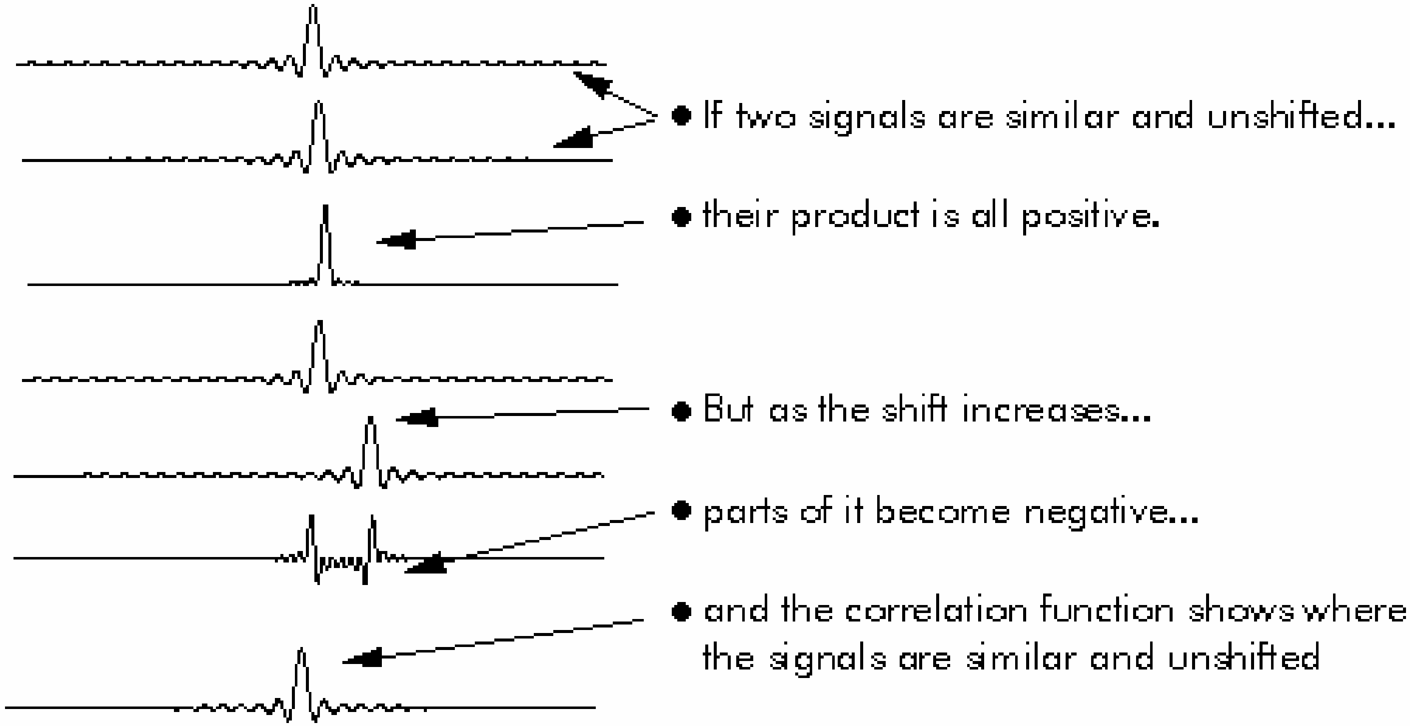
Key idea: Reject the noise (N) outside this space as irrelevant:
=> maximize S/N

Irrelevance Theorem: Noise Outside Signal Space



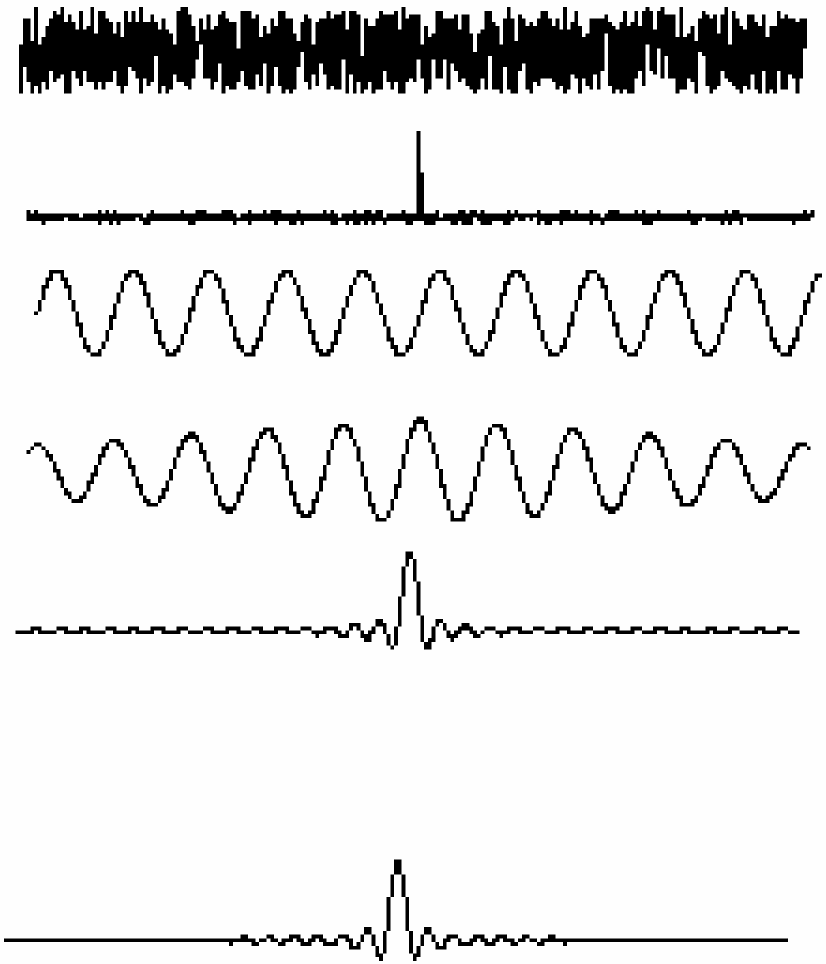
- Noise PSD is flat ("white") => total noise power infinite across the spectrum.
- We care only about the noise projected in the finite signal dimensions (eg: the bandwidth of interest).

Aside: Correlation Effect



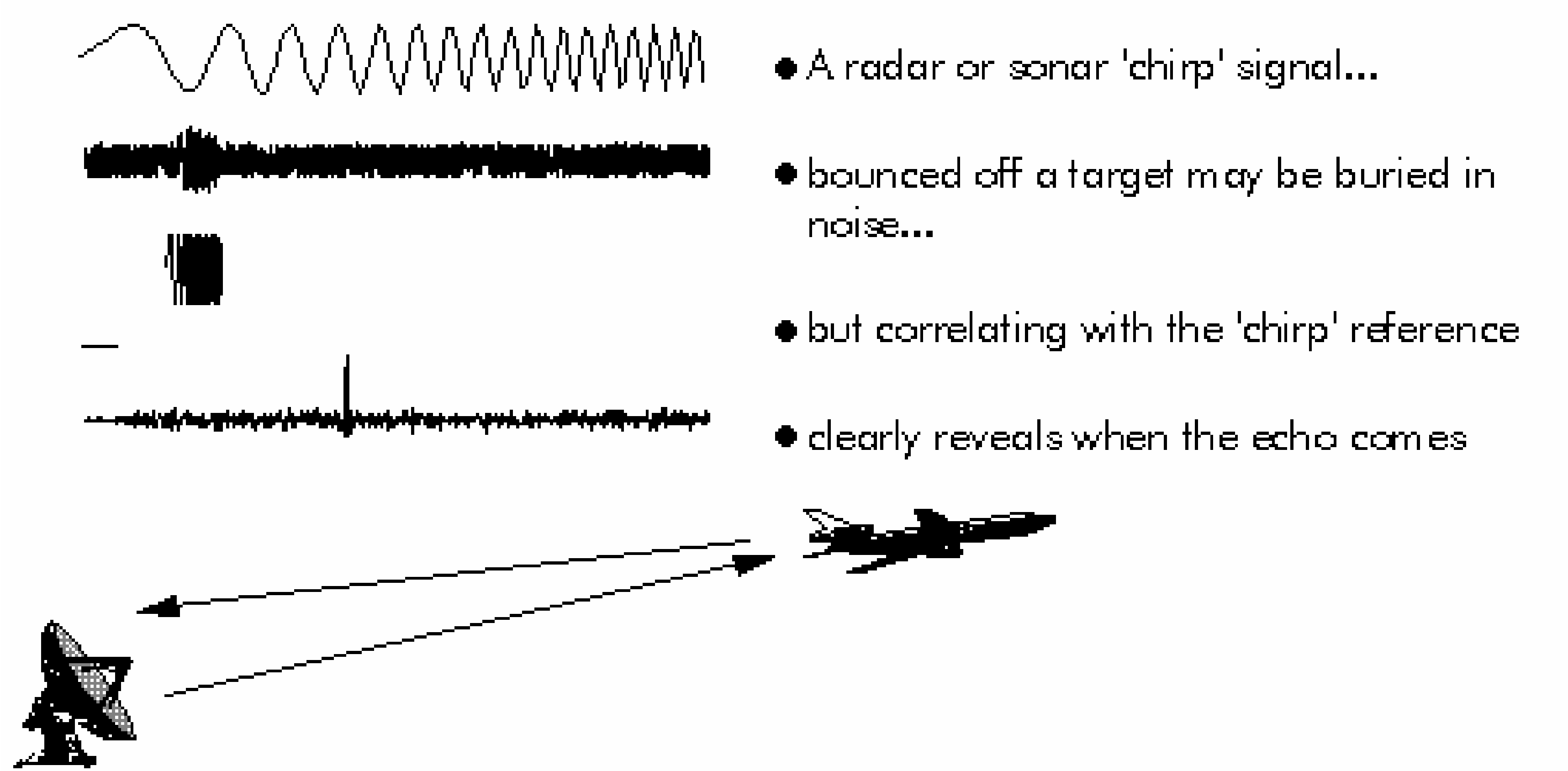
- Correlation is a maximum when two signals are similar in shape, and are in phase (or 'unshifted' with respect to each other).
- Correlation is a measure of the *similarity* between two signals as a function of *time shift (“lag”, τ)* between them
- When the two signals are similar in shape and unshifted with respect to each other, their product is all positive. This is like *constructive interference*,
- The *breadth* of the correlation function - where it has significant value - shows for *how long* the signals remain similar.

Aside: Autocorrelation

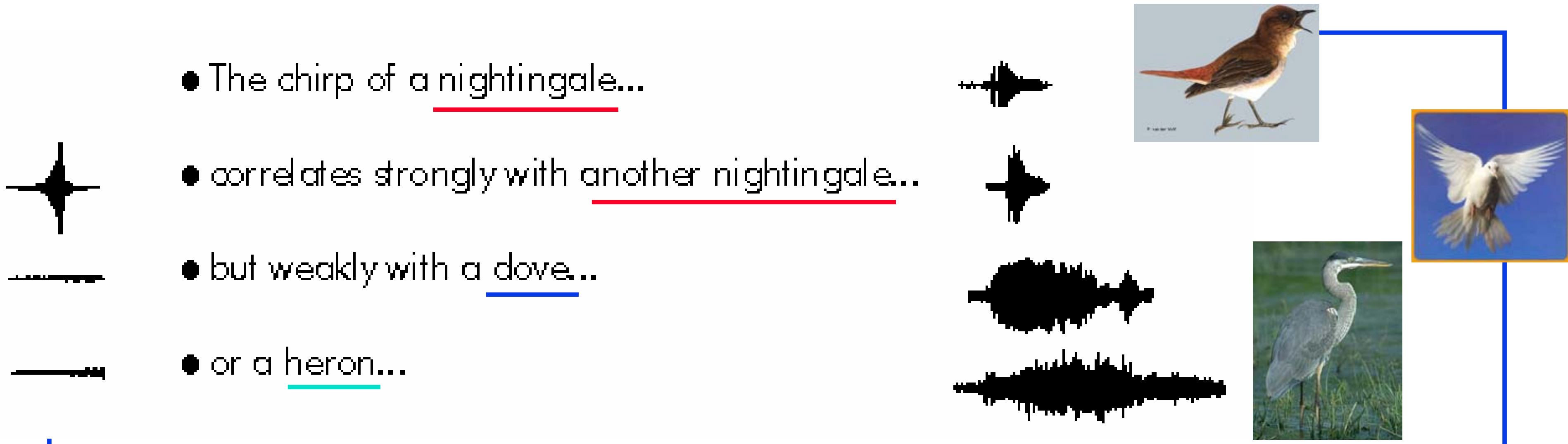


- Random noise is similar to itself, and in phase, only with no time shift at all
- so its correlation function is a spike
- periodic signals go in and out of phase as they are time shifted
- so their correlation functions are periodic
- signals that last only a short while are only similar while they last
- so their correlation functions are short

Aside: Cross-Correlation &Radar



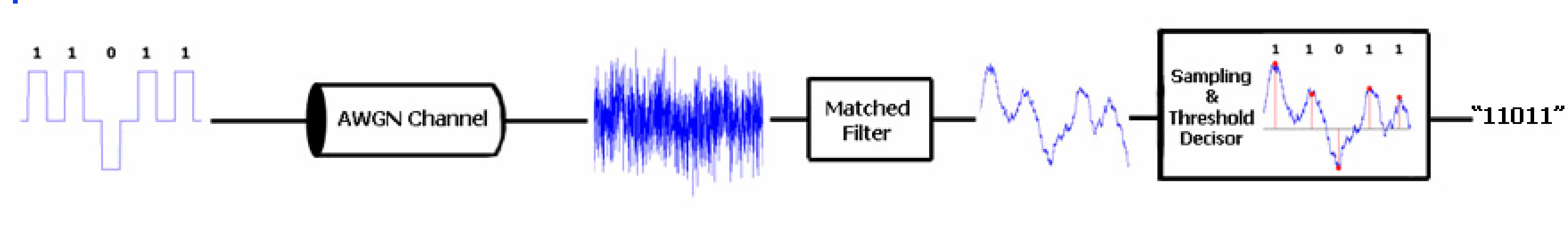
- Figure: shows how the signal can be located within the noise.
- A copy of the known reference signal is correlated with the unknown signal.
- The correlation will be high when the reference is similar to the unknown signal.
- A large value for correlation shows the degree of confidence that the reference signal is detected.
- The large value of the correlation indicates when the reference signal occurs.



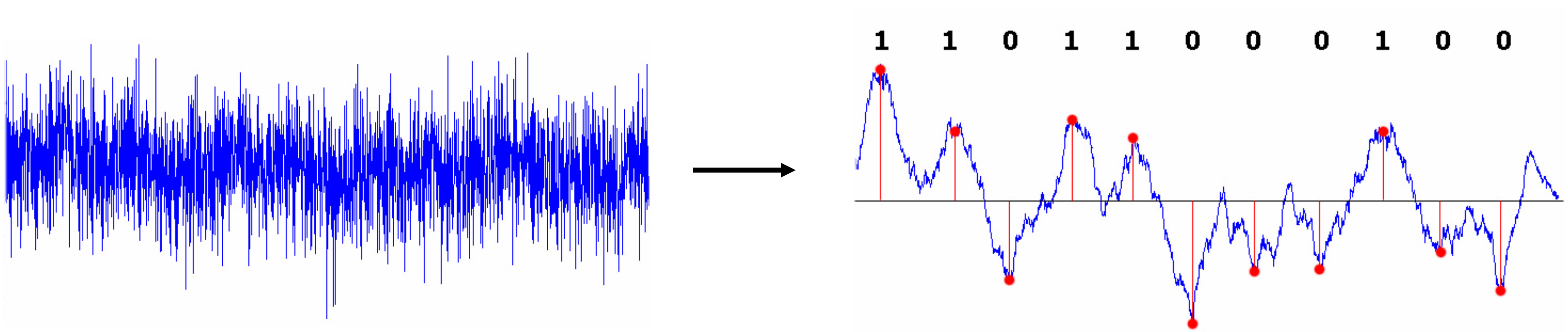
- A copy of a known <u>reference signal</u> is correlated with the unknown signal.
- The <u>correlation will be high if the reference is similar to the unknown signal</u>.
- The unknown signal is correlated with a <u>number of known reference functions</u>.
- A large value for correlation shows the <u>degree of similarity</u> to the reference.
- The largest value for correlation is the <u>most likely match</u>.

- Same principle in communications: reference signals corresponding to symbols
- The ideal communications channel may have attenuated, phase shifted the reference signal, and added noise

Matched Filter: back to cartoon…



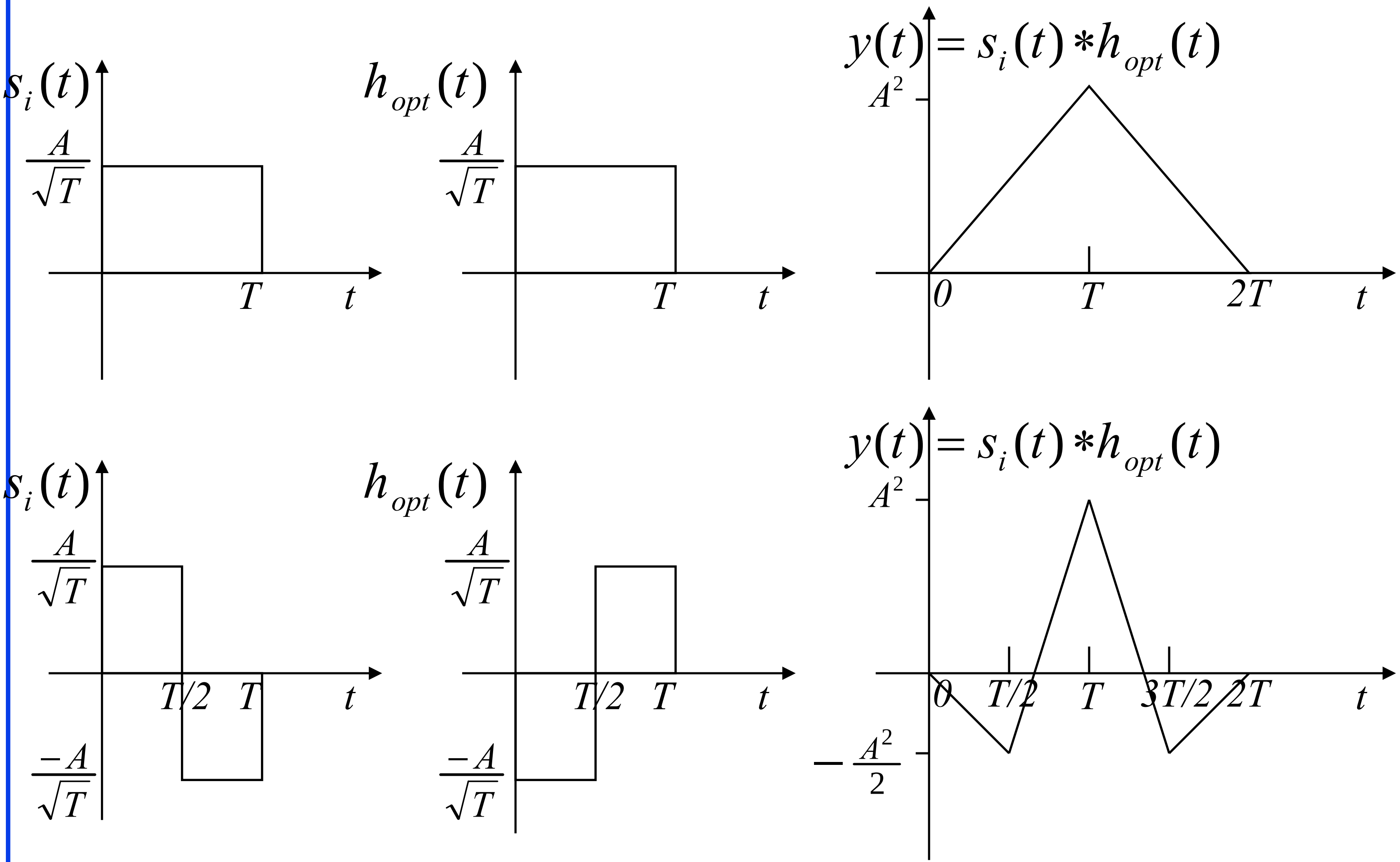
- Consider the received signal as a vector **r,** and the transmitted signal vector as **s**
- Matched filter "projects" the **r** onto signal space spanned by **s** ("matches" it)



Filtered signal can now be safely sampled by the receiver at the correct sampling instants, resulting in a correct interpretation of the binary message

Matched filter is the filter that <u>maximizes the signal-to-noise ratio</u> it can be shown that it also <u>minimizes the BER: *it is a simple projection operation*</u>

Example of matched filter (real signals)



Properties of the Matched Filter

1. The Fourier transform of a matched filter output with the matched signal as input is, except for a time delay factor, proportional to the ESD of the input signal.

$$\boxed{Z(f) = |S(f)|^2 \exp(-j2\pi fT)}$$

2. The output signal of a matched filter is proportional to a shifted version of the autocorrelation function of the input signal to which the filter is matched.

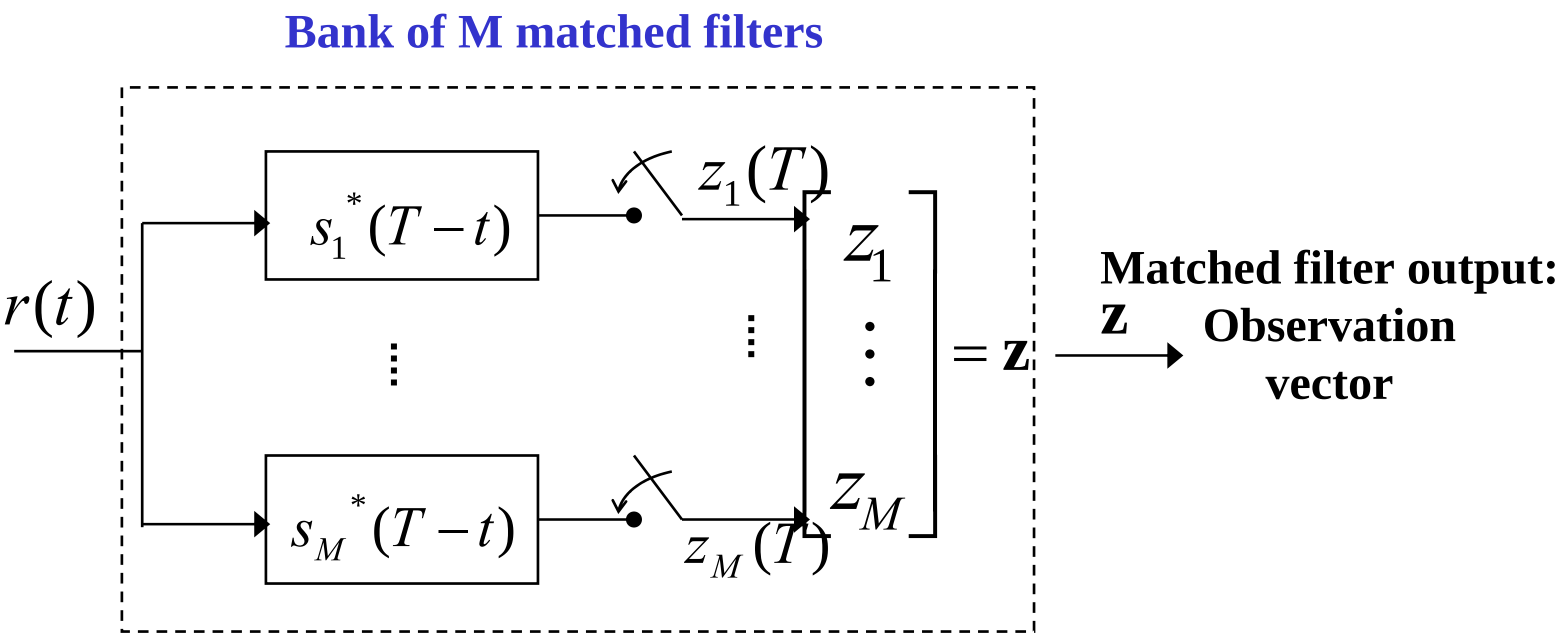
$$\boxed{z(t) = R_s(t-T) \Rightarrow z(T) = R_s(0) = E_s}$$

3. The output SNR of a matched filter depends only on the ratio of the signal energy to the PSD of the white noise at the filter input.

$$\boxed{\max\left(\frac{S}{N}\right)_T = \frac{E_s}{N_0/2}}$$

4. Two matching conditions in the matched-filtering operation:
   - spectral phase matching that gives the desired output peak at time *T*.
   - spectral amplitude matching that gives optimum SNR to the peak value.

