5.3 DF With Noise Whitening

- The MMSE-DFE incorporated temporal decision feedback and linear processing for suppression of present and future MAI. An effective alternative was presented in [Duel 93]: whiten the noise and make the MAI causal only.

- For simplicity, consider synchronous, one shot (so times n-1, n+1 vanish). Then model is $K \times K$ set

$$\underline{\zeta} = C^{+} R C A \underline{b} + \eta \quad , \quad R_{\eta} = N_0 C^{+} R C$$

- Now apply a transformation to whiten the noise components. Why? To reduce coupling among decision variables — otherwise we might discard measurements that could reduce noise on next decision.
  - Many ways to whiten. [Duel 93] selected Gram-Schmidt orthogonalization (Appendix P), which is equivalent to Cholesky decomposition of noise covariance matrix.
  - Cholesky: $C^{+} R C = F^{+} F$, where F is

- Premultiply by the upper triangular $(F^+)^{-1}$:

$$\underline{z} = (F^+)^{-1}\underline{\zeta} = (F^+)^{-1}\underbrace{C^+ R C}_{F^+F} A \underline{b} + \underbrace{(F^+)^{-1}\underline{\eta}}_{\alpha}$$

$$= F A \underline{b} + \underline{\alpha}$$

and $\underline{\alpha}$ is white, with

$$R_\alpha = (F^+)^{-1} R_\eta F^{-1}$$
$$= N_0 (F^+)^{-1} F^+ F F^{-1}$$
$$= N_0 I$$

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_K \end{bmatrix} = \begin{bmatrix} F_{11} & & & \bigcirc \\ F_{21} & F_{22} & & \\ \cdots & \cdots & \ddots & \\ F_{K1} & \cdots & \cdots & F_{KK} \end{bmatrix} \begin{bmatrix} A_1 b_1 \\ A_2 b_2 \\ \vdots \\ A_K b_K \end{bmatrix} + \begin{bmatrix} \alpha_1 \\ \vdots \\ \vdots \\ \alpha_K \end{bmatrix}$$

This system is causal with white noise.

- Decisions:

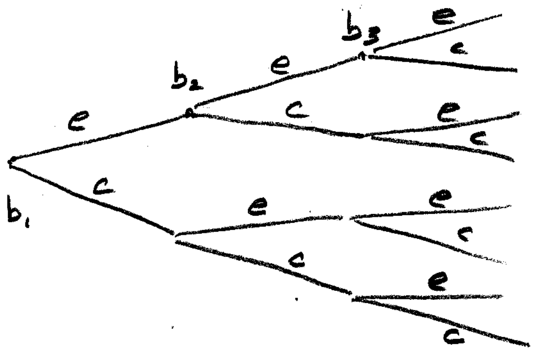
$$\tilde{b}_1 = \operatorname{sgn}[z_1]$$
$$\tilde{b}_2 = \operatorname{sgn}[z_2 - F_{21} A_1 \tilde{b}_1]$$
$$\tilde{b}_3 = \operatorname{sgn}[z_3 - F_{31} A_1 \tilde{b}_1 - F_{32} A_2 \tilde{b}_2]$$

etc

- Error probabilities (with DA, i.e., perfect feedback)

$$P_{e_1} = Q\left(\frac{F_{11}A_1}{N_0}\right), \quad P_{e_2} = Q\left(\frac{F_{22}A_2}{N_0}\right), \ldots \quad P_{e_K} = Q\left(\frac{F_{KK}A_K}{N_0}\right)$$

With errors, it's easiest to use a tree, since conditional:



Exact calculation can get out of hand, but tree pruning (e.g. M-algorithm) should keep it manageable.

- Order is important!

– Consider first equation (where $\underline{w}_1$ = col 1 of F^{-1})

$$z_1 = \underline{w}_1^+ \underline{\xi} = F_{11} A_1 b_1 \qquad + \alpha_1 \qquad \text{(equiv to decorrelator)}$$

This is complete zero forcing of K-1 other users — and ZF always carries a price in SNR.

strategy: Order the users so $[C^+RC]_{1,1} A_1$ is largest. That user is most likely to survive ZF.

