Coded Modulation [Ch. 17]:

- Orthogonal modulation had $\mathcal{O}(K_b)$ complexity MLWD but performance no better than BPSK.
- For better performance, map a *sequence* of info bits onto a *sequence* of symbols, then transmit using linear stream modulation. If clever, still have $\mathcal{O}(K_b)$ MLWD.
- Fitz calls it "orthogonal modulation with memory."
- This idea subsumes most coding+modulation schemes.
- We focus on performance, spectral efficiency, and demodulator design rather than on code design.

Basic Idea:

- A sequence of K_b bits $\{I^{(k)}\}_{k=1}^{K_b}$ is mapped to a sequence of N_f constellation labels $\{J^{(l)}\}_{l=1}^{N_f}$.
- Each label $J^{(l)}$ is mapped to symbol $\tilde{D}_z^{(l)} = a(J^{(l)})$.
- The symbol sequence $\{\tilde{D}_z^{(l)}\}_{l=1}^{N_f}$ is M_s-ary stream modulated: $X_z(t) = \sum_{l=1}^{N_f} \tilde{D}_z^{(l)} \sqrt{E_b} u\big(t - (l-1)T\big)$.

Fundamental Goals:

- Out of $M_s^{N_f}$ possible symbol sequences, choose 2^{K_b} sequences with large minimum Euclidean distance.
- Ensure that the bit-sequence to symbol-sequence mapping facilitates $\mathcal{O}(K_b)$ MLWD. (Use FSM!)

Outline:

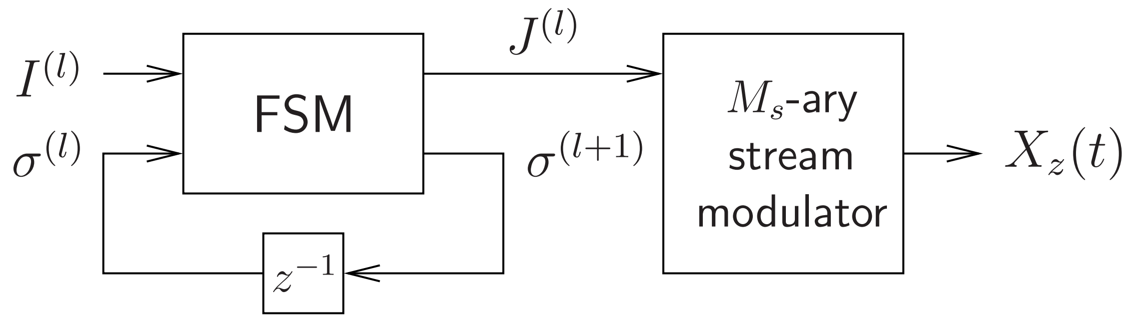
1. Rate-1 mappings (i.e., $N_f \approx K_b$).
   - System description and WEP union bound.
   - $\mathcal{O}(K_b)$ MLWD: Viterbi decoding [slight detour].
   - Spectral characteristics.
2. Arbitrary rate mappings: convolutional and trellis codes.
   - Basic idea and WEP union bound.
   - Spectral characteristics.

Assumptions:

- Bits $\{I^{(l)}\}_{l=1}^{K_b}$ are independent and equally likely.
- Symbol mapping ensures $\mathrm{E}\left[|\tilde{D}_z^{(l)}|^2\right] = R = \frac{\#\text{ bits}}{\text{symbol}}$.

Coded Modulation for $R = 1$:

- K_b bits $\{I^{(l)}\}$ mapped onto K_b constellation labels $\{J^{(l)}\}$ using a *finite state machine* (FSM).