Implementation of matched filter receiver



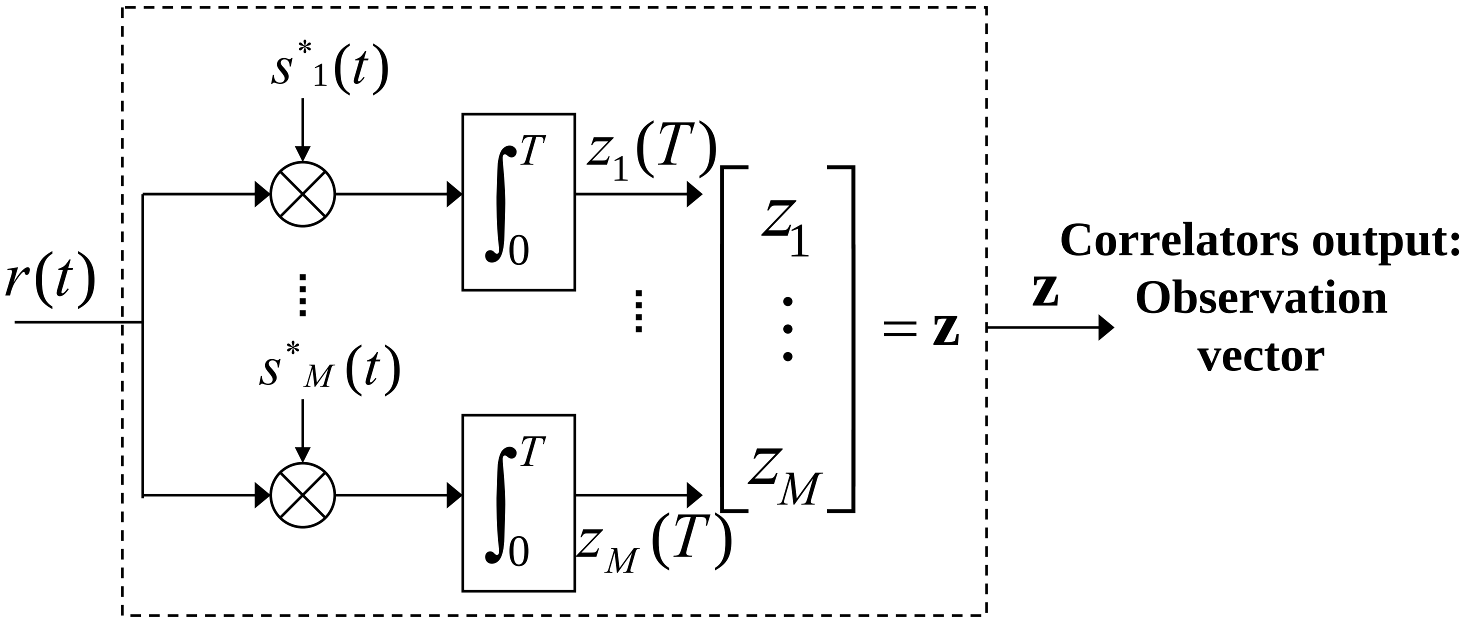
$$z_i = r(t) * s^*_i(T-t) \quad i = 1, \ldots, M$$

$$\mathbf{z} = (z_1(T), z_2(T), \ldots, z_M(T)) = (z_1, z_2, \ldots, z_M)$$

Note: we are projecting along the basis directions of the signal space

Implementation of correlator receiver

Bank of M correlators

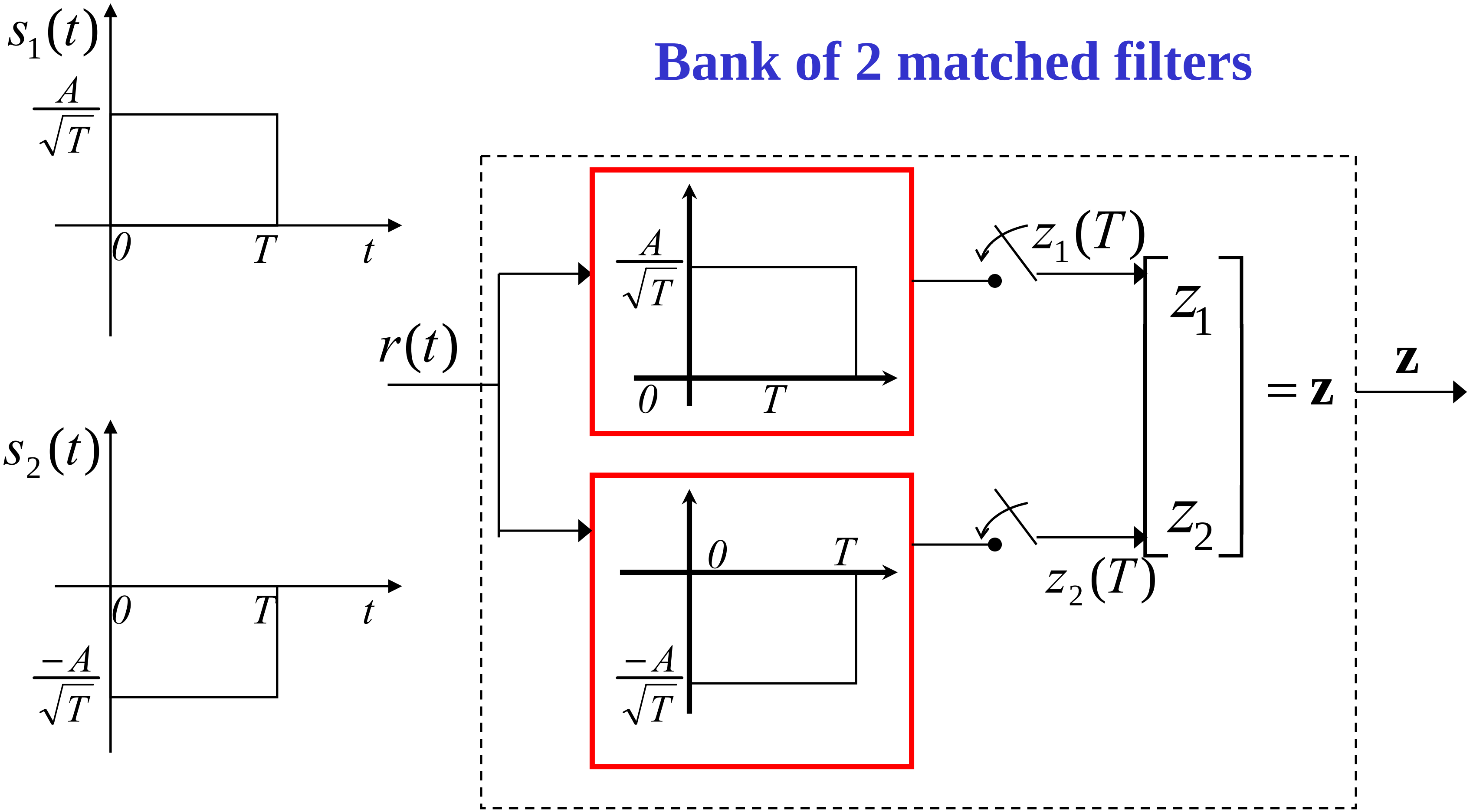


$$\mathbf{z} = (z_1(T), z_2(T), \ldots, z_M(T)) = (z_1, z_2, \ldots, z_M)$$

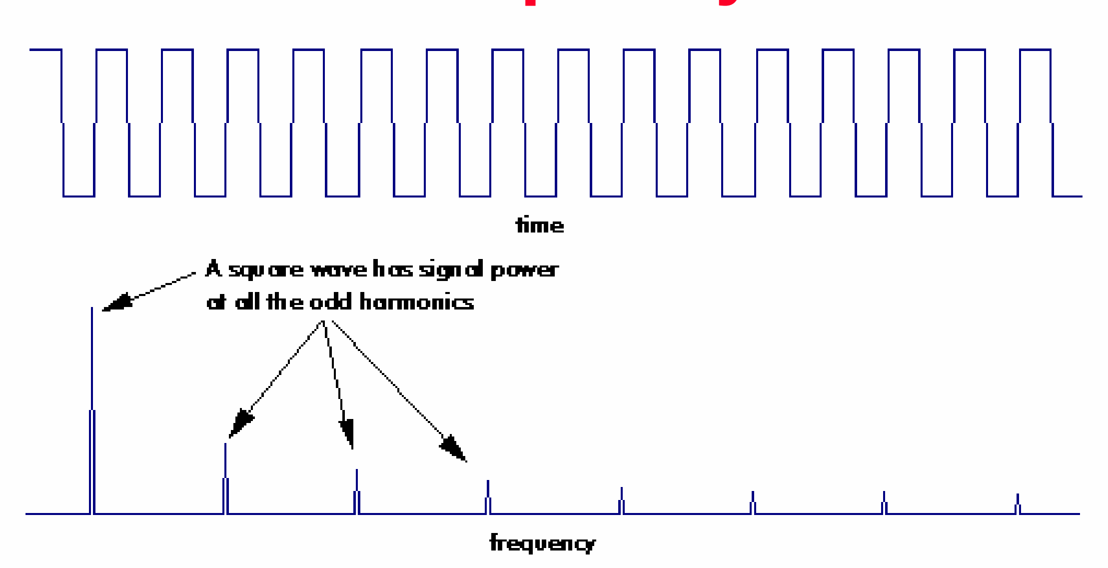
$$z_i = \int_0^T r(t) s_i(t) dt \qquad i = 1, \ldots, M$$

Note: In <u>previous</u> slide we "filter" i.e. convolute in the boxes shown.

Implementation example of matched filter receivers

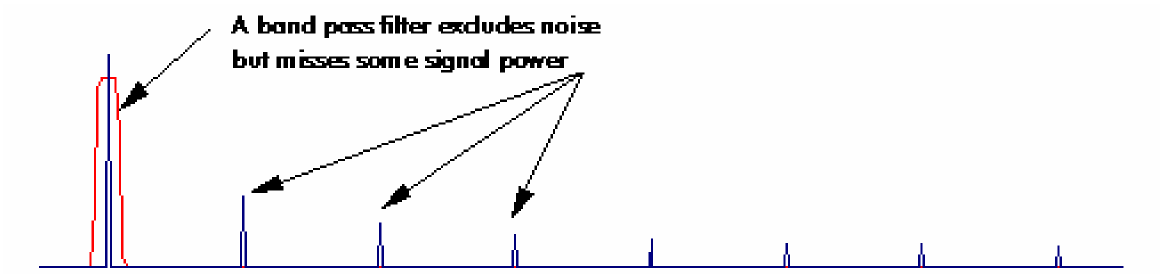


Matched Filter: Frequency domain View



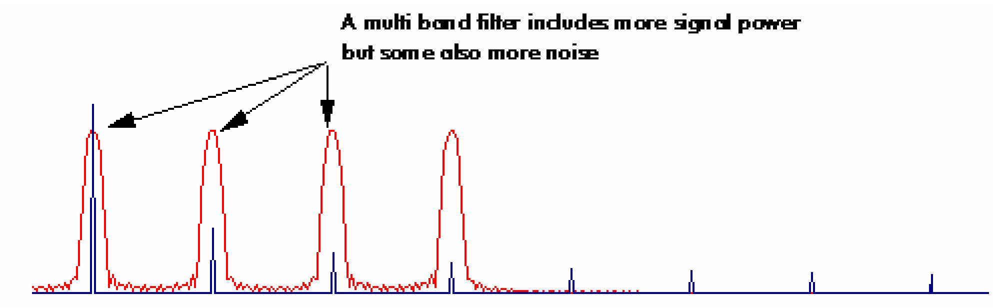
Simple Bandpass Filter:

excludes noise, but misses some signal power

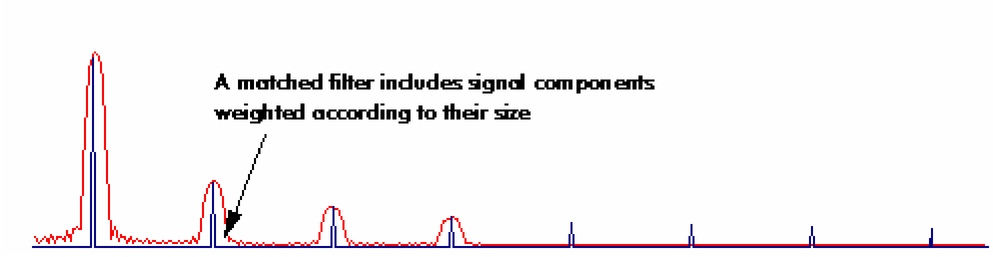


Matched Filter: Frequency Domain View (Contd)

Multi-Bandpass Filter: includes more signal power, but adds more noise also!

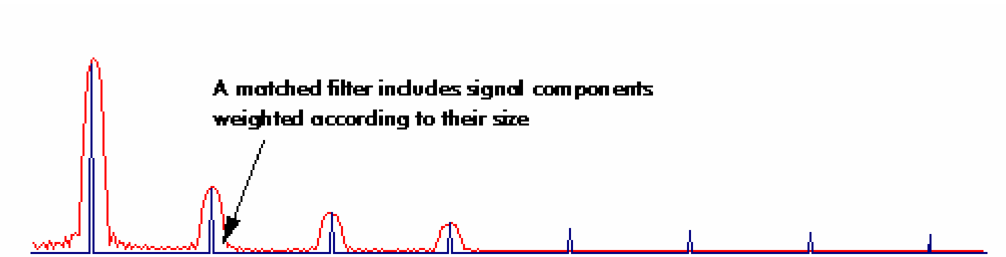


Matched Filter: includes more signal power, weighted according to size
=> maximal noise rejection!

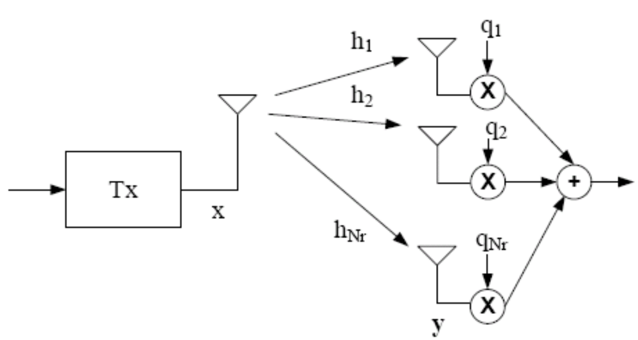


Maximal Ratio Combining (MRC) viewpoint

- Generalization of this f-domain picture, for combining multi-tap signal



Weight each branch with a complex factor $q_i = |q_i|e^{j\phi_i}$ and then adding up the N_r branches



$$y(t) = x(t) \sum_{i=1}^{N_r} |q_i||h_i| \exp\{j(\phi_i + \theta_i)\}.$$

phase of the combining coefficient $\phi_i = -\theta_i$

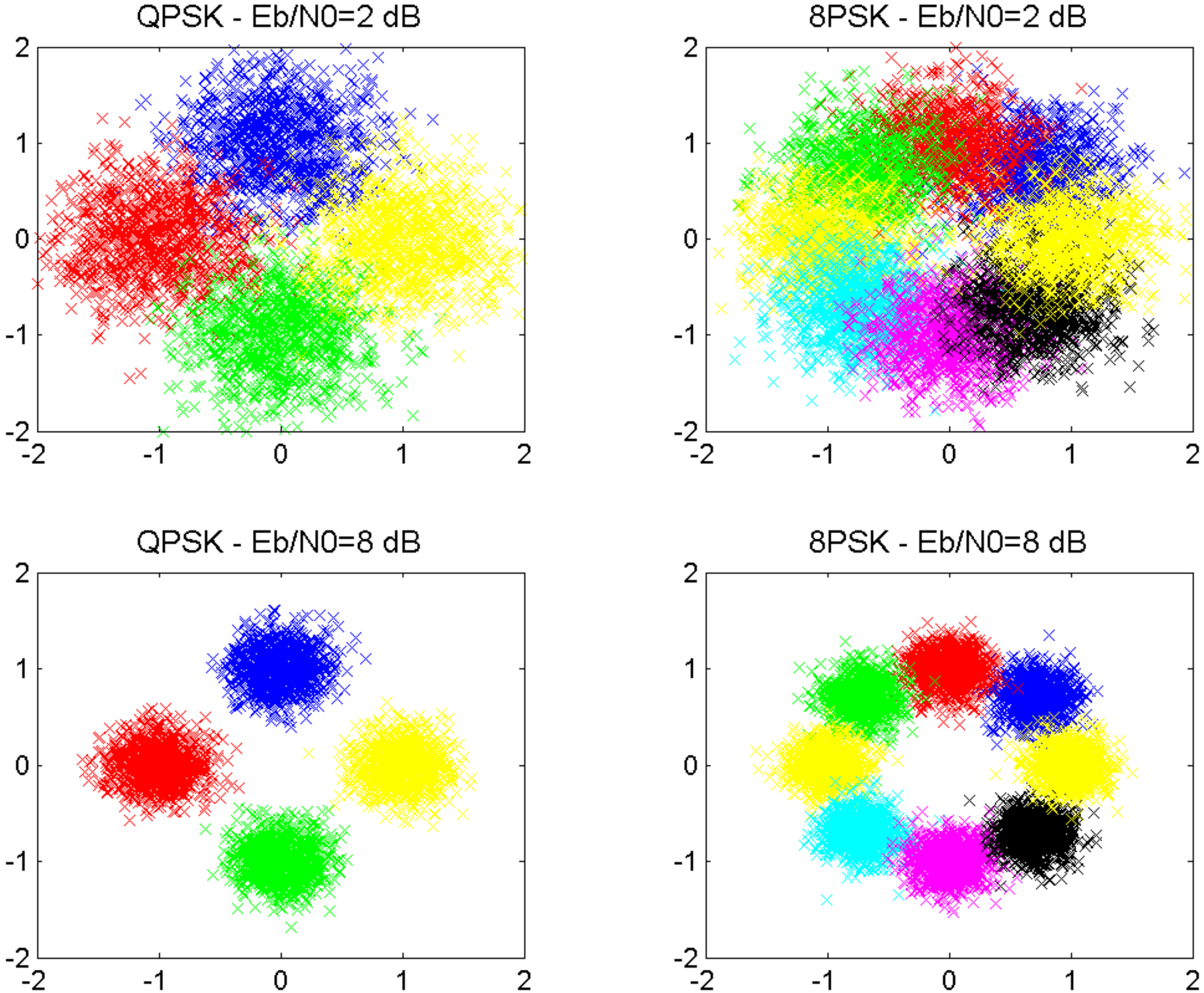
maximizing combining values as $|q_i^*|^2 = |h_i|^2/\sigma^2$

each branch is multiplied by its SNR

SNR:

$$\boxed{\gamma_{MRC} = \frac{\mathcal{E}_x \sum_{i=1}^{N_r} |h_i|^2}{\sigma^2} = \sum_{i=1}^{N_r} \gamma_i.}$$

Examples of matched filter output for bandpass modulation schemes

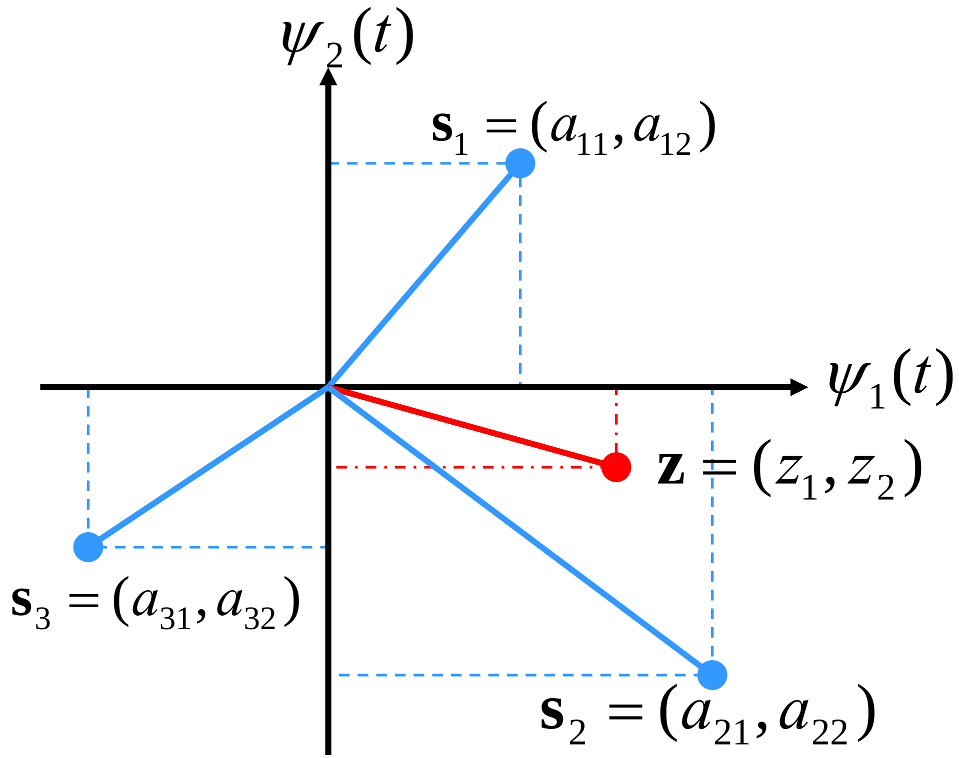


Signal Space Concepts

Signal space: Overview

- What is a signal space?
  - Vector representations of signals in an N-dimensional orthogonal space
- Why do we need a signal space?
  - It is a means to convert signals to vectors and vice versa.
  - It is a means to calculate signals energy and Euclidean distances between signals.
- Why are we interested in Euclidean distances between signals?
  - For detection purposes: The received signal is transformed to a received vectors.
  - The signal which has the minimum distance to the received signal is estimated as the transmitted signal.

Schematic example of a signal space



Transmitted signal alternatives

$$s_1(t) = a_{11}\psi_1(t) + a_{12}\psi_2(t) \Leftrightarrow \mathbf{s}_1 = (a_{11}, a_{12})$$
$$s_2(t) = a_{21}\psi_1(t) + a_{22}\psi_2(t) \Leftrightarrow \mathbf{s}_2 = (a_{21}, a_{22})$$
$$s_3(t) = a_{31}\psi_1(t) + a_{32}\psi_2(t) \Leftrightarrow \mathbf{s}_3 = (a_{31}, a_{32})$$

Received signal at matched filter output

$$z(t) = z_1\psi_1(t) + z_2\psi_2(t) \Leftrightarrow \mathbf{z} = (z_1, z_2)$$

Signal space

- To form a signal space, first we need to know the <u>inner product</u> between two signals (functions):
  - Inner (scalar) product:

$$<x(t), y(t)>= \int_{-\infty}^{\infty} x(t)y^*(t)dt$$

= cross-correlation between x(t) and y(t)

  - Properties of inner product:

$$<ax(t), y(t)>= a<x(t), y(t)>$$

$$<x(t), ay(t)>= a^*<x(t), y(t)>$$

$$<x(t)+y(t), z(t)>=<x(t), z(t)>+<y(t), z(t)>$$

Signal space …

- The distance in signal space is measure by calculating the norm.
- What is **norm**?
  - Norm of a signal:

$$\left\|x(t)\right\| = \sqrt{< x(t), x(t) >} = \sqrt{\int_{-\infty}^{\infty} \left|x(t)\right|^2 dt} = \sqrt{E_x}$$

= "length" of x(t)

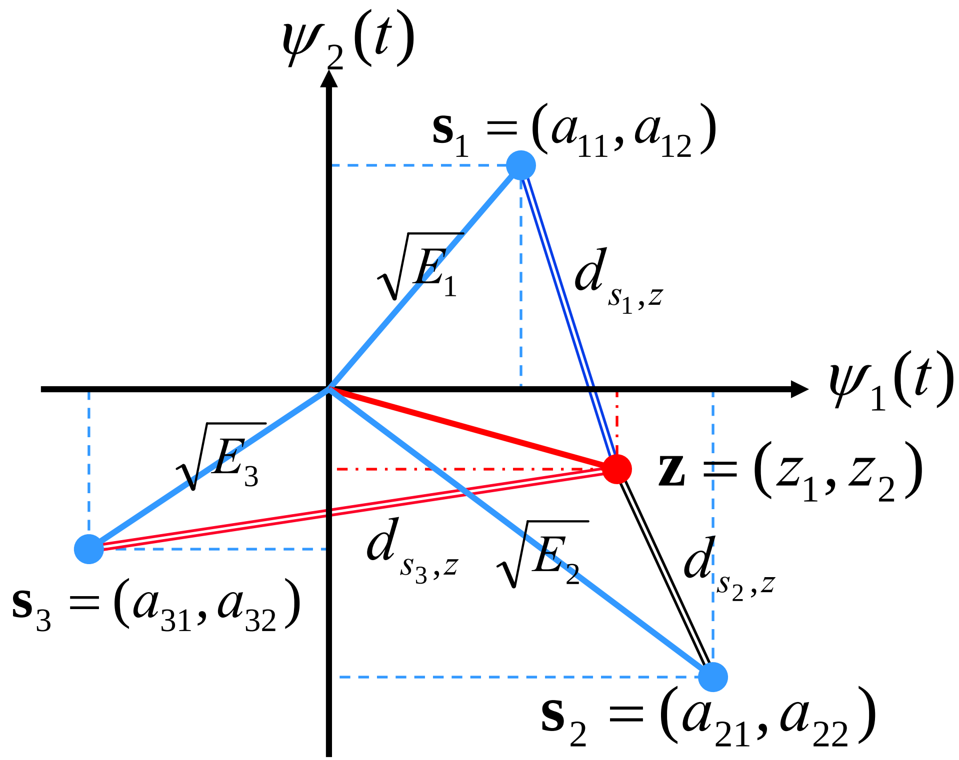
$$\left\|ax(t)\right\| = \left|a\right| \left\|x(t)\right\|$$

  - Norm between two signals:

$$d_{x,y} = \left\|x(t) - y(t)\right\|$$

- We refer to the norm between two signals as the Euclidean distance between two signals.

Example of distances in signal space



The Euclidean distance between signals *z(t)* and *s(t)*:

$$d_{s_i,z} = \|s_i(t) - z(t)\| = \sqrt{(a_{i1} - z_1)^2 + (a_{i2} - z_2)^2}$$

$$i = 1,2,3$$

Orthogonal signal space

- N-dimensional orthogonal signal space is characterized by N linearly independent functions $\{\psi_j(t)\}_{j=1}^{N}$ called basis functions. The basis functions must satisfy the orthogonality condition

$$\boxed{<\psi_i(t),\psi_j(t)>=\int_0^T \psi_i(t)\psi_j^*(t)dt = K_i\delta_{ji} \qquad \begin{array}{c} 0 \le t \le T \\ j,i=1,\ldots,N \end{array}}$$

where

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

- If all $K_i = 1$, the signal space is orthonormal.
  -

Constructing Orthonormal basis from non-orthonormal set of vectors:

- Gram-Schmidt procedure

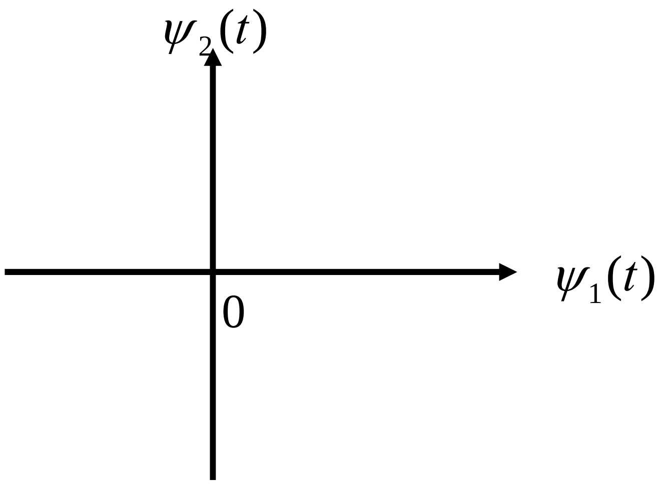
Example of an orthonormal bases

- Example: *2-dimensional* orthonormal signal space

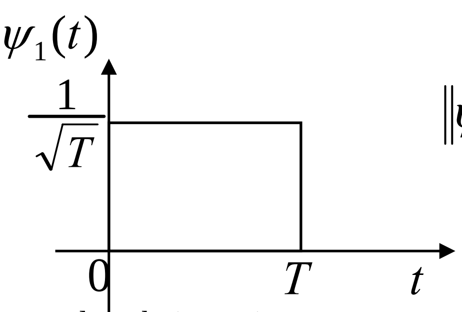
$$\begin{cases} \psi_1(t) = \sqrt{\dfrac{2}{T}}\cos(2\pi t/T) & 0 \le t < T \\ \psi_2(t) = \sqrt{\dfrac{2}{T}}\sin(2\pi t/T) & 0 \le t < T \end{cases}$$

$$<\psi_1(t), \psi_2(t)> = \int_0^T \psi_1(t)\psi_2(t)\,dt = 0$$

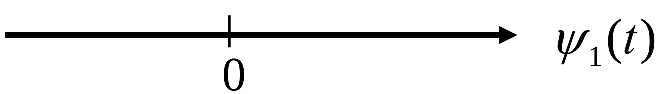
$$\|\psi_1(t)\| = \|\psi_2(t)\| = 1$$



- Example: *1-dimensional* orthonornal signal space



$$\|\psi_1(t)\| = 1$$



Sine/Cosine Bases: Note!

- *Approximately* orthonormal!

$$\phi_1(t) = \sqrt{\frac{2}{T}}\cos(2\pi f_c t)$$

$$\phi_2(t) = \sqrt{\frac{2}{T}}\sin(2\pi f_c t).$$

$$\int_0^T \phi_1^2(t)dt = \frac{2}{T}\int_0^T .5[1+\cos(4\pi f_c t)]dt = 1 + \frac{\sin(4\pi f_c T)}{4\pi f_c T}.$$

The numerator in the second term of (5.8) is bounded by one and for $f_c T >> 1$ the denominator of this term is very large. Thus, this second term can be neglected. Similarly,

$$\int_0^T \phi_1(t)\phi_2(t)dt = \frac{2}{T}\int_0^T .5\sin(4\pi f_c t)dt = \frac{-\cos(4\pi f_c T)}{4\pi f_c T} \approx 0,$$

where the approximation is taken as an equality for $f_c T >> 1$.

