# Communication of a Single Bit: [Ch. 13]

Problem Setup:

- $I \in \{0, 1\}$ "information bit"

  $\Pr(I = i) = \pi_i$ "prior probability"

- $I = i \quad \Rightarrow \quad X_z(t) = x_i(t)$ "modulation"

  $x_i(t)$ has time support only on $[0, T_p]$

  $W_b = 1/T_p$ bits/sec "bit rate"

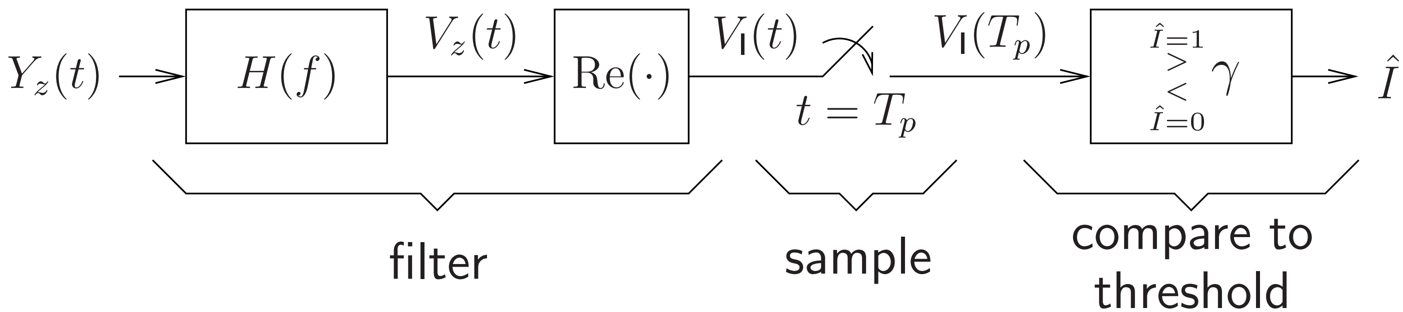
  $E_b = \pi_0 E_0 + \pi_1 E_1, \quad E_i = \int_0^{T_p} |x_i(t)|^2 dt$ "bit energy"

- $Y_z(t) = X_z(t) + W_z(t)$ "AWGN channel"

  $W_z(t)$ is circular complex Gaussian with $S_{W_z}(f) = N_0$

To minimize *bit error probability* (BEP)...

- We assume the following structure




  as justified in a detection theory course (e.g., ECE-806).

- We optimize $\gamma$, $H(f)$, and $\{x_0(t), x_1(t)\}$, in that order.

Hypothesis testing theory says that the *maximum a-posteriori* (MAP) detector gives min BEP:

$$\hat{i} = \arg\max_i \underbrace{\Pr\big(I = i \,\big|\, V_{\mathsf{l}}(T_p) = v_{\mathsf{l}}\big)}_{\text{"posterior probability"}}.$$

Using Bayes rule

$$\Pr\big(I = i \,\big|\, V_{\mathsf{l}}(T_p) = v_{\mathsf{l}}\big) \;=\; \frac{f_{V_{\mathsf{l}}(T_p)|I}\big(v_{\mathsf{l}}\big|i\big)\,\pi_i}{f_{V_{\mathsf{l}}(T_p)}\big(v_{\mathsf{l}}\big)}$$

$$\Rightarrow \hat{i} \;=\; \arg\max_i f_{V_{\mathsf{l}}(T_p)|I}\big(v_{\mathsf{l}}\big|i\big)\,\pi_i.$$

So what is $f_{V_{\mathsf{l}}(T_p)|I}(v_{\mathsf{l}}|i)$?

Note that

$$V_{\mathsf{I}}(T_p)\Big|_{I=i} = m_i(T_p) + N_{\mathsf{I}}(T_p),$$

where

$$m_i(t) = \mathrm{Re}\int_{-\infty}^{\infty} x_i(\tau)h(t-\tau)d\tau$$

$$N_{\mathsf{I}}(T_p) \sim \mathcal{N}(0, \sigma^2_{N_{\mathsf{I}}}) \text{ for some } \sigma^2_{N_{\mathsf{I}}},$$

thus

$$f_{V_{\mathsf{I}}(T_p)|I}(v_{\mathsf{I}}|i) = \frac{1}{\sqrt{2\pi\sigma^2_{N_{\mathsf{I}}}}}\exp\left[-\frac{\big(v_{\mathsf{I}} - m_i(T_p)\big)^2}{2\sigma^2_{N_{\mathsf{I}}}}\right].$$

In the case that $\pi_0 = \pi_1$ ("equal priors"),

$$\hat{i} = \arg\max_i \underbrace{f_{V_{\mathsf{I}}(T_p)|I}(v_{\mathsf{I}}|i)}_{\text{"likelihood function"}} \qquad \text{"Maximum Likelihood" (ML)}$$

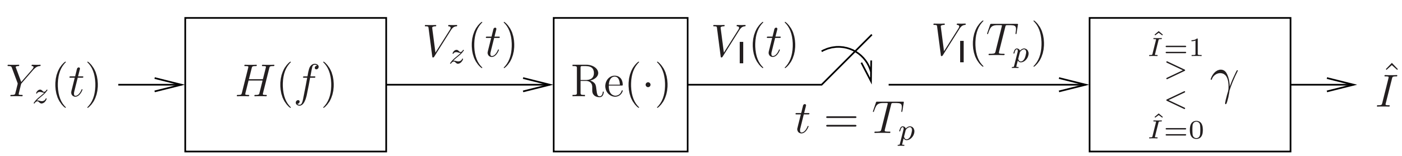
$$= \arg\min_i \big(v_{\mathsf{I}} - m_i(T_p)\big)^2,$$

and so the min-BEP detector minimizes Euclidean distance!