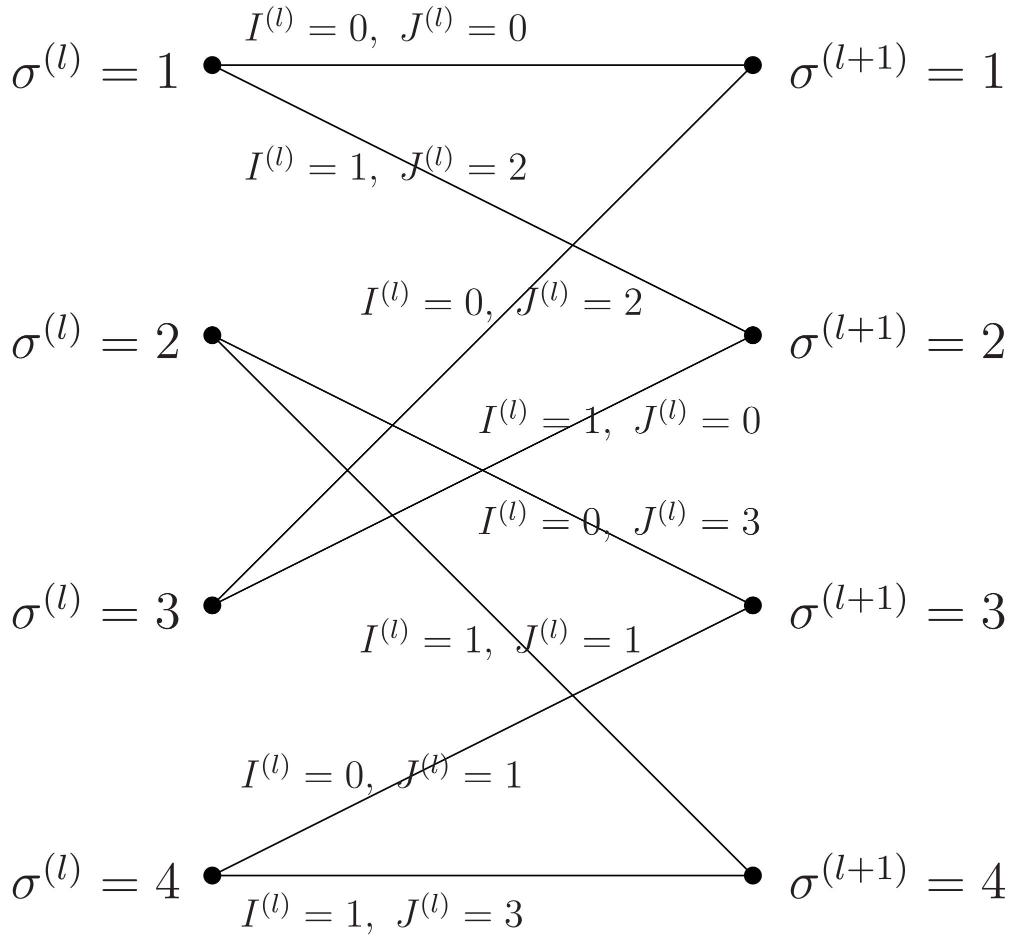


- FSM characterized by N_s modulation states $\sigma^{(l)} \in \Omega_\sigma$:

$$\begin{aligned} \sigma^{(l+1)} &= g_1\big(\sigma^{(l)}, I^{(l)}\big) \\ J^{(l)} &= g_2\big(\sigma^{(l)}, I^{(l)}\big) \end{aligned}$$

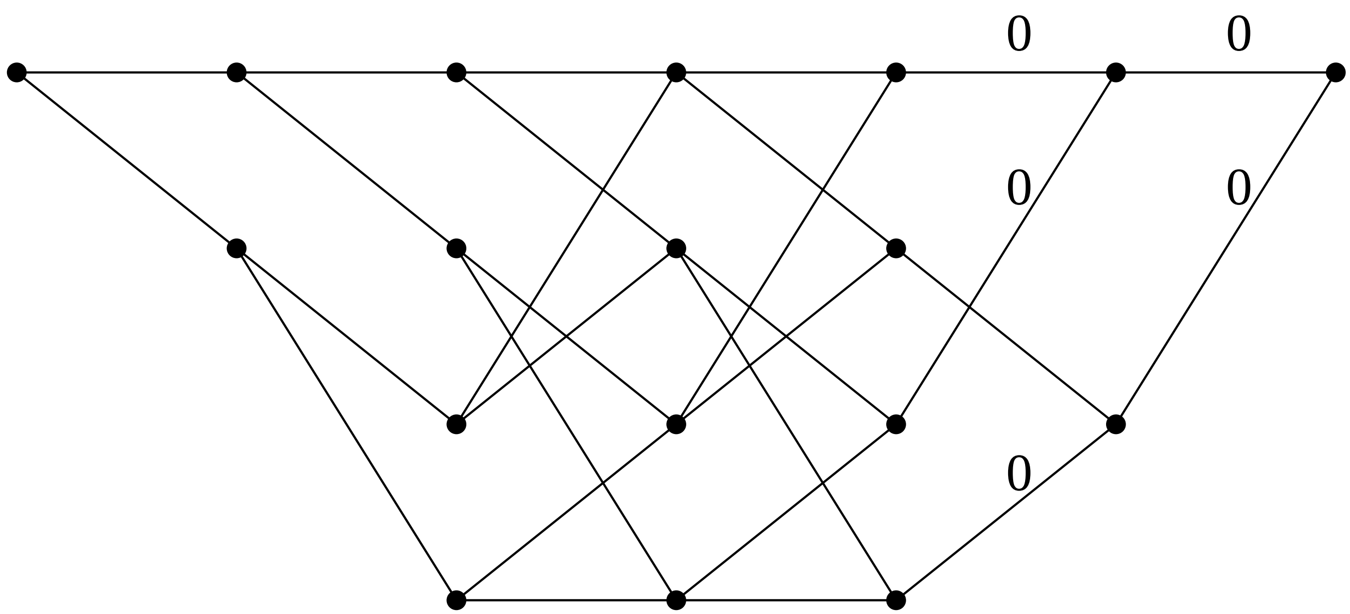
Larger N_s means more freedom in sequence design but higher demodulation complexity.

A FSM is well described by a *trellis diagram*:



This Ho-Cavers-Varaldi trellis code has $N_s = 4$ and $M_s = 4$.

Expanded trellis diagram for $K_b = 4$ bits:



Here, $\nu_c = 2$ additional zero-bits are used to return to initial state ("termination"). Thus, for $R = 1$, have frame length $N_f = K_b + \nu_c$. Note $R_{\text{eff}} = \frac{K_b}{K_b+\nu_c} \approx 1 = R$ for large K_b.

Also note: $\#$ of valid paths through trellis $= 2^{K_b} = 16$.

<u>MLWD</u>:

Orthogonal modulation leads to a decoupled ML metric:

$$\begin{aligned}\underline{\hat{I}} &= \arg \max_{i\in\{0,\ldots,2^{K_b}-1\}} T_i \\ &= \arg\min_i \sum_{k=1}^{N_f} \left| Q^{(k)} - \sqrt{E_b}\, \tilde{d}_i^{(k)} \right|^2\end{aligned}$$

