4.2.4 Orthogonal, Biorthogonal and Simplex Signals

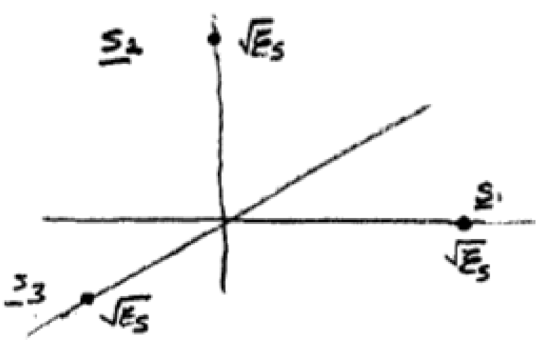
- In PAM, QAM and PSK, we had only one basis function. For orthogonal, biorthogonal and simplex signals, however, we use more than one orthogonal basis function, so N-dimensional Examples: the Fourier basis; time-translated pulses; the Walsh-Hadamard basis.
    - o We've become used to SER getting worse quickly as we add bits to the symbol – but with these orthogonal signals it actually gets *better*.
    - o The drawback is bandwidth occupancy; the number of dimensions in bandwidth W and symbol time T_s is

      $$N \approx 2W\,T_s$$

    - o So we use these sets when the power budget is tight, but there's plenty of bandwidth.

Orthogonal Signals

- With orthogonal signals, we select only one of the orthogonal basis functions for transmission:



$$\underline{s}_1 = \begin{bmatrix} \sqrt{E_s} \\ 0 \\ 0 \end{bmatrix} \quad \underline{s}_2 = \begin{bmatrix} 0 \\ \sqrt{E_s} \\ 0 \end{bmatrix} \quad \underline{s}_3 = \begin{bmatrix} 0 \\ 0 \\ \sqrt{E_s} \end{bmatrix}$$

The number of signals M equals the number of dimensions N.

- Examples of orthogonal signals are frequency-shift keying (FSK), pulse position modulation (PPM), and choice of Walsh-Hadamard functions (note that with Fourier basis, it's FSK, *not* OFDM).

- Energy and distance:

  - The signals are equidistant, as can be seen from the sketch or from

$$\mathbf{s}_i - \mathbf{s}_j = \begin{bmatrix} \vdots \\ \sqrt{E_s} \\ \vdots \\ -\sqrt{E_s} \\ \vdots \end{bmatrix} \begin{array}{l} \\ \leftarrow \text{ location } i \\ \\ \leftarrow \text{ location } j \\ \\ \end{array}$$

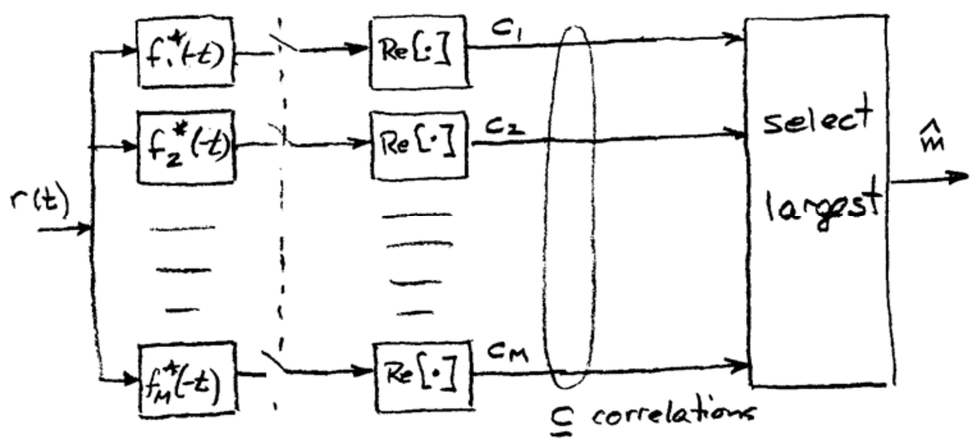
so $\left\|\mathbf{s}_i - \mathbf{s}_j\right\| = \sqrt{2E_s} = \sqrt{2\log_2(M)E_b} = \sqrt{2kE_b} = d_{\min}, \ \forall i, j$

- Effect of adding bits:

  - For a fixed energy per bit, adding more bits *increases* the minimum distance – strong contrast to PAM, QAM, PSK.

