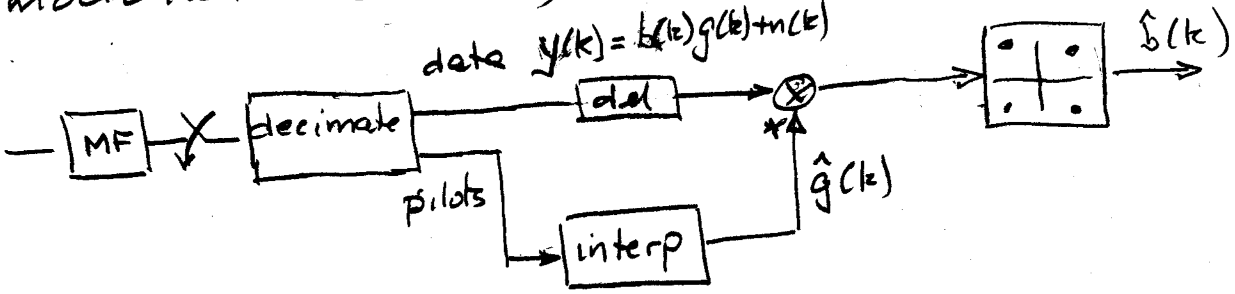
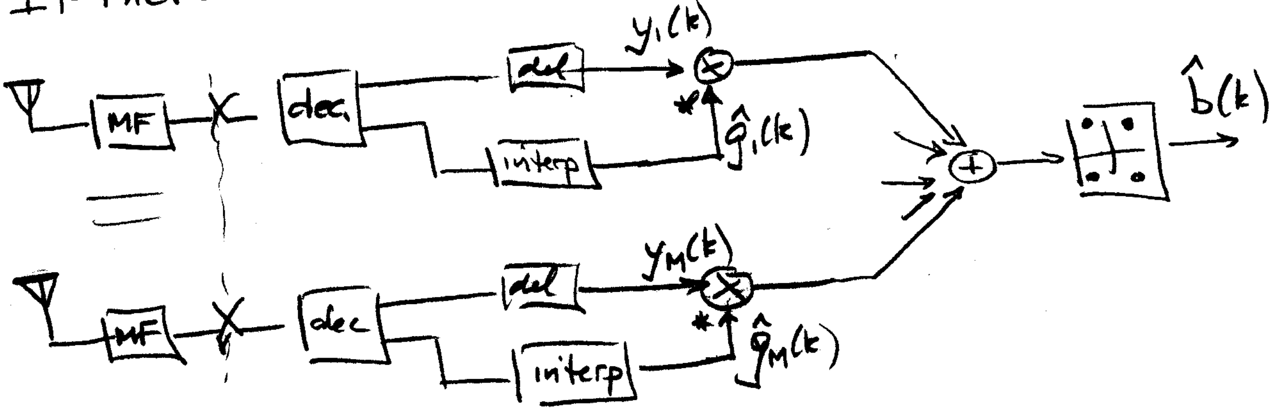
4.5 Channel Estimation From Training Sequences

- We have seen the value of accurate channel estimates in detection of single users. Channel estimation is even more important in multi user detection. How to get those estimates?

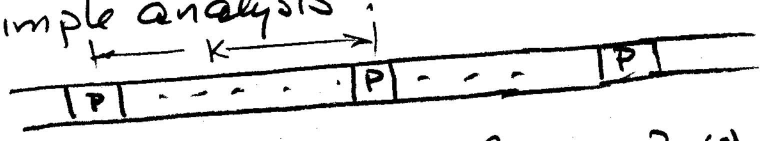
- The easiest example is pilot symbol assisted modulation (PSAM). Every K^{th} symbol is known.



If there are diversity Rx antennas, then



– simple analysis:



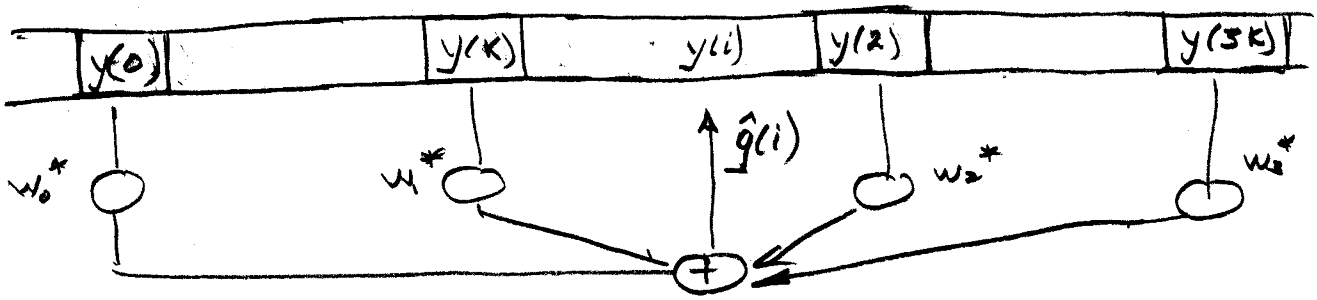
$R_g(k)$ gain autocorr'n f'n, $R_g(0) = \sigma_g^2$