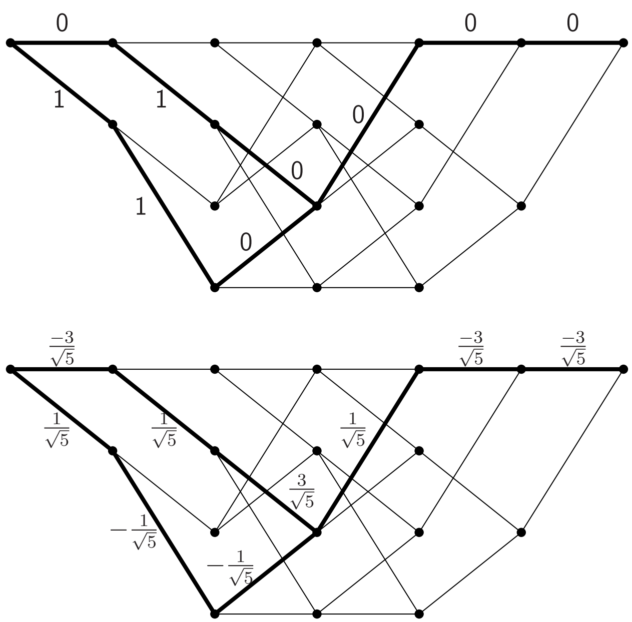
Hence, MLWD $\Leftrightarrow$ finding closest sequence $\{\tilde{d}_i^{(k)}\}_{k=1}^{N_f}$

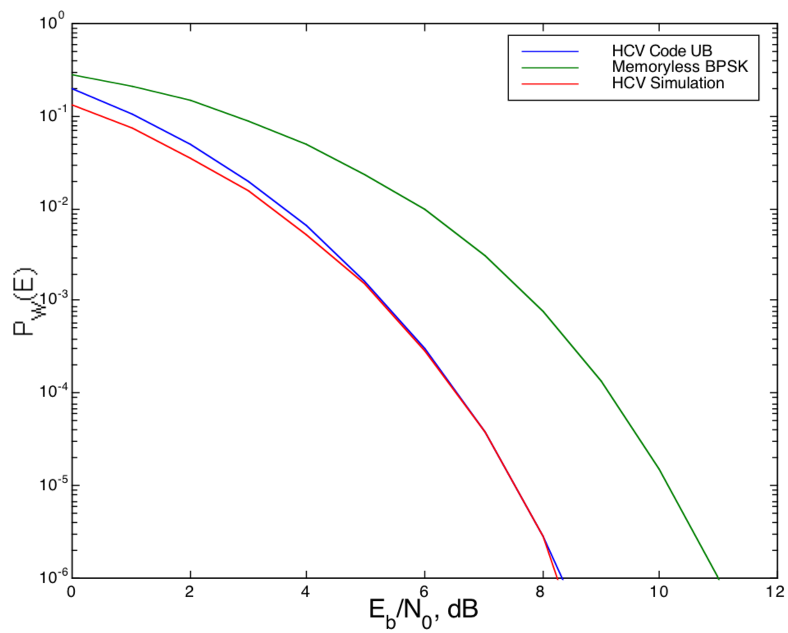
Union bound on WEP can be calculated via:

$$\Delta_E(i,j) \;=\; \int |x_i(t) - x_j(t)|^2 dt \;=\; E_b \sum_{k=1}^{N_f} \left| \tilde{d}_i^{(k)} - \tilde{d}_j^{(k)} \right|^2$$

Example: HCV code with 4-PAM modulation and $K_b = 4$.

- $J^{(l)} \in [0,1,2,3] \quad \rightarrow \quad \tilde{D}_z^{(l)} \in \left[\frac{-3}{\sqrt{5}}, \frac{-1}{\sqrt{5}}, \frac{1}{\sqrt{5}}, \frac{3}{\sqrt{5}}\right]$
- $\frac{1}{2}2^{K_b}(2^{K_b}-1) = 120$ distances in error spectrum.
- $\Delta_E\big([0100],[1100]\big) = 7.2E_b \gg 4E_b$ (recall BPSK).





Spectral Characteristics:

- Though the info bits $\{I^{(k)}\}$ are iid, the coded symbols $\{\tilde{D}_z^{(l)}\}$ will be correlated.
- Correlated symbols lead to a "shaping" of the power spectrum.
- In some cases, the spectrum becomes more compact, which may be reason enough to use modulation with memory.
- In the sequel, we develop tools to analyze the spectrum.

Energy spectrum (averaged per bit):

$$D_{X_z}(f) \;=\; \frac{1}{K_b}\,\mathrm{E}\left\{G_{X_z}(f)\right\}$$

where

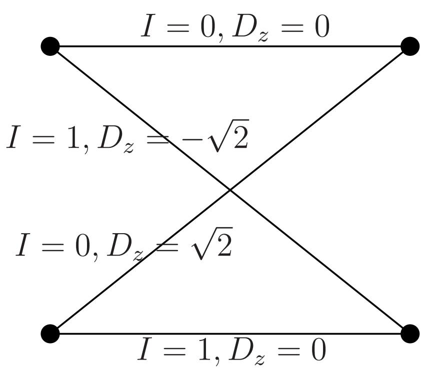
$$x_i(t) \;=\; \sqrt{E_b}\sum_{k=1}^{N_f}\tilde{d}_i^{(k)}u\big(t-T(k-1)\big)$$

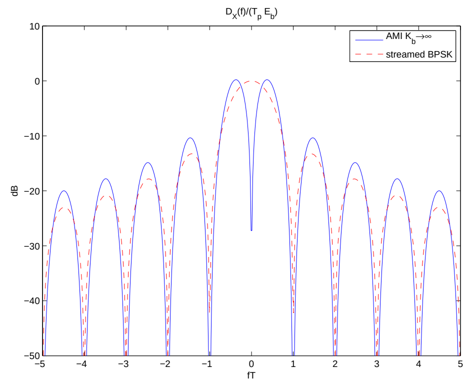
$$X_i(f) \;=\; \sqrt{E_b}\sum_{k=1}^{N_f}\tilde{d}_i^{(k)}U(f)e^{-j2\pi fT(k-1)}$$

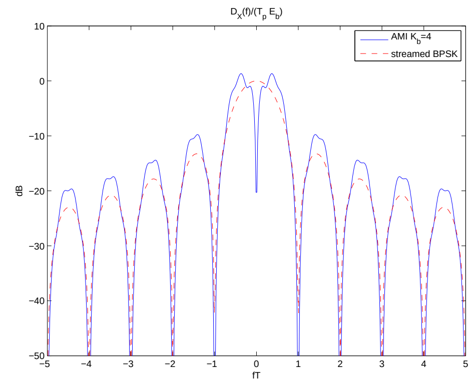
$$D_{X_z}(f) \;=\; \frac{E_b}{K_b 2^{K_b}}\sum_{i=0}^{2^{K_b}-1}\left|\sum_{k=1}^{N_f}\tilde{d}_i^{(k)}U(f)e^{-j2\pi fT(k-1)}\right|^2$$

Possible to evaluate above expression for small K_b.