- These are the in-phase & quadrature-phase dimensions of complex baseband equivalent representations.

Example: BPSK

Example 5.1:
Binary phase shift keying (BPSK) modulation transmits the signal $s_1(t) = \alpha \cos(2\pi f_c t), 0 \le t \le T$, to send a 1 bit and the signal $s_2(t) = -\alpha \cos(2\pi f_c t), 0 \le t \le T$, to send a 0 bit. Find the set of orthonormal basis functions and coefficients $\{s_{ij}\}$ for this modulation.

Solution: There is only one basis function for $s_1(t)$ and $s_2(t)$, $\phi(t) = \sqrt{2/T}\cos(2\pi f_c t)$, where the $\sqrt{2/T}$ is needed for normalization. The coefficients are then given by $s_1 = \alpha\sqrt{T/2}$ and $s_2 = -\alpha\sqrt{T/2}$.

- Note: two symbols, but only one dimension in BPSK.

Signal space …

- Any arbitrary finite set of waveforms $\{s_i(t)\}_{i=1}^{M}$
  where each member of the set is of duration T, can be expressed as a linear combination of N orthogonal waveforms
  where $\{\psi_j(t)\}_{j=1}^{N}$ $\quad N \leq M$

$$\boxed{s_i(t) = \sum_{j=1}^{N} a_{ij}\psi_j(t) \qquad i = 1,\dots,M \quad N \leq M}$$

where

$$a_{ij} = \frac{1}{K_j} < s_i(t), \psi_j(t) > = \frac{1}{K_j}\int_0^T s_i(t)\psi_j^*(t)dt \qquad \begin{matrix} j = 1,\dots,N \\ i = 1,\dots,M \end{matrix} \qquad 0 \leq t \leq T$$

$$\mathbf{s}_i = (a_{i1}, a_{i2}, \dots, a_{iN})$$

Vector representation of waveform

$$E_i = \sum_{j=1}^{N} K_j |a_{ij}|^2$$

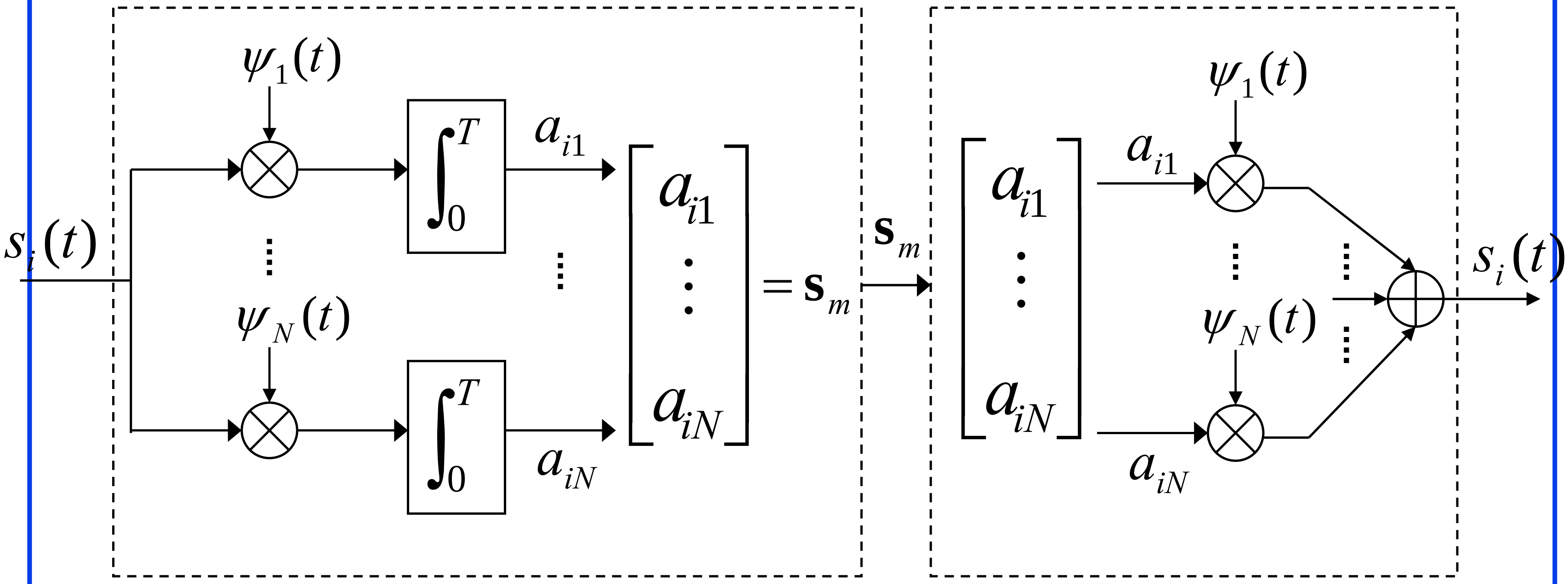
Waveform energy

Signal space …

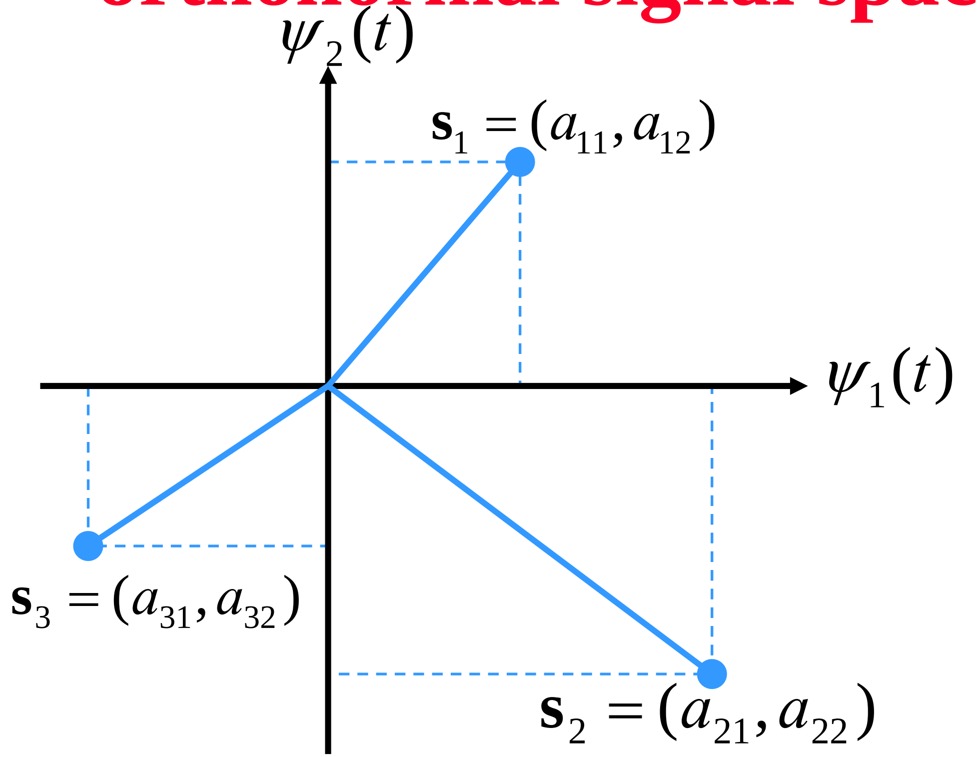
$$s_i(t) = \sum_{j=1}^{N} a_{ij}\psi_j(t)$$

$$\mathbf{s}_i = (a_{i1}, a_{i2}, \ldots, a_{iN})$$

Waveform to vector conversion **Vector to waveform conversion**



Example of projecting signals to an orthonormal signal space



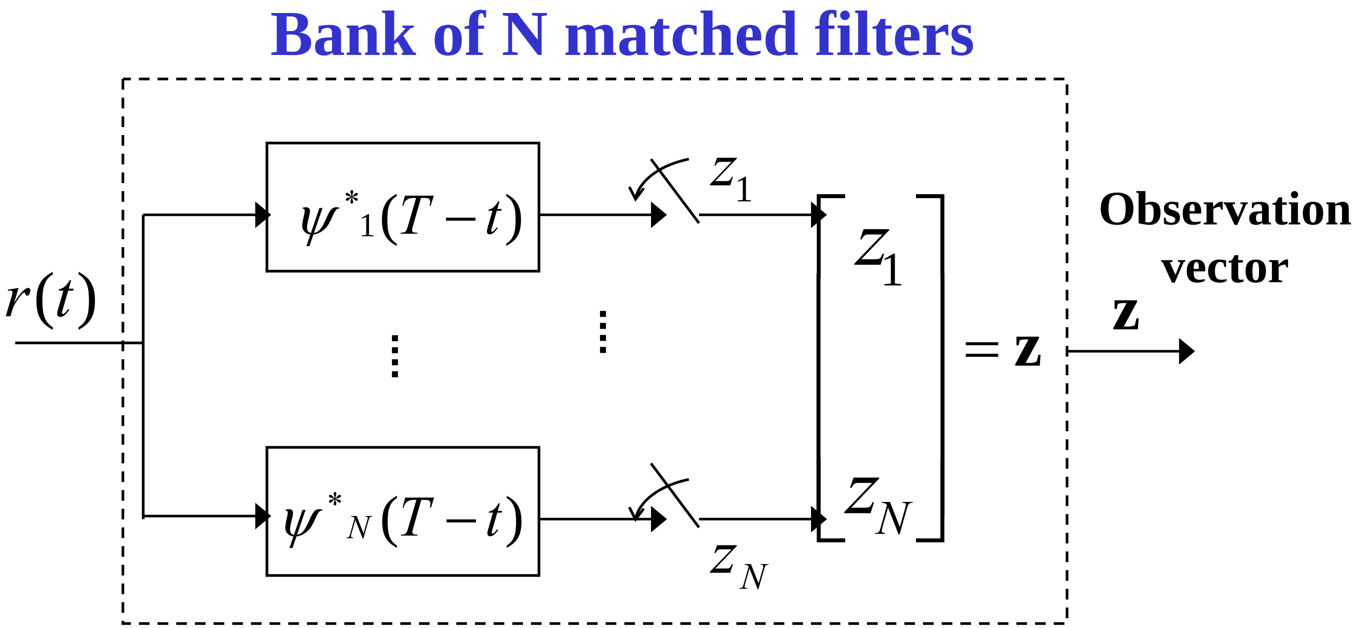
Transmitted signal alternatives

$$
\begin{cases}
s_1(t) = a_{11}\psi_1(t) + a_{12}\psi_2(t) \Leftrightarrow \mathbf{s}_1 = (a_{11}, a_{12}) \\
s_2(t) = a_{21}\psi_1(t) + a_{22}\psi_2(t) \Leftrightarrow \mathbf{s}_2 = (a_{21}, a_{22}) \\
s_3(t) = a_{31}\psi_1(t) + a_{32}\psi_2(t) \Leftrightarrow \mathbf{s}_3 = (a_{31}, a_{32})
\end{cases}
$$

$$a_{ij} = \int_0^T s_i(t)\psi_j(t)dt \quad j = 1,\dots,N \quad i = 1,\dots,M \quad 0 \le t \le T$$

Matched filter receiver (revisited)

(note: we match to the ***basis directions****)*



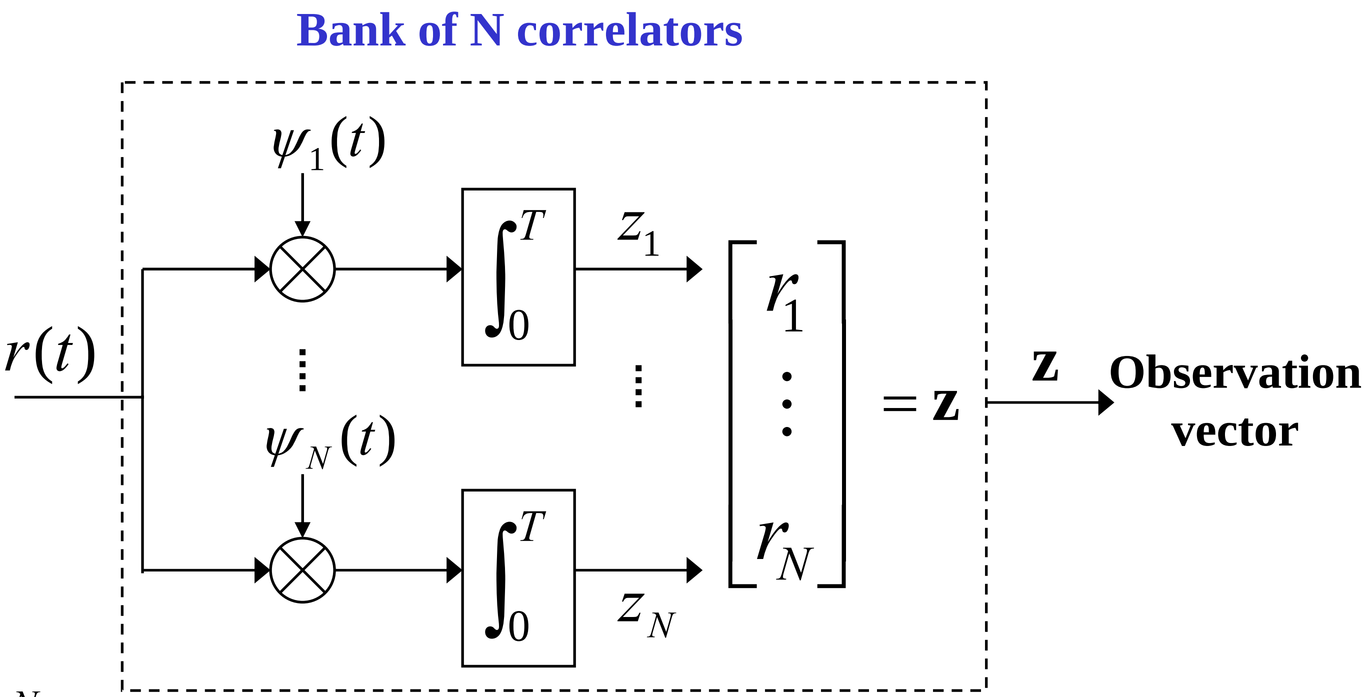
$$s_i(t) = \sum_{j=1}^{N} a_{ij}\psi_j(t) \qquad i = 1,\ldots,M$$

$$\mathbf{z} = (z_1, z_2, \ldots, z_N)$$

$$N \le M$$

$$z_j = r(t) * \psi_j(T-t) \qquad j = 1,\ldots,N$$

Correlator receiver (revisited)



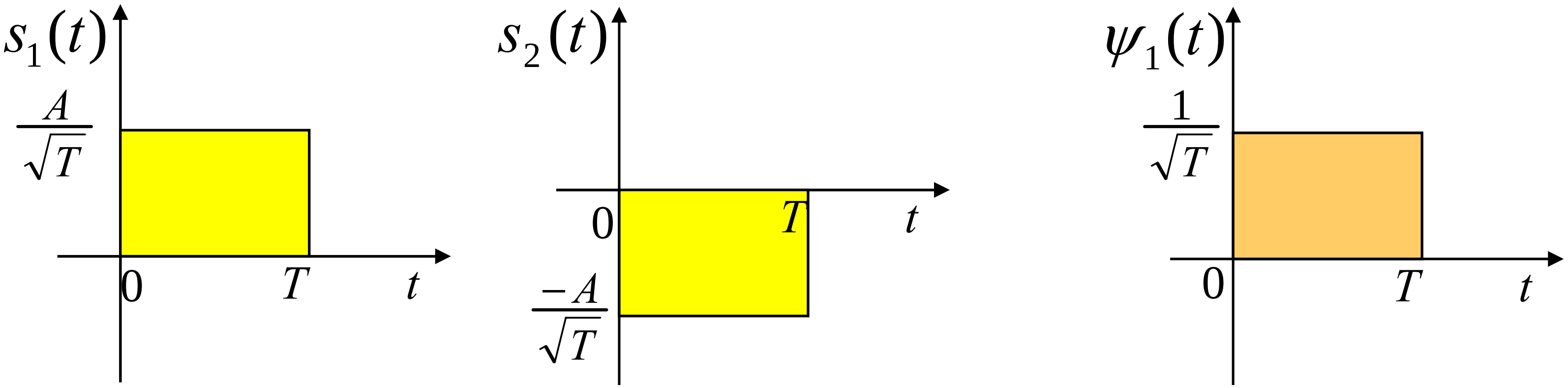
$$s_i(t) = \sum_{j=1}^{N} a_{ij}\psi_j(t) \qquad i = 1,\ldots,M$$

$$\mathbf{z} = (z_1, z_2, \ldots, z_N)$$

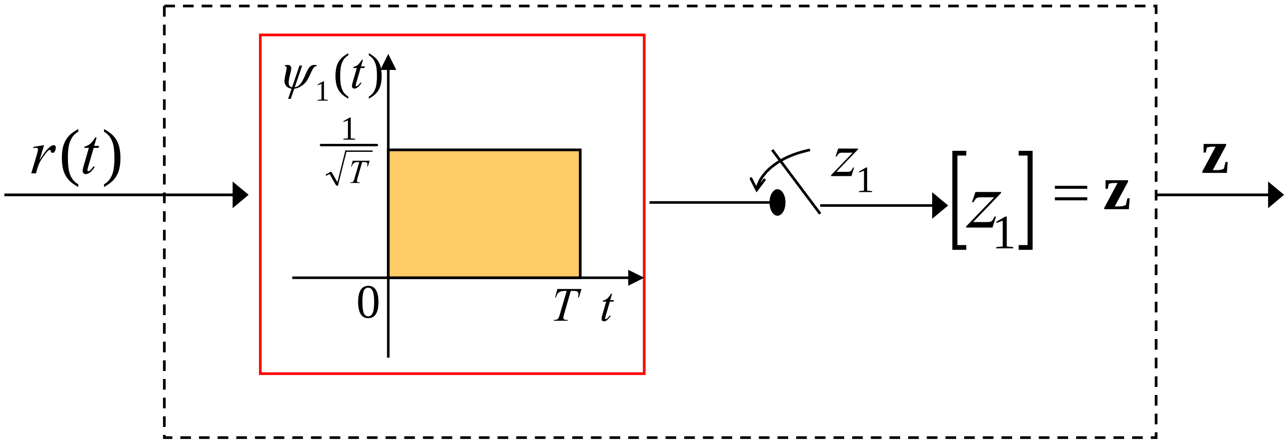
$$z_j = \int_0^T r(t)\psi_j(t)dt \qquad j = 1,\ldots,N$$

$$N \le M$$

Example of matched filter receivers using basic functions



1 matched filter



- *Number of matched filters (or correlators) is reduced by 1* compared to using matched filters (correlators) to the transmitted signal!

White noise in Orthonormal Signal Space

- AWGN *n(t)* can be expressed as

$$n(t) = \hat{n}(t) + \tilde{n}(t)$$

Noise projected on the signal space (colored):impacts the detection process.

Noise outside on the signal space (irrelevant)

$$\begin{cases} \hat{n}(t) = \sum_{j=1}^{N} n_j \psi_j(t) & \\ n_j = < n(t), \psi_j(t) > & j = 1, \ldots, N \\ < \tilde{n}(t), \psi_j(t) > = 0 & j = 1, \ldots, N \end{cases}$$

Vector representation of $\hat{n}(t)$

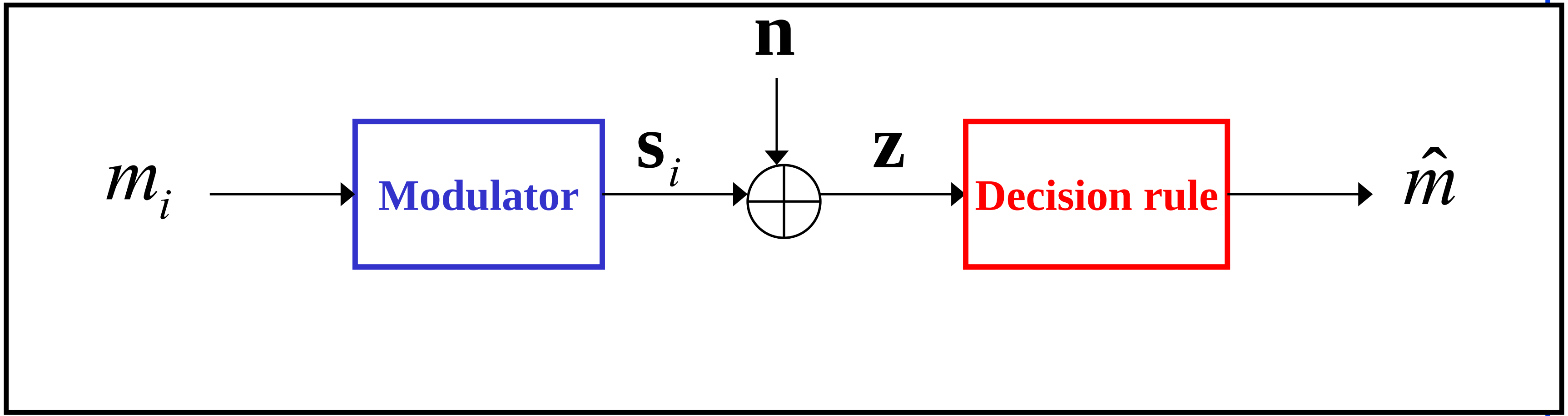
$$\mathbf{n} = (n_1, n_2, \ldots, n_N)$$

$\{n_j\}_{j=1}^{N}$ independent zero-mean Gaussain random variables with variance $\text{var}(n_j) = N_0 / 2$

Detection: Maximum Likelihood & Performance Bounds

Detection of signal in AWGN

- Detection problem:
  - Given the observation vector $\mathbf{z}$, perform a mapping from $\mathbf{z}$ to an estimate $\hat{m}$ of the transmitted symbol, m_i, such that the average probability of error in the decision is minimized.



Statistics of the observation Vector

- AWGN channel model: $\mathbf{z} = \mathbf{s}_i + \mathbf{n}$
  - Signal vector $\mathbf{s}_i = (a_{i1}, a_{i2}, \ldots, a_{iN})$ is deterministic.
  - Elements of noise vector $\mathbf{n} = (n_1, n_2, \ldots, n_N)$ are <u>*i.i.d Gaussian*</u> random variables with zero-mean and variance $N_0/2$. The noise vector pdf is

$$p_{\mathbf{n}}(\mathbf{n}) = \frac{1}{(\pi N_0)^{N/2}} \exp\left(-\frac{\|\mathbf{n}\|^2}{N_0} \right)$$

  - The elements of observed vector $\mathbf{z} = (z_1, z_2, \ldots, z_N)$ are independent Gaussian random variables. Its pdf is

$$p_{\mathbf{z}}(\mathbf{z} \mid \mathbf{s}_i) = \frac{1}{(\pi N_0)^{N/2}} \exp\left(-\frac{\|\mathbf{z} - \mathbf{s}_i\|^2}{N_0} \right)$$

Detection

- Optimum decision rule (*maximum a posteriori probability*):

$$\begin{aligned}&\text{Set } \hat{m} = m_i \text{ if}\\ &\Pr(m_i \text{ sent} \mid \mathbf{z}) \geq \Pr(m_k \text{ sent} \mid \mathbf{z}), \text{ for all } k \neq i\\ &\text{where } k = 1,\ldots,M.\end{aligned}$$

- Applying Bayes' rule gives:

$$\begin{aligned}&\text{Set } \hat{m} = m_i \text{ if}\\ &p_k \frac{p_{\mathbf{z}}(\mathbf{z} \mid m_k)}{p_{\mathbf{z}}(\mathbf{z})}, \text{ is maximum for all } k = i\end{aligned}$$

Detection …

- Partition the signal space into M decision regions, such that $Z_1, \ldots, Z_M$

> Vector $\mathbf{z}$ lies inside region Z_i if
>
> $$\ln\left[p_k \frac{p_{\mathbf{z}}(\mathbf{z} \mid m_k)}{p_{\mathbf{z}}(\mathbf{z})}\right], \text{ is maximum for all } k = i.$$
>
> That means
>
> $$\hat{m} = m_i$$

Detection (ML rule)

- For equal probable symbols, the optimum decision rule (maximum posteriori probability) is simplified to:

$$\boxed{\begin{array}{l}\text{Set } \hat{m} = m_i \text{ if} \\ p_{\mathbf{z}}(\mathbf{z} \mid m_k), \text{ is maximum for all } k = i\end{array}}$$

or equivalently:

$$\boxed{\begin{array}{l}\text{Set } \hat{m} = m_i \text{ if} \\ \ln[p_{\mathbf{z}}(\mathbf{z} \mid m_k)], \text{ is maximum for all } k = i\end{array}}$$

which is known as *maximum likelihood*.

