6.C Approximations of ML for Narrowband Signals

- We have seen that ML-MUD preserves the order of diversity, irrespective of the number of users. It "makes the other users disappear," except for a small dB performance penalty.

- But its computational load can be huge, even if synchronous users, flat fading.

- This motivates a continuing search for detectors that approach ML performance, without the complexity burden. A common feature: they use the Finite Alphabet property, the structure of the interferers; they don't simply treat it as more noise characterised by 2nd order stats.

SIC

- In Successive Interference Cancellation schemes, the receiver makes user data decisions sequentially, in some permuted order

$(k_1, k_2, \cdots, k_K)$ $\qquad k_i, \; i = 1..K$

After each decision, that user signal is subtracted out, and the next one is detected.

- Receive $\underline{y}^{(0)} = CA\underline{b} + \underline{\nu}$

$$= \underline{c}_1 A_1 b_1 + \underline{c}_2 A_2 b_2 + \cdots + \underline{c}_K A_K b_K + \underline{\nu}$$

$$= \underline{c}_{k_1} A_{k_1} b_{k_1} + \underline{c}_{k_2} A_{k_2} b_{k_2} + \cdots + \underline{c}_{k_K} A_{k_K} b_{k_K} + \underline{\nu}$$

  - Stage 1: $i=1$, decide $\hat{b}_{k_1}$ from $\underline{y}^{(0)}$

    Then $\underline{y}^{(1)} = \underline{y}^{(0)} - \hat{\underline{c}}_{k_1} \hat{A}_{k_1} \hat{b}_{k_1}$

    If channel estimates and decision are correct, we are left with a $(K-1)$-user problem.

  - Stage i: decide $\hat{b}_{k_i}$ from $\underline{y}^{(i-1)}$

    then $\underline{y}^{(i)} = \underline{y}^{(i-1)} - \hat{\underline{c}}_{k_i} \hat{A}_{k_i} \hat{b}_{k_i}$

  It should get easier with every stage, since fewer users, same number of antennas (unlike ZF, MMSE).

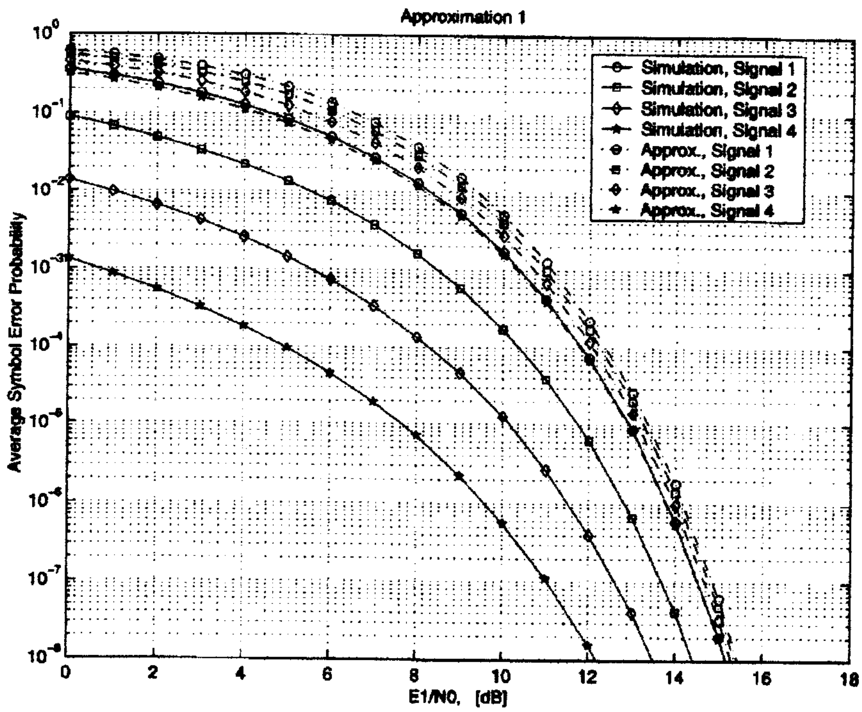
- Questions:
  - How to make the decision (single user, ZF, MMSE etc)?
  - What order of detection?
  - Error propagation? In BPSK, a decision error <u>doubles</u> the effect of that bit, instead of removing it.
  - Channel estimation error? Even if decisions are correct, subtraction leaves

    $$\left(\underline{c}_{k_i} A_{k_i} - \hat{\underline{c}}_{k_i} \hat{A}_{k_i}\right) b_{k_i}$$ as a form of noise.