Example: Alternate Mark Inversion with $K_b = 4$ and rectangular $u(t)$, with streamed BPSK for comparison.







$$
\begin{aligned}
[I^{(1)}, I^{(2)}, I^{(3)}, I^{(4)}] &\rightarrow [\tilde{D}_z^{(1)}, \tilde{D}_z^{(2)}, \tilde{D}_z^{(3)}, \tilde{D}_z^{(4)} | \tilde{D}_z^{(5)}] \\
[0, 0, 0, 0] &\rightarrow [0, 0, 0, 0, 0] \\
[0, 0, 0, 1] &\rightarrow [0, 0, 0, -\sqrt{2}, \sqrt{2}] \\
\vdots \quad & \quad \vdots
\end{aligned}
$$

Find all 2^{K_b} sequences & average the corresponding spectra.

In the case of large K_b, need a different approach...

$$
\begin{aligned}
D_{X_z}(f) &= \frac{E_b}{K_b}\,\mathrm{E}\left|\sum_{l=1}^{N_f}\tilde{D}_z^{(l)}U(f)e^{-j2\pi fT(l-1)}\right|^2 \\
&= \frac{E_b}{K_b}\sum_{l=1}^{N_f}\sum_{k=1}^{N_f}\underbrace{\mathrm{E}\left[\tilde{D}_z^{(l)}\tilde{D}_z^{(k)*}\right]}_{R_{\tilde{D}}[l-k]}|U(f)|^2 e^{-j2\pi fT(l-k)} \\
&= E_b|U(f)|^2\frac{1}{K_b}\sum_{m=-N_f+1}^{N_f-1}(N_f-|m|)R_{\tilde{D}}[m]e^{-j2\pi fTm} \\
&= E_b|U(f)|^2\sum_{m=-\infty}^{\infty}R_{\tilde{D}}[m]e^{-j2\pi fTm} \quad \text{as } K_b\to\infty \\
&= E_b|U(f)|^2 S_{\tilde{D}}(e^{j2\pi fT})
\end{aligned}
$$

To find $R_{\tilde{D}}[m]$, note

$$R_{\tilde{D}}[m] = \mathrm{E}[\tilde{D}_z^{(l)} \tilde{D}_z^{(l-m)*}] = \sum_i \sum_j d_i d_j^* P_{\tilde{D}_z^{(l)}, \tilde{D}_z^{(l-m)}}(d_i, d_j)$$

- The <u>trellis edge</u> $S^{(l)}$ connecting state $\sigma^{(l)}$ to $\sigma^{(l+1)}$ completely determines the symbol $\tilde{D}_z^{(l)}$.
- $S^{(l)}$ can be represented by an integer in $\{1, \ldots, 2N_s\}$.

Thus we note that

$$P_{S^{(l)}, S^{(l-m)}}(s_i, s_j) \quad \text{specifies} \quad P_{\tilde{D}_z^{(l)}, \tilde{D}_z^{(l-m)}}(d_i, d_j)$$

To characterize $P_{S^{(l)}, S^{(l-m)}}(\cdot, \cdot)$, we use the fact that

$$P_{S^{(l)}, S^{(l-m)}}(s_i, s_j) = P_{S^{(l)} | S^{(l-m)}}(s_i | s_j) P_{S^{(l-m)}}(s_j)$$

Assume uniform $P_{S^{(l-m)}}(\cdot)$. To find $P_{S^{(l)}|S^{(l-m)}}(\cdot|\cdot)$, note

$$P_{S^{(l)}}(s_i) = \sum_{s_j=1}^{2N_s} P_{S^{(l)},S^{(l-1)}}(s_i, s_j) = \sum_{s_j=1}^{2N_s} \underbrace{P_{S^{(l)}|S^{(l-1)}}(s_i|s_j)}_{\triangleq\, [\boldsymbol{S}]_{j,i}} P_{S^{(l-1)}}(s_j)$$

where $[\boldsymbol{S}]_{j,i}$ are easily determined from $g_1(I^{(l)}, \sigma^{(l)})$.

Defining the row vector $\underline{P}_{S^{(l)}} \triangleq [P_{S^{(l)}}(1), \ldots, P_{S^{(l)}}(2N_s)]$,

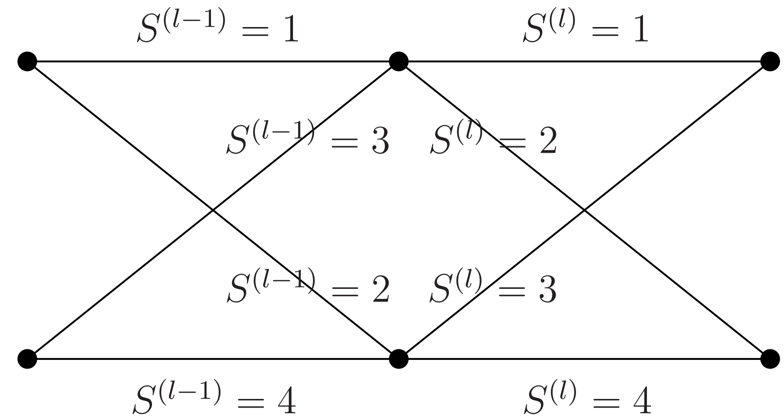
$$\begin{aligned} \underline{P}_{S^{(l)}} &= \underline{P}_{S^{(l-1)}}\boldsymbol{S}, \qquad \underline{P}_{S^{(l-1)}} = \underline{P}_{S^{(l-2)}}\boldsymbol{S} \\ \Rightarrow \underline{P}_{S^{(l)}} &= \underline{P}_{S^{(l-m)}}\boldsymbol{S}^m \end{aligned}$$