σ_n^2 noise sample variance

$b(iK) = 1$, known, arbitrarily 1

$y(k) = g(k)\,b(k) + n(k)$ describes all samples.

– design interpolator:



$$\hat{g}(i) = \underline{w}_i^{+} \underline{y} = [w_0^* \, w_1^* \, w_2^* \, w_3^*] \begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ y(3) \end{bmatrix}$$

Classic linear estimation:

$$\min \overline{|g(i) - \hat{g}(i)|^2} = \overline{|g(i)|^2} - \underline{w}_i^{+} \overline{\underline{y}\, g^*(i)} - \overline{g(i)\, \underline{y}^{+}}\, \underline{w} + \underline{w}_i^{+} \overline{\underline{y}\,\underline{y}^{+}}\, \underline{w}$$

$$= \sigma_g^2 - \underline{w}_i^{+} \underline{p}_i - \underline{p}_i^{+} \underline{w}_i + \underline{w}_i^{+} R_y \underline{w}_i$$

where

$$R_y = \begin{bmatrix} \sigma_g^2 + \sigma_n^2 & R_g(1) & R_g(2) & R_g(3) \\ R_g^*(1) & \sigma_g^2 + \sigma_n^2 & R_g(1) & R_g(2) \\ R_g^*(2) & R_g^*(1) & \sigma_g^2 + \sigma_n^2 & R_g(1) \\ R_g^*(3) & R_g^*(2) & R_g^*(1) & \sigma_g^2 + \sigma_n^2 \end{bmatrix}$$

$$\underline{p}_i = \begin{bmatrix} R_g(i) \\ R_g(i-K) \\ R_g(i-2K) \\ R_g(i-3K) \end{bmatrix}$$

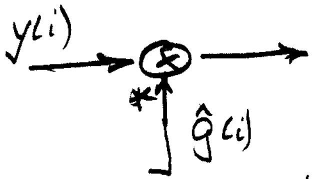
with solution

$$\underline{w}_i = R_y^{-1} \underline{p}_i \qquad \hat{g}(i) = \underline{w}_i^{+} \underline{y} = \underline{p}_i^{+} R_y^{-1} \underline{y}$$

and error variance $E[|g(i) - \hat{g}(i)|^2] = \sigma_g^2 - \underline{w}_i^{+} R_y^{-1} \underline{w}_i = \sigma_e^2$

e is uncorrelated with pilot samples $\underline{y}$
and with $\hat{g}$

— Now calculate the error rate.



We know that the corr'n coeff $\rho = \dfrac{\sigma_{y\hat{g}}^2}{\sigma_y \sigma_{\hat{g}}}$ determines performance. Assume BPSK

$$\sigma_y^2 = \tfrac{1}{2}E\left[\left|b(i)g(i)+n(i)\right|^2\right] = \sigma_g^2 + \sigma_n^2$$

$$\sigma_{\hat{g}}^2 = \tfrac{1}{2}E\left[\underline{p}_i^{\dagger} R_y^{-1} \underline{y}\,\underline{y}^{\dagger} R_y^{-1} \underline{p}_i\right] = \underline{p}_i^{\dagger} R_y \underline{p}_i$$

$$\sigma_{y\hat{g}}^2 = \tfrac{1}{2}E\left[\left(b(i)\left(\hat{g}(i)+e(i)\right)+n(i)\right)\hat{g}^*(i)\right] = b(i)\,\sigma_{\hat{g}}^2$$

$$= b(i)\,\underline{p}_i^{\dagger} R_y \underline{p}$$

so

$$\rho = \frac{b(i)\,\underline{p}_i^{\dagger} R_y \underline{p}_i}{\sqrt{\left(\sigma_g^2+\sigma_n^2\right)\underline{p}_i^{\dagger} R_y^{-1} \underline{p}_i}}$$

for $b(i) = +1$ then $BER = \tfrac{1}{2}(1-\rho)$

— Performance:

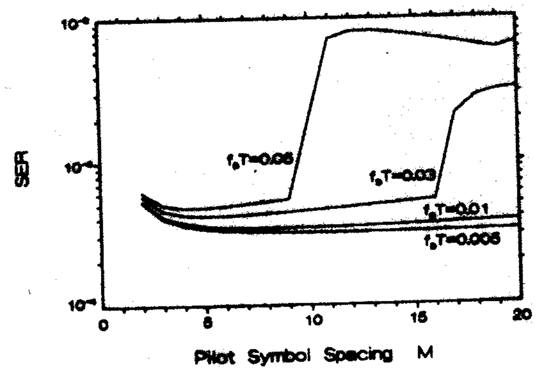


Fig. 3. Effect of frame size on BPSK (γ_b = 30 dB, K = 63).

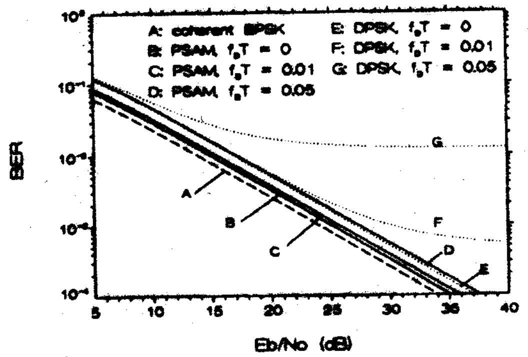


Fig. 5. BER performance of optimized PSAM for BPSK.

Cavers
Trans VT
Nov 91

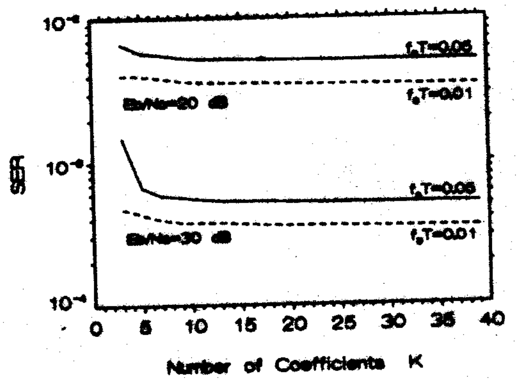
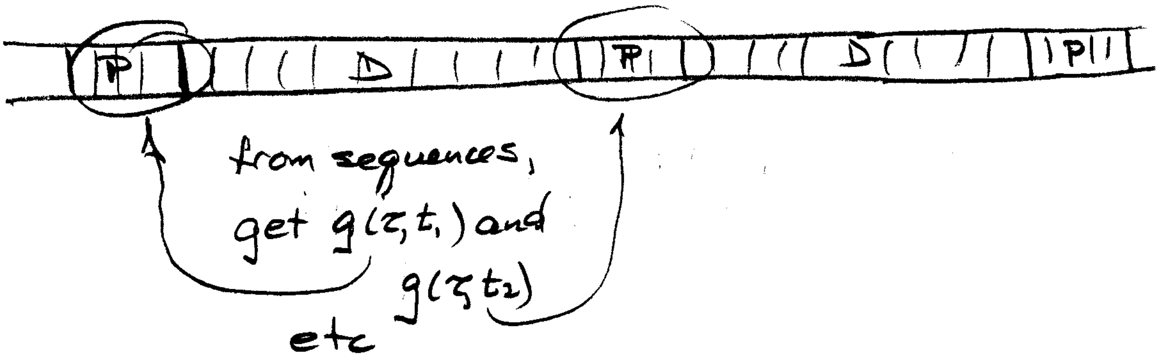
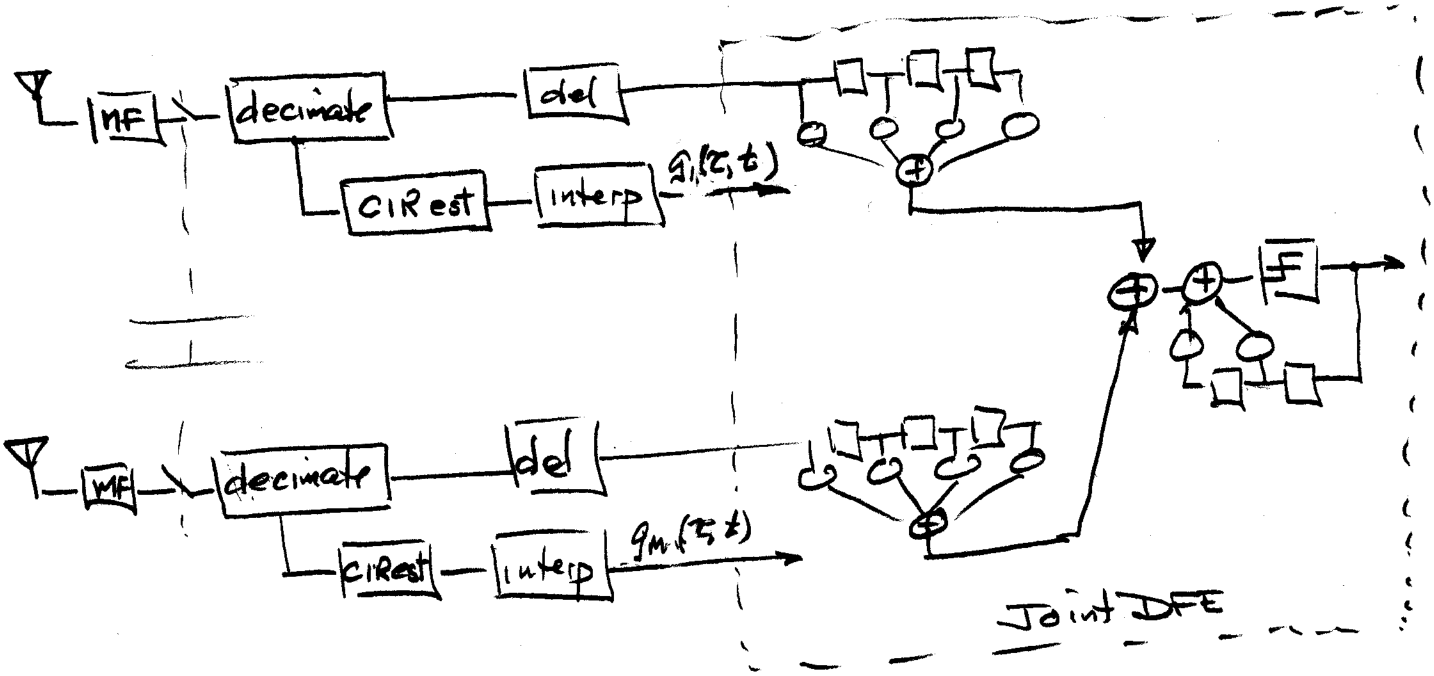