This implies the simple threshold test:

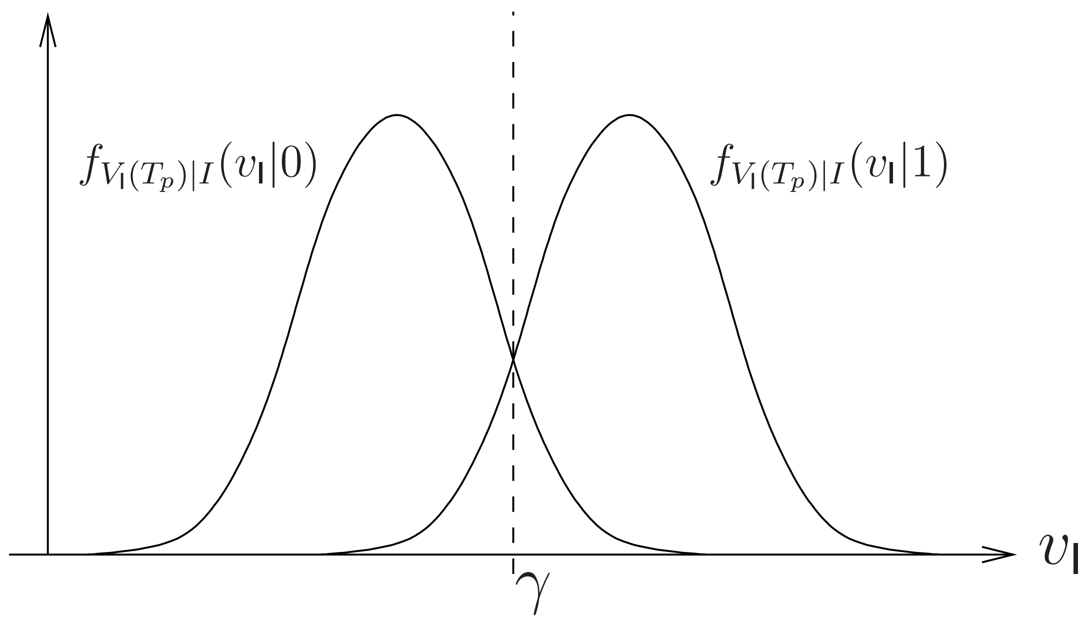
$$V_{\mathsf{I}}(T_p) \underset{\hat{I}=0}{\overset{\hat{I}=1}{\gtrless}} \gamma \quad \text{where} \quad \gamma = \frac{m_1(T_p) + m_0(T_p)}{2}.$$

which helps to justify the last stage of processing in



Analysis of bit error probability:

$$\begin{aligned}
\text{BEP} &= \Pr(\hat{I} \neq I) = \sum_{i \in \{0,1\}} \Pr(\hat{I} \neq I | I = i) \Pr(I = i) \\
&= \Pr(\hat{I} = 1 | I = 0)\pi_0 + \Pr(\hat{I} = 0 | I = 1)\pi_1 \\
&= \Pr\big(V_\mathsf{I}(T_p) > \gamma | I = 0\big)\pi_0 + \Pr\big(V_\mathsf{I}(T_p) < \gamma | I = 1\big)\pi_1 .
\end{aligned}$$



How do we compute Gaussian tail probabilities?

Say $X \sim \mathcal{N}(\mu, \sigma^2)$. Consider Gaussian CDF:

$$\begin{aligned} F_X(x) &= \Pr(X < x) \\ &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt \\ &= \frac{1}{2} + \frac{1}{2}\operatorname{erf}\left(\frac{x-\mu}{\sqrt{2}\sigma}\right), \end{aligned}$$

where $\operatorname{erf}(\cdot)$ is a well-tabulated function:

$$\operatorname{erf}(x) \triangleq \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \triangleq 1 - \operatorname{erfc}(x).$$

For the case that $\boxed{\pi_0 = \pi_1}$, we can use the properties $\gamma = \frac{m_1(T_p)+m_0(T_p)}{2}$ and $V_\mathsf{I}(T_p)\big|_{I=i} \sim \mathcal{N}(m_i(T_p), \sigma^2_{N_\mathsf{I}})$ to find the BEP of our receiver in the case that $I = 1$:

$$\begin{aligned}
P\big(\hat{I} = 0 | I = 1\big) &= \Pr\big(V_\mathsf{I}(T_p) < \gamma | I = 1\big) \\
&= \int_{-\infty}^{\gamma} \frac{1}{\sqrt{2\pi\sigma^2_{N_\mathsf{I}}}} \exp\left(-\frac{(v - m_1(T_p))^2}{2\sigma^2_{N_\mathsf{I}}}\right) dv \\
&= \frac{1}{2} + \frac{1}{2}\,\mathrm{erf}\left(\frac{\gamma - m_1(T_p)}{\sqrt{2}\,\sigma_{N_\mathsf{I}}}\right) \\
&= \frac{1}{2} - \frac{1}{2}\,\mathrm{erf}\left(\frac{m_1(T_p) - m_0(T_p)}{2\sqrt{2}\,\sigma_{N_\mathsf{I}}}\right) \\
&= \frac{1}{2}\,\mathrm{erfc}\left(\frac{m_1(T_p) - m_0(T_p)}{2\sqrt{2}\,\sigma_{N_\mathsf{I}}}\right).
\end{aligned}$$

Since $P(\hat{I} = 1 | I = 0)$ can be found to be identical,

$$\text{BEP} = \frac{1}{2}\,\mathrm{erfc}\left(\frac{m_1(T_p) - m_0(T_p)}{2\sqrt{2}\,\sigma_{N_\mathsf{I}}}\right).$$

Then, defining the *effective SNR* $\eta$:

$$\eta \triangleq \left(\frac{m_1(T_p) - m_0(T_p)}{2\sqrt{2}\,\sigma_{N_\mathsf{I}}}\right)^2,$$

we get the simple expression

$$\text{BEP} = \frac{1}{2}\,\mathrm{erfc}(\sqrt{\eta}) \approx \frac{1}{2}e^{-\eta} \text{ at high } \eta.$$

Notice: Minimizing BEP $\Leftrightarrow$ maximizing $\eta$. From

$$\sigma_{N_\mathsf{I}}^2 = \frac{N_0}{2}\int_{-\infty}^{\infty}|H(f)|^2 df$$

$$\begin{aligned} m_1(T_p) - m_0(T_p) &= \mathrm{Re}\left(\mathcal{F}^{-1}\Big\{H(f)\big[X_1(f) - X_0(f)\big]\Big\}_{t=T_p}\right) \\ &= \mathrm{Re}\int_{-\infty}^{\infty} H(f) \underbrace{\big[X_1(f) - X_0(f)\big]}_{\triangleq B_{10}(f) \text{ “effective signal”}} e^{j2\pi f T_p} df \end{aligned}$$

we find

$$\eta = \frac{\left(\mathrm{Re}\int_{-\infty}^{\infty} H(f) B_{10}(f) e^{j2\pi f T_p} df\right)^2}{4N_0 \int_{-\infty}^{\infty}|H(f)|^2 df}.$$

So how can $H(f)$ be chosen to maximize $\eta$?