From the definition of $[\boldsymbol{S}]_{j,i}$ above, we can now see that

$$P_{S^{(l)}|S^{(l-m)}}(s_i|s_j) = [\boldsymbol{S}^m]_{j,i}.$$

AMI Example: Assume $\underline{P}_{S^{(l-m)}} = [\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}]$.

$$\boldsymbol{S} = \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \end{pmatrix}$$



For any $m > 1$, notice that all entries of $\boldsymbol{S}^m$ equal 0.25!

- $\rightsquigarrow$ Edges $S^{(l)}$ & $S^{(l-m)}$ independent for $m > 1$.
- $\rightsquigarrow$ Symbols $\tilde{D}_z^{(l)}$ & $\tilde{D}_z^{(l-m)}$ uncorrelated for $m > 1$.

It can be shown (see next page) that

$$\begin{aligned} \{R_{\tilde{D}}[0], R_{\tilde{D}}[1], R_{\tilde{D}}[2], R_{\tilde{D}}[3], \dots\} &= \{1, -\tfrac{1}{2}, 0, 0, \dots\} \\ \Rightarrow S_{\tilde{D}}(e^{j2\pi fT}) = \textstyle\sum_{m=-\infty}^{\infty} R_{\tilde{D}}[m] e^{-j2\pi fTm} &= 1 - \cos(2\pi fT) \end{aligned}$$

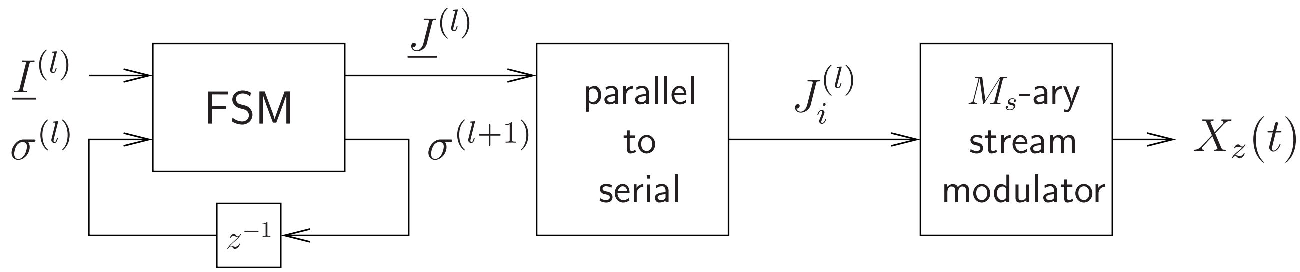
So how did we figure out that $-\frac{1}{2} = R_{\tilde{D}}[1] = R_{\tilde{D}}[-1]^*$?

$$\underbrace{P_{S^{(l-1)}}(s_j)}_{\frac{1}{4}\forall j} \quad \underbrace{P_{S^{(l)}|S^{(l-1)}}(s_i|s_j)}_{j \downarrow \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}} = \underbrace{P_{S^{(l)},S^{(l-1)}}(s_i, s_j)}_{\begin{pmatrix} \frac{1}{8} & \frac{1}{8} & 0 & 0 \\ 0 & 0 & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & 0 & 0 \\ 0 & 0 & \frac{1}{8} & \frac{1}{8} \end{pmatrix}}$$

$$\begin{aligned} R_{\tilde{D}}[1] &= \underbrace{\sum_i \sum_j d_j^* \; P_{\tilde{D}_z^{(l)} \tilde{D}_z^{(l-1)}}(d_i, d_j) \; d_i}_{\begin{pmatrix} 0 & -\sqrt{2} & \sqrt{2} & 0 \end{pmatrix}^* \begin{pmatrix} \frac{1}{8} & \frac{1}{8} & 0 & 0 \\ 0 & 0 & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & 0 & 0 \\ 0 & 0 & \frac{1}{8} & \frac{1}{8} \end{pmatrix} \begin{pmatrix} 0 \\ -\sqrt{2} \\ \sqrt{2} \\ 0 \end{pmatrix}} \\ &= -\frac{1}{2} \end{aligned}$$

$S^{(l)}$	$\tilde{D}_z^{(l)}$
1	0
2	$-\sqrt{2}$
3	$\sqrt{2}$
4	0

Coded Modulation for General R:



- K_b bits parsed into N_b blocks of K_m bits $(K_b = N_b K_m)$
- The FSM accepts $\underline{I}^{(l)}$, a block of K_m bits, and produces $\underline{J}^{(l)}$, a block of N_m constellation labels.

$$\begin{aligned} \sigma^{(l+1)} &= g_1(\sigma^{(l)}, \underline{I}^{(l)}) \\ \underline{J}^{(l)} &= g_2(\sigma^{(l)}, \underline{I}^{(l)}) \end{aligned}$$

- Label sequence drives M_s-ary linear stream modulation:

$$\tilde{D}_z^{((l-1)N_m+i)} = a\big(J_i^{(l)}\big) \text{ for } i = 1, \ldots, N_m$$