- "Onion peeling." "Delayering" (when combined with coding)

Whitening SIC (not a standard term)

- This is the simplest approach, inherited from similar studies in CDMA.
- Sort them in order of decreasing power. Detect the strongest one first, since it has all the other users as interference.
- Detection with a single-user detector is not particularly smart for narrowband systems with no signature sequences to help suppress interference.
- Still, AWGN with powers in geometric ratios can be analysed [Jaus02]. Realism?



This is for 4PSK, $K = 4$ users, successive users have power ratio $\alpha = 2.41$

SIC With Whitening

- We can use the approach in [Duel93], developed for CDMA, in narrowband.

- We have $\underline{y} = CA\underline{b} + \underline{\nu}$ as MF outputs and sufficient stats

$$\underline{\zeta} = A^+C^+CA\underline{b} + A^+C^+\underline{\nu}$$
$$= R\underline{b} + \underline{\eta} \qquad R_\eta = N_o R$$

- Now transform to whiten noise components (reduces coupling). Use Cholesky decomposition.

$R = L^+L$, where L is

Premult by the upper triangular $(L^+)^{-1}$:

$$\underline{z} = (L^+)^{-1}\underline{\zeta} = (L^+)^{-1}\underbrace{R}_{L^+L}b + (L^+)^{-1}\underline{\eta}$$

$$= L\underline{b} + \underline{\alpha} \qquad R_\alpha = (L^+)^{-1}R_\eta L^{-1} = N_o (L^+)^{-1}L^+LL^{-1}$$
$$= N_o I \quad \text{(this is why "whitening")}$$

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_K \end{bmatrix} = \begin{bmatrix} L_{11} & & & \\ L_{21} & L_{22} & & \\ L_{31} & L_{32} & L_{33} & \\ & & & \ddots \\ L_{K1} & \cdots & \cdots & L_{KK} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_K \end{bmatrix} + \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_K \end{bmatrix}$$

This is "causal" with white noise.

- Make decisions in SIC fashion

$$\hat{b}_1 = \text{sgn}[z_1]$$
$$\hat{b}_2 = \text{sgn}[z_2 - L_{21}\hat{b}_1]$$
etc

- Note first decision is zero forcing, so diversity order is $M-(K-1)$. Need at least as many antennas as users.

  For second decision, user 1 has been removed by subtraction (we hope), so ZF nulls out only $K-2$ interferers, so diversity order increases by one. And so on.

- Order users by decreasing strength, as shown by diagonal elements of R. So:
  - calculate R
  - reorder cols and rows so diagonal elements decrease, top left to bottom right.
  - Cholesky decomp $L^+L = R$
  - Solve for $b_1, \cdots b_K$ using decision feedback (DF) — i.e. SIC.

- Note Cholesky decomposition of $R = \underbrace{A^+C^+}_{F^+}\underbrace{CA}_{F}$ can be accomplished by Gram-Schmidt orthogonalisation of cols of F.

$$\underbrace{\left[\underline{f}_1 \middle| \underline{f}_2 \middle| \underline{f}_3 \middle| \cdots \middle| \underline{f}_K\right]}_{M\times K} = \underbrace{\left[\underline{\phi}_1 \middle| \underline{\phi}_2 \middle| \cdots \middle| \underline{\phi}_K\right]}_{M\times K} \underbrace{\begin{bmatrix} u_{11} & u_{12} & & u_{1K} \\ 0 & u_{22} & & u_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0\;0 & u_{KK} \end{bmatrix}}_{K\times K}$$

or $F = \Phi U$

where Φ is unitary $\Phi^{+}\Phi = I$

and U is upper triangular.