  - But adding each bit doubles the number of signals *M*, which equals the number of dimensions *N* – and that doubles the bandwidth!

- Error analysis for orthogonal signals. Equal energy, equiprobable signals, so receiver is



  - Error probability is same for all signals. If s_1 was sent, then the correlation vector is

$$\mathbf{c} = \begin{bmatrix} \sqrt{E_s} + n_1 \\ n_2 \\ \vdots \\ n_M \end{bmatrix}$$ with real components.

- Assume s_1 was sent. The probability of a *correct* symbol decision, conditioned on the value of the received c_1, is

$$P_{cs}(c_1) = P\left[(n_2 < c_1) \wedge (n_3 < c_1) \wedge \ldots \wedge (n_M < c_1)\right]$$

$$= \left(1 - Q\left(\frac{c_1}{\sqrt{N_0/2}}\right)\right)^{M-1}$$

- Then the *unconditional* probability of correct symbol detection is

$$P_{cs} = \int_{-\infty}^{\infty} \left(1 - Q\left(\frac{c_1}{\sqrt{N_0/2}}\right)\right)^{M-1} p_{c_1}(c_1)\, dc_1$$

$$= \int_{-\infty}^{\infty} \left(1 - Q\left(\frac{c_1}{\sqrt{N_0/2}}\right)\right)^{M-1} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{N_0/2}} \exp\left(-\frac{1}{2}\frac{2}{N_0}\left(c_1 - \sqrt{E_s}\right)^2\right) dc_1$$

$$= \int_{-\infty}^{\infty} \left(1 - Q(u)\right)^{M-1} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(u - \sqrt{2\gamma_s}\right)^2\right) du$$

Needs a numerical evaluation.

- The unconditional probability of symbol error (SER) is

$$P_{es} = 1 - P_{cs}$$

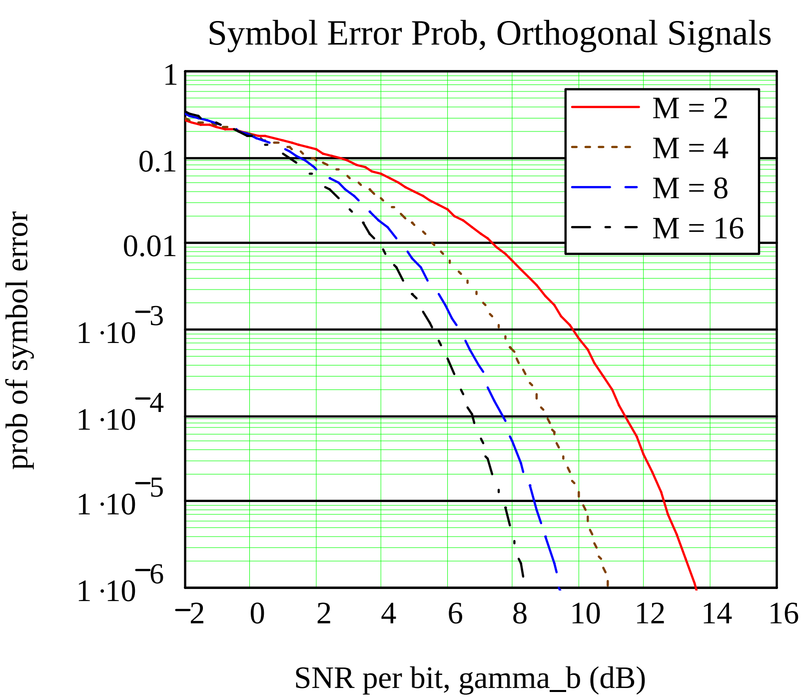
- Now for the *bit* error rate.
  - All errors equally likely, at $\frac{P_{es}}{M-1} = \frac{P_{es}}{2^k - 1}$. No point in Gray coding.
  - Can get i bit errors in $\binom{k}{i}$ equiprobable ways, so

$$P_{eb} = \frac{P_s}{2^k - 1} \sum_{i=1}^{k} i \binom{k}{i} = \frac{P_s}{2^k - 1} \sum_{i=1}^{k} k \binom{k-1}{i-1}$$

$$= k \frac{2^{k-1}}{2^k - 1} P_s \approx \frac{k}{2} P_s$$

About half the bits are in error in the average symbol error.

- The SER for orthogonal signals is striking:



The SER *improves* as we add bits!