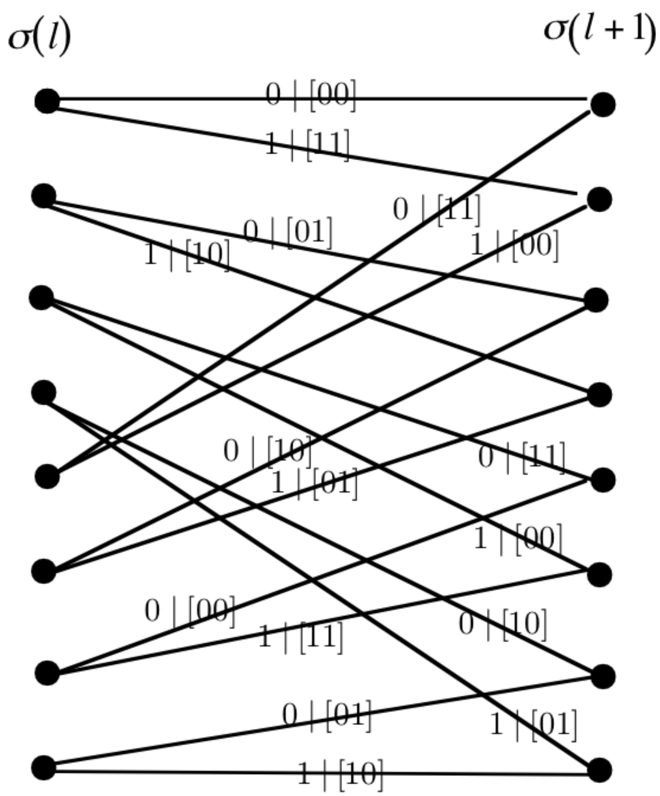
- Total # of symbols in frame is $N_f = N_b N_m + \nu_c$.
- Effective rate (bits/channel-use) is

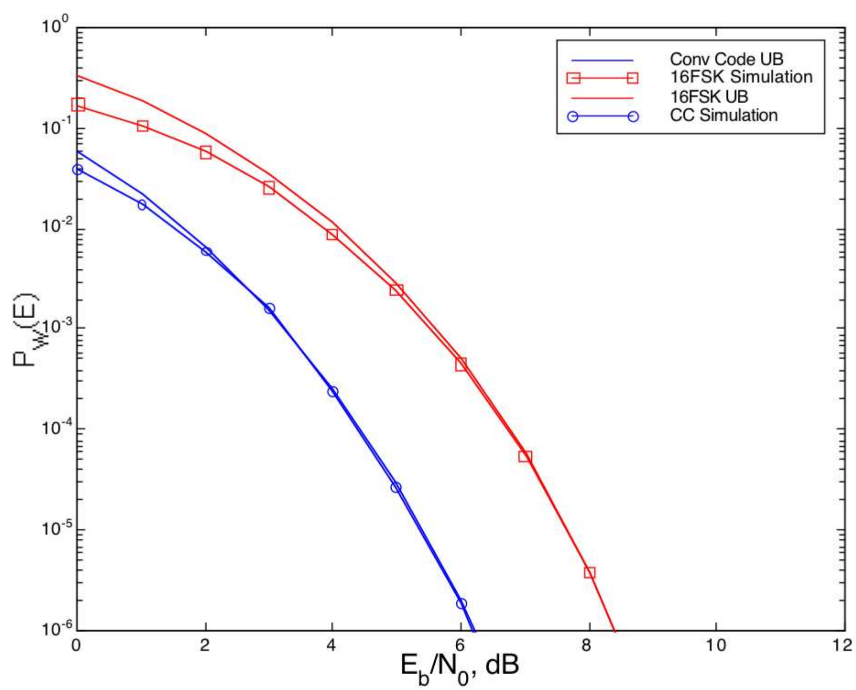
$$R_{\text{eff}} = \frac{K_b}{N_b N_m + \nu_c} = \frac{N_b K_m}{N_b N_m + \nu_c} \approx \frac{K_m}{N_m} \triangleq R.$$

- Might choose $R > 1$ or $R < 1$ depending on desired performance/spectral-efficiency tradeoff.
- Symbols always normalized so that $\mathrm{E}\big|\tilde{D}_z^{(l)}\big|^2 = R$.

$R < 1$ Example: Convolutional Code

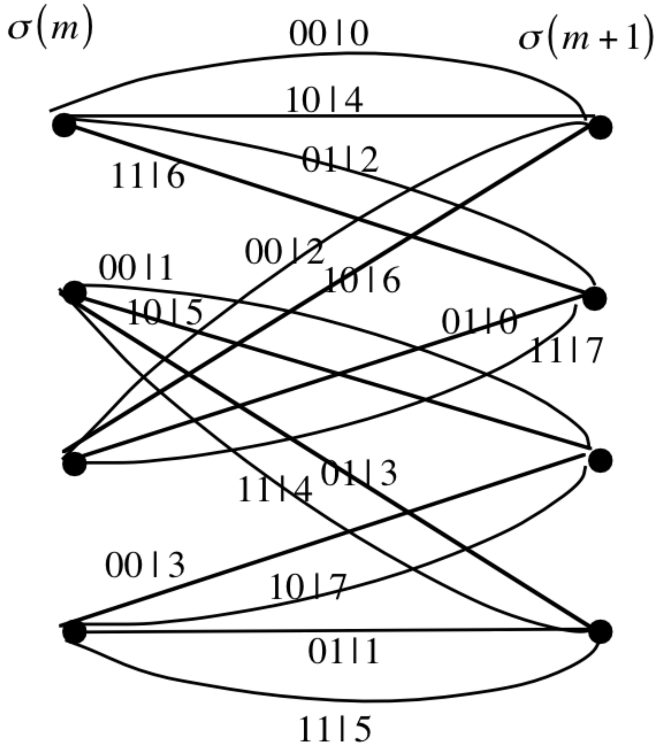
- BPSK ($M_s = 2$), $R = \frac{1}{2}$ ($K_m = 1, N_m = 2$), $N_s = 8$
- ≈2 dB better than 16-FSK at similar spectral efficiency.