Using the Cauchy-Schwarz inequality:

$$\frac{|\int_{-\infty}^{\infty} H(f)Y^*(f)df|^2}{\int_{-\infty}^{\infty}|H(f)|^2df \cdot \int_{-\infty}^{\infty}|Y(f)|^2df} \leq 1 \qquad \text{with equality iff } H(f) = CY(f)$$

$$\Rightarrow \quad \frac{|\int_{-\infty}^{\infty} H(f)Y^*(f)df|^2}{4N_0\int_{-\infty}^{\infty}|H(f)|^2df} \leq \frac{1}{4N_0}\int_{-\infty}^{\infty}|Y(f)|^2df,$$

and the fact that

$$(\mathrm{Re}(X))^2 \leq |X|^2 \qquad \text{with equality iff } \mathrm{Im}(X) = 0 \quad ,$$

we see, by choosing $Y(f) = B_{10}^*(f)e^{-j2\pi fT_p}$, that

$$\eta \leq \frac{1}{4N_0}\int_{-\infty}^{\infty}|B_{10}(f)|^2df.$$

where equality in the previous expression occurs when

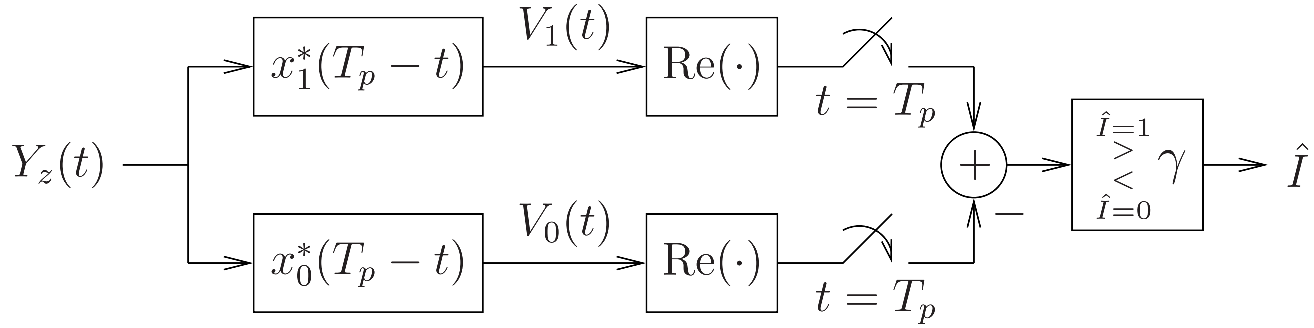
$$\begin{aligned} H(f) &= C B_{10}^*(f) e^{-j2\pi f T_p} \\ C &\in \mathbb{R}. \end{aligned}$$

For $\eta$ maximization, we choose $C = 1$ (w.l.o.g.), and so

$$\begin{aligned} \eta_{\text{max}} &= \frac{1}{4N_0} \int_{-\infty}^{\infty} |B_{10}(f)|^2 df \\ H_{\text{max}}(f) &= B_{10}^*(f) e^{-j2\pi f T_p} \\ h_{\text{max}}(t) &= \mathcal{F}^{-1}\left\{ B_{10}^*(f) e^{-j2\pi f T_p} \right\} \;=\; b_{10}^*(T_p - t) \\ &= x_1^*(T_p - t) - x_0^*(T_p - t), \end{aligned}$$

which is known as a "matched filter" front-end.

For more insight into matched filtering, redraw receiver as




In the equal priors case (i.e., ML decision making), we saw

$$\gamma = \frac{1}{2}\big(m_1(T_p) + m_0(T_p)\big).$$

When $h(t) = h_{\mathsf{max}}(t)$, we find that

$$\begin{aligned} m_i(T_p) &= \mathrm{Re} \int_0^{T_p} x_i(\tau) h(T_p - \tau) d\tau \\ &= \mathrm{Re} \int_0^{T_p} x_i(\tau) \big[x_1^*(\tau) - x_0^*(\tau)\big] d\tau. \end{aligned}$$

Notice that

$$m_1(T_p) \;=\; \mathrm{Re}\underbrace{\int |x_1(\tau)|^2 d\tau}_{E_1} - \mathrm{Re}\underbrace{\int x_1(\tau)x_0^*(\tau)d\tau}_{\sqrt{E_0E_1}\rho_{10}}$$

$$m_0(T_p) \;=\; -\mathrm{Re}\underbrace{\int |x_0(\tau)|^2 d\tau}_{E_0} + \mathrm{Re}\underbrace{\int x_0(\tau)x_1^*(\tau)d\tau}_{\sqrt{E_0E_1}\rho_{10}^*},$$

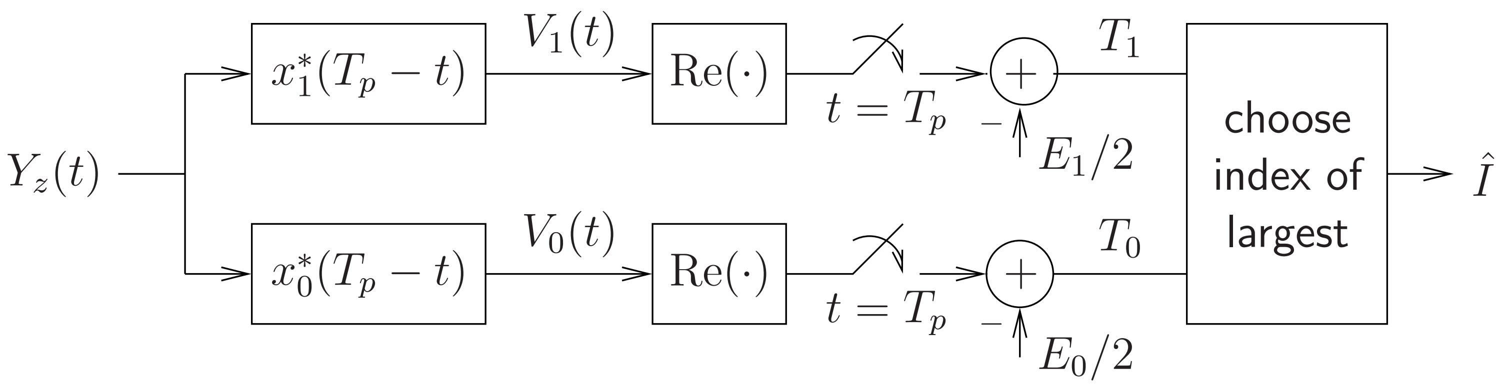
where $|\rho_{10}| \leq 1$ is the *signal correlation coefficient*. Thus, the ML threshold becomes

$$\gamma \;=\; \frac{1}{2}\big(E_1 - E_0\big) + \frac{1}{2}\sqrt{E_0E_1}\,\underbrace{\mathrm{Re}\big(\rho_{10}^* - \rho_{10}\big)}_{=0}.$$

We can then rewrite the ML threshold comparison as

$$V_{\mathsf{I}}(T_p) \underset{\hat{I}=0}{\overset{\hat{I}=1}{\gtrless}} \gamma \quad \Leftrightarrow \quad \mathrm{Re}\, V_1(T_p) - \frac{E_1}{2} \underset{\hat{I}=0}{\overset{\hat{I}=1}{\gtrless}} \mathrm{Re}\, V_0(T_p) - \frac{E_0}{2},$$

which yields the ML block diagram:




Here, $T_0$ and $T_1$ are referred to as "ML metrics."