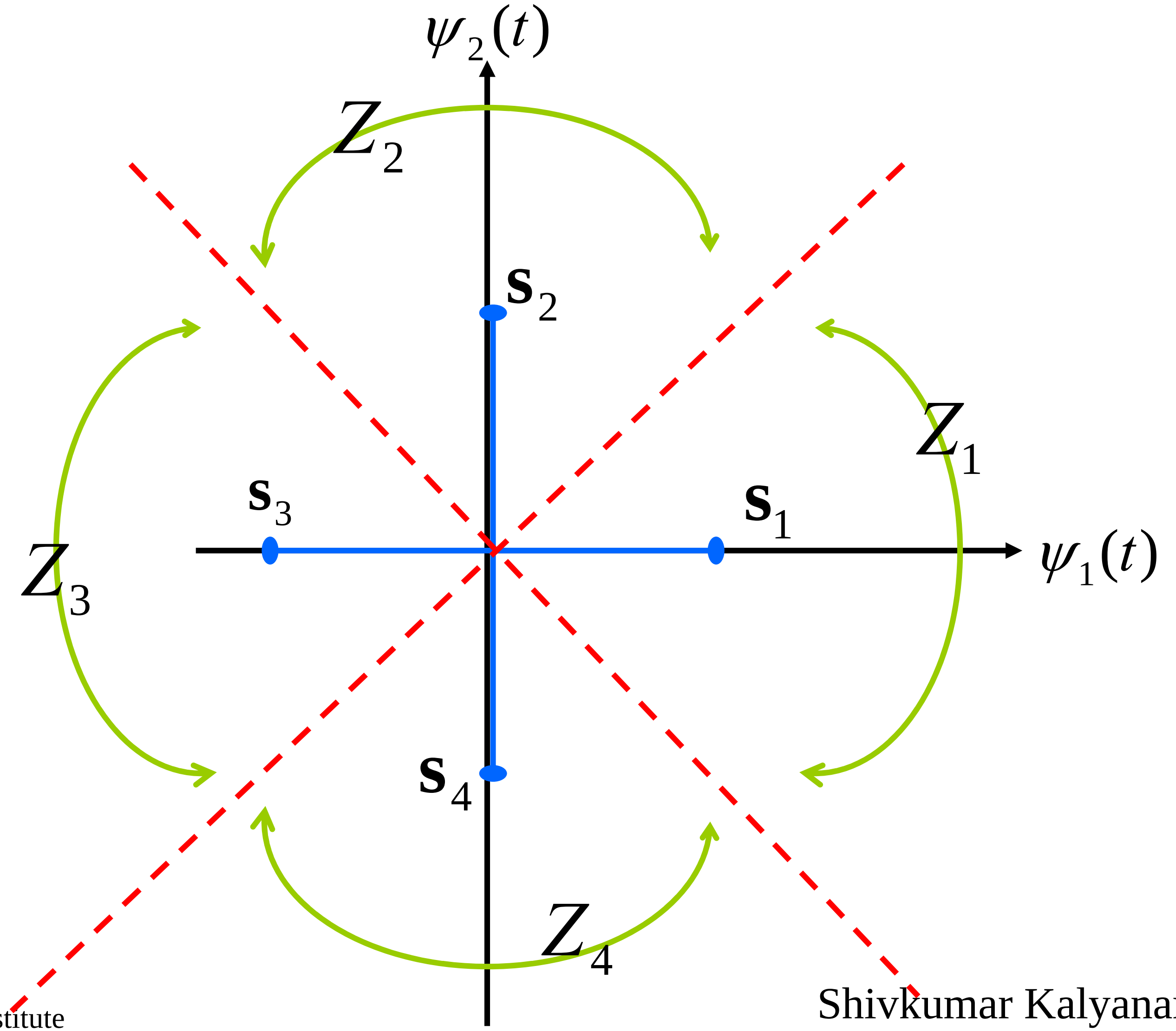
Detection (ML)…

- Partition the signal space into M decision regions, $Z_1, \ldots, Z_M$
- Restate the maximum likelihood decision rule as follows:

Vector $\mathbf{z}$ lies inside region Z_i if
$\ln[p_{\mathbf{z}}(\mathbf{z} \mid m_k)]$, is maximum for all $k = i$
That means
$\hat{m} = m_i$

Vector $\mathbf{z}$ lies inside region Z_i if
$\left\| \mathbf{z} - \mathbf{s}_k \right\|$, is minimum for all $k = i$

Schematic example of ML decision regions



Probability of symbol error

- Erroneous decision: For the transmitted symbol $\mathbf{s}_i$ or equivalently signal vector m_i, an error in decision occurs if the observation vector $\mathbf{z}$ does not fall inside region Z_i.
  - Probability of erroneous decision for a transmitted symbol

$$\boxed{\Pr(\hat{m} \neq m_i) = \Pr(m_i \text{ sent})\Pr(\mathbf{z} \text{ does not lie inside } Z_i \,|\, m_i \text{ sent})}$$

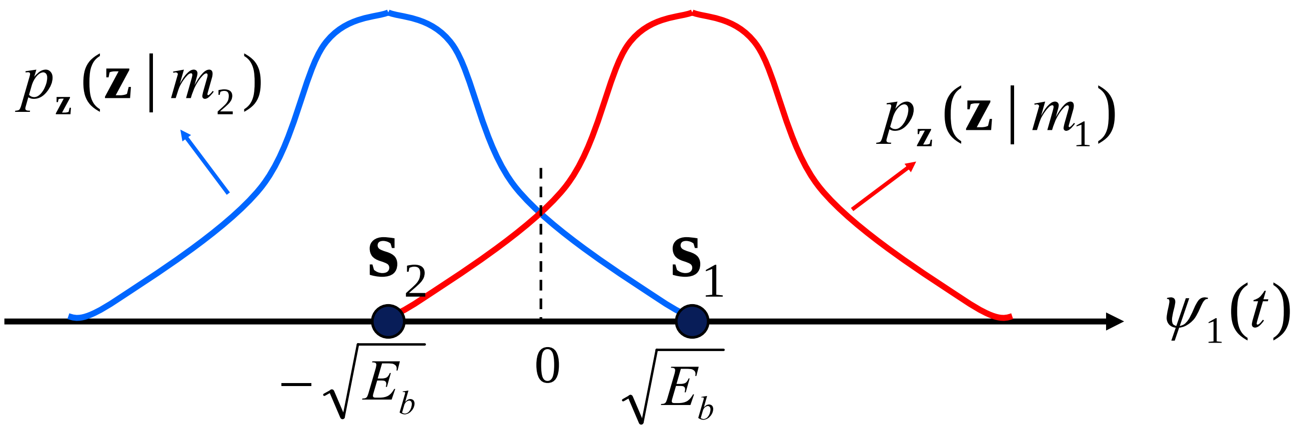
  - Probability of correct decision for a transmitted symbol

$$\boxed{\Pr(\hat{m} = m_i) = \Pr(m_i \text{ sent})\Pr(\mathbf{z} \text{ lies inside } Z_i \,|\, m_i \text{ sent})}$$

$$P_c(m_i) = \Pr(\mathbf{z} \text{ lies inside } Z_i \,|\, m_i \text{ sent}) = \int_{Z_i} p_{\mathbf{z}}(\mathbf{z} \,|\, m_i)\, d\mathbf{z}$$

$$P_e(m_i) = 1 - P_c(m_i)$$

Example for binary PAM



$$P_e(m_1) = P_e(m_2) = Q\left(\frac{\|\mathbf{s}_1 - \mathbf{s}_2\|/2}{\sqrt{N_0/2}}\right)$$

$$P_B = P_E(2) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

Average prob. of symbol error …

- Average probability of symbol error :

$$\boxed{P_E(M) = \sum_{i=1}^{M} \Pr(\hat{m} \neq m_i)}$$

- For equally probable symbols:

$$\boxed{\begin{aligned} P_E(M) &= \frac{1}{M}\sum_{i=1}^{M} P_e(m_i) = 1 - \frac{1}{M}\sum_{i=1}^{M} P_c(m_i) \\ &= 1 - \frac{1}{M}\sum_{i=1}^{M} \int_{Z_i} p_{\mathbf{z}}(\mathbf{z} \mid m_i)\, d\mathbf{z} \end{aligned}}$$

Union bound

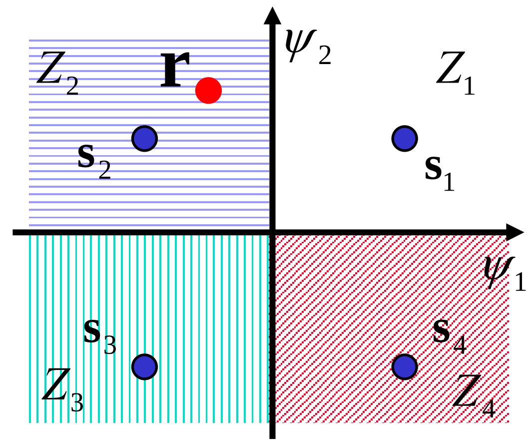
Union bound

The probability of a finite union of events is upper bounded by the sum of the probabilities of the individual events.

$$\boxed{P_e(m_i) \le \sum_{\substack{k=1 \\ k \ne i}}^{M} P_2(\mathbf{s}_k, \mathbf{s}_i)} \Rightarrow \boxed{P_E(M) \le \frac{1}{M} \sum_{i=1}^{M} \sum_{\substack{k=1 \\ k \ne i}}^{M} P_2(\mathbf{s}_k, \mathbf{s}_i)}$$

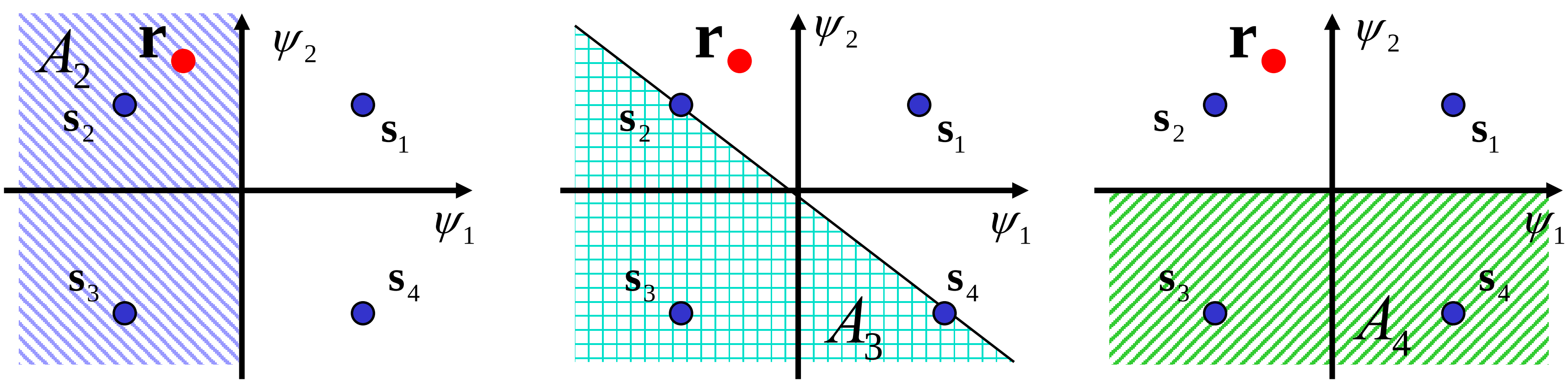
Example of union bound

$$P_e(m_1) = \int_{Z_2 \cup Z_3 \cup Z_4} p_{\mathbf{r}}(\mathbf{r} \mid m_1)\,d\mathbf{r}$$



Union bound:

$$P_e(m_1) \le \sum_{k=2}^{4} P_2(\mathbf{s}_k, \mathbf{s}_1)$$



$$P_2(\mathbf{s}_2, \mathbf{s}_1) = \int_{A_2} p_{\mathbf{r}}(\mathbf{r} \mid m_1)\,d\mathbf{r}$$

$$P_2(\mathbf{s}_3, \mathbf{s}_1) = \int_{A_3} p_{\mathbf{r}}(\mathbf{r} \mid m_1)\,d\mathbf{r}$$

$$P_2(\mathbf{s}_4, \mathbf{s}_1) = \int_{A_4} p_{\mathbf{r}}(\mathbf{r} \mid m_1)\,d\mathbf{r}$$

Upper bound based on minimum distance

$$P_2(\mathbf{s}_k, \mathbf{s}_i) = \Pr(\mathbf{z} \text{ is closer to } \mathbf{s}_k \text{ than } \mathbf{s}_i \text{, when } \mathbf{s}_i \text{ is sent})$$

$$= \int_{d_{ik}}^{\infty} \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{u^2}{N_0}\right) du = Q\left(\frac{d_{ik}/2}{\sqrt{N_0/2}}\right)$$

$$d_{ik} = \left\| \mathbf{s}_i - \mathbf{s}_k \right\|$$

$$P_E(M) \le \frac{1}{M} \sum_{i=1}^{M} \sum_{\substack{k=1 \\ k \ne i}}^{M} P_2(\mathbf{s}_k, \mathbf{s}_i) \le (M-1) Q\left(\frac{d_{\min}/2}{\sqrt{N_0/2}}\right)$$

Minimum distance in the signal space: $d_{\min} = \min_{\substack{i,k \\ i \ne k}} d_{ik}$

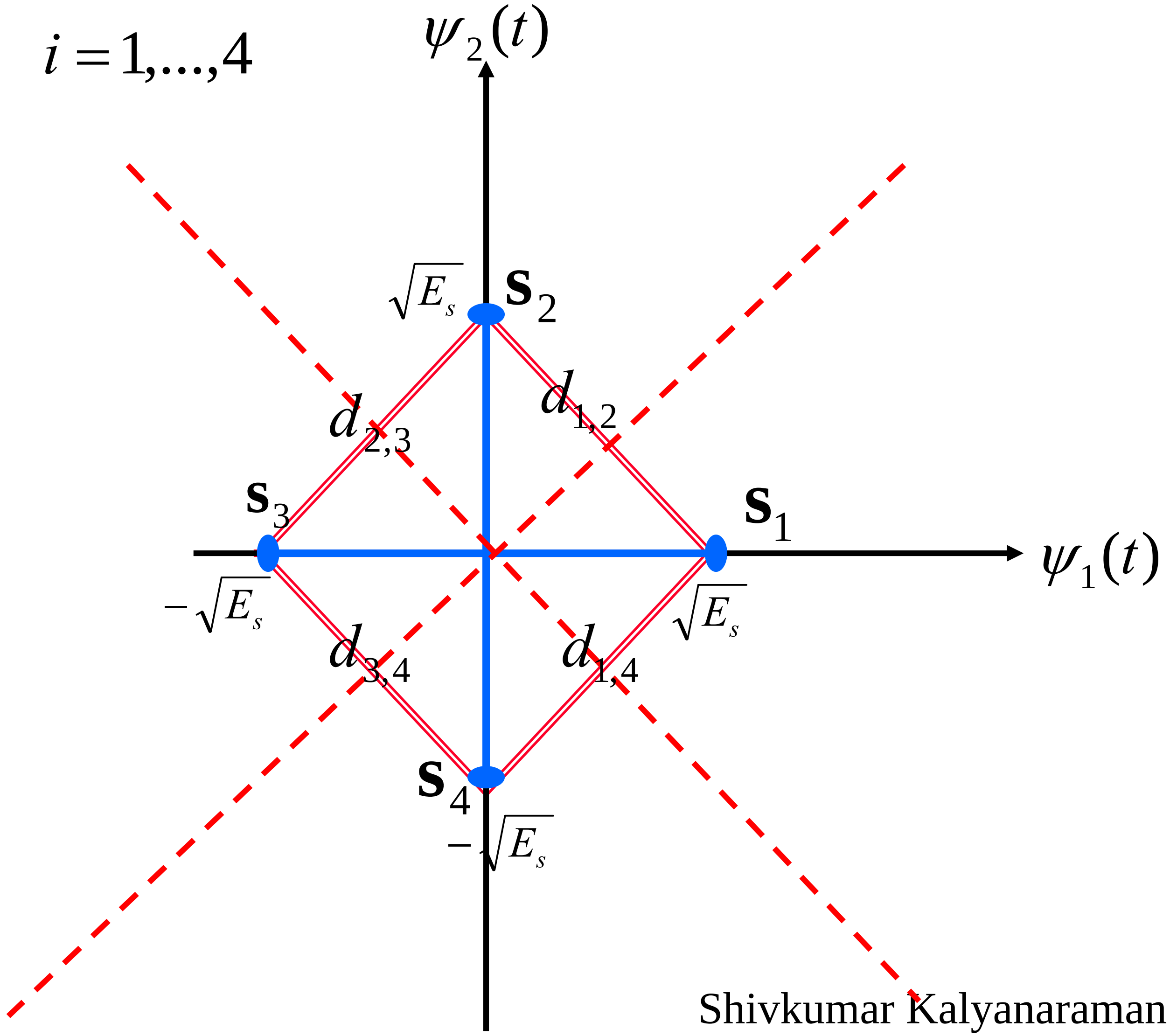
Example of upper bound on av. Symbol error prob. based on union bound

$$\|\mathbf{s}_i\| = \sqrt{E_i} = \sqrt{E_s}, \quad i = 1,\ldots,4$$

$$d_{i,k} = \sqrt{2E_s}$$

$$i \neq k$$

$$d_{\min} = \sqrt{2E_s}$$



Example 5.3:
Consider a signal constellation in $\mathcal{R}^2$ defined by $s_1 = (A, 0)$, $s_2 = (0, A)$, $s_3 = (-A, 0)$ and $s_4 = (0, -A)$. Assume $A/\sqrt{N_0} = 4$. Find the minimum distance and the union bound (5.40), looser bound (5.43), closed form bound (5.44), and nearest neighbor approximation (5.45) on P_e for this constellation set.

Solution: The constellation is as depicted in Figure 5.3 with the radius of the circle equal to A. By symmetry, we need only consider the error probability associated with one of the constellation points, since it will be the same for the others. We focus on the error associated with transmitting constellation point s_1. The minimum distance to this constellation point is easily computed as $d_{min} = d_{12} = d_{23} = d_{34} = d_{14} = \sqrt{A^2 + A^2} = \sqrt{2A^2}$. The distance to the other constellation points are $d_{13} = d_{24} = 2A$. By symmetry, $P_e(m_i \text{ sent}) = P_e(m_j \text{ sent}), j \neq i$, so the union bound simplifies to

$$P_e \leq \sum_{j=2}^{4} Q\left(\frac{d_{1j}}{\sqrt{2N_0}}\right) = 2Q(A/\sqrt{N_0}) + Q(\sqrt{2}A/\sqrt{N_0}) = 2Q(4) + Q(\sqrt{32}) = 3.1679 * 10^{-5}.$$

The looser bound yields

$$P_e \leq 3Q(4) = 9.5014 * 10^{-5}$$

which is roughly a factor of 3 looser than the union bound. The closed-form bound yields

$$P_e \leq \frac{3}{\pi}\exp\left[\frac{-.5A^2}{N_0}\right] = 3.2034 * 10^{-4},$$

which differs from the union bound by about an order of magnitude. Finally, the nearest neighbor approximation yields

$$P_e \approx 2Q(4) = 3.1671 * 10^{-5},$$

which, as expected, is approximately equal to the union bound.

Eb/No figure of merit in digital communications

- SNR or S/N is the average signal power to the average noise power. SNR should be modified in terms of bit-energy in digital communication, because:
  - Signals are transmitted within a symbol duration and hence, are ***energy signal (zero power).***
  - A metric at the ***bit-level*** facilitates comparison of different DCS transmitting different number of bits per symbol.

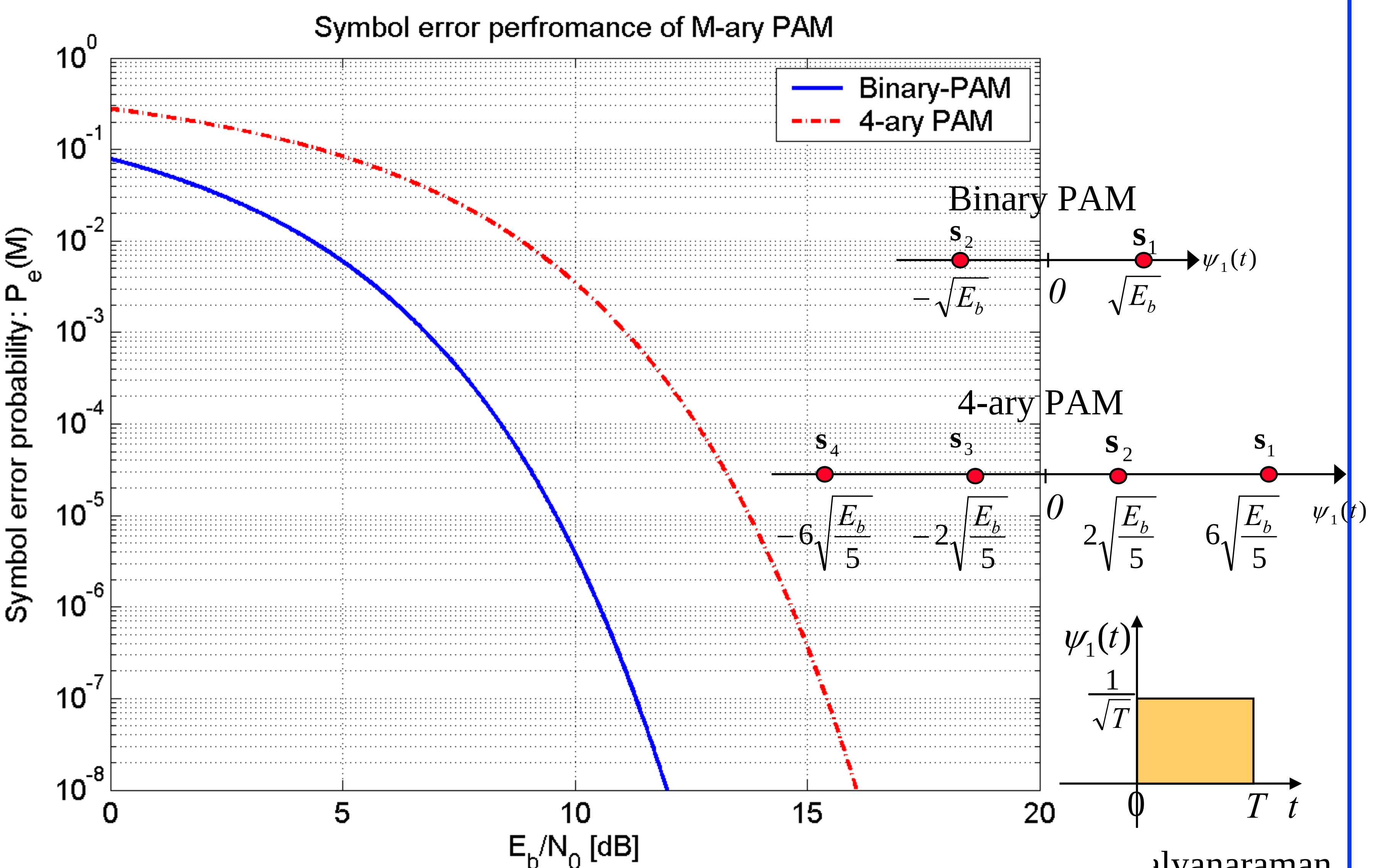
$$\boxed{\frac{E_b}{N_0} = \frac{ST_b}{N/W} = \frac{S}{N}\frac{W}{R_b}}$$

R_b : Bit rate

W : Bandwidth

Note: S/N = Eb/No x spectral efficiency

Example of Symbol error prob. For PAM signals

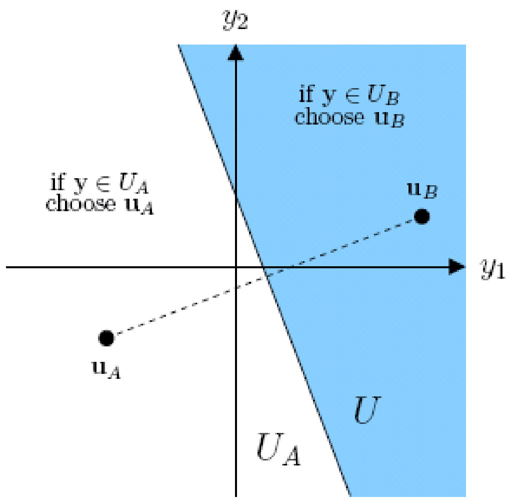


Maximum Likelihood (ML) Detection: *Vector* Case

Nearest Neighbor Rule: $\|\mathbf{y}-\mathbf{u}_A\| < \|\mathbf{y}-\mathbf{u}_B\|,$

By the *isotropic* property of the Gaussian noise, we expect the error probability to be the same for both the transmit symbols u_A, u_B.

Error probability: $\mathbb{P}\left\{(\mathbf{u}_A-\mathbf{u}_B)^t\,\mathbf{w} < -\frac{\|\mathbf{u}_A-\mathbf{u}_B\|^2}{2}\right\}$



$$(\mathbf{u}_A-\mathbf{u}_B)^t\,\mathbf{w} \sim \mathcal{N}\left(0, \|\mathbf{u}_A-\mathbf{u}_B\|^2 N_0/2\right).$$

Project the received vector **y** along the difference vector direction $\mathbf{u_A}$- $\mathbf{u_B}$ is a "*sufficient statistic*".

Noise outside these finite dimensions is *irrelevant* for detection. (rotational invariance of detection problem)

ps: Vector norm is a natural extension of "magnitude" or length

Extension to M-PAM (Multi-Level Modulation)

The above argument for binary detection generalizes naturally to the case when the transmit vector can be one of M vectors $\mathbf{u}_1, \ldots, \mathbf{u}_M$. The projection of $\mathbf{y}$ onto the subspace spanned by $\mathbf{u}_1, \ldots, \mathbf{u}_M$ is a sufficient statistic for the detection problem. In the special case when the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_M$ are collinear, i.e. $\mathbf{u}_i = \mathbf{h}x_i$ for some vector $\mathbf{h}$ (for example, when we are transmitting from a PAM constellation), then a projection onto the direction $\mathbf{h}$ provides a sufficient statistic.

- **Note: h refers to the constellation shape/direction**

<u>MPAM</u>: vectors $\mathbf{u}_1, \ldots, \mathbf{u}_M$ are collinear, i.e. $\mathbf{u}_i = \mathbf{h}x_i$

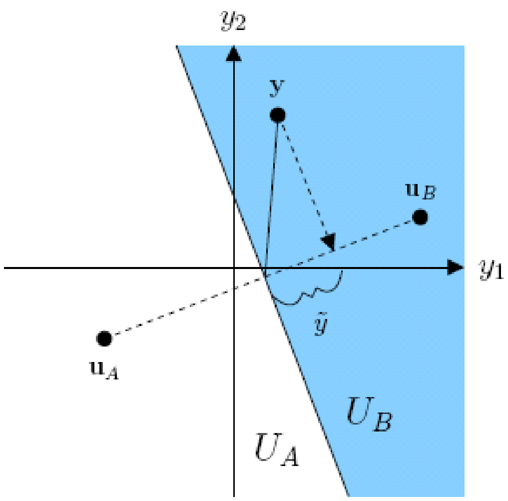
Complex Vector Space Detection

$$\mathbf{y} = \mathbf{u} + \mathbf{w}, \quad \text{where } \mathbf{w} \sim \mathcal{CN}\left(0, N_0\mathbf{I}\right).$$

$$\mathbf{v} := \left(\mathbf{u}_A - \mathbf{u}_B\right)/\|\mathbf{u}_A - \mathbf{u}_B\|.$$

$$\tilde{y} := \mathbf{v}^*\left(\mathbf{y} - \frac{1}{2}\left(\mathbf{u}_A + \mathbf{u}_B\right)\right) = x\|\mathbf{u}_A - \mathbf{u}_B\| + w$$

$$\boxed{\Re[\tilde{\mathbf{y}}] = x\|\mathbf{u}_A - \mathbf{u}_B\| + \Re[w]}$$



Error probability: $Q\left(\frac{\|\mathbf{u}_A - \mathbf{u}_B\|}{2\sqrt{N_0/2}}\right)$

- Note: Instead of $\mathbf{v}^T$, use **v*** for complex vectors ("*transpose and conjugate*") for inner products…

Complex Detection: Summary

Binary Signals:

The transmit vector $\mathbf{u}$ is either $\mathbf{u}_A$ or $\mathbf{u}_B$ and we wish to detect $\mathbf{u}$ from received vector

$$\mathbf{y} = \mathbf{u} + \mathbf{w}, \tag{A.52}$$

where $\mathbf{w} \sim \mathcal{CN}(0, N_0\mathbf{I})$. The ML detector picks the transmit vector closest to $\mathbf{y}$ and the error probability is:

$$Q\left(\frac{||\mathbf{u}_A - \mathbf{u}_B||}{2\sqrt{N_0/2}}\right). \tag{A.53}$$

Collinear Signals:

The transmit symbol x is equally likely to take one of a finite set of values in $\mathcal{C}$ (the *constellation* points) and the received vector is

$$\mathbf{y} = \mathbf{h}x + \mathbf{w}, \tag{A.54}$$

where $\mathbf{h}$ is a fixed vector.

Projecting $\mathbf{y}$ onto the unit vector $\mathbf{v} := \mathbf{h}/||\mathbf{h}||$ yields a scalar sufficient statistic:

$$\mathbf{v}^*\mathbf{y} = ||\mathbf{h}||x + w. \tag{A.55}$$

Here $w \sim \mathcal{CN}(0, N_0)$.