Fig. 4. Effect of interpolator size on BPSK (M = 7).

- If the channel has delay spread, sound the channel with pilot <u>sequences</u>



Use in a DFE or Viterbi equaliser



- see N Lo, D. Falconer, A. Sheikh "Adaptive Equalization and Diversity Combining for Mobile Radio Using Interpolated Channel Estimates" IEEE Trans Veh Tech, vol 40 no 3 pp 636-645, Aug 91.

Note the diversity effect and reduction of ISI floor

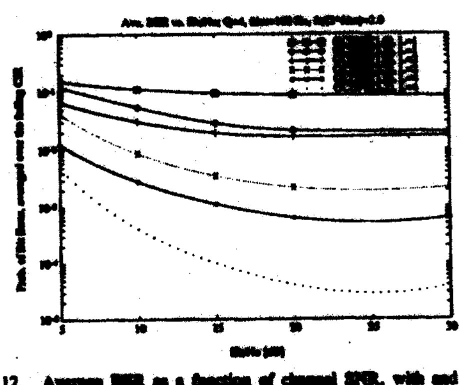


Fig. 12. Average BER as a function of channel SNR, with and without equalization.

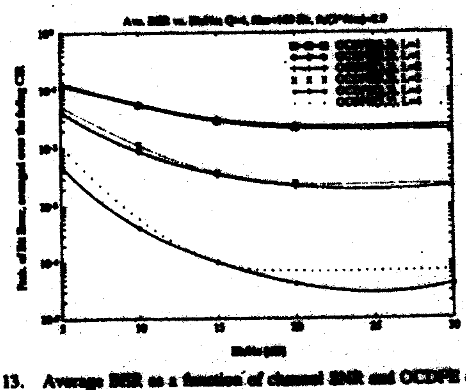
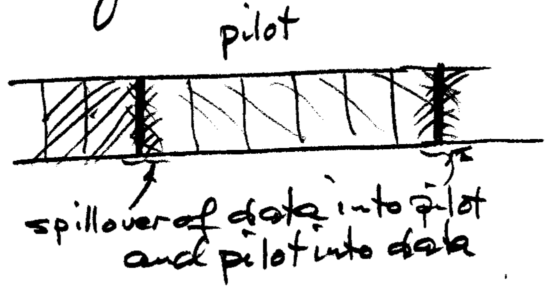
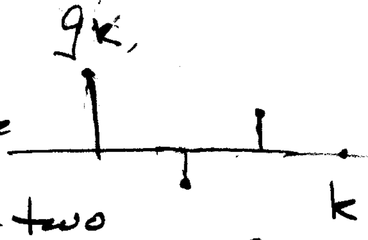


Fig. 13. Average BER as a function of channel SNR and OCDFE configuration.