Notice: Each front-end filter is matched to a particular transmission waveform.

Finally, we design $\{x_0(t), x_1(t)\}$ to minimize BEP. Since

$$\eta_{\text{max}} = \frac{1}{4N_0}\int_{-\infty}^{\infty} |B_{10}(f)|^2 df = \frac{1}{4N_0}\int_{-\infty}^{\infty} |b_{10}(t)|^2 dt$$

$$= \frac{1}{4N_0}\underbrace{\int_0^{T_p} |x_1(t) - x_0(t)|^2 dt}_{\Delta_E(1,0)},$$

where $\Delta_E(1,0)$ denotes the *Euclidean distance between* $x_1(t)$ *and* $x_0(t)$, ML demodulation achieves (under equal priors $\pi_0 = \pi_1$)

$$\text{BEP} = \frac{1}{2}\,\text{erfc}\left(\sqrt{\frac{\Delta_E(1,0)}{4N_0}}\right).$$

Here, minimizing BEP $\Leftrightarrow$ maximizing $\Delta_E(1,0)$, where

$$\Delta_E(1,0) = \underbrace{\int_0^{T_p} |x_1(t)|^2 dt}_{E_1} + \underbrace{\int_0^{T_p} |x_0(t)|^2 dt}_{E_0} - 2\,\mathrm{Re} \underbrace{\int_0^{T_p} x_1(t) x_0^*(t) dt}_{\sqrt{E_1 E_0}\,\rho_{10}}.$$

Assuming fixed average bit energy $E_b = \frac{1}{2}(E_0 + E_1)$, we find

$$\Delta_E(1,0) = 2E_b - 2\sqrt{E_1(2E_b - E_1)}\,\mathrm{Re}\,\rho_{10}.$$

which is maximized by choosing

1. $\mathrm{Re}\,\rho_{10}$ as small as possible,
2. $E_1 = E_b = E_0$ when $\mathrm{Re}\,\rho_{10} \le 0$, else $E_1 \in \{0, 2E_b\}$.

Recall that $-1 \le \mathrm{Re}\,\rho_{10} \le 1$.

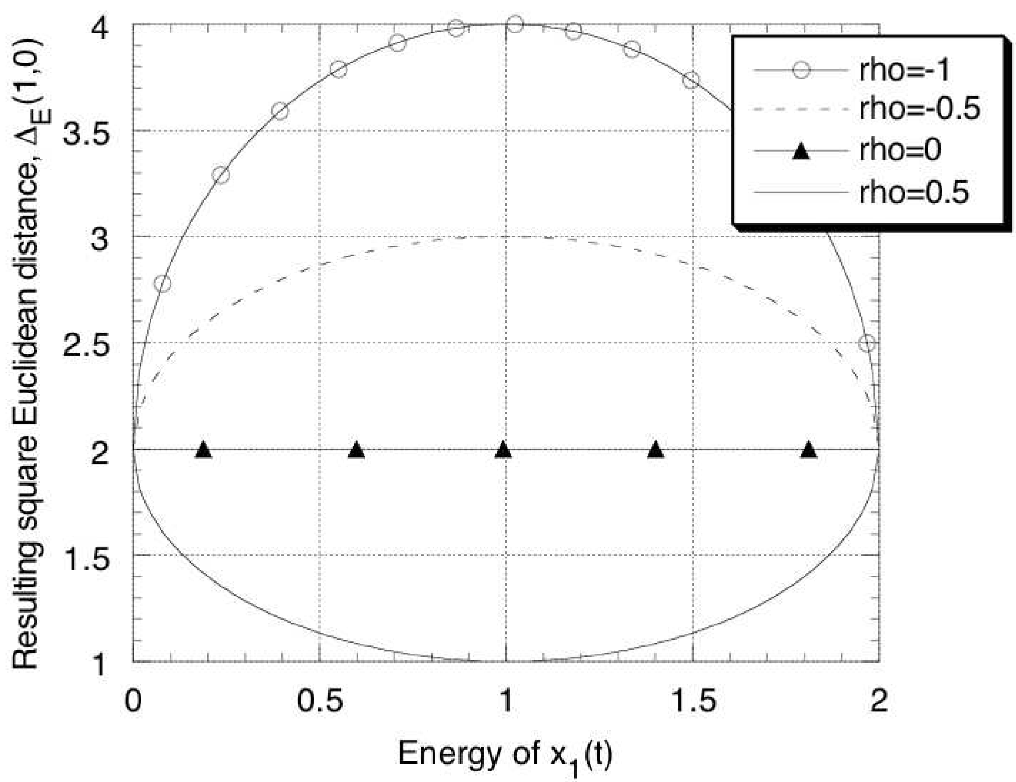
When only $E_b$ is constrained, *antipodal signaling*, i.e.,

$$\rho_{10} = -1 \quad \& \quad E_1 = E_b = E_0$$

yields the minimal BEP, i.e., BEP $= \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right)$.

Illustration of $\{E_1, \operatorname{Re}\rho_{10}\}$ design when $E_b = 1$:

(“rho” $= \operatorname{Re}\rho_{10}$.)



Example 1: Binary Phase Shift Keying (BPSK):

$$x_0(t) = \begin{cases} \sqrt{\frac{E_b}{T_p}} & t \in [0, T_p] \\ 0 & t \notin [0, T_p] \end{cases}$$

$$x_1(t) = \begin{cases} \sqrt{\frac{E_b}{T_p}} \exp(j\theta) & t \in [0, T_p] \\ 0 & t \notin [0, T_p] \end{cases},$$

where $\theta$ is a design parameter.

Note that $E_0 = E_1 = E_b$ and that

$$\rho_{10} = \exp(j\theta),$$

so that $\theta = \pi$ yields $\mathrm{Re}\,\rho_{10} = -1$; antipodal BPSK.