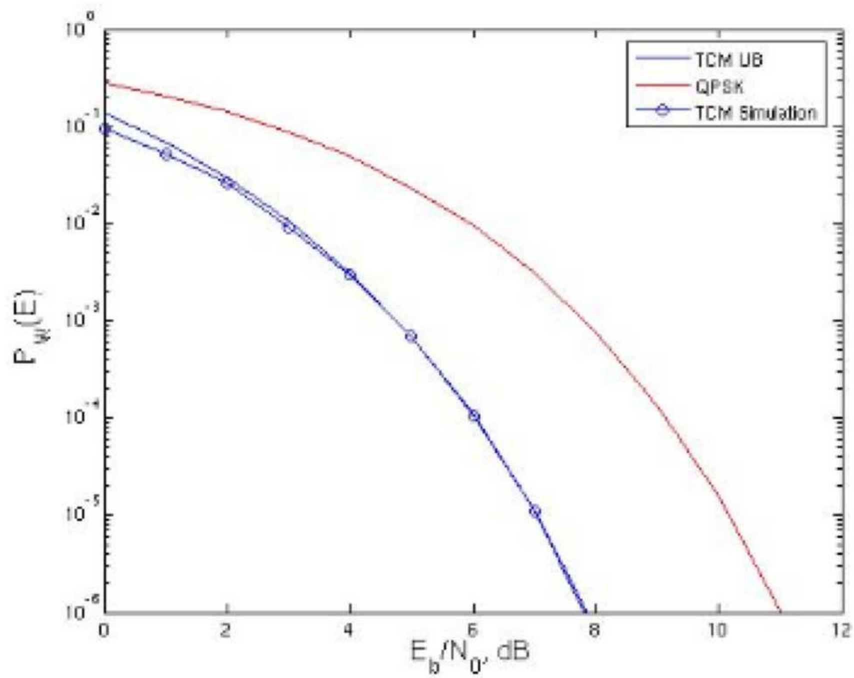




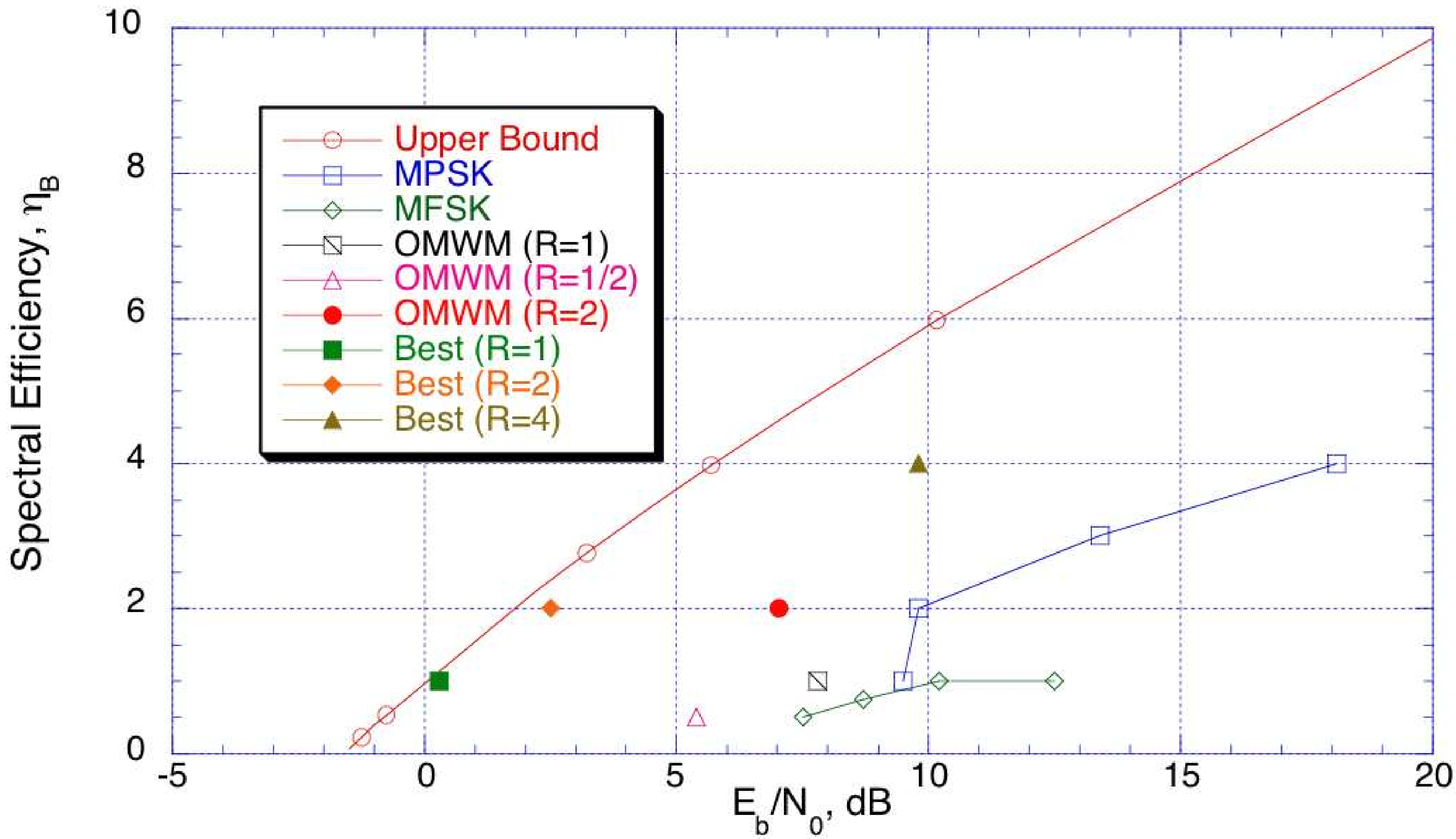
$R > 1$ Example: Trellis Code

- 8-PSK ($M_s = 8$), $R = 2$ ($K_m = 2, N_m = 1$), $N_s = 4$.
- ≈3 dB better than QPSK at same spectral efficiency.