If further the constellation is real-valued, then

$$\Re[\mathbf{v}^*\mathbf{y}] = ||\mathbf{h}||x + \Re[w] \tag{A.56}$$

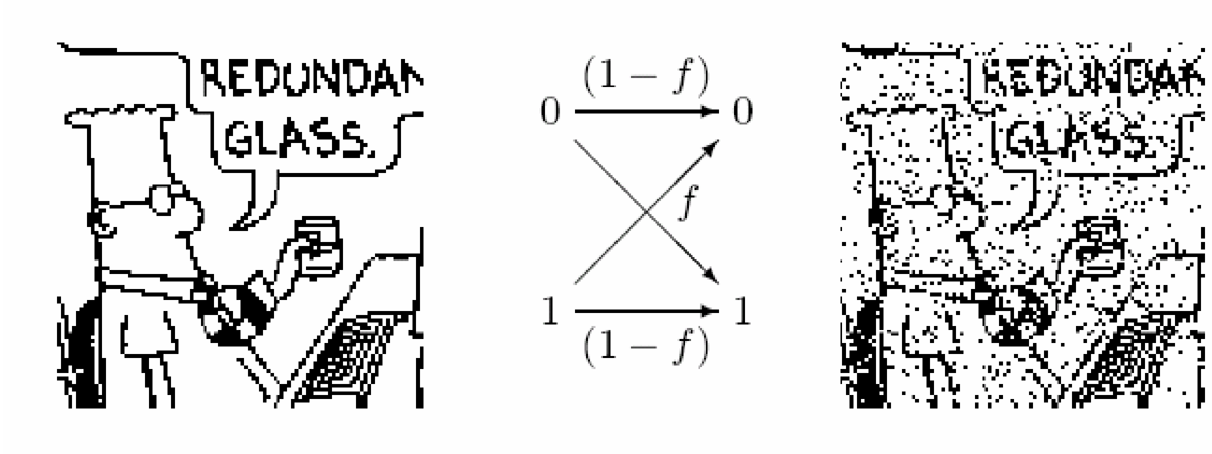
is sufficient. Here $\Re[w] \sim \mathcal{N}(0, N_0/2)$.

With antipodal signalling, $x = \pm a$, the ML error probability is simply

$$Q\left(\frac{a||\mathbf{h}||}{\sqrt{N_0/2}}\right). \tag{A.57}$$

Via a translation, the binary signal detection problem in the first part of the summary can be reduced to this antipodal signalling scenario.

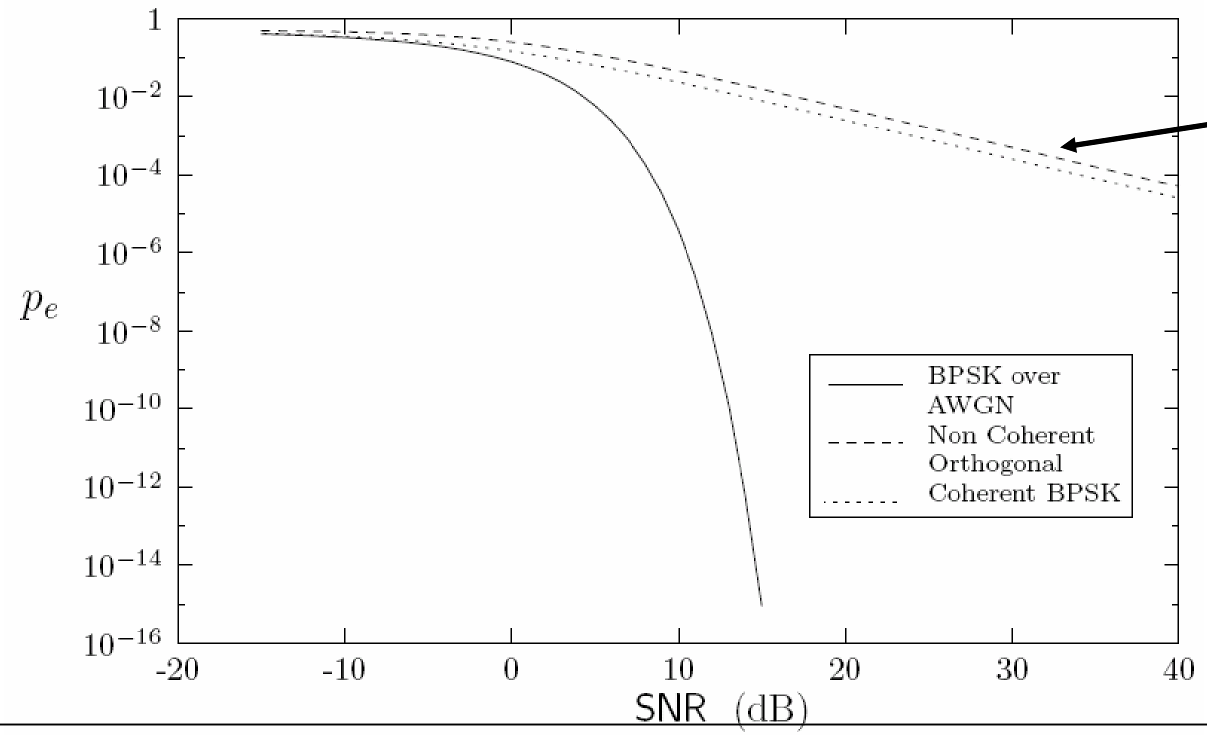
Detection Error => BER



- If the bit error is i.i.d (discrete memoryless channel) over the sequence of bits, then you can model it as a **binary symmetric channel (BSC)**
  - BER is modeled as a uniform probability *f*
  - As BER (*f*) increases, the effects become increasingly intolerable
  - *f* tends to increase rapidly with lower SNR: "*waterfall*" curve (Q-function)

$$\mathbb{P}\left\{y < \frac{u_A + u_B}{2} \middle| u = u_A\right\} = \mathbb{P}\left\{w > \frac{|u_A - u_B|}{2}\right\} = Q\left(\frac{|u_A - u_B|}{2\sqrt{N_0/2}}\right)$$

SNR vs BER: AWGN vs Rayleigh



Need diversity techniques to deal with Rayleigh (even 1-tap, flat-fading)!

- Observe the "waterfall" like characteristic (essentially plotting the Q(x) function)!
- Telephone lines: SNR = 37dB, but low b/w (3.7kHz)
- Wireless: Low SNR = 5-10dB, higher bandwidth (upto 10 Mhz, MAN, and 20Mhz LAN)
- Optical fiber comm: High SNR, high bandwidth ! But cant process w/ complicated codes, signal processing etc

Better performance through diversity

Diversity ⇔ the receiver is provided with multiple copies of the transmitted signal. The multiple signal copies should experience *uncorrelated fading* in the channel.

In this case the probability that *all* signal copies fade simultaneously is reduced dramatically with respect to the probability that a *single* copy experiences a fade.

As a rough rule:

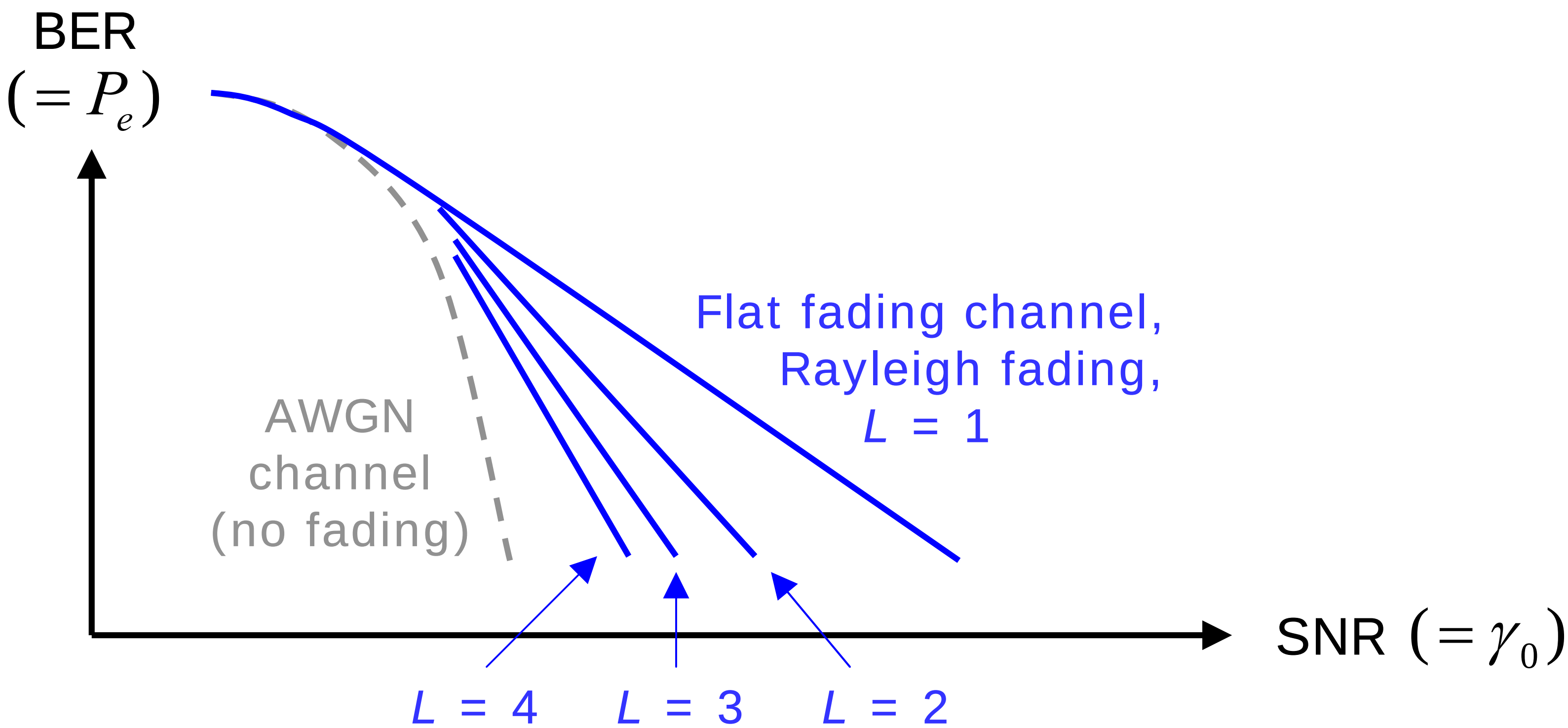
P_e is proportional to $\dfrac{1}{\gamma_0^{\,L}}$

Diversity of *L*:th order

BER

Average SNR

BER vs. SNR (diversity effect)



We will explore this story later… slide set part II

Modulation Techniques

What is Modulation?

- Encoding information in a manner suitable for transmission.
  - Translate baseband source signal to bandpass signal
  - Bandpass signal: “modulated signal”
- How?
  - Vary amplitude, phase or frequency of a carrier
- Demodulation: extract baseband message from carrier

Digital vs Analog Modulation

- Cheaper, faster, more power efficient
- Higher data rates, power error correction, impairment resistance:
  - Using coding, modulation, diversity
  - Equalization, multicarrier techniques for ISI mitigation
- More efficient multiple access strategies, better security: CDMA, encryption etc

Goals of Modulation Techniques

- **High Bit Rate**
- **High Spectral Efficiency** (*max Bps/Hz*)
- **High Power Efficiency** (*min power to achieve a target BER*)
- **Low-Cost/Low-Power Implementation**
- **Robustness to Impairments**

Modulation: representation

- **Any modulated signal can be represented as**

$$s(t) = \underset{\text{amplitude}}{A(t)} \cos [\omega_c t + \underset{\text{phase or frequency}}{\phi(t)}]$$

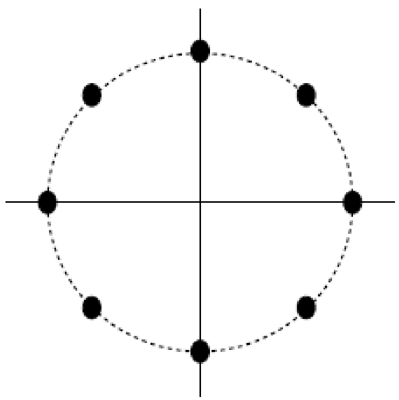
$$= \underbrace{A(t) \cos \phi(t)}_{\textit{in-phase}} \cos \omega_c t \quad - \underbrace{A(t) \sin \phi(t)}_{\textit{quadrature}} \sin \omega_c t$$

- **Linear versus nonlinear modulation ⇒ impact on spectral efficiency**
  - **Linear**: Amplitude or phase
  - **Non-linear**: frequency: spectral broadening
- **Constant envelope versus non-constant envelope**

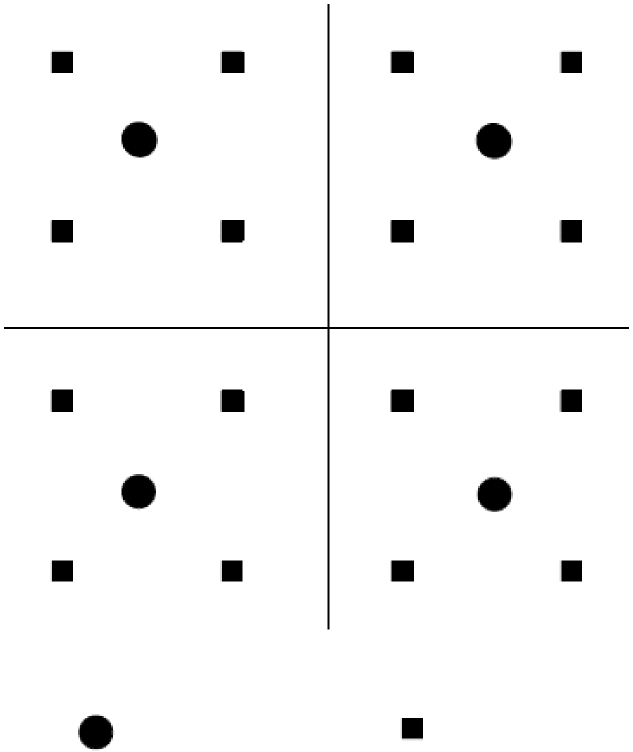
⇒ hardware implications with impact on power efficiency

(=> reliability: i.e. target BER at lower SNRs)

Complex Vector Spaces: Constellations



MPSK



4QAM and 16QAM Constellations.

Circular

Square

- Each signal is encoded (modulated) as a vector in a signal space

Linear Modulation Techniques

$$s(t) = \underbrace{\left[\sum_n a_n\, g(t-nT)\right]}_{\text{I(t), in-phase}} \cos \omega_c t - \underbrace{\left[\sum_n b_n\, g(t-nT)\right]}_{\text{Q(t), quadrature}} \sin \omega_c t$$

LINEAR MODULATIONS

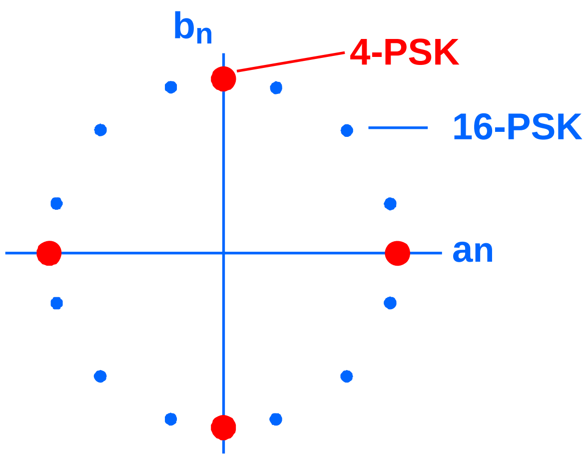
- **M-ARY QUADRATURE *AMPLITUDE* MOD. (M-QAM)** ← Square Constellations
  - $M \neq 4$
  - $M = 4$ (4-QAM = 4-PSK)
- **M-ARY *PHASE* SHIFT KEYING (M-PSK)** ← Circular Constellations
  - $M = 4$ (4-QAM = 4-PSK)
  - $M \neq 4$

M=4 (4-QAM = 4-PSK)

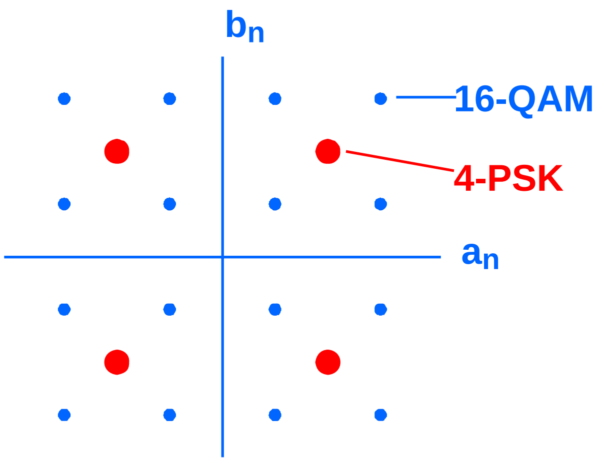
- CONVENTIONAL 4-PSK (QPSK)
- OFFSET 4-PSK (OQPSK)
- DIFFERENTIAL 4-PSK (DQPSK, π/4-DQPSK)

M-PSK and M-QAM

M-PSK (Circular Constellations)



M-QAM (Square Constellations)



Tradeoffs

- Higher-order modulations (M large) are ***more spectrally efficient*** but ***less power efficient (i.e. BER higher)***.
- M-QAM is more spectrally efficient than M-PSK but also more sensitive to system nonlinearities.

Bandwidth vs. Power Efficiency

MPSK:

Bandwidth and Power Efficiency of M-ary PSK Signals

M	2	4	8	16	32	64
$\eta_B = R_b/B$ *	0.5	1	1.5	2	2.5	3
E_b/N_o for BER=10^{-6}	10.5	10.5	14	18.5	23.4	28.5

* B: First null bandwidth of M-ary PSK signals

MQAM:

Bandwidth and Power Efficiency of QAM [Zie92]

M	4	16	64	256	1024	4096
η_B	1	2	3	4	5	6
E_b/N_o for BER = 10^{-6}	10.5	15	18.5	24	28	33.5

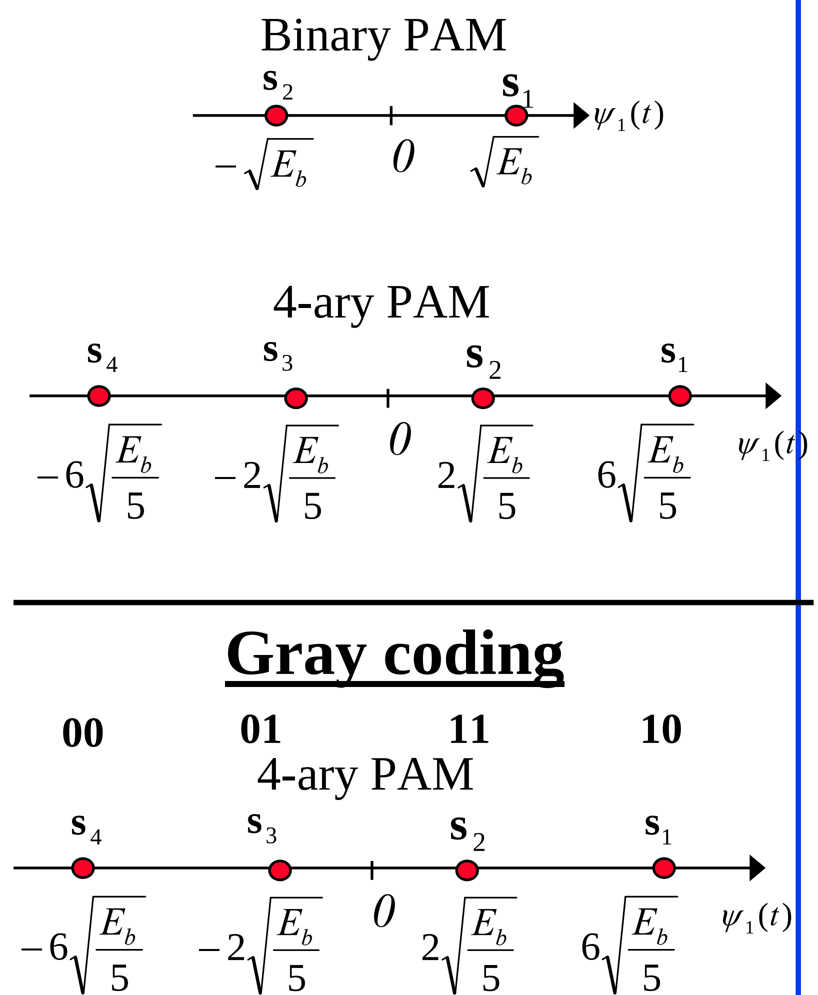
MFSK:

Bandwidth and Power Efficiency of Coherent M-ary FSK [Zie92]

M	2	4	8	16	32	64
η_B	0.4	0.57	0.55	0.42	0.29	0.18
E_b/N_o for BER = 10^{-6}	13.5	10.8	9.3	8.2	7.5	6.9

MPAM & Symbol Mapping

- Note: the average energy per-bit is constant
- Gray coding used for mapping bits to symbols
  - Why? Most likely error is to confuse with neighboring symbol.
  - Make sure that the neighboring symbol has only 1-bit difference (hamming distance = 1)



MPAM: Details

$$s_i(t) = \Re\{A_i g(t) e^{j2\pi f_c t}\} = A_i g(t)\cos(2\pi f_c t), \quad 0 \le t \le T_s >> 1/f_c,$$

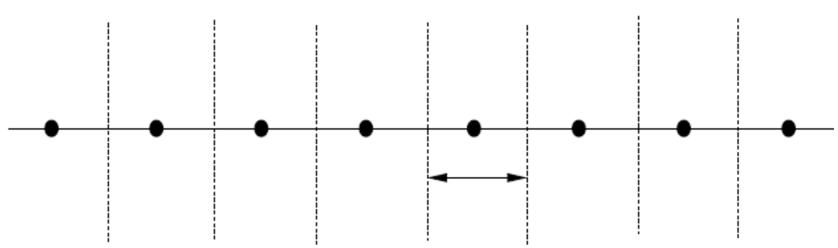
Unequal energies/symbol:

$$E_{s_i} = \int_0^{T_s} s_i^2(t)dt = \int_0^{T_s} A_i^2 g^2(t)\cos^2(2\pi f_c t)dt = A_i^2 \qquad \overline{E_s} = \frac{1}{M}\sum_{i=1}^{M} A_i^2.$$

Example 5.4:
For $g(t) = \sqrt{2/T_s}, 0 \le t < T_s$ a rectangular pulse shape, find the average energy of 4PAM modulation.

Solution: For 4PAM the A_i values are $A_i = \{-3d, -d, d, 3d\}$, so the average energy is

$$\overline{E_s} = \frac{d^2}{4}(9+1+1+9) = 5d^2.$$

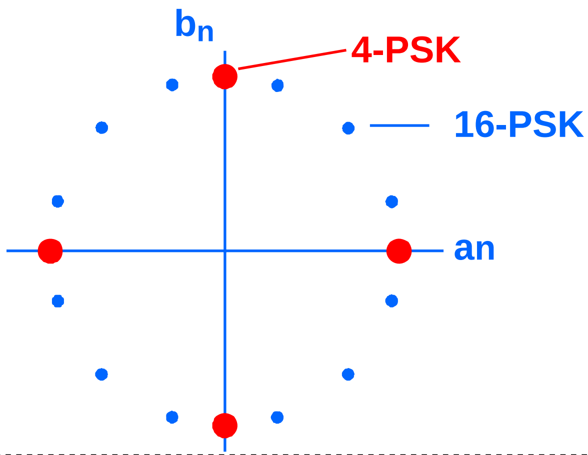


Decision Regions

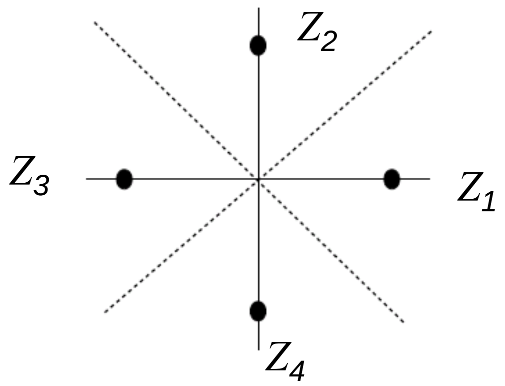
M-PSK (Circular Constellations)

MPSK:

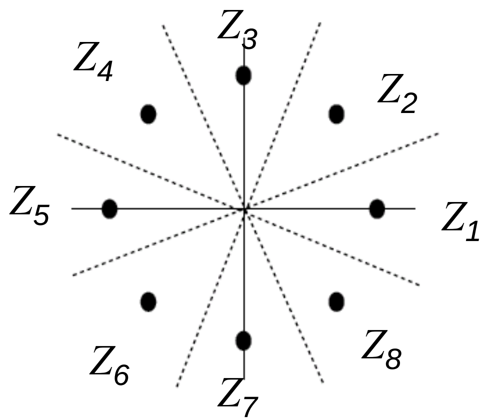
$$
\begin{aligned}
s_i(t) &= \Re\{Ag(t)e^{j2\pi(i-1)/M}e^{j2\pi f_c t}\}, \quad 0 \le t \le T_s \\
&= Ag(t)\cos\left[2\pi f_c t + \frac{2\pi(i-1)}{M}\right] \\
&= Ag(t)\cos\left[\frac{2\pi(i-1)}{M}\right]\cos 2\pi f_c t - Ag(t)\sin\left[\frac{2\pi(i-1)}{M}\right]\sin 2\pi f_c t.
\end{aligned}
$$



MPSK: Decision Regions & Demod'ln

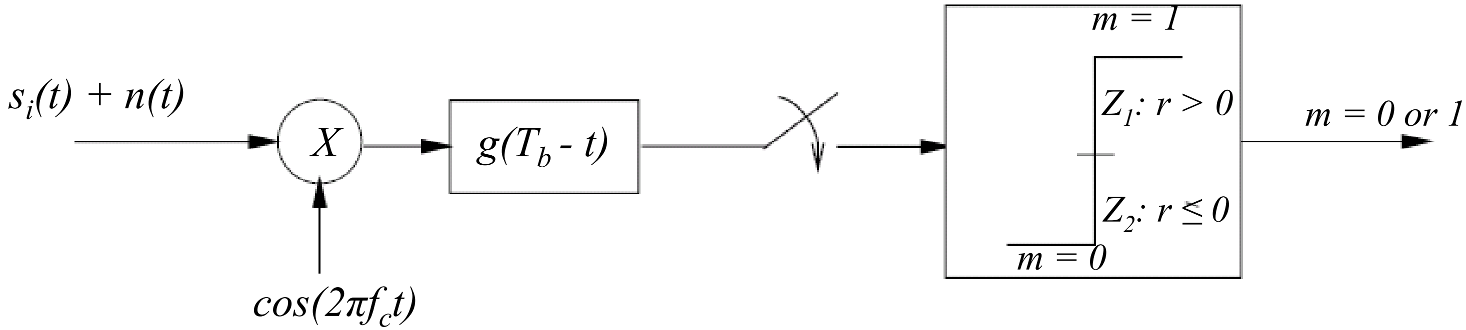


4PSK



8PSK

4PSK: 1 bit/complex dimension or 2 bits/symbol

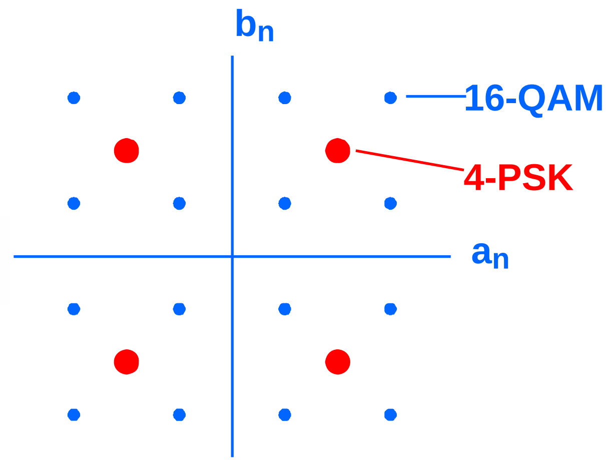


Coherent Demodulator for BPSK.