But the bandwidth doubles with each additional bit.

- Why does SER drop as M increases?

  - All points are separated by the same $d = \sqrt{2\log_2(M)E_b}$. Since d^2 increases linearly with $k = \log_2(M)$, the *pairwise* error probability (between two specific points) *decreases* exponentially with *k*, the number of bits per symbol.

  - Offsetting this is the number *M*-1 of neighbours of any point: *M*-1 ways of getting it wrong, and *M increases* exponentially with *k*.

  - Which effect wins? Can't get much analytical traction from the integral expression two pages back. So fall back to a union bound.

- Union bound analysis:

  - Without loss of generality, assume $\mathbf{s}_1$ was transmitted.

  - Define $E_m,\ m = 2,\ldots,M$ as the event that $\mathbf{r}$ is closer to $\mathbf{s}_m$ than to $\mathbf{s}_1$.

- Note that E_m is a *pairwise* error. It does not imply that the receiver's decision is *m*. In fact, we have events $E_2, \ldots, E_M$ and they are not mutually exclusive, since **r** could lie closer to two or more points than to $\mathbf{s}_1$:

- The overall error event $E = E_2 \cup E_3 \cup \ldots \cup E_M$, so the SER is bounded by the sum of the pairwise event probabilities[1]

$$P_{es} = P[E] = P[E_2 \cup E_3 \cup \ldots \cup E_M] \le \sum_{m=2}^{M} P[E_m]$$

$$= (M-1)Q\left(\frac{d/2}{\sqrt{N_0/2}}\right) = (M-1)Q\left(\sqrt{\log_2(M)\gamma_b}\right)$$

$$\le 2^k \frac{1}{2} e^{-k\gamma_b/2} < e^{-k\left(\frac{\gamma_b}{2} - \ln(2)\right)}$$

[1] A general expansion is

$$P[A_1 \cup A_2 \cup \ldots \cup A_M] = \sum_{i=1}^{M} P[A_i] - \sum_{i=i}^{M} \sum_{\substack{j=1 \\ j\neq i}}^{M} P[A_i \cap A_j]$$

$$+ \sum_{i=1}^{M} \sum_{\substack{j=1 \\ j\neq i}}^{M} \sum_{\substack{k=1 \\ k\neq j \\ k\neq i}}^{M} P[A_i \cap A_j \cap A_k] - \cdots \text{ etc.}$$

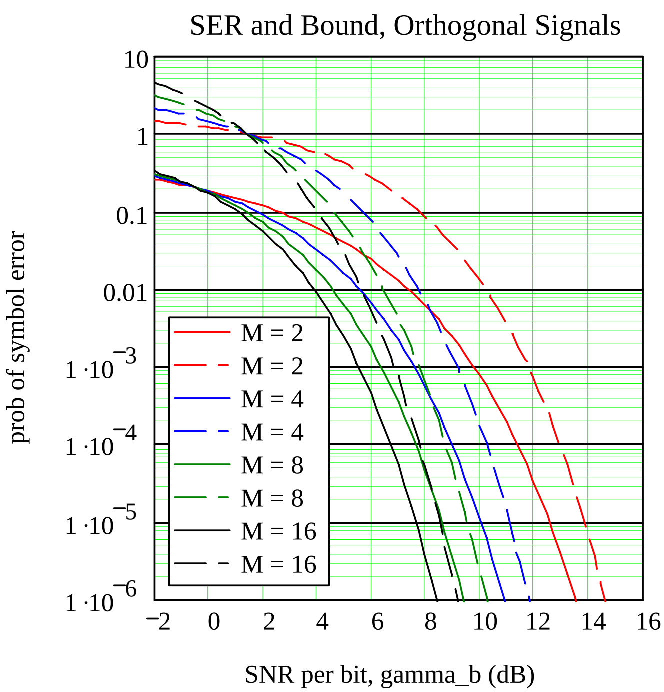
- From the bound, if $\gamma_b > 2\ln(2) = 1.386$, then loading more bits onto a symbol causes the upper bound on SER to drop exponentially to zero.

- Conversely, if $\gamma_b < 2\ln(2)$, increasing *k* causes the upper bound on SER to rise exponentially (SER itself saturates at 1).

- Too bad the dimensionality, the number of correlators and the bandwidth also increase exponentially with *k*.