Then

$$R = F^{+}F = U^{+}\Phi^{+}\Phi U = \underset{\triangle}{U^{+}}\;\underset{\nabla}{U}$$

V-BLAST

- V-BLAST (Vertical - Bell Labs Layered Space Time) was described in [Gold99]. It is closely related to whitening SIC, since it nulls remaining interferers, makes a decision, then removes the decided bit's contribution by subtraction, thereby reducing the order by one.
- The primary difference is the detection order. Whitening chooses on the basis of received power. V-BLAST chooses on the basis of SNR, given the remaining users.
- Here's how it works. We start with the first bit decision. We have

$$\underline{y} = F\underline{b} + \underline{v} \qquad F = CA$$

If we null interferers, a full ZF solution is

$$\hat{\underline{b}} = W^{\dagger}\underline{y} = W^{\dagger}F\underline{b} + W^{\dagger}\underline{v}$$

where $W^{\dagger} = (F^{\dagger}F)^{-1}F^{\dagger}$, the pseudoinv of F

so $\hat{\underline{b}} = \underline{b} + \underline{\alpha} \qquad \underline{\alpha} = W^{\dagger}\underline{v}$

- Which bit first?
  - They all have unit gain in $\hat{\underline{b}}$
  - The noise variance $\sigma_{\alpha_k}^2 = N_0 \|\underline{w}_k\|^2$ col k of W
  - so choose the bit with min $\sigma_{\alpha_k}^2$, that is, the best SNR. Index k_1

- So $\hat{b}_{k_1} = \underline{w}_{k_1}^{\dagger} \underline{y}$, $\tilde{b}_{k_1} = \operatorname{sgn}(\hat{b}_{k_1})$

- Now, deflate the set:

$$\underline{y} = \underline{y} - F^{\langle k_1 \rangle} \tilde{b}_{k_1}$$

$$F = \tilde{F}_{k_1} \qquad \text{(remove column } k_1\text{)}$$

and repeat, now with K-1 users

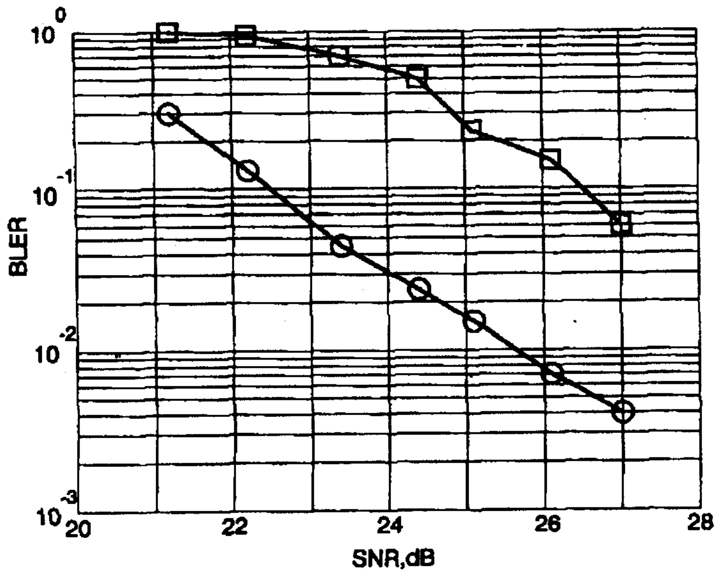


Fig. 2 *Block error rate (BLER) against SNR for M = 8, N = 12 V-BLAST laboratory prototype operating at 25.9bit/s/Hz*

□ nulling alone (no symbol cancellation) i.e. ZF
○ nulling with optimally-ordered symbol cancellation

MMSE VBLAST

- Although VBLAST was originally presented as successive use of ZF, it is possible to use other detectors, as well, as the authors [Gold 99] noted.
- MMSE is the obvious alternative. The strongest user, for any realisation of C, fares better than with ZF. This should also improve the detection performance of users remaining after IC.
- In the case of power separation among users, MMSE should be a big improvement:
  - the strongest signal, which is most likely to be selected for first detection, has <u>much</u> better error rate than if it were ZF (why?)
  - and if it is successfully subtracted out, then the next signal enjoys an additional order of diversity (as well as the MMSE advantage wrt other, weaker signals.