Spectral characteristics:

$$G_{X_z}(f) \triangleq |X_z(f)|^2, \qquad \text{“energy spectrum”}$$

$$D_{X_z}(f) \triangleq \frac{1}{K_b}\,\mathrm{E}\{G_{X_z}(f)\} \qquad \text{“average energy spectrum per bit”}$$

$$= \pi_0 G_{X_0}(f) + \pi_1 G_{X_1}(f).$$

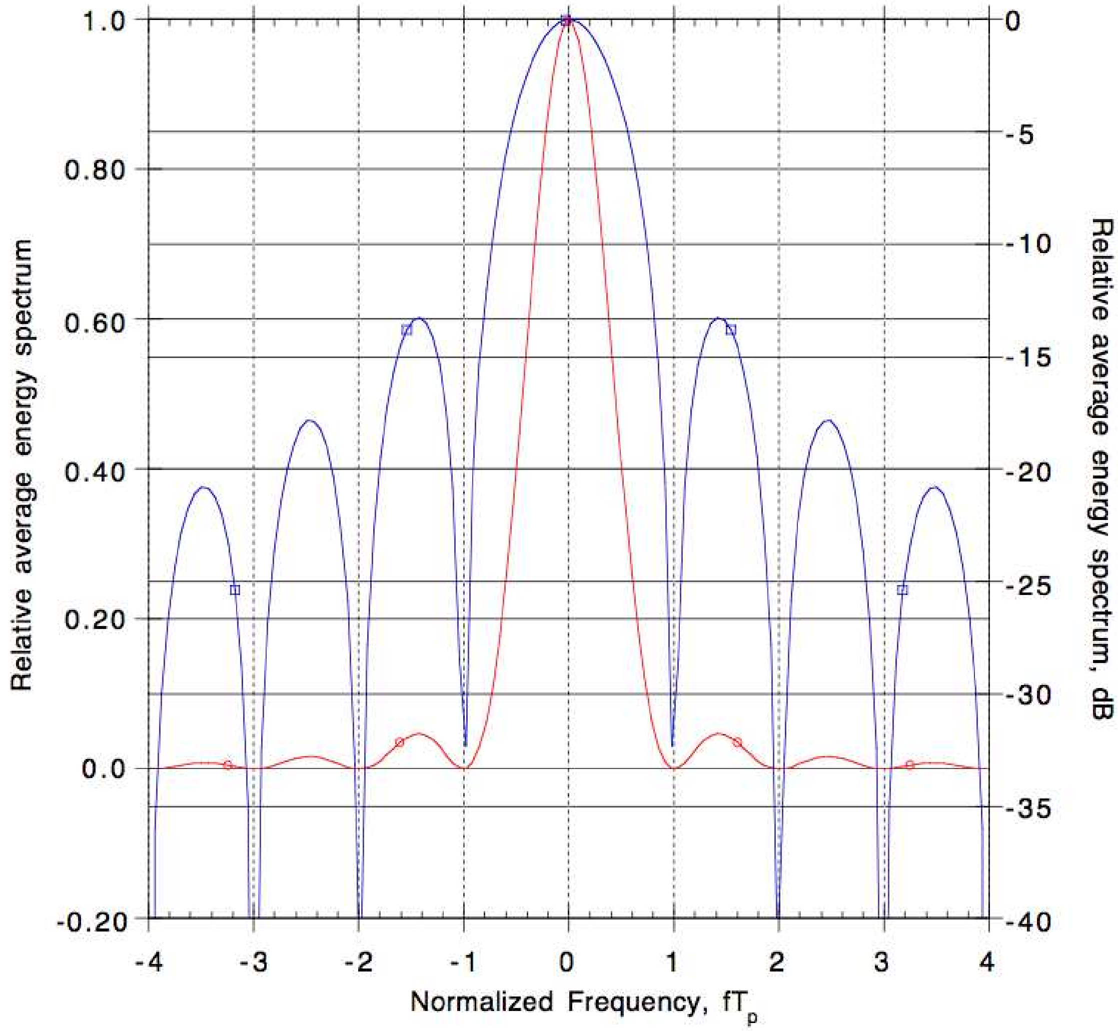
For BPSK, it's easy to show that

$$G_{X_i}(f) = \left(\frac{\sin(\pi f T_p)}{\pi f T_p}\right)^2 E_b T_p = D_{X_z}(f)$$

From the plot of $D_{X_z}(f)$ on the next page, we see that

$$B_T \approx 1/T_p \qquad \rightsquigarrow \eta_B = 1\,\text{bit/sec/Hz}.$$

# BPSK average energy spectrum per bit, $D_{X_z}(f)$:



Example 2: Binary Frequency Shift Keying (BFSK):

$$x_0(t) = \begin{cases} \sqrt{\frac{E_b}{T_p}} \exp(j2\pi f_d t) & t \in [0, T_p] \\ 0 & t \notin [0, T_p] \end{cases}$$

$$x_1(t) = \begin{cases} \sqrt{\frac{E_b}{T_p}} \exp(-j2\pi f_d t) & t \in [0, T_p] \\ 0 & t \notin [0, T_p] \end{cases},$$

where *frequency deviation* $f_d$ is a design parameter.

Note that $E_0 = E_1 = E_b$ and that

$$\rho_{10} = \frac{1}{T_p} \int_0^{T_p} \exp(-j4\pi f_d t) dt.$$

Recall that $\Delta_E(1,0) = E_0 + E_1 - 2\sqrt{E_0 E_1}\,\mathrm{Re}(\rho_{10})$.

We integrate to find that

$$\rho_{10} = \frac{\sin(4\pi f_d T_p)}{4\pi f_d T_p} - j\,\frac{\cos(4\pi f_d T_p) - 1}{4\pi f_d T_p},$$