- This scheme is called "block orthogonal coding." Thresholds are characteristic of coded systems.

- Remember that this analysis is based on bounds…

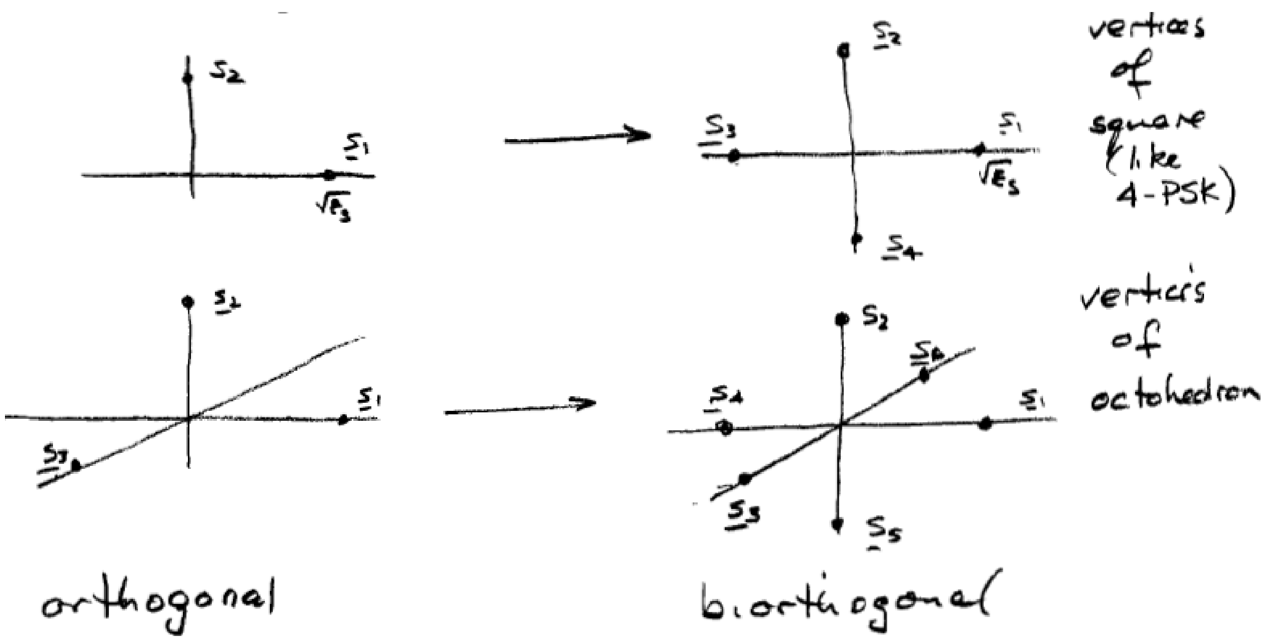


True SERs are solid lines, bounds are dashed.

Upper bound is useful to show decreasing SER, but is too loose for a good approximation.

Biorthogonal Signals

- If you have coherent detection, why waste the other side of the axis? Double the number of signal points (add one bit) with $\pm\sqrt{E_s}$. Now $M = 2N$, with no increase in bandwidth. Or keep the same M and cut the bandwidth in half.



- Energy and distance:

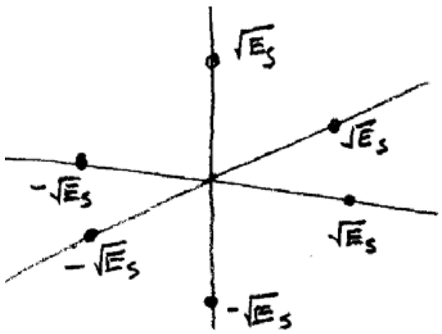
All signal points are equidistant from $\mathbf{s}_i$

$$\left\| \mathbf{s}_i - \mathbf{s}_k \right\| = \sqrt{2E_s} = d_{\min} \quad \text{(the same as orthogonal signals)}$$

except one – the reflection through the origin – which is farther away

$$\left\| \mathbf{s}_i - {}^{-}\mathbf{s}_i \right\| = 2\sqrt{E_s}$$

- Symbol error rate:



The probability of error is messy, but the union bound is easy. A signal is equidistant from all other signals but its own complement.