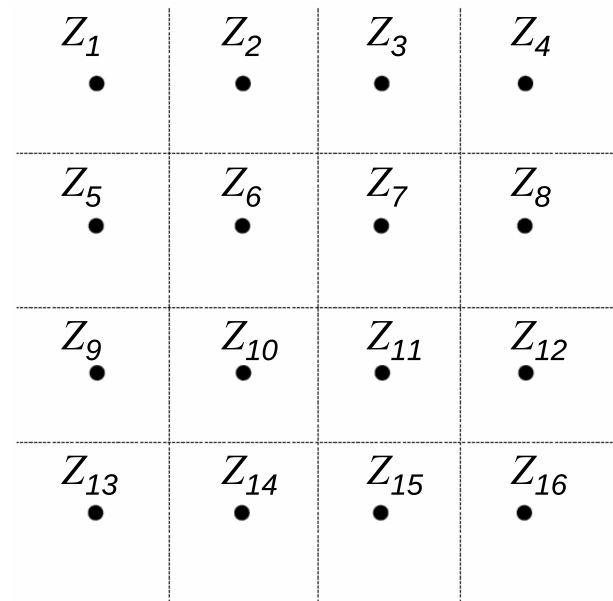
M-QAM (Square Constellations)

MQAM:

$$s_i(t) = \Re\{A_i e^{j\theta_i} g(t) e^{j2\pi f_c t}\}$$
$$= A_i \cos(\theta_i) g(t) \cos(2\pi f_c t) - A_i \sin(\theta_i) g(t) \sin(2\pi f_c t), \quad 0 \le t \le T_s.$$



- Unequal symbol energies: $E_{s_i} = \int_0^{T_s} s_i^2(t) = A_i^2$
- MQAM with square constellations of size L^2 is equivalent to MPAM modulation with constellations of size L on each of the in-phase and quadrature signal components
- For square constellations it takes approximately 6 dB more power to send an additional 1 bit/dimension or 2 bits/symbol while maintaining the same minimum distance between constellation points
- Hard to find a Gray code mapping where all adjacent symbols differ by a single bit



16QAM: Decision Regions

Non-Coherent Modulation: DPSK

- Information in MPSK, MQAM carried in signal phase.
- Requires coherent demodulation: i.e. phase of the transmitted signal carrier φ_0 must be matched to the phase of the receiver carrier φ
- More cost, susceptible to carrier phase drift.
- Harder to obtain in fading channels

- **<u>Differential modulation</u>**: do not require phase reference.
  - More general: modulation w/ memory: depends upon prior symbols transmitted.
  - Use prev symbol as the a phase reference for current symbol
  - Info bits encoded as the differential phase between current & previous symbol
  - Less sensitive to carrier phase drift (f-domain) ; more sensitive to doppler effects: decorrelation of signal phase in time-domain

Example 5.5:

Find the sequence of symbols transmitted using DPSK for the bit sequence 101110 starting at the kth symbol time, assuming the transmitted symbol at the $(k-1)$th symbol time was $s(k-1) = Ae^{j\pi}$.

Solution: The first bit, a 1, results in a phase transition of π, so $s(k) = A$. The next bit, a 0, results in no transition, so $s(k+1) = A$. The next bit, a 1, results in another transition of π, so $s(k+1) = Ae^{j\pi}$, and so on. The full symbol sequence corresponding to 101110 is $A, A, Ae^{j\pi}, A, Ae^{j\pi}, Ae^{j\pi}$.

Differential Modulation (Contd)

- **DPSK**: Differential BPSK
- A 0 bit is encoded by no change in phase, whereas a 1 bit is encoded as a phase change of π.
  - If symbol over time $[(k-1)T_s, kT_s)$ has phase $\theta(k-1) = e^{j\theta i}$, $\theta_i = 0, \pi$,
  - then to encode a 0 bit over $[kTs, (k+1)Ts)$, the symbol would have
    - phase: $\theta(k) = e^{j\theta i}$ and…
  - … to encode a 1 bit the symbol would have
    - phase $\theta(k) = e^{j(\theta i+\pi)}$.

- **DQPSK**: gray coding:

Bit Sequence	Phase Transition
00	0
01	$\pi/2$
10	$-\pi/2$
11	π

Table 5.1: Mapping for D-QPSK with Gray Encoding

Quadrature Offset

- Phase transitions of 180^o can cause *large amplitude transitions (through zero point).*
  - Abrupt phase transitions and large amplitude variations can be distorted by nonlinear amplifiers and filters

- Avoided by offsetting the quadrature branch pulse $g(t)$ by half a symbol period
  - Usually abbreviated as O-MPSK, where the O indicates the offset
- QPSK modulation with quadrature offset is referred to as O-QPSK
- O-QPSK has the same spectral properties as QPSK for linear amplification,..
  - … but has higher spectral efficiency under nonlinear amplification,
  - since the maximum phase transition of the signal is 90 degrees

- Another technique to mitigate the amplitude fluctuations of a 180 degree phase shift used in the IS-54 standard for digital cellular is $\pi/4$-QPSK
  - Maximum phase transition of 135 degrees, versus 90 degrees for offset QPSK and 180 degrees for QPSK

Offset QPSK waveforms

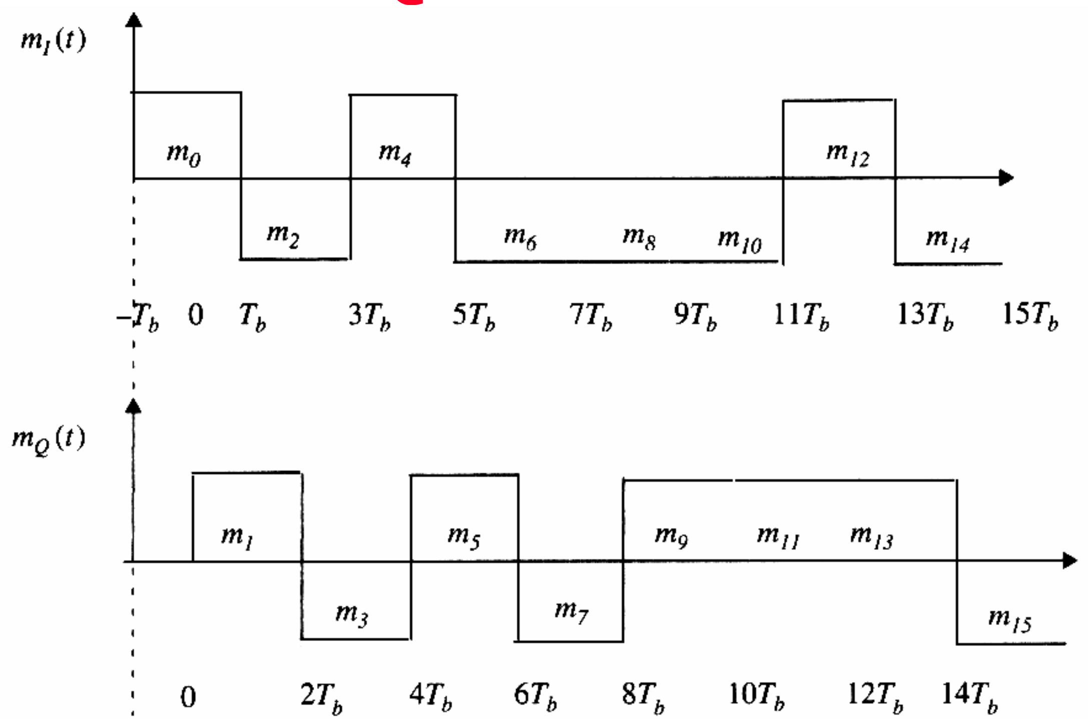
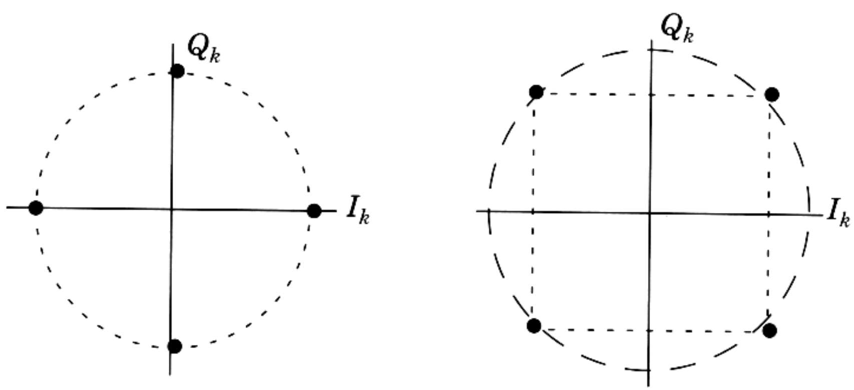


Figure 6.30 The time offset waveforms that are applied to the in-phase and quadrature arms of an OQPSK modulator. Notice that a half-symbol offset is used.



(a) (b)

Frequency Shift Keying (FSK)

- **Continuous Phase FSK (CPFSK)**
  - digital data encoded in the frequency shift
  - typically implemented with frequency modulator to maintain continuous phase

    $$s(t) = A \cos [\omega_c t + 2 \pi k_f \int_{-\infty}^{t} d(\tau)\, d\tau]$$

  - nonlinear modulation but constant-envelope
- **Minimum Shift Keying (MSK)**
  - minimum bandwidth, sidelobes large
  - can be implemented using I-Q receiver
- **Gaussian Minimum Shift Keying (GMSK)**
  - reduces sidelobes of MSK using a premodulation filter
  - used by RAM Mobile Data, CDPD, and HIPERLAN

Minimum Shift Keying (MSK) spectra

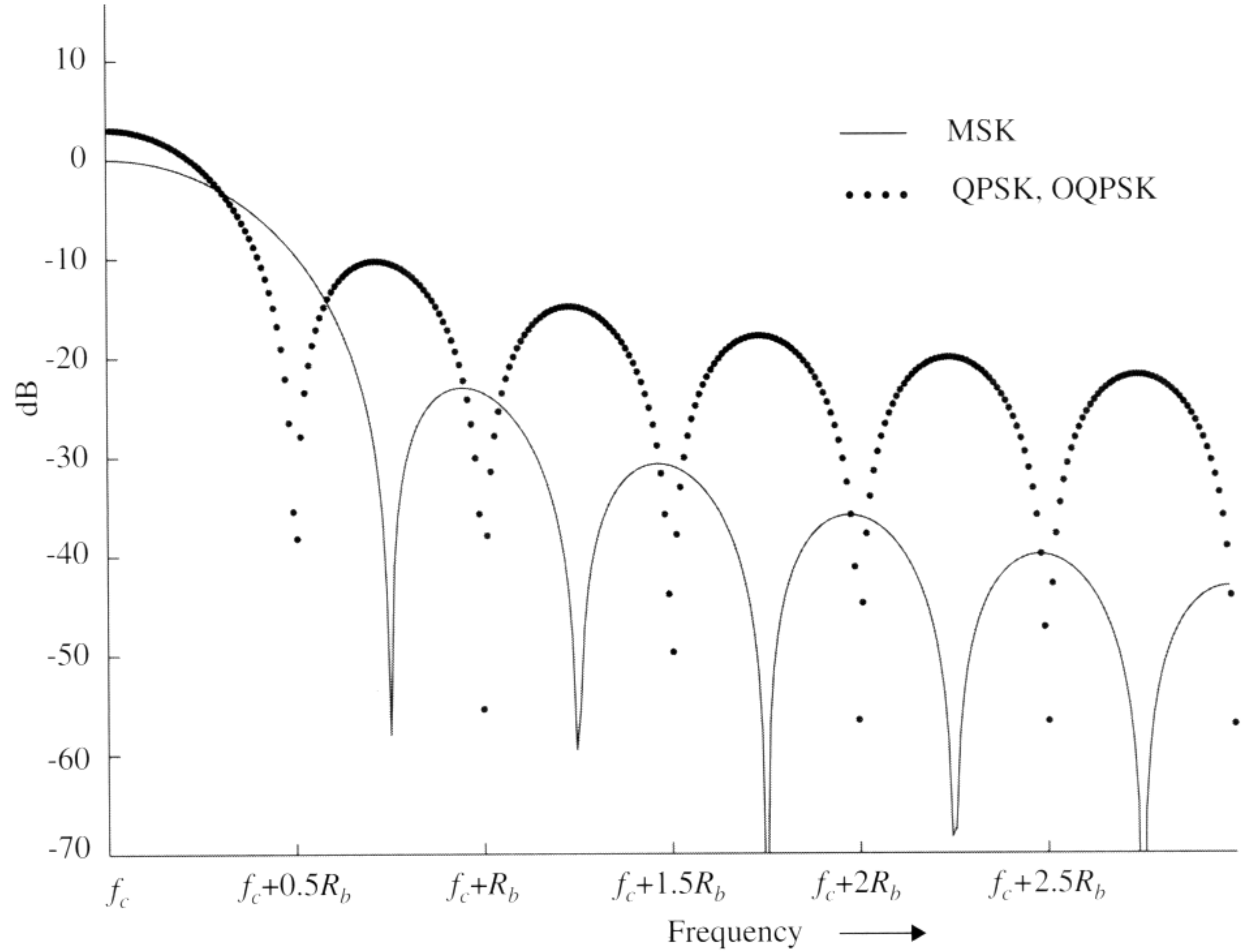
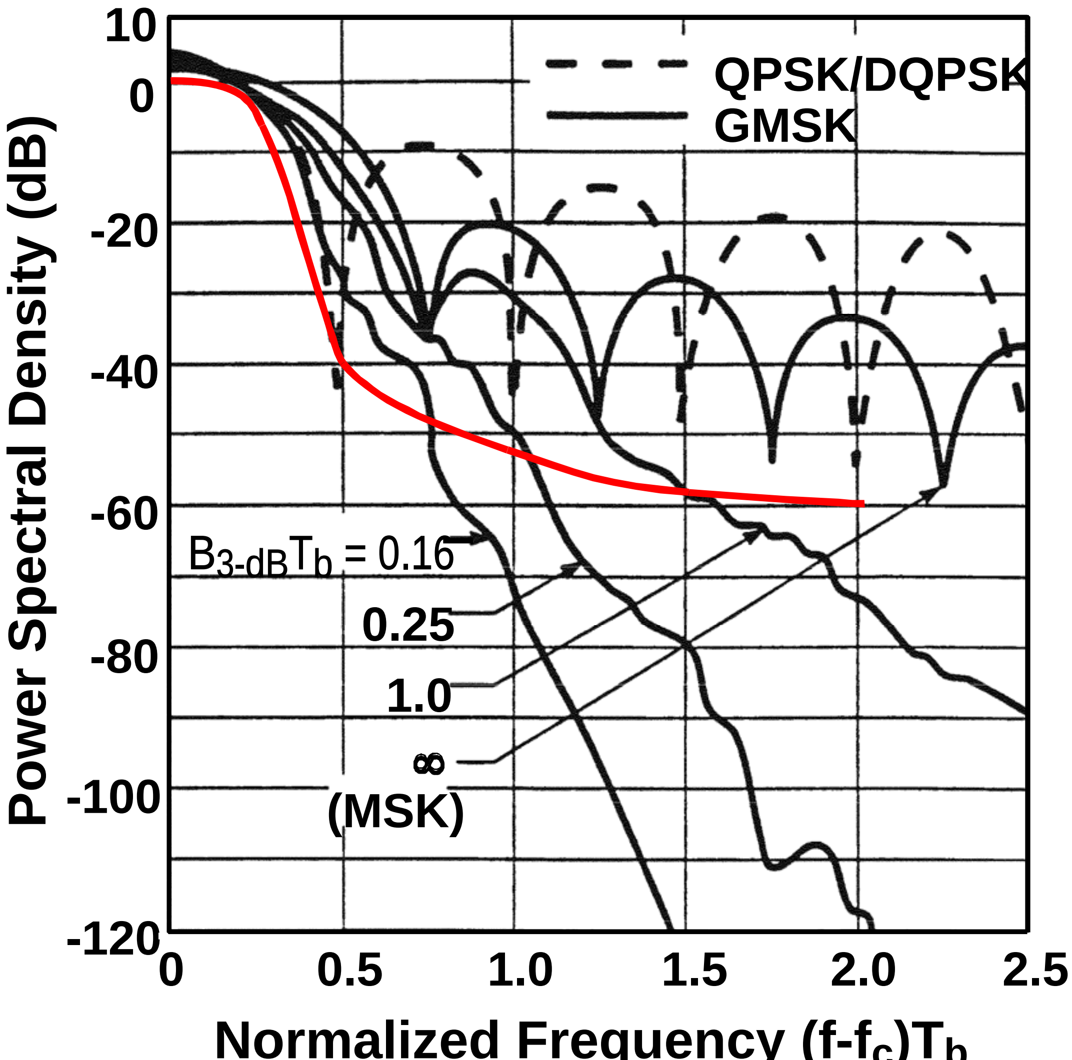
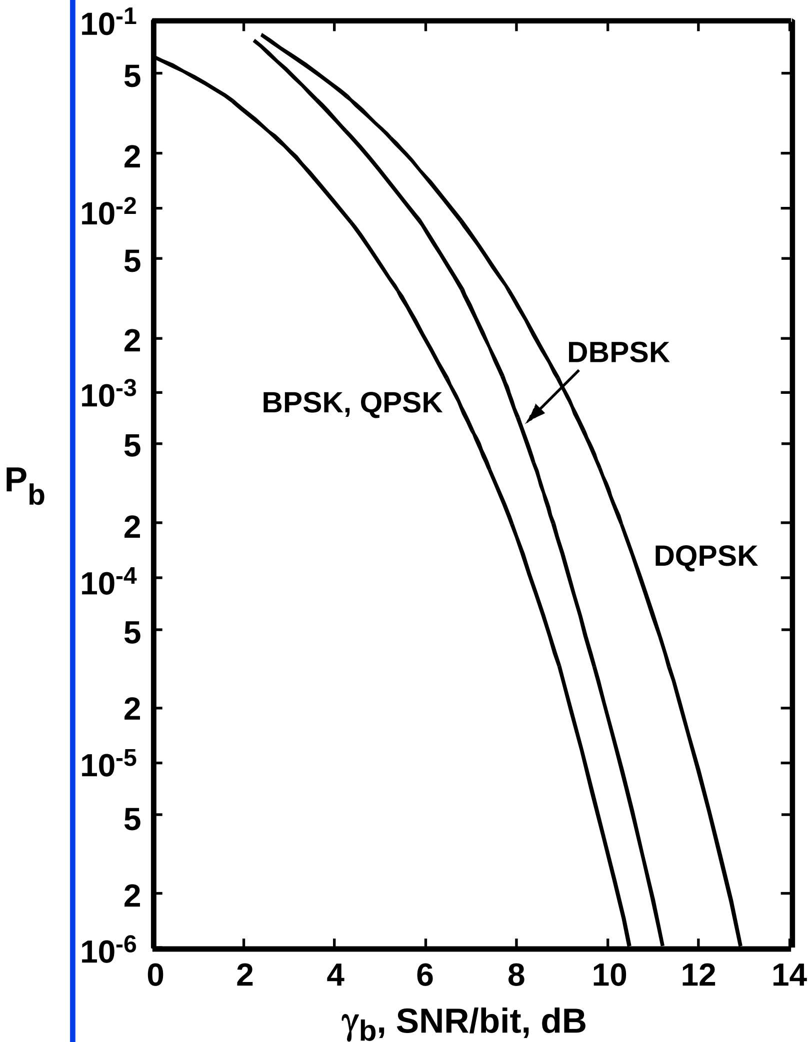


Figure 6.38 Power spectral density of MSK signals as compared to QPSK and OQPSK signals.

Spectral Characteristics



Bit Error Probability (BER): AWGN

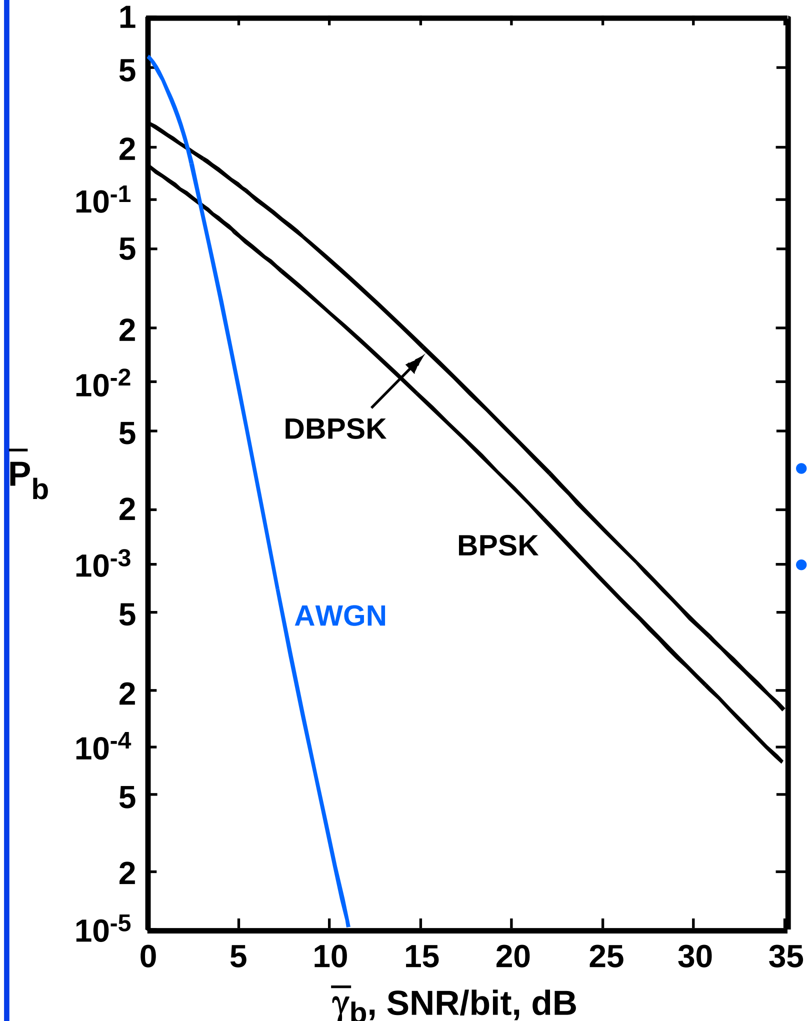


For $P_b = 10^{-3}$

BPSK 6.5 dB
QPSK 6.5 dB
DBPSK ~8 dB
DQPSK ~9 dB

- QPSK is more spectrally efficient than BPSK with the same performance.
- M-PSK, for M>4, is more spectrally efficient but requires more SNR per bit.
- There is ~3 dB power penalty for differential detection.

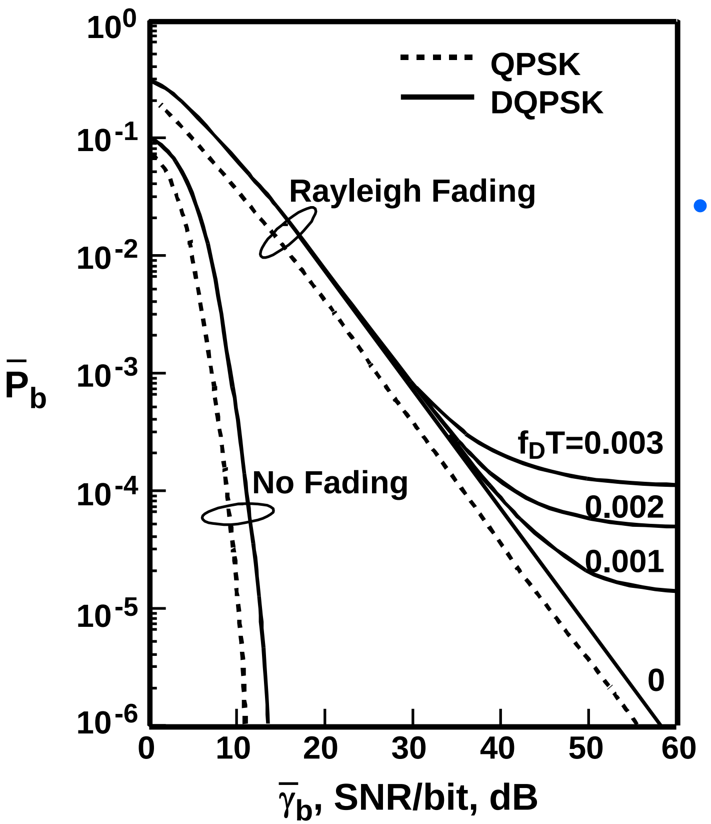
Bit Error Probability (BER): Fading Channel



- $\bar{P}_b$ is inversely proportion to the average SNR per bit.
- Transmission in a fading environment requires about 18 dB more power for $\bar{P}_b = 10^{-3}$.

Bit Error Probability (BER): Doppler Effects

- Doppler causes an irreducible error floor when differential detection is used $\Rightarrow$ decorrelation of reference signal.