- Which bit first?
  - We have $\underline{\hat{b}} = W^{+}\underline{y}$
    where $W = F(F^{+}F + N_0 I)^{-1}$
  - Estimation error $\underline{e} = \underline{b} - \underline{\hat{b}}$ and we choose the bit with the minimum error variance.
    Those variances are the diagonal elements of $\sigma_b^2 I - \sigma_b^2 (F^{+}F + \frac{1}{\sigma_b^2} I)^{-1} F^{+}F$
  - So choose bit k_1, where
    $$k_1 = \arg\max_{k} \left[\left[(F^{+}F + 2I)^{-1} F^{+}F \right]_{kk} \right]$$

Note $A_i = \sqrt{2\Gamma_i}$ so $N_0 \to 1$

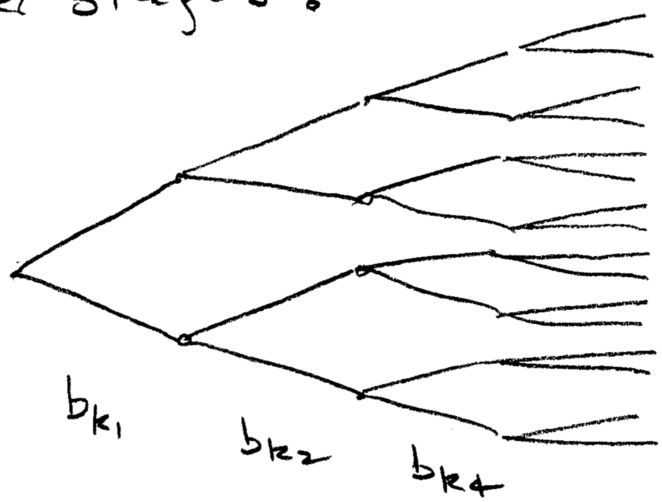
- Decision $\hat{b}_{k_1} = \underline{w}_{k_1}^{+} \underline{y}$, $\tilde{b}_{k_1} = \mathrm{sgn}(\hat{b}_{k_1})$

- Deflation as in ZF VBLAST.

SIC With Tree or Trellis Search

- So far, we have using simple DF, and hoping that error propagation doesn't play too large a role.

- Why not carry alternative decisions into later stages? Now we have a tree

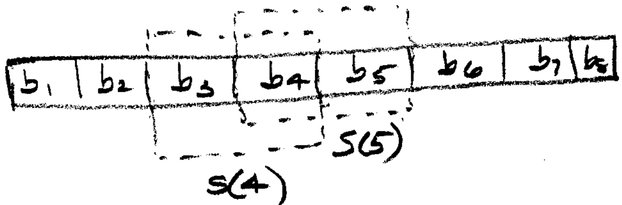


Examining all paths out to the leaves at level K is equivalent to full ML. Is there an alternative?

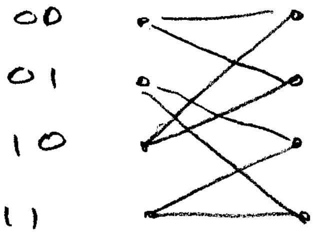
- Suboptimal tree search algorithms look attractive. Sequential decoding [Fano, mid 60's]. Or "breadth first" [Ande84, Simm90]: retain a limited number of survivor paths; extend only the survivors; harvest the best of the extensions as the new set of survivors

- Typical schemes:
  - M-algorithm: keep the M paths with the best metrics (M is not antennas).
  - T-algorithm: keep those paths for which the metric is with a factor T of the best metric.

- Alternatively, think in Viterbi terms. Keep a finite number of adjacent bits as the state, so it defines a trellis instead of tree.



A transition (eg $s(4) \rightarrow s(5)$) identifies three successive bits. But $y(5)$ also depends on b_1, b_2, so use per-survivor path decisions.



— This is better than hard DF for all past bits.

- See [Wei97] for this approach in CDMA context.

- Note that a full tree search is equivalent to full ML.

- Also note that we can precede ML by any invertible linear transformation (whitening, MMSE, inverse) and the decisions and error rate are unchanged. In fact, we can precede ML by <u>any</u> invertible transformation, linear or not. [Appendix R]