- So the union bound on SER is

$$\begin{aligned} P_s &\le (M-2)Q\left(\sqrt{\gamma_s}\right) + Q\left(\sqrt{2\gamma_s}\right) \\ &\le (M-2)Q\left(\sqrt{\gamma_s}\right) \quad \text{second term redundant (Sec'n 4.5.4)} \\ &= (M-2)Q\left(\sqrt{\log_2(M)\,\gamma_b}\right) \end{aligned}$$

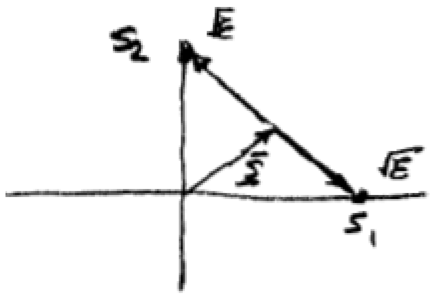
which is slightly less than orthogonal signaling, for the same M and same γ_b. The major benefit is that biorthogonal needs only half the bandwidth of orthogonal, since it has half the number of dimensions.

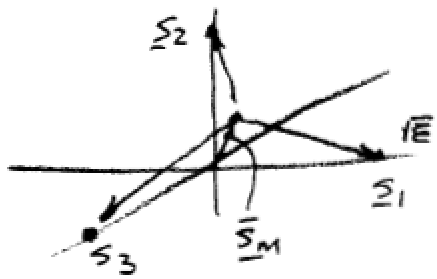
- And the bit error probability is

$$P_b \approx \frac{\log_2(M)}{2} P_s$$

Simplex Signals

- The orthogonal signals can be seen as a mean values, shared by all, and signal-dependent increments:





The mean signal (the centroid) is $\bar{\mathbf{s}} = \dfrac{1}{M}\sum_{m=1}^{M}\mathbf{s}_m$.

  - The mean does not contribute to distinguishability of signals – so why not subtract it?

$$\mathbf{s}'_m = \mathbf{s}_m - \bar{\mathbf{s}} \qquad\qquad \mathbf{s}'_m = \sqrt{E_s}\begin{bmatrix} -1/M \\ -1/M \\ \vdots \\ 1-1/M \\ \vdots \\ -1/M \end{bmatrix}\begin{array}{l} \\ \\ \\ \leftarrow \text{ location } m \\ \\ \\ \end{array}$$

  where $\sqrt{E_s}$ is the original signal energy.

- Removing the mean lowers the dimensionality by one. After rotation into tidier coordinates:

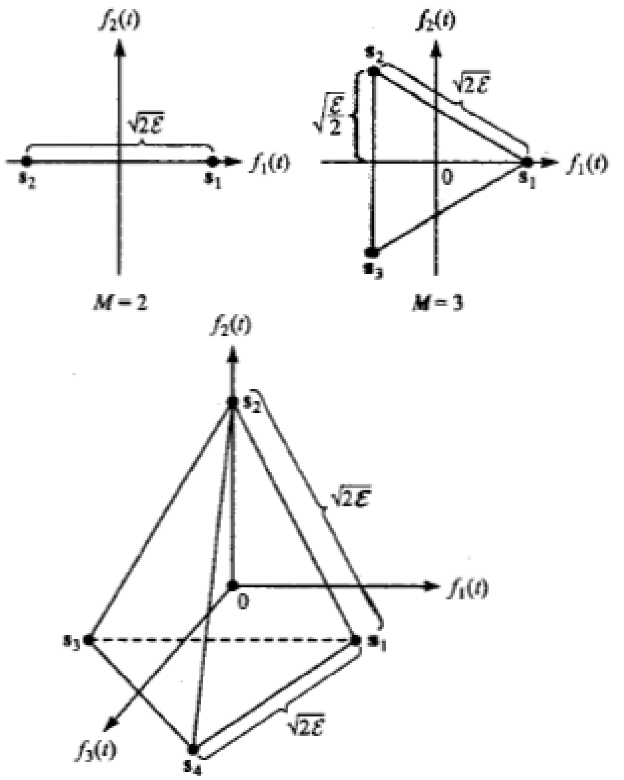


FIGURE 4.3–10
Signal space diagrams for M-ary simplex signals.

Vertices of
- line segment
- equilateral triangle
- tetrahedron.

- They still have equal energy (though no longer orthogonal). That energy is

$$E_s' = \left\| \mathbf{s}_m' \right\|^2 = E_s \left(M \frac{1}{M^2} + 1 - \frac{2}{M} \right) = E \left(1 - \frac{1}{M} \right)$$

The energy saving is small for larger M.