- The irreducible $\bar{P}_b$ depends on the data rate and the Doppler. For f_D = 80 Hz,

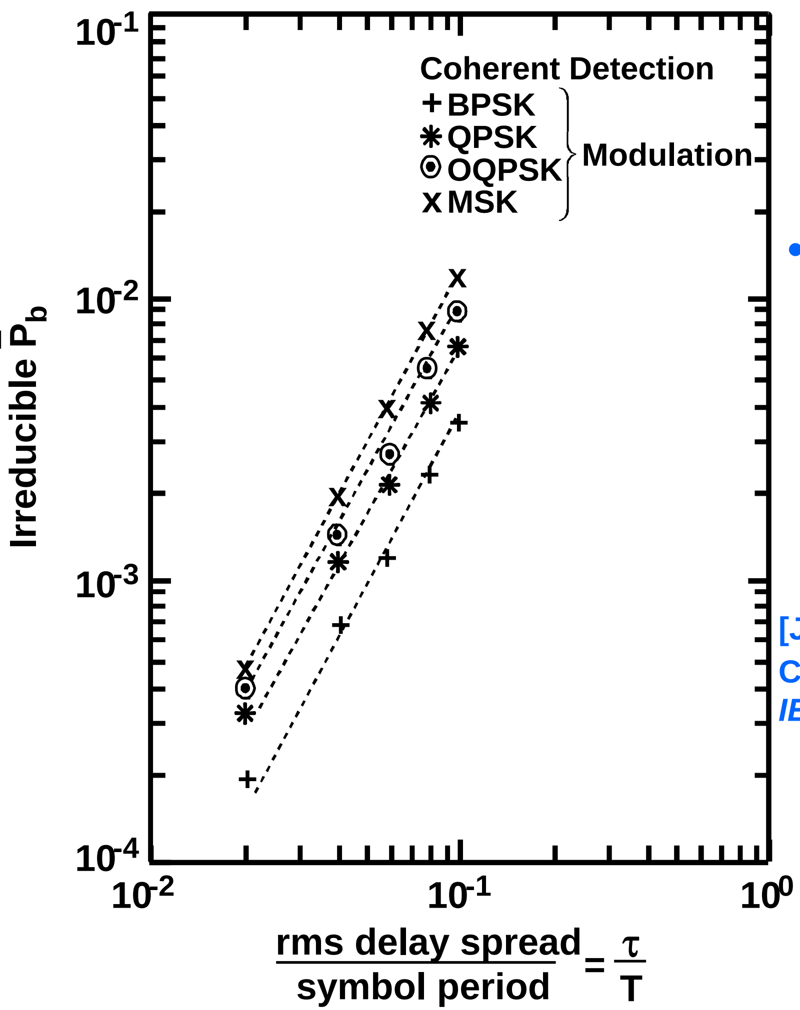
data rate	T	$\bar{P}_{b_{floor}}$
10 kbps	10^{-4}s	3×10^{-4}
100 kbps	10^{-5}s	3×10^{-6}
1 Mbps	10^{-6}s	3×10^{-8}

The implication is that Doppler is ***not an issue for high-speed wireless data.***

[M. D. Yacoub, *Foundations of Mobile Radio Engineering* , CRC Press, 1993]

Bit Error Probability (BER): Delay Spread

- ISI causes an irreducible error floor.



- The rms delay spread imposes a limit on the maximum bit rate in a multipath environment.

For example, for QPSK,

	τ	Maximum Bit Rate
Mobile (rural)	25 μsec	8 kbps
Mobile (city)	2.5 μsec	80 kbps
Microcells	500 nsec	400 kbps
Large Building	100 nsec	2 Mbps

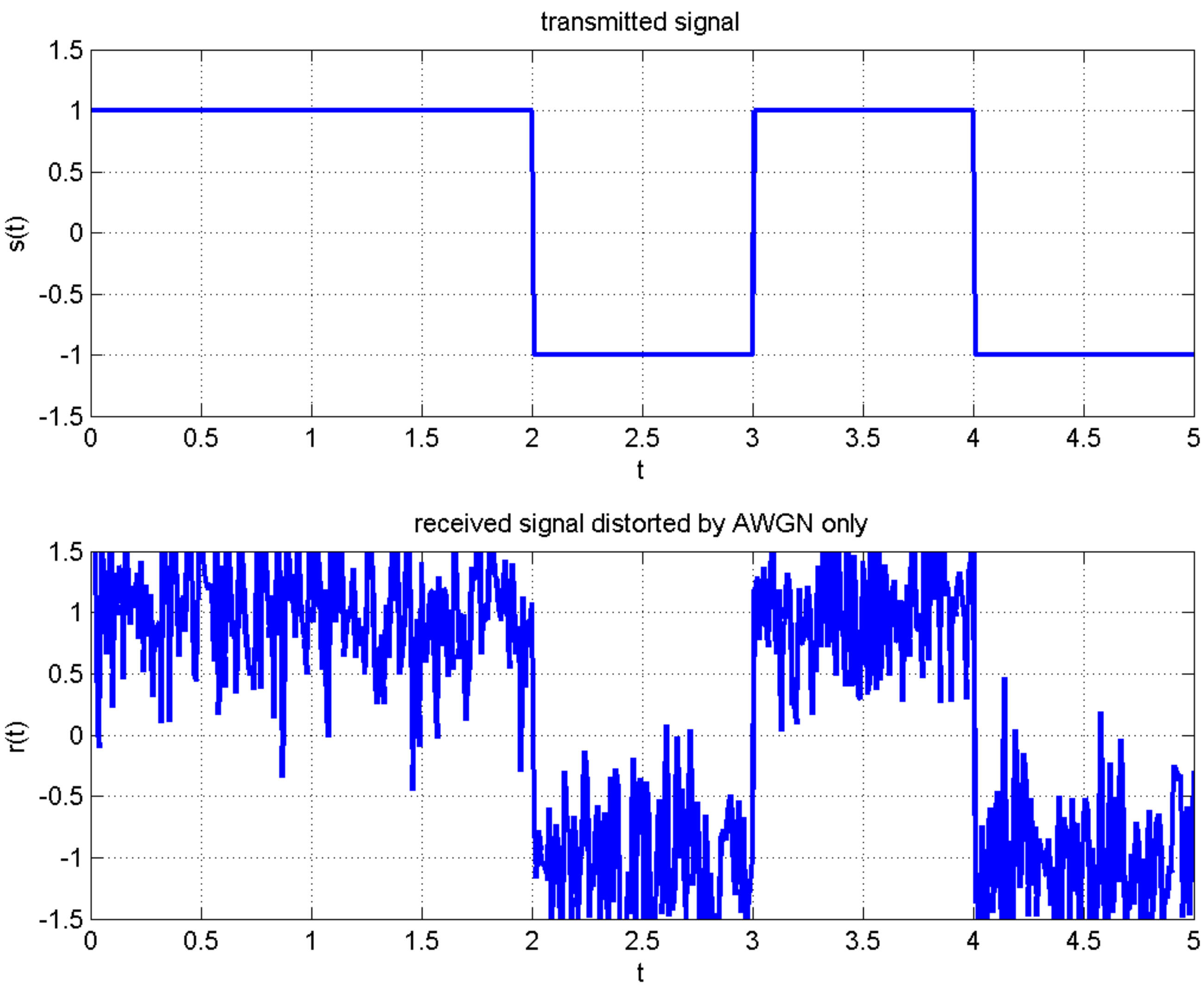
[J. C.-I. Chuang, "The Effects of Time Delay Spread on Portable Radio Communications Channels with Digital Modulation," *IEEE JSAC*, June 1987]

Summary of Modulation Issues

- Tradeoffs
  - linear versus nonlinear modulation
  - constant envelope versus non-constant envelope
  - coherent versus differential detection
  - power efficiency versus spectral efficiency
- Limitations
  - flat fading
  - doppler
  - delay spread

Pulse Shaping

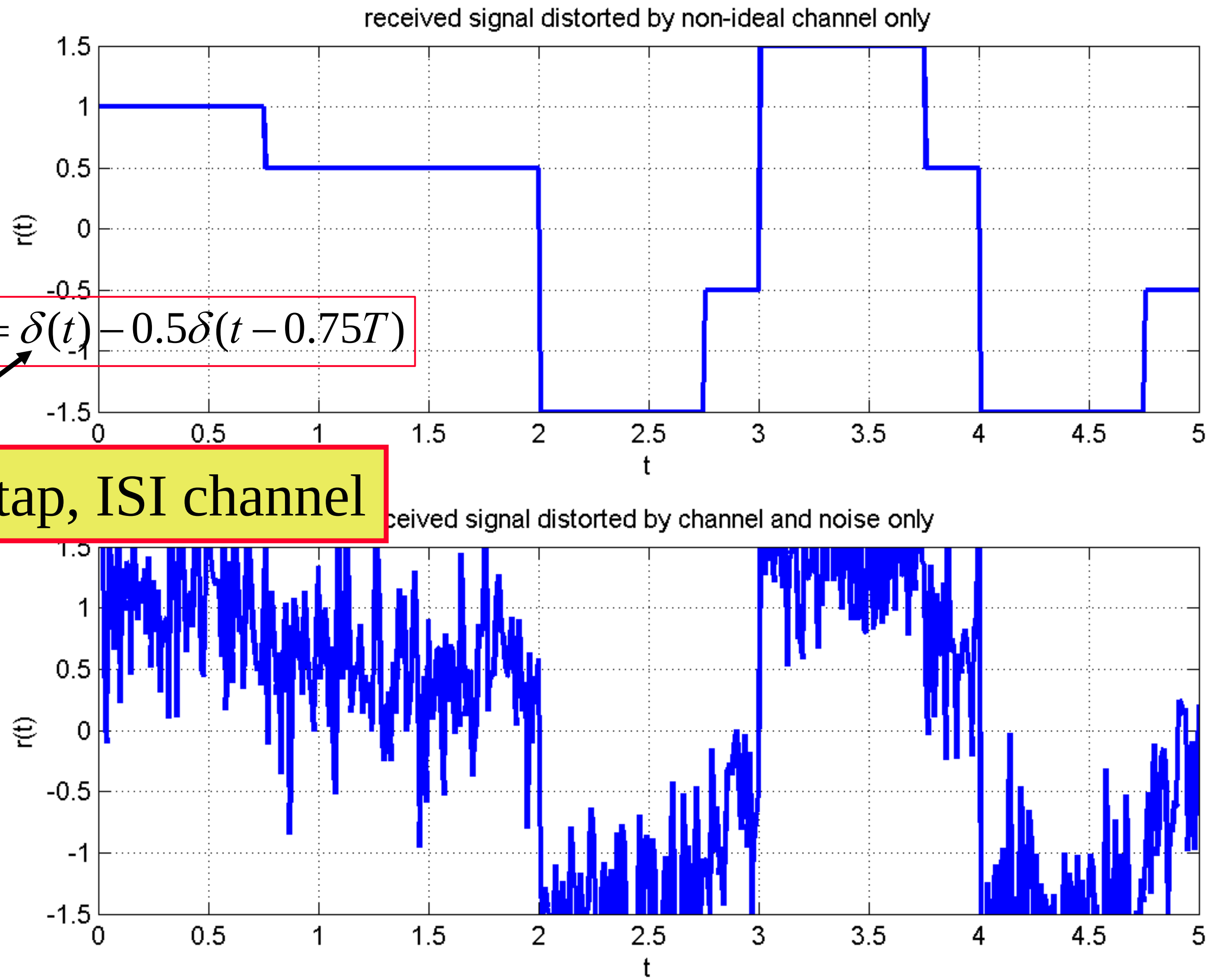
Recall: Impact of AWGN only



Impact of AWGN & Channel Distortion

$$h_c(t) = \delta(t) - 0.5\delta(t - 0.75T)$$

Multi-tap, ISI channel



ISI Effects: Band-limited Filtering of Channel

- ISI due to filtering effect of the communications channel (e.g. wireless channels)
  - Channels behave like band-limited filters

$$H_c(f) = \underbrace{\left|H_c(f)\right|} e^{j\underbrace{\theta_c(f)}}$$

Non-constant amplitude

Amplitude distortion

Non-linear phase

Phase distortion

Inter-Symbol Interference (ISI)

- ISI in the detection process due to the filtering effects of the system
- Overall equivalent system transfer function

$$H(f) = H_t(f)H_c(f)H_r(f)$$

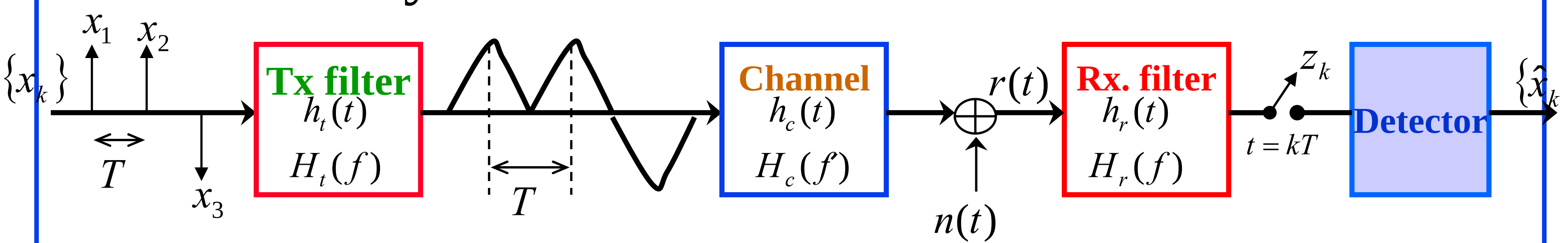
  - creates echoes and hence time dispersion
  - causes ISI at <u>sampling time</u>

ISI effect

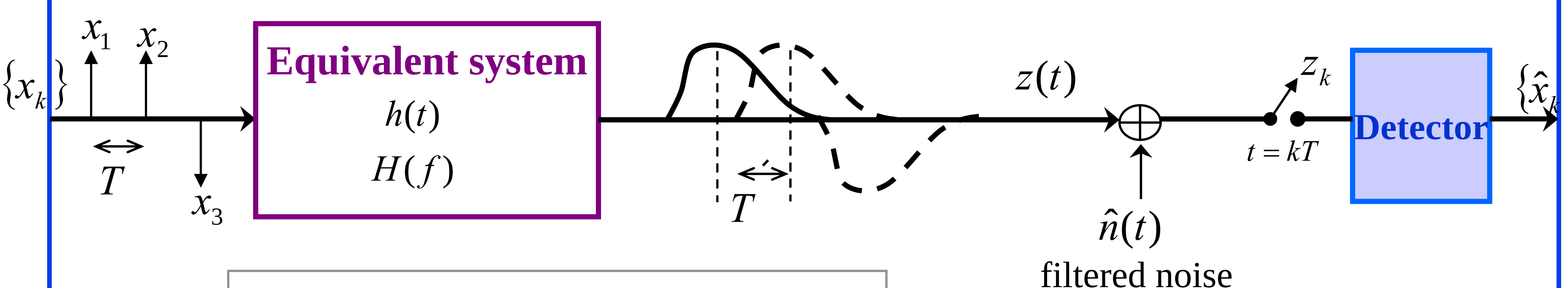
$$z_k = s_k + n_k + \sum_{i \neq k} \alpha_i s_i$$

Inter-symbol interference (ISI): Model

- Baseband system model



- Equivalent model



$$H(f) = H_t(f) H_c(f) H_r(f)$$

Nyquist bandwidth constraint

- Nyquist bandwidth constraint (*on equivalent system*):
  - The theoretical minimum required system bandwidth to detect R_s [symbols/s] without ISI is $R_s/2$ [Hz].
  - Equivalently, a system with bandwidth $W=1/2T=R_s/2$ [Hz] can support a maximum transmission rate of $2W=1/T=Rs$ [symbols/s] without ISI.

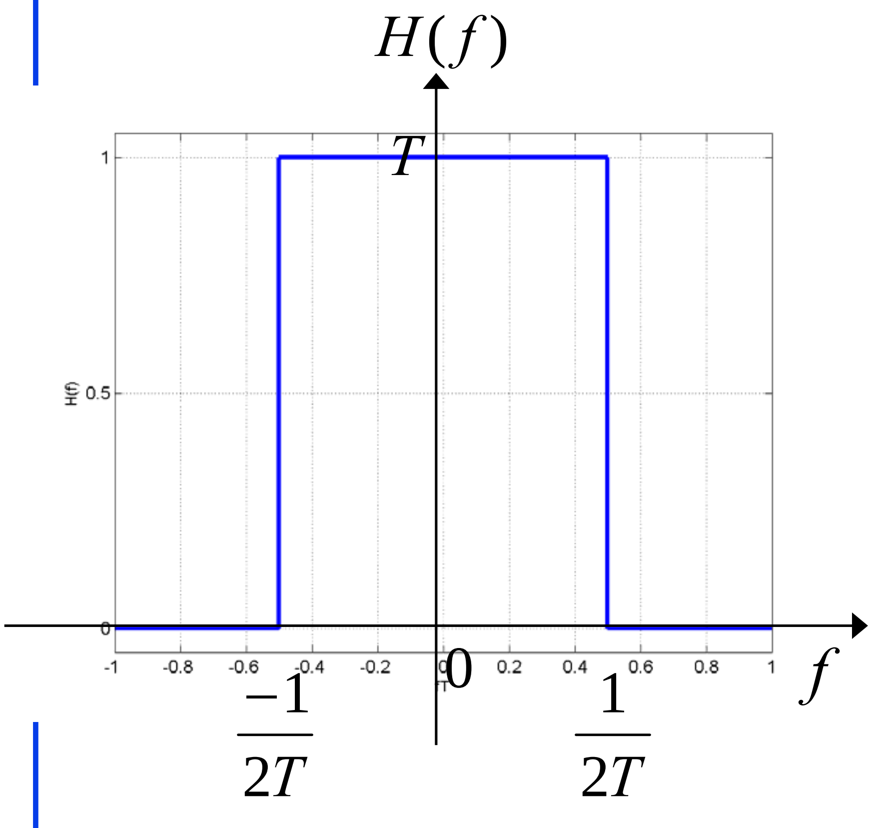
$$\boxed{\frac{1}{2T}=\frac{R_s}{2}\leq W \Rightarrow \frac{R_s}{W}\geq 2 \quad \text{[symbol/s/Hz]}}$$

- Bandwidth efficiency, R/W [bits/s/Hz] :
  - An important measure in DCs representing data throughput per hertz of bandwidth.
  - Showing how efficiently the bandwidth resources are used by signaling techniques.

Equiv System: Ideal Nyquist pulse (filter)

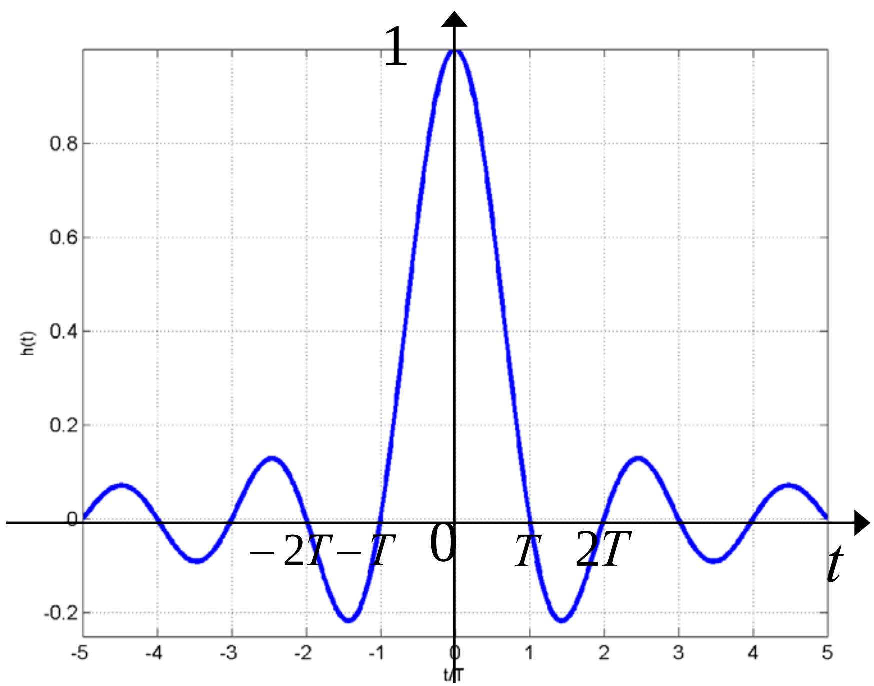
Ideal Nyquist filter

Ideal Nyquist pulse



$$W = \frac{1}{2T}$$

$$h(t) = \mathrm{sinc}(t/T)$$



Rensselaer Polytechnic Institute

Shivkumar Kalyanaraman

Nyquist pulses (filters)

- Nyquist pulses (filters):
  - Pulses (filters) which result in no ISI at the <u>sampling time</u>.
- Nyquist filter:
  - Its transfer function in frequency domain is obtained by convolving a rectangular function with any real even-symmetric frequency function
- Nyquist pulse:
  - Its shape can be represented by a sinc(t/T) function multiply by another time function.
- Example of Nyquist filters: Raised-Cosine filter

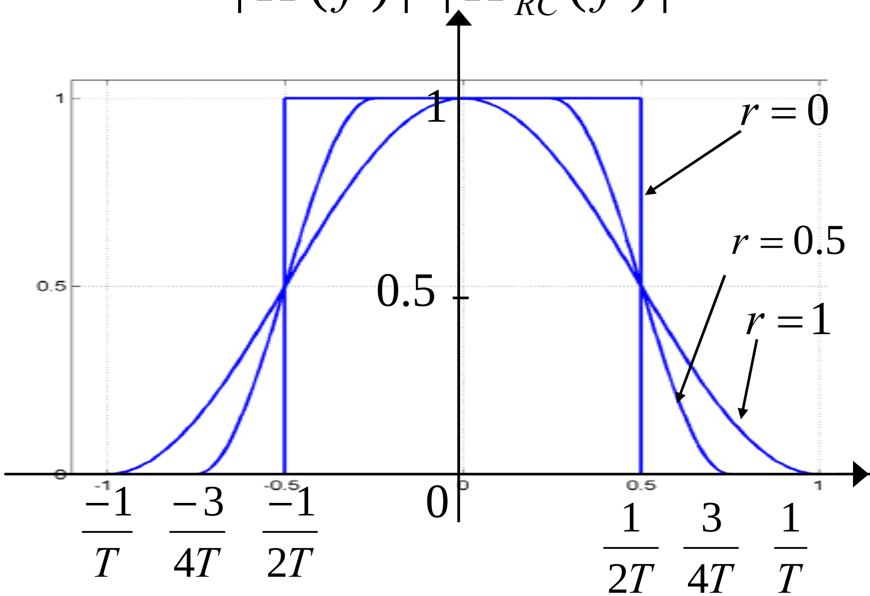
Pulse shaping to reduce ISI

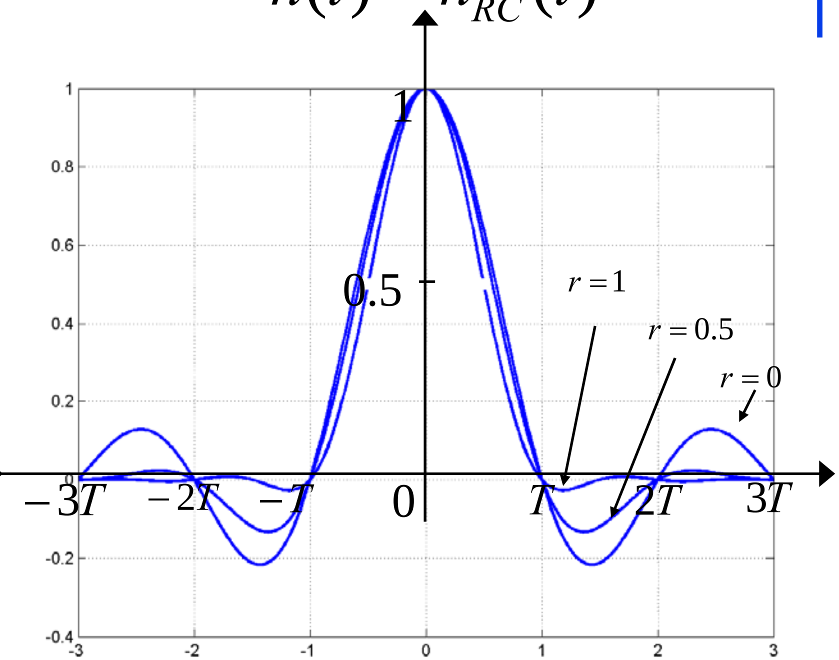
- Goals and trade-off in pulse-shaping
  - Reduce ISI
  - Efficient bandwidth utilization
  - Robustness to timing error (small side lobes)

Raised Cosine Filter: Nyquist Pulse Approximation

$$|H(f)| = |H_{RC}(f)|$$

$$h(t) = h_{RC}(t)$$





Baseband $W_{sSB} = (1+r)\dfrac{R_s}{2}$

Passband $W_{DSB} = (1+r)R_s$

Raised Cosine Filter

- Raised-Cosine Filter
  - A Nyquist pulse (No ISI at the sampling time)

$$H(f) = \begin{cases} 1 & \text{for } |f| < 2W_0 - W \\ \cos^2\left[\dfrac{\pi}{4}\dfrac{|f| + W - 2W_0}{W - W_0}\right] & \text{for } 2W_0 - W < |f| < W \\ 0 & \text{for } |f| > W \end{cases}$$

$$h(t) = 2W_0(\mathrm{sinc}(2W_0 t))\frac{\cos[2\pi(W - W_0)t]}{1 - [4(W - W_0)t]^2}$$

Excess bandwidth: $\boxed{W - W_0}$

Roll-off factor $\boxed{r = \dfrac{W - W_0}{W_0}}$

$0 \le r \le 1$

Pulse Shaping and Equalization Principles

No ISI at the sampling time

$$H_{\mathrm{RC}}(f) = H_t(f)H_c(f)H_r(f)H_e(f)$$

- *Square-Root Raised Cosine (SRRC)* filter and Equalizer

$$H_{\mathrm{RC}}(f) = H_t(f)H_r(f)$$

$$H_r(f) = H_t(f) = \sqrt{H_{\mathrm{RC}}(f)} = H_{\mathrm{SRRC}}(f)$$

Taking care of ISI caused by tr. filter

$$H_e(f) = \frac{1}{H_c(f)}$$

Taking care of ISI caused by channel

Pulse Shaping & Orthogonal Bases

With the basis set $\phi_1(t) = \sqrt{2/T}\cos(2\pi f_c t)$ and $\phi_2(t) = \sqrt{2/T}\sin(2\pi f_c t)$ the basis function representation (5.3) corresponds to the complex baseband representation of $s_i(t)$ in terms of its in-phase and quadrature components with an extra factor of $\sqrt{2/T}$:

$$s_i(t) = s_{i1}\sqrt{\frac{2}{T}}\cos(2\pi f_c t) + s_{i2}\sqrt{\frac{2}{T}}\sin(2\pi f_c t). \tag{5.10}$$

Note that the carrier basis functions may have an initial phase offset ϕ_0. The basis set may also include a baseband pulse-shaping filter $g(t)$ to improve the spectral characteristics of the transmitted signal:

$$s_i(t) = s_{i1}g(t)\cos(2\pi f_c t) + s_{i2}g(t)\sin(2\pi f_c t). \tag{5.11}$$

In this case the pulse shape $g(t)$ must maintain the orthonormal properties (5.5) of basis functions, i.e. we must have

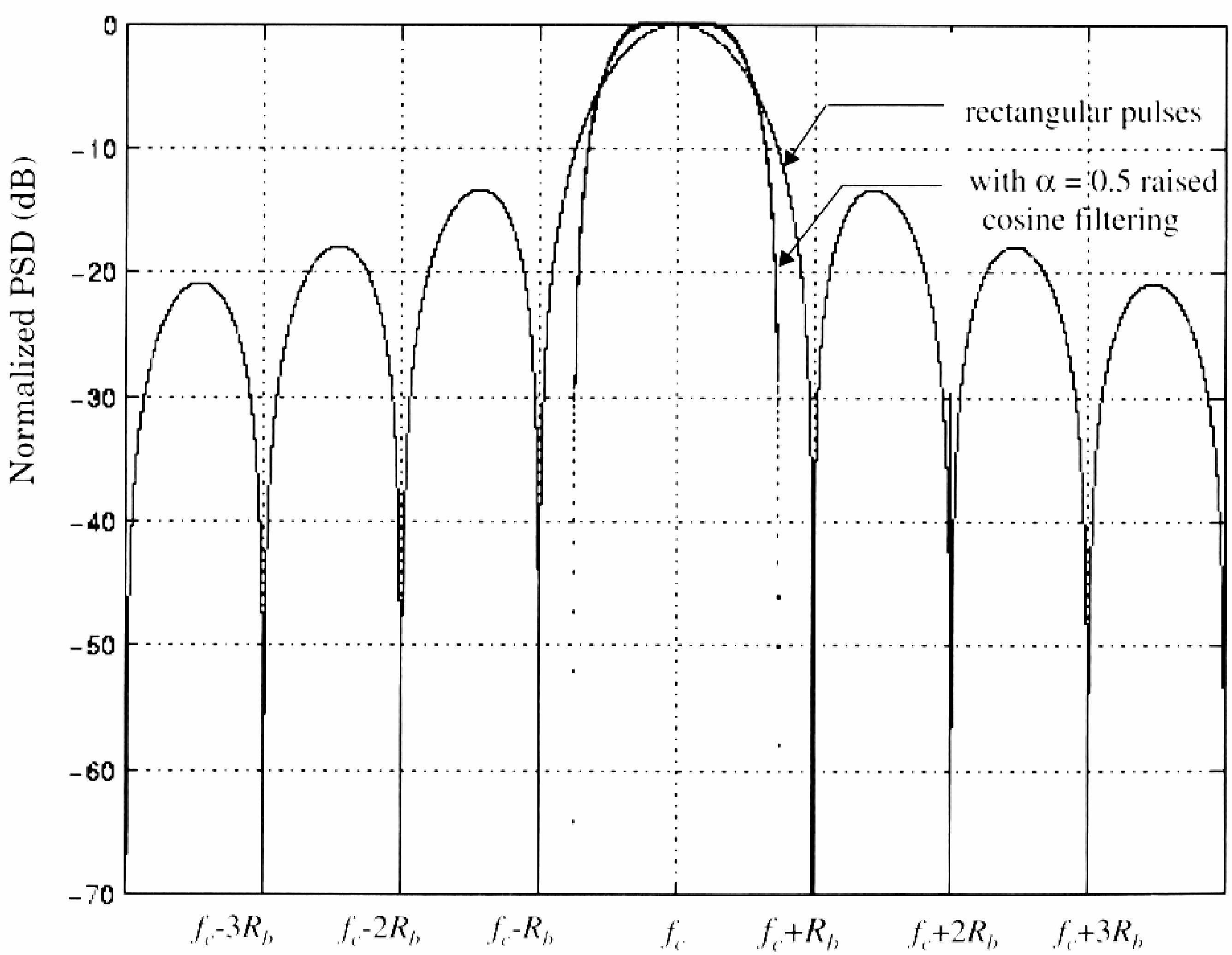
$$\int_0^T g^2(t)\cos^2(2\pi f_c t)dt = 1 \tag{5.12}$$

and

$$\int_0^T g^2(t)\cos(2\pi f_c t)\sin(2\pi f_c t) = 0, \tag{5.13}$$

where the equalities may be approximations for $f_c T >> 1$ as in (5.8) and (5.9) above. If the bandwidth of $g(t)$ satisfies $B << f_c$ then $g^2(t)$ is roughly constant over T_c, so (5.13) is approximately true since the sine and cosine functions are orthogonal over one period $T_c = 1/f_c$. The simplest pulse shape that satisfies (5.12) and (5.13) is the rectangular pulse shape $g(t) = \sqrt{2/T}, 0 \le t < T$.