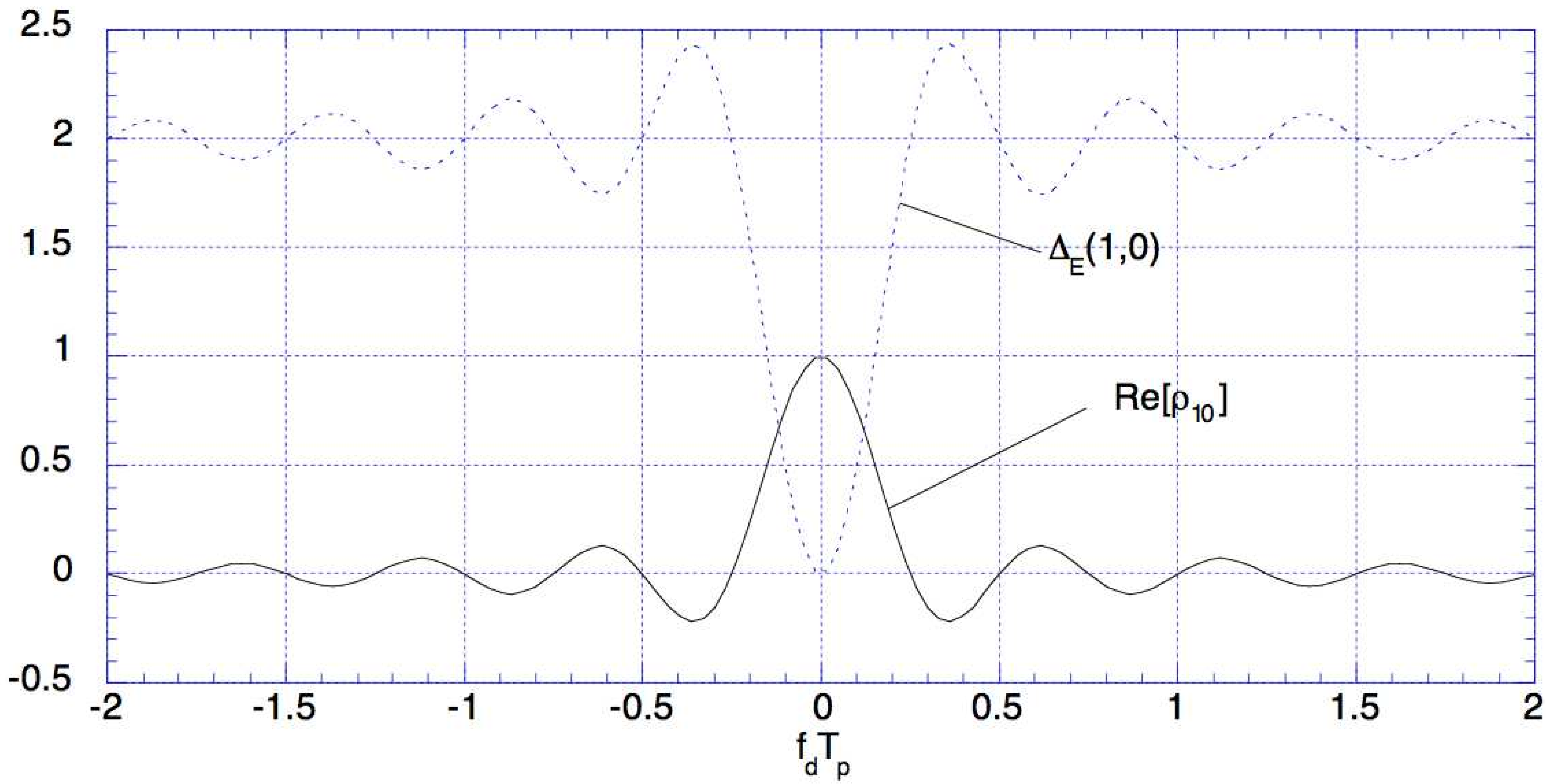
implying that

$$\Delta_E(1,0) = 2\left(1 - \frac{\sin(4\pi f_d T_p)}{4\pi f_d T_p}\right) E_b,$$

and thus

$$\begin{aligned} \text{BEP} &= \frac{1}{2}\,\text{erfc}\left(\sqrt{\frac{\Delta_E(1,0)}{4N_0}}\right) \\ &= \frac{1}{2}\,\text{erfc}\left(\sqrt{\frac{E_b}{2N_0}\left(1 - \frac{\sin(4\pi f_d T_p)}{4\pi f_d T_p}\right)}\right). \end{aligned}$$

## $\Delta_E(1,0)$ and $\mathrm{Re}\,\rho_{10}$ versus frequency deviation $f_d$:



For BFSK, can show that energy spectra equal

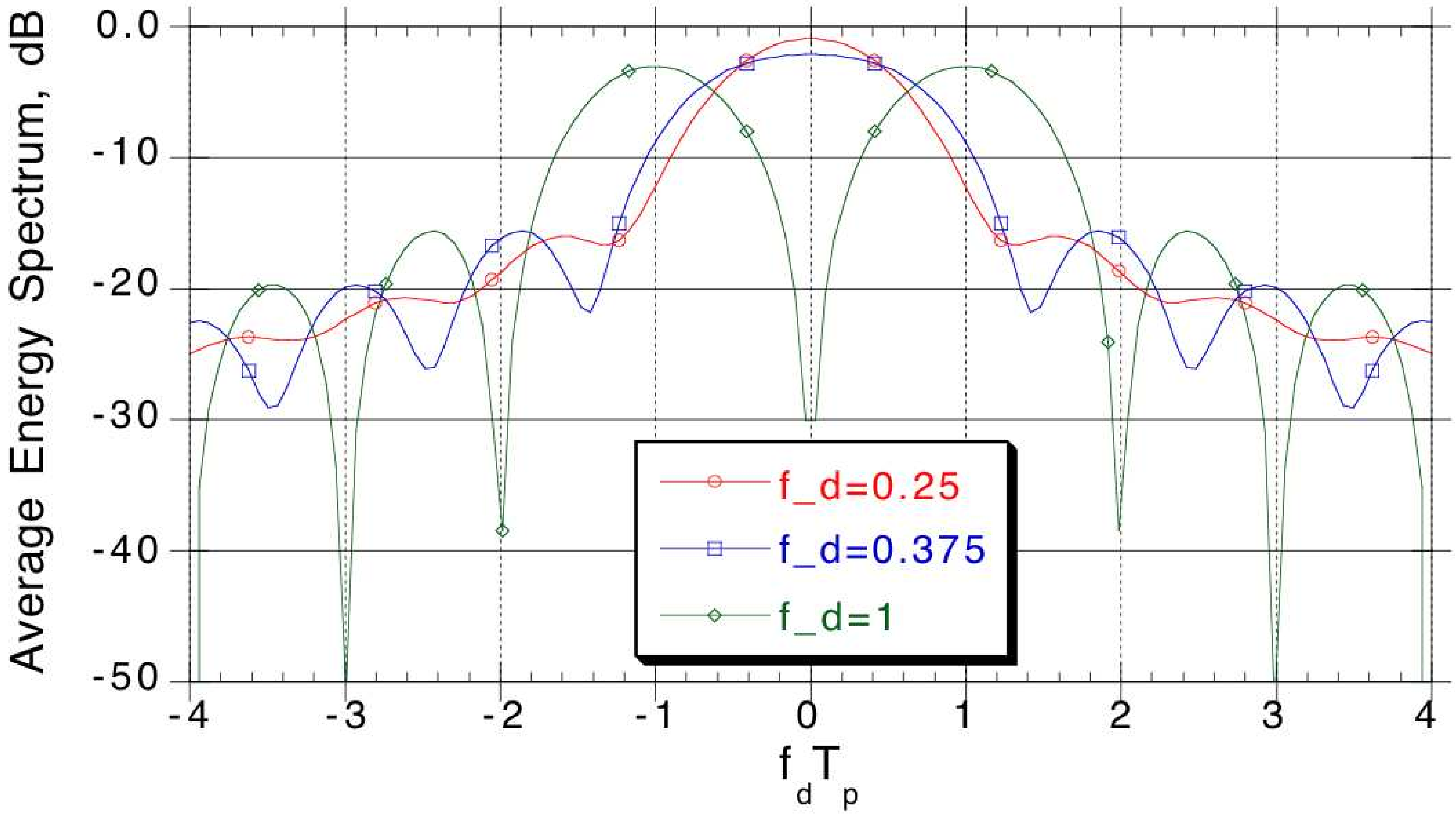
$$G_{X_0}(f) = \left(\frac{\sin(\pi(f - f_d)T_p)}{\pi(f - f_d)T_p}\right)^2 E_b T_p$$