- The correlation among signals is:

$$\mathrm{Re}\left[\rho_{mn}\right] = \frac{\left(\mathbf{s}'_m, \mathbf{s}'_n\right)}{\left\|\mathbf{s}'_m\right\| \cdot \left\|\mathbf{s}'_n\right\|} = \frac{-\dfrac{1}{M}}{1-\dfrac{1}{M}} = -\frac{1}{M-1} \quad \text{for } m \neq n$$

A uniform negative correlation.

- The SER is easy. Translation doesn't change the error rate, so the SER is that of orthogonal signals with a $M/(M-1)$ SNR boost.

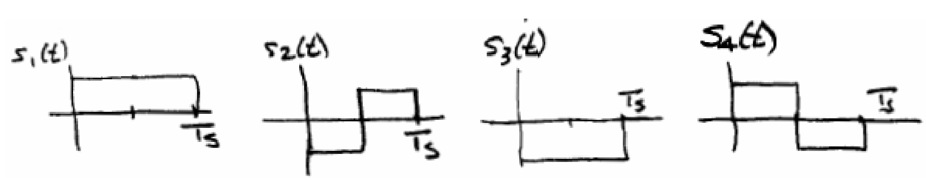
4.2.5 Vertices of a Hypercube and Generalizations

- More signals defined on a multidimensional space. Unlike orthogonal, we will now allow more than one dimension to be used in a symbol.

- Vertices of a hypercube is straightforward: two-level PAM (binary antipodal) on each of the N dimensions:

  - With the Fourier basis, it is BPSK on each frequency at once – a simple OFDM

  - With the time translate basis, it is a classical NRZ transmission:



- Signal vectors:

$$\mathbf{s}_m = \begin{bmatrix} s_{m1} \\ s_{m2} \\ \vdots \\ s_{mN} \end{bmatrix} \qquad \text{with} \quad s_{mn} = \pm\sqrt{E_s / N}$$

for $m = 1, \ldots M$ and $M = 2^N$.

- In space, it looks like this

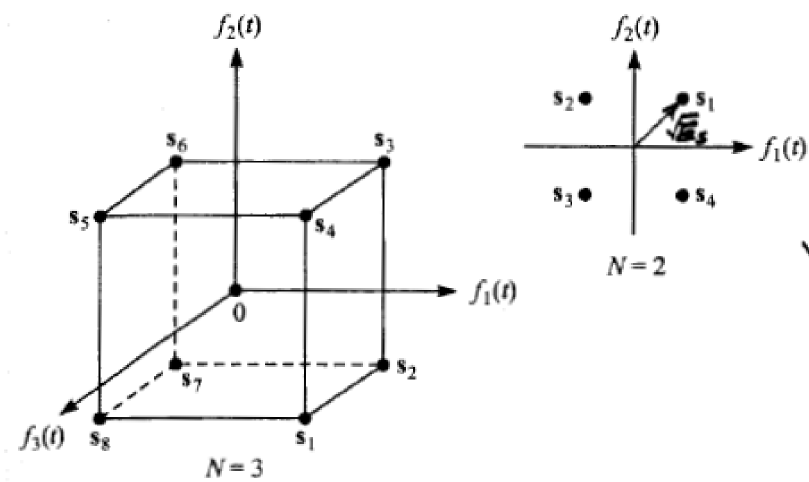


FIGURE 4.3–11
Signal space diagrams for signals generated from binary codes.

"vertices of a hypercube"

- Every point has the same distance from the origin, hence the same energy

$$\|\mathbf{s}_m\|^2 = E_s = \log_2(M)\, E_b = N\, E_b$$

- Minimum distance occurs for differences in a single coordinate:

$$\mathbf{s}_m = \sqrt{\frac{E_s}{N}} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{s}_n = \sqrt{\frac{E_s}{N}} \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$$

$$d_{\min} = \left\| \mathbf{s}_m - \mathbf{s}_n \right\| = 2\sqrt{\frac{E_s}{N}} = 2\sqrt{E_b}$$

More generally, for d_H disagreements (Hamming distance),

$$\left\| \mathbf{s}_m - \mathbf{s}_n \right\| = 2\sqrt{\frac{d_H E_s}{N}} = 2\sqrt{d_H E_b}$$