- Pilot sequences in more detail:



since



so it is common to make one or two symbols at the beginning and end "sacrificial," unused.

- An alternative estimation method is ML (instead of the MMSE we just examined), which leads to LS. Here's how. Suppose training seq is $\underline{p}$. Then we receive

$$\underline{y} = \begin{bmatrix} y(1) \\ \vdots \\ y(I) \end{bmatrix} = \begin{bmatrix} \underline{p} & 0 & 0 \\ & \underline{p} & 0 \\ & & \underline{p} \\ 0 & & \\ 0 & 0 & \end{bmatrix} \begin{bmatrix} g_0 \\ g_1 \\ g_2 \end{bmatrix} + \begin{bmatrix} n(1) \\ \vdots \\ n(I) \end{bmatrix}$$

$$\underline{y} = P\underline{g} + \underline{n}$$

To minimise the sum of squared errors

$\|\underline{y} - P\underline{g}\|^2$ note: deterministic, not MMSE

then $\hat{\underline{g}} = (P^+P)^{-1}P^+\underline{y} = P^{\#}\underline{y}$ pseudo inverse

Proof: $\underline{e} = \underline{y} - P\hat{\underline{g}}$; $J = \underline{e}^+\underline{e} = \underline{y}^+\underline{y} - \underline{y}^+P\hat{\underline{g}} - \hat{\underline{g}}^+P^+\underline{y} + \hat{\underline{g}}^+P^+P\hat{\underline{g}}$;

$\nabla_{\underline{g}} J = \underline{0}$; $P^+P\hat{\underline{g}} = P^+\underline{y}$; $\hat{\underline{g}} = (P^+P)^{-1}P^+\underline{y}$

The columns of P must be LI — close to orthog is even better (good corr'n properties).

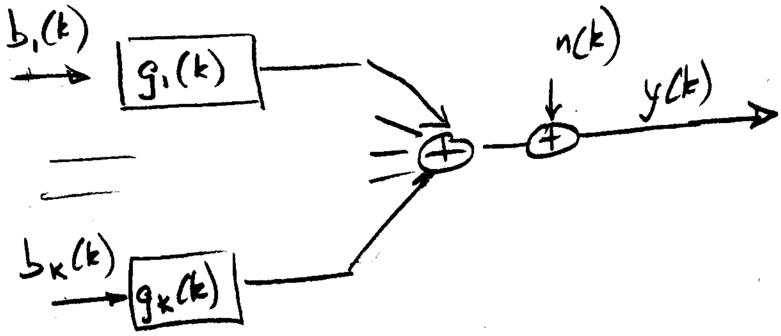
- The resulting error, if $P\underline{g}$ is a good model and $\underline{n}$ is white noise, σ_n^2, is

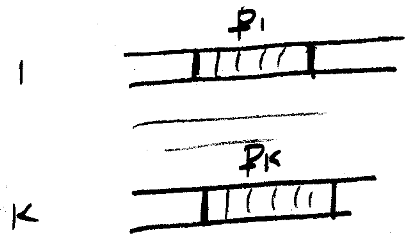
$$\underline{\epsilon} = \hat{\underline{g}} - \underline{g} \qquad \text{(different from } \underline{e} = \hat{\underline{y}} - \underline{y}\text{)}$$

$$R_{\epsilon} = \sigma_n^2 (P^T P)^{-1}$$ so you really don't want LD in the cols of P!

From this, you can calculate corr'n coeffs.

- Last example: multiple users





sequences have sacrificial symbols for delay spread and misalignment

Now
$$\underline{y} = \begin{bmatrix} p_1 & 0 & 0 & & p_K & 0 & 0 \\ & p_1 & & \cdots & & p_K & \\ & & p_1 & & & & p_K \\ 0 & 0 & & & 0 & 0 & \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ g_K \end{bmatrix} + \begin{bmatrix} n(1) \\ \vdots \\ n(N) \end{bmatrix}$$

again $\underline{y} = P\underline{g} + \underline{n}$

Longer sequences, since all cols of P must be linearly independent.