– Consider the second equation. It's at the mercy of the decision inherited from the first equation. Another reason to make first user the strongest. Second user must cancel K-2 other users, so it should be 2nd strongest $[C^+RC]_{2,2} A_2$

– Order by decreasing strength. Final user has no MAI, like the previous ones (assuming no decision errors) — and the full benefit of diversity, as we'll see in examining Cholesky:

$$z_K = \cancel{F_{K1} A_1} b_1 + \cancel{F_{K2} A_2} b_2 + \cdots + \cancel{F_{K,K-1} A_{K-1}} b_{K-1} + F_{KK} A_K b_K + \alpha_K$$

Equivalent to single-user performance.

– [Duel93] notes the attraction of whitening and DF in near-far situations, but cautions that incorrect ranking makes error propagation significant, sometimes to the point that decorrelation is preferable.

- In order to gain the full benefit of decision feedback, we need $F_{k,i} A_i$ $i = 1..k-1$ to be known accurately, especially for $k = K$. If errors, then even correct decisions cause reappearance of some MAI:

$$\tilde{b}_k = \mathrm{sgn}\left[z_k - \sum_{i=1}^{k-1} \hat{F}_{ki} \hat{A}_i \tilde{b}_i\right]$$

$$= \mathrm{sgn}\left[F_{kk} A_k b_k + \underbrace{\sum_{i=1}^{k-1} F_{ki} A_i b_i - \sum_{i=1}^{k-1} \hat{F}_{ki} \hat{A}_i \tilde{b}_i}_{\sum_{i=1}^{k-1} (F_{ki} A_i - \hat{F}_{ki} \hat{A}_i) b_i} + \alpha_k\right]$$

  - No discussion of errors in C, R available, but [Duel93] suggests for $\hat{A}_k$'s the use of a parallel decorrelator using $(CRC)^{-1}$.

- Performance in toy systems.

  [Duel93] AWGN, synchronous, $K = 2$, pulse cross corr'n coeff $r = 0.7$.

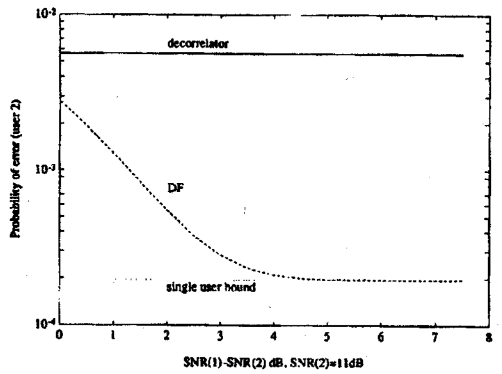


Fig. 2. Error probabilities for two-user channel, $r = 0.7$. (Example 1).

Weaker user $\gamma_2 = 11\,\mathrm{dB}$

Plotted against $(\gamma_1/\gamma_2)_{dB}$.

- decorrelator (ZF) much worse than single user bound since $r = 0.7$
- as interferer becomes stronger, weaker user approaches single user since almost always correct decisions fed back.