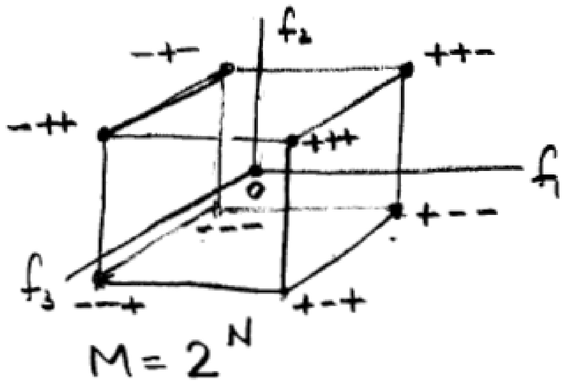
- What are the effects on $d_{\min}$ and bandwidth if we increase the number of bits, keeping E_b fixed?

- It's easy to generalize from binary to PAM, PSK, QAM, etc. on each of the dimensions:

  - With the Fourier basis, it is OFDM

  - With time translates, it is the usual serial transmission.

- Error analysis:



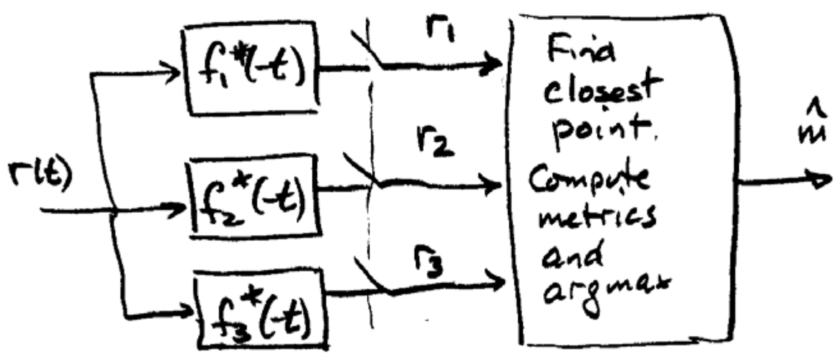
All points have same energy

$E_s = N\,E_b$

Euclidean distance for Hamming

distance h is $d = 2\sqrt{h\,E_b}$

  - If labeling is done independently on different dimensions (see sketch), then the independence of the noise causes it to decompose to N independent detectors. So the following structure

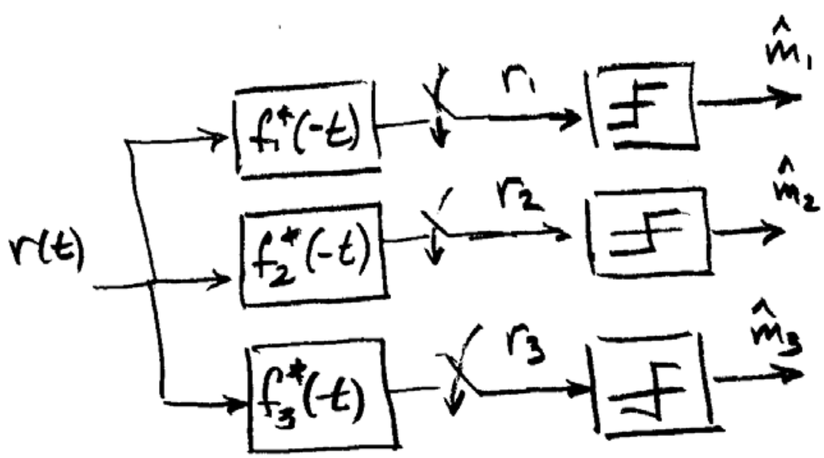


$P_s \le N\,Q\left(\sqrt{2\gamma_b}\right)$

$= \log_2(M)Q\left(\sqrt{2\gamma_b}\right)$

P_b depends on labels

becomes



P_s is meaningless

$P_b = Q\left(\sqrt{2\gamma_b}\right)$, just binary antipodal

- Generalization: user multilevel signals (PAM) in each dimension. Again, if labeling is independent by dimension, then it's just independent and parallel use, like QAM. Not too exciting. Yet. These signals form a finite lattice.

- Generalization: use a subset of the cube vertices, with points selected for greater minimum distance. This is a binary block code.
    - Advantage: greater Euclidean distance
    - Disadvantage: lower data rate