$$G_{X_1}(f) = \left(\frac{\sin(\pi(f + f_d)T_p)}{\pi(f + f_d)T_p}\right)^2 E_b T_p,$$

so that average energy spectrum per bit equals

$$D_{X_z}(f) = \frac{1}{2}\left(\frac{\sin(\pi(f - f_d)T_p)}{\pi(f - f_d)T_p}\right)^2 E_b T_p + \frac{1}{2}\left(\frac{\sin(\pi(f + f_d)T_p)}{\pi(f + f_d)T_p}\right)^2 E_b T_p.$$

BFSK average energy spectrum per bit for various $f_d$:



BFSK design strategies:

1. Choose $f_d$ to minimize $\operatorname{Re}\rho_{10}$ (i.e., maximize $\Delta_E(1,0)$):

   $f_d \approx 0.375/T_p, \quad \operatorname{Re}\rho_{10} \approx -0.21,$ $\quad \rightsquigarrow$ 2.2dB SNR loss.

   $B_T \approx 1.5/T_p$ $\quad \rightsquigarrow \eta_B \triangleq \dfrac{W_b}{B_T} \approx 0.67$ bits/sec/Hz.

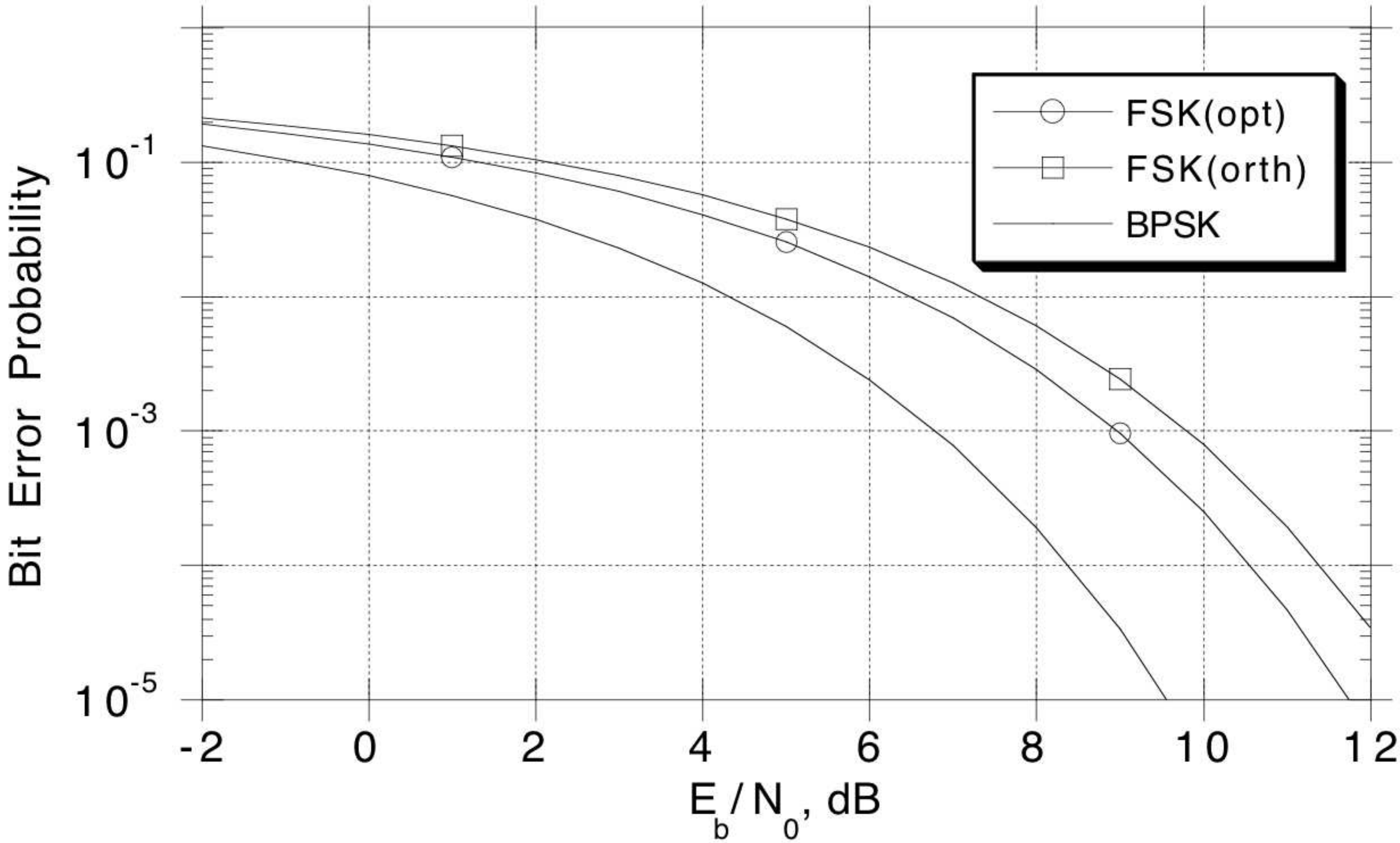
2. Choose $f_d$ to yield orthogonal signals (i.e., $\rho_{10} = 0$):

   $f_d = 0.25/T_p$ (“MSK”) , $\quad \rightsquigarrow$ 3dB SNR loss.