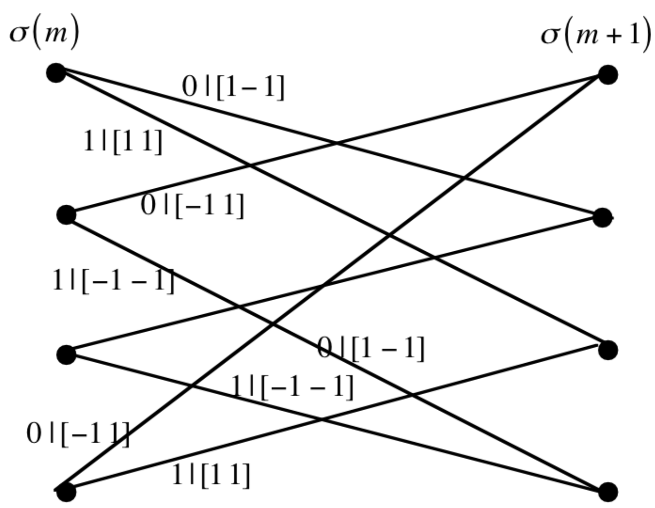


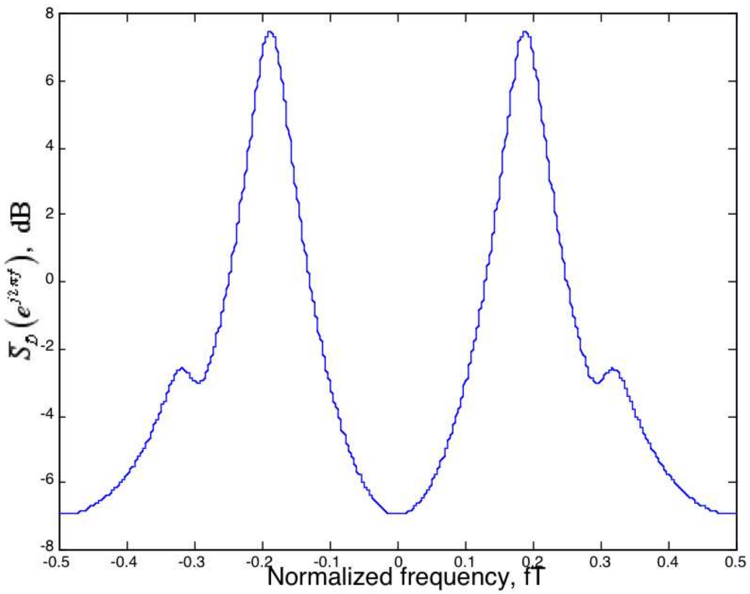
Spectral efficiency of coded modulation schemes:



Spectral Shaping Example: the Miller Code

- $N_s = 4$, BPSK ($M_s = 2$), $R = \frac{1}{2}$ ($K_m = 1, N_m = 2$)
- Run Length Limited: ≤ 4 same symbols in a row.
- Used in magnetic recording, since low frequencies interfere with servo mechanism of read/write head.





Analyzing Spectral Characteristics when $N_m > 1$

Notice: N_m = # of symbols/block = # symbols/edge.

$$D_{X_z}(f) = \frac{E_b}{K_b} \sum_{l=1}^{N_f} \sum_{k=1}^{N_f} \mathrm{E}\left[\tilde{D}_z^{(l)} \tilde{D}_z^{(k)*}\right] |U(f)|^2 e^{-j2\pi fT(l-k)}$$

$$\mathrm{E}\left[\tilde{D}_z^{(l)} \tilde{D}_z^{(k)*}\right] = R_{\tilde{D}}\big[l-k, \langle l-1\rangle_{N_m} + 1\big] \quad \text{"cyclostationary"}$$

As $K_b \to \infty$, the same techniques used before yield

$$D_{X_z}(f) = E_b |U(f)|^2 \sum_{m=-\infty}^{\infty} \underbrace{\frac{R}{N_m} \sum_{l=1}^{N_m} R_{\tilde{D}}[m,l]}_{\triangleq \overline{R}_{\tilde{D}}[m]} e^{-j2\pi fTm}$$

$$= E_b |U(f)|^2 \overline{S}_{\tilde{D}}(e^{j2\pi fT})$$

To find $R_{\tilde{D}}[m,l]$, we start with the definition

$$R_{\tilde{D}}[m,l] = \sum_i \sum_j d_i d_j^* P_{\tilde{D}_z^{((q-1)N_m+l)},\tilde{D}_z^{((q-1)N_m+l-m)}}(d_i,d_j), \ \forall q$$

Noting that the edge determines the symbol-block:

$$S^{(q)} \quad \rightarrow \quad \{\tilde{D}_z^{((q-1)N_m+l)}\}_{l=1}^{N_m}$$

1. Use, as before, $[\boldsymbol{S}^p]_{j,i} = P_{S^{(q)}|S^{(q-p)}}(i|j)$ and the uniform $P_{S^{(q-p)}}(\cdot)$ assumption to find $P_{S^{(q)},S^{(q-p)}}(\cdot,\cdot)$.
2. Use $P_{S^{(q)},S^{(q-p)}}(\cdot,\cdot)$ and the trellis description to find $R_{\tilde{D}}[m,l]$.