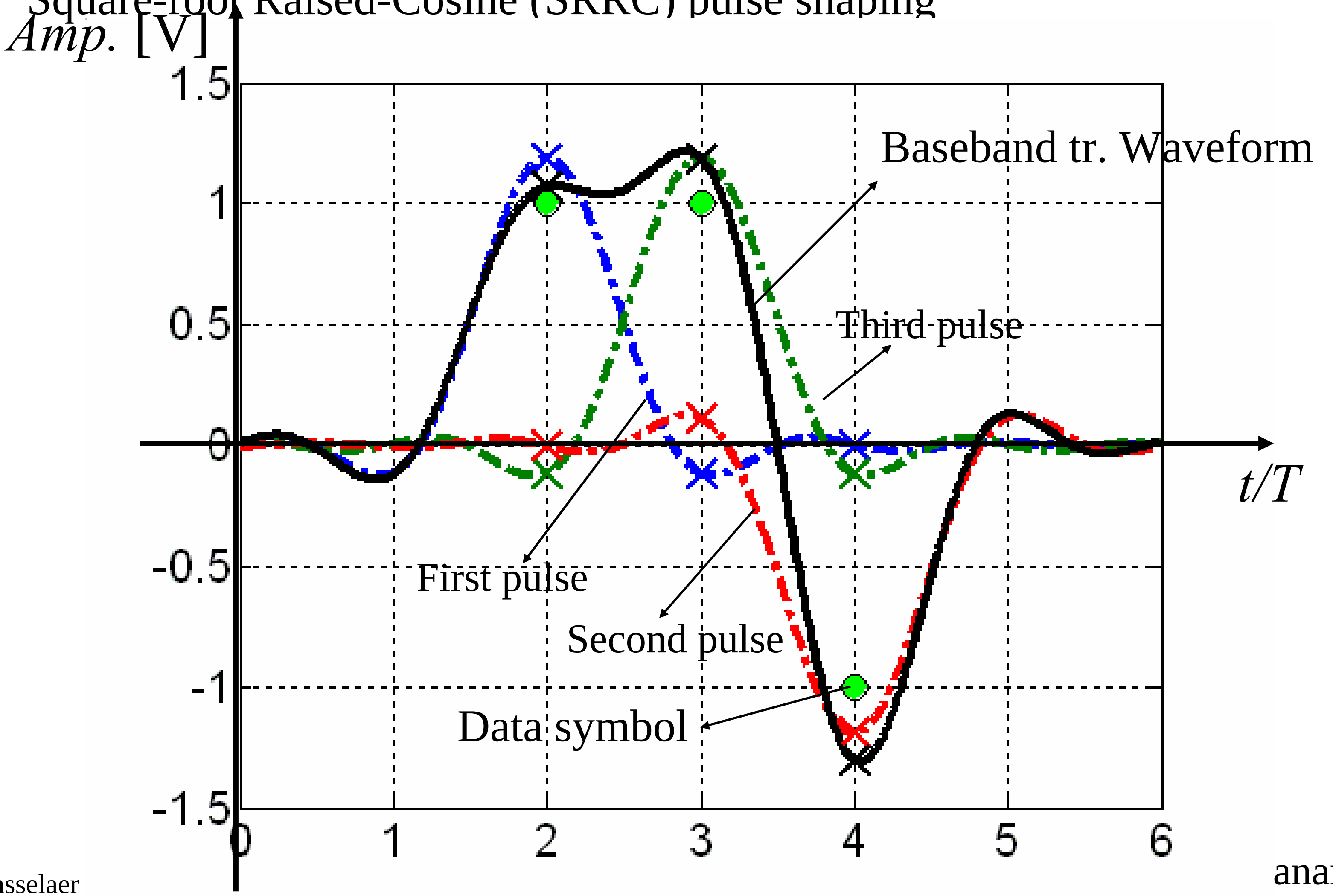
Virtue of pulse shaping



PSD of a BPSK signal

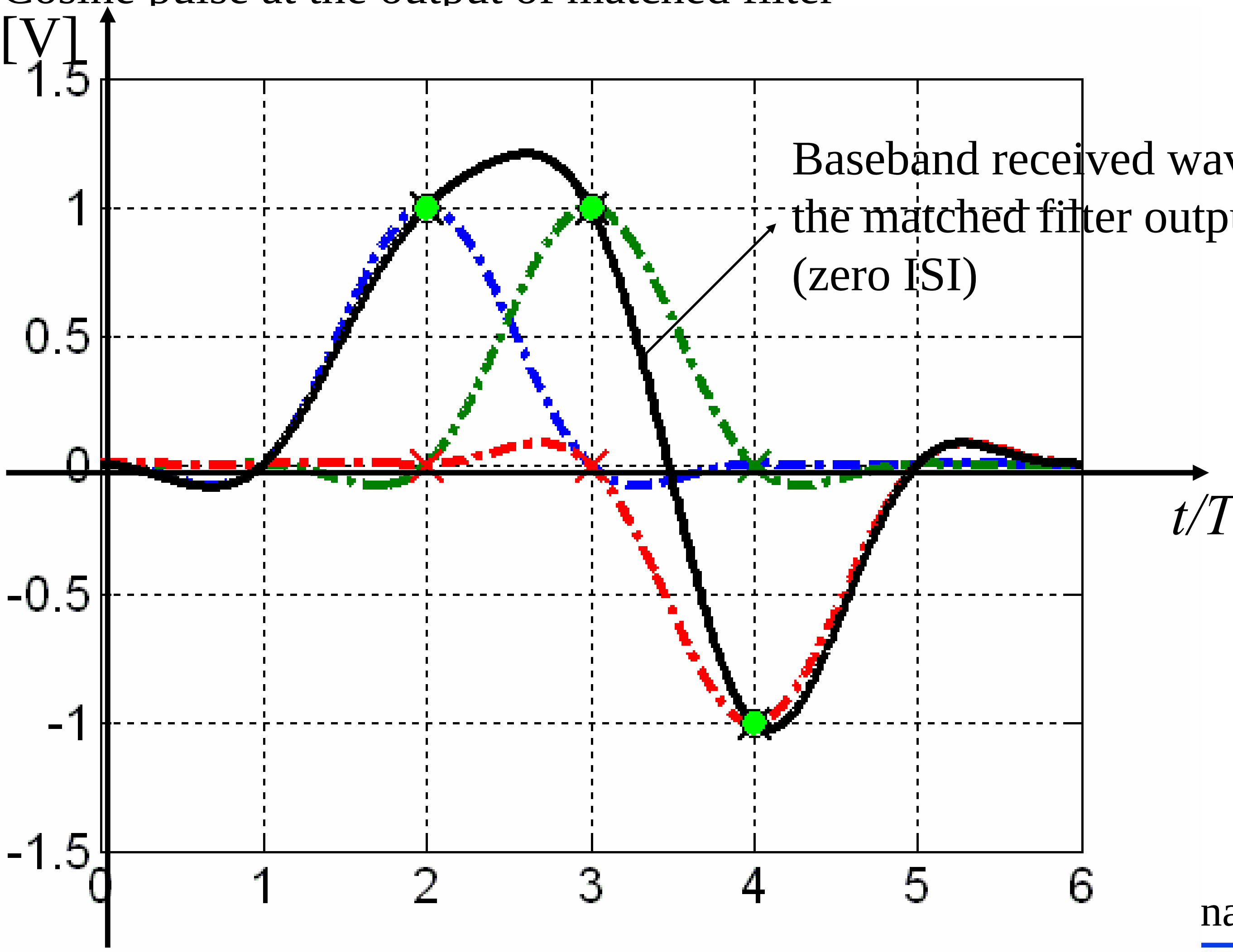
Example of pulse shaping

- Square-root Raised-Cosine (SRRC) pulse shaping



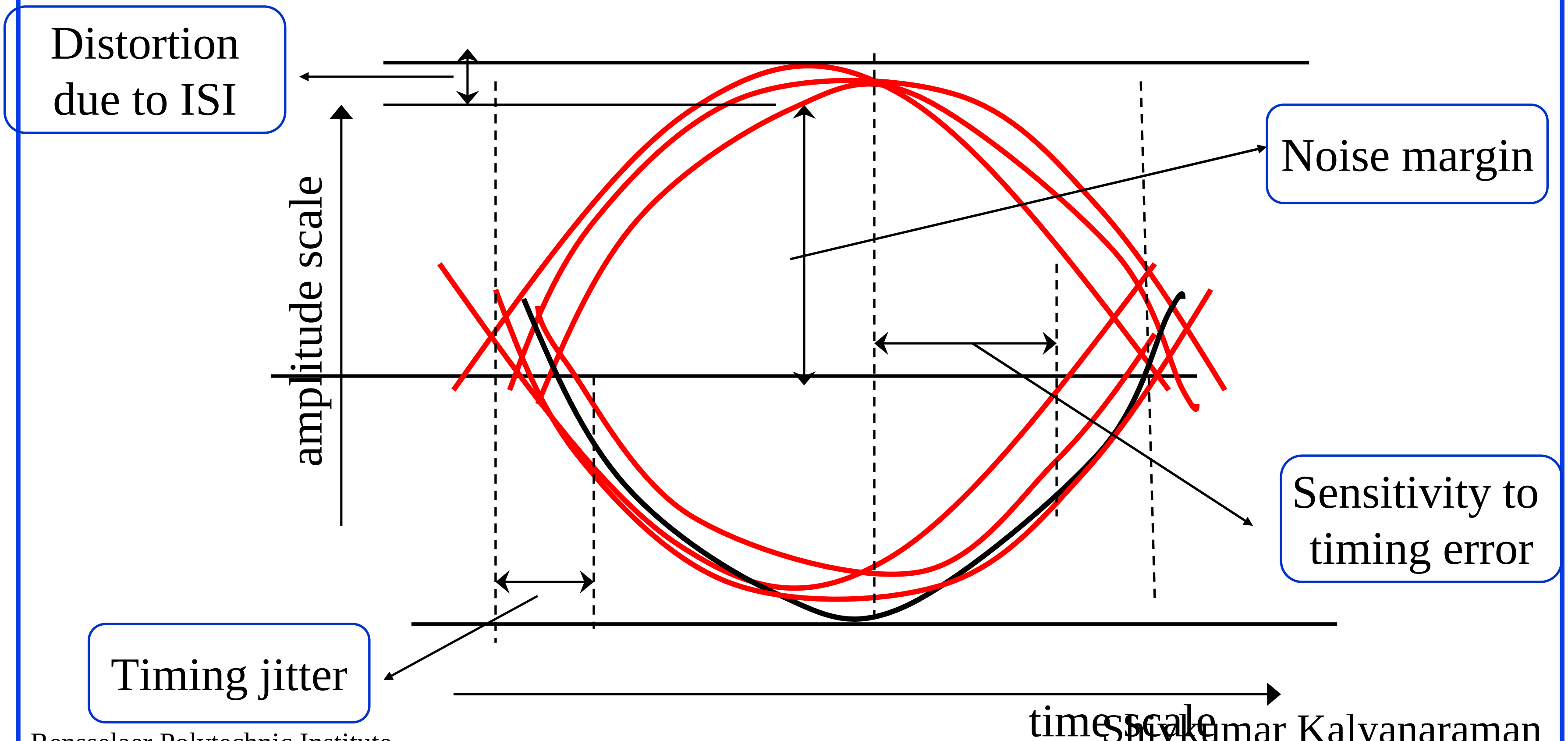
Example of pulse shaping …

- Raised Cosine pulse at the output of matched filter



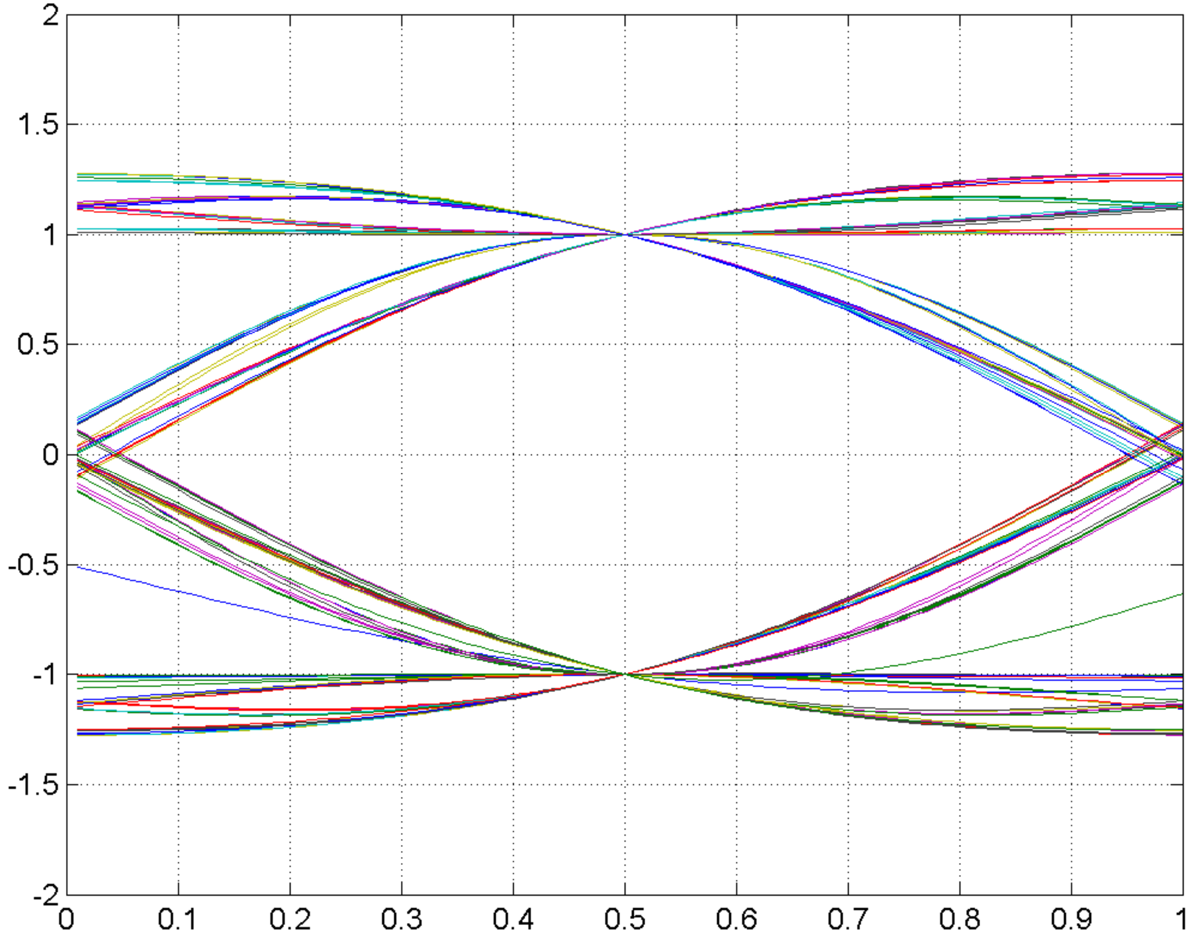
Eye pattern

- Eye pattern: Display on an oscilloscope which sweeps the system response to a baseband signal at the rate *1/T* (*T* symbol duration)



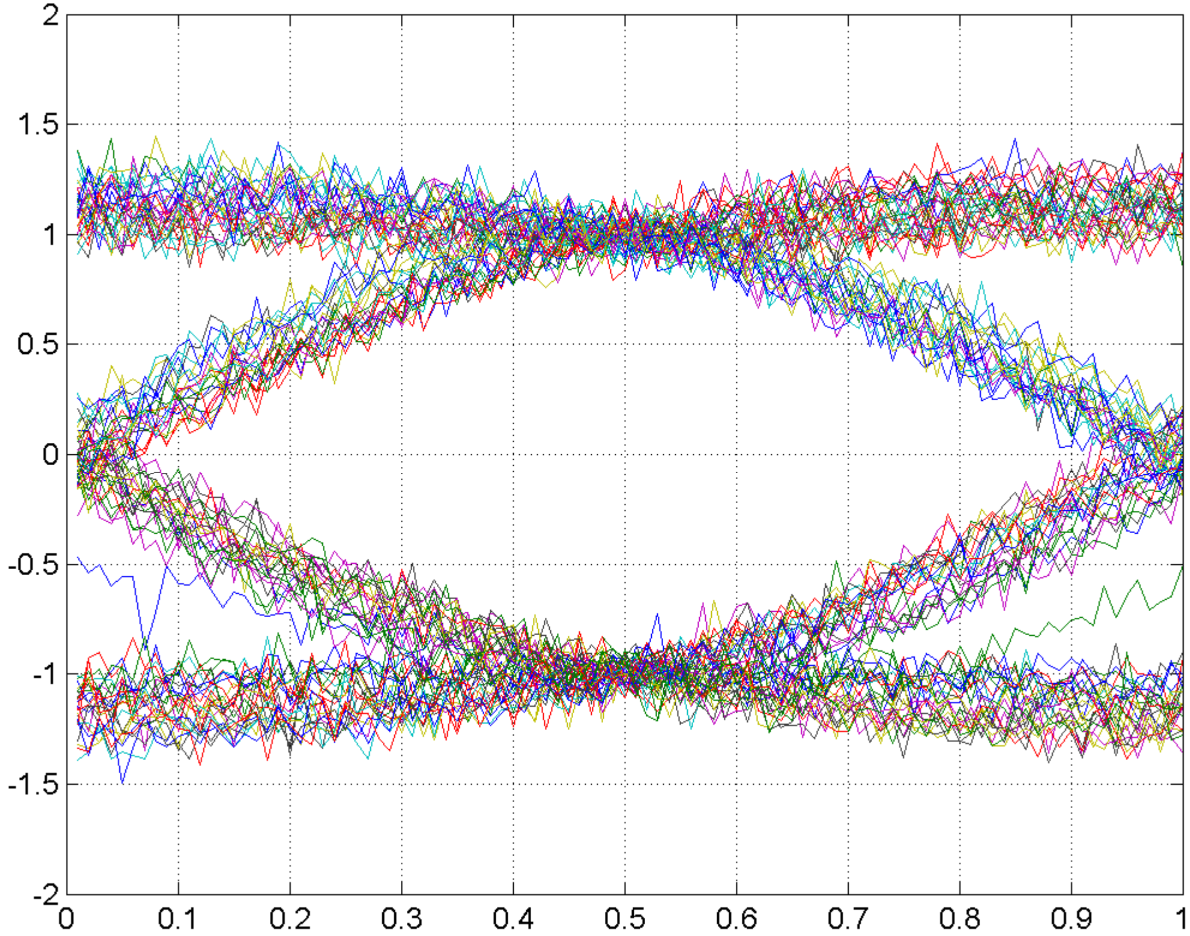
Example of eye pattern:
Binary-PAM, SRRC pulse

- Perfect channel (no noise and no ISI)



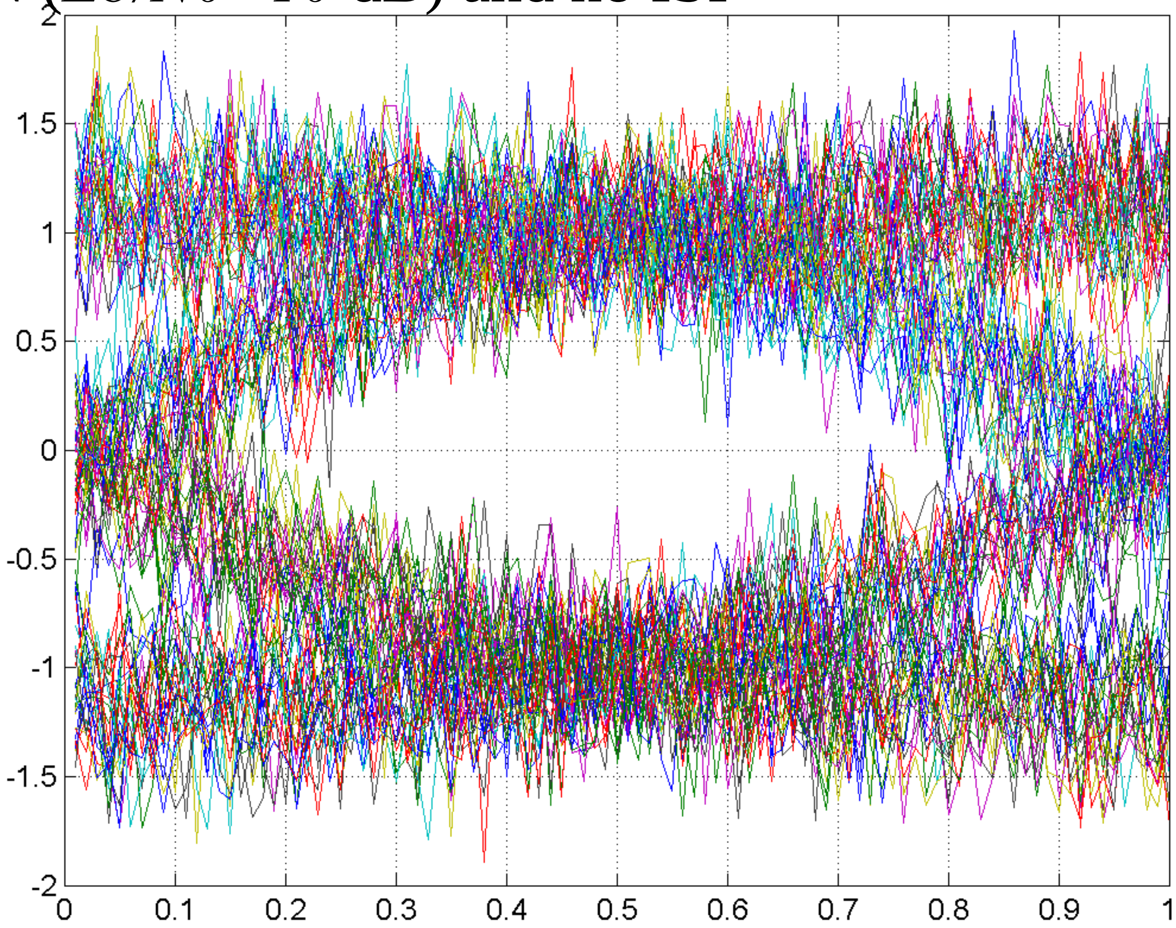
Example of eye pattern:
Binary-PAM, SRRC pulse …

- AWGN (*Eb/N0=20* dB) and no ISI



Example of eye pattern:
Binary-PAM, SRRC pulse …

- AWGN (*Eb/N0=10* dB) and no ISI



Summary

- Digital Basics
- Modulation & Detection, Performance, Bounds
- Modulation Schemes, Constellations
- Pulse Shaping

Extra Slides

Bandpass Modulation: I, Q Representation

$$s(t) = \alpha(t)\cos\left(\phi(t) + \phi_0\right)\cos\left(2\pi f_c t\right) - \alpha(t)\sin\left(\phi(t) + \phi_0\right)\sin\left(2\pi f_c t\right)$$

$$s(t) = s_I(t)\cos\left(2\pi f_c t\right) - s_Q(t)\sin\left(2\pi f_c t\right)$$

$$s_I(t) = \alpha(t)\cos\left(\phi(t) + \phi_0\right)$$ **In-phase component**

$$s_Q(t) = \alpha(t)\sin\left(\phi(t) + \phi_0\right)$$ **Quadrature component**

Equivalent lowpass representation:

$$s(t) = \mathrm{Re}\left[u(t)e^{j2\pi f_c t}\right], \quad u(t) = s_I(t) + j\, s_Q(t)$$

Analog: Frequency Modulation (FM) vs Amplitude Modulation (AM)

- FM: all information in the phase or frequency of carrier
  - Non-linear or rapid improvement in reception quality beyond a minimum received signal threshold: *"capture" effect*.
  - Better noise immunity & resistance to fading
  - Tradeoff bandwidth (modulation index) for improved SNR: 6dB gain for 2x bandwidth
  - Constant envelope signal: efficient (70%) class C power amps ok.
- AM: linear dependence on quality & power of rcvd signal
  - Spectrally efficient but susceptible to noise & fading
  - Fading improvement using in-band pilot tones & adapt receiver gain to compensate
  - Non-constant envelope: Power inefficient (30-40%) Class A or AB power amps needed: ½ the talk time as FM!

Example Analog: Amplitude Modulation