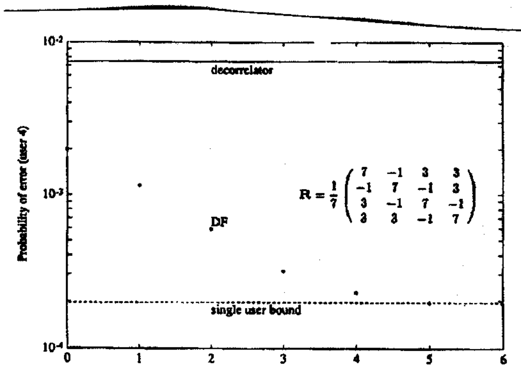


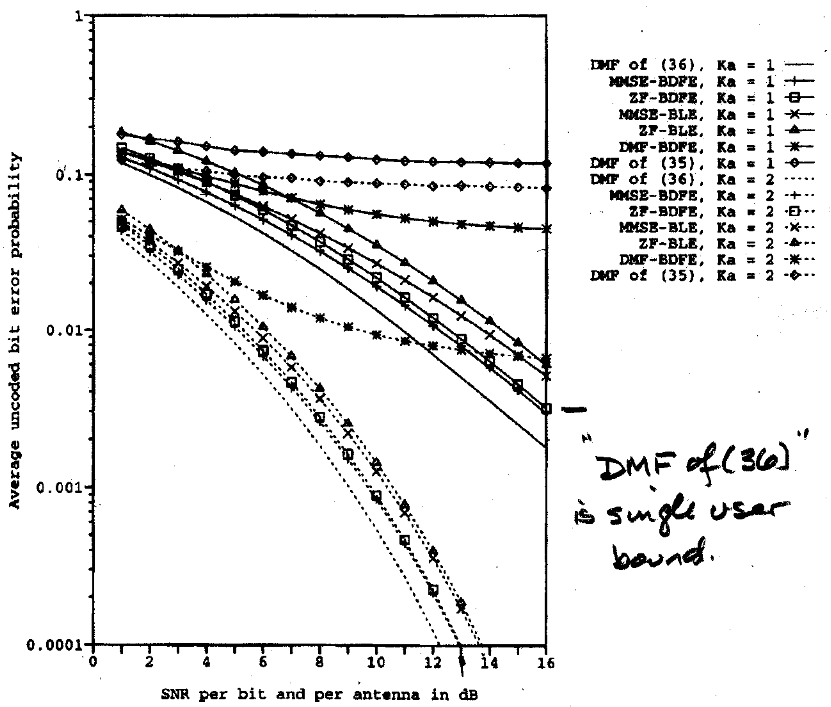
Fig. 3. Error probabilities for the first four-user channel of Example 2.

$K = 4$ users, synchronous
$N_c = 7$ Gold codes
weakest user (4) has $\gamma = 11$ dB.
other users equipower.
plotted against $(\gamma_{strong}/\gamma_{weak})_{dB}$

- Again decorrelator poor, since eigen value ratio is large
- Weakest user converges to single user

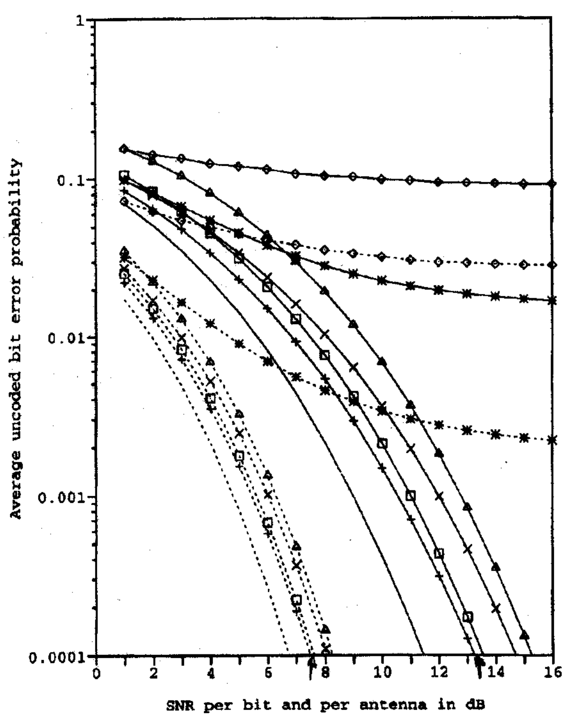
• [Jung95] has more realistic results. Asynchronous, fading, selective. $K = 8$ users, $N_c = 14$ chips, $N = 20$, perfect CSI. ZF-BDFE is an unwindowed, unranked version of causal whitening with DF. MMSE-BDFE is a variant that's still causal, but does mmse, not ZF (gives first users a break). Very close to single user bound.

$L = 2$



"DMF of (36)" is single user bound.

$L = 27$



Caveats: perfect CSI; not heavily loaded.