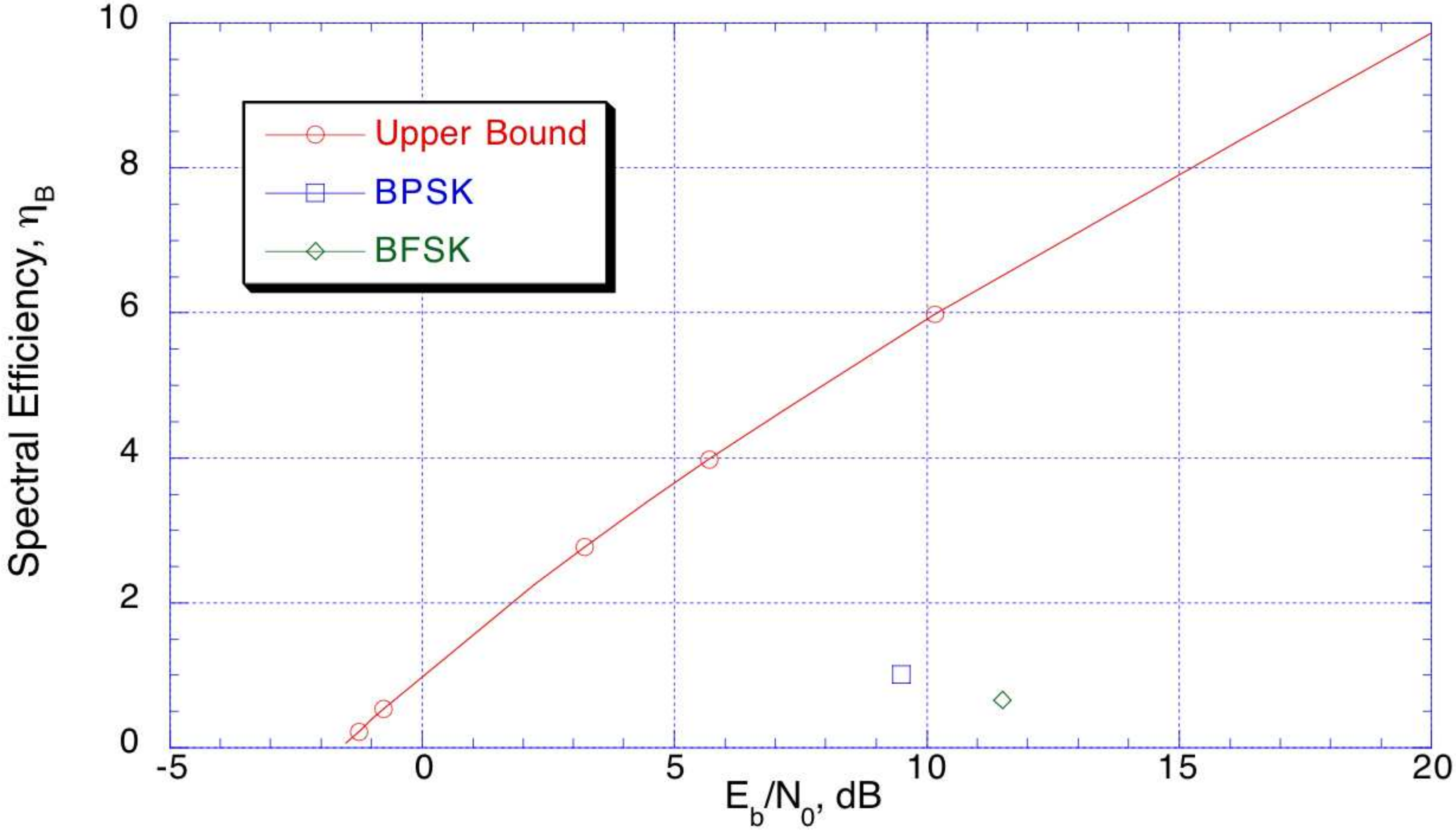
   $B_T \approx 1.0/T_p$ $\quad \rightsquigarrow \eta_B \triangleq \dfrac{W_b}{B_T} \approx 1.0$ bits/sec/Hz.

where “loss” is relative to optimal (i.e., antipodal) signaling.

Comparison of BEP for binary modulations:



Comparison of spectral efficiencies to Shannon bound:




(considering WEP$= 10^{-5}$ as “reliable”)

Now consider the **non-equal prior** case (i.e., $\pi_0 \neq \pi_1$):

$$\hat{i} = \arg\max_i f_{V_\mathsf{I}(T_p)|I}(v_\mathsf{I}|i)\pi_i \qquad \text{“MAP” (as before)}$$

$$= \arg\max_i \exp\left(-\frac{[v_\mathsf{I} - m_i(T_p)]^2}{2\sigma_{N_\mathsf{I}}^2}\right)\pi_i.$$

Detection theory shows that the same $h_{\mathsf{max}}(t)$ applies to both equal and non-equal priors cases, so that

$$m_i(T_p) = \mathrm{Re}\int_0^{T_p} x_i(\tau)h_{\mathsf{max}}(T_p - \tau)d\tau$$

$$= \mathrm{Re}\int_0^{T_p} x_i(\tau)\big[x_1^*(\tau) - x_0^*(\tau)\big]d\tau$$

$$\Rightarrow m_1(T_p) = E_1 - \sqrt{E_0 E_1}\,\mathrm{Re}\,\rho_{10}$$

$$\Rightarrow m_0(T_p) = -E_0 + \sqrt{E_0 E_1}\,\mathrm{Re}\,\rho_{10}$$

and

$$\begin{aligned}
\sigma_{N_\mathsf{I}}^2 &= \frac{N_0}{2}\int_{-\infty}^{\infty}|H_{\mathsf{max}}(f)|^2 df \\
&= \frac{N_0}{2}\int_{-\infty}^{\infty}|h_{\mathsf{max}}(t)|^2 dt \\
&= \frac{N_0}{2}\int_{-\infty}^{\infty}|x_1^*(t)-x_0^*(t)|^2 dt \\
&= \frac{N_0}{2}\Delta_E(1,0) \\
&= \frac{N_0}{2}\big(E_1+E_0-2\sqrt{E_0E_1}\,\mathrm{Re}\,\rho_{10}\big)
\end{aligned}$$

Plugging these expressions for $m_i(T_p)$ and $\sigma^2_{N_\mathsf{I}}$ into the decision rule for $\hat{I}$, and remembering that

$$\begin{aligned} V_\mathsf{I}(T_p) &= \operatorname{Re} V_1(T_p) - \operatorname{Re} V_0(T_p) \\ T_i &= \operatorname{Re} V_i(T_p) - \frac{E_i}{2}, \end{aligned}$$

we can simplify the rule and write (after a bit of work)

$$\hat{I} = \arg\max_i \; \exp\left(\frac{T_i}{N_0/2}\right)\pi_i.$$