Adaptive Minimum-BER Linear Multiuser Detection

S. Chen†, A.K. Samingan†, B. Mulgrew‡ and L. Hanzo†

† Department of Electronics and Computer Science
University of Southampton, Southampton SO17 1BJ, U.K.

‡ Department of Electronics and Electrical Engineering
University of Edinburgh, Edinburgh EH9 3JL, U.K.

Presented at ICASSP2001, May 7-11, 2001, Salt Lake City, USA
System Model

Downlink synchronous, $N$-user and $M$-chip per bit

\[
\mathbf{r}(k) = \mathbf{P} \begin{bmatrix} b(k) \\ b(k-1) \\ \vdots \\ b(k-L+1) \end{bmatrix} + \mathbf{n}(k) = \bar{\mathbf{r}}(k) + \mathbf{n}(k)
\]
where the user bit vector $\mathbf{b}(k) = [b_1(k) \cdots b_N(k)]^T$, $L$ is the ISI span, the Gaussian noise vector $\mathbf{n}(k) = [n_1(k) \cdots n_M(k)]^T$ with zero mean vector and

$$E[\mathbf{n}(k)\mathbf{n}^T(k)] = \sigma_n^2 \mathbf{I},$$

the $M \times LN$ system matrix

$$\mathbf{P} = \mathbf{H} \begin{bmatrix} \bar{\mathbf{S}} \mathbf{A} & 0 & \cdots & 0 \\ 0 & \bar{\mathbf{S}} \mathbf{A} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \bar{\mathbf{S}} \mathbf{A} \end{bmatrix},$$

the user unit-length signature sequence matrix $\bar{\mathbf{S}} = [\bar{s}_1 \cdots \bar{s}_N]$, the diagonal user signal amplitude matrix $\mathbf{A} = \text{diag}\{A_1 \cdots A_N\}$, and the $M \times LM$ CIR matrix $\mathbf{H}$

$$\mathbf{H} = \begin{bmatrix} h_0 & h_1 & \cdots & h_{n_h - 1} \\ h_0 & h_1 & \cdots & h_{n_h - 1} \\ \vdots & \ddots & \ddots & \vdots \\ h_0 & h_1 & \cdots & h_{n_h - 1} \end{bmatrix}.$$
Linear Detector

Linear detector for user $i$

$$\hat{b}_i(k) = \text{sgn}(y(k)) \text{ with } y(k) = \mathbf{w}^T \mathbf{r}(k)$$

where $\mathbf{w} = [w_1 \cdots w_M]^T$ is the detector weight vector.

- MMSE solution most widely used, with LMS adaptive implementation.
- There are $N_b = 2^{LN}$ combinations of $[\mathbf{b}^T(k) \mathbf{b}^T(k - 1) \cdots \mathbf{b}^T(k - L + 1)]^T$:

$$\mathbf{b}^{(j)} = \begin{bmatrix} b^{(j)}(k) \\ b^{(j)}(k - 1) \\ \vdots \\ b^{(j)}(k - L + 1) \end{bmatrix}, \ 1 \leq j \leq N_b$$
with \( b^{(j)}_i \) the \( i \)th element of \( b^{(j)}(k) \).

- \( \bar{r}(k) \) only takes value from the noise-free signal state set:
  \[
  r_j = Pb^{(j)}, \ 1 \leq j \leq N_b
  \]

- The detector \( y(k) = y'(k) + n'(k) \), with \( y'(k) \) only takes value from the set:
  \[
  y_j = w^T r_j, \ 1 \leq j \leq N_b
  \]

\( n'(k) \) is Gaussian with zero mean and variance \( \sigma^2_n w^T w \).
Motivations for Adaptive MBER

- MMSE can be inferior to MBER

Two equal power users with chip codes (+1, +1) and (+1, −1)

Transfer function of CIR
\[ H(z) = 1 + 0.8z^{-1} + 0.6z^{-2} \]

SNR₁ = 25 dB
BER surface for user 1

MMSE solution: \( \log_{10}(BER) = -3.88 \)
MBER solutions: \( \log_{10}(BER) = -5.56 \)
• LMS-style stochastic gradient adaptation

★ Two existing stochastic gradient adaptive MBER algorithms


   Difference approximation for gradient of one-bit error measure, no need for noise pdf assumption, complexity \(O(M^2)\), very low convergence rate for small BER

2. Approximate or Adaptive MBER, AMBER, (Globecom’98, pp.3590–3595)

   Like signed-error LMS but modified to continue updating weights in vicinity of decision boundary, very simple with a complexity \(O(M)\)

★ Our approach, LBER, based on kernel density estimation of BER from training data

   Also a complexity \(O(M)\), simpler than DMBER but more complex than AMBER
Theoretical MBER Solution

Define the signed decision variable

\[ y_s(k) = \text{sgn}(b_i(k))y(k) = \text{sgn}(b_i(k)) \left( y'(k) + n'(k) \right) \]

with p.d.f.:

\[
p_y(y_s) = \frac{1}{N_b \sqrt{2\pi \sigma_n} \sqrt{w^T w}} \sum_{j=1}^{N_b} \exp \left( -\frac{(y_s - \text{sgn}(b_i^{(j)})y_j)^2}{2\sigma_n^2 w^T w} \right)
\]

Thus error probability of linear detector:

\[ P_E(w) = \text{Prob}\{ \text{sgn}(b_i(k))y(k) < 0 \} = \int_{-\infty}^{0} p_y(y_s) \, dy_s = \frac{1}{N_b} \sum_{j=1}^{N_b} Q(c_j(w)) \]
where

\[ Q(y) = \frac{1}{\sqrt{2\pi}} \int_{y}^{\infty} \exp \left( -\frac{x^2}{2} \right) \, dx \quad \text{and} \quad c_j(w) = \frac{\text{sgn}(b^{(j)}_i) y_j}{\sigma_n \sqrt{w^T w}} = \frac{\text{sgn}(b^{(j)}_i) w^T r_j}{\sigma_n \sqrt{w^T w}} \]

Gradient

\[ \nabla P_E(w) = \frac{1}{N_b \sqrt{2\pi \sigma_n}} \left( \frac{ww^T - w^T wI}{(w^T w)^{3/2}} \right) \sum_{j=1}^{N_b} \exp \left( -\frac{y_j^2}{2\sigma_n^2 w^T w} \right) \text{sgn}(b^{(j)}_i) r_j \]

By normalizing \( w \) to unit length,

\[ \nabla P_E(w) = \frac{1}{N_b \sqrt{2\pi \sigma_n}} \sum_{j=1}^{N_b} \exp \left( -\frac{y_j^2}{2\sigma_n^2} \right) \text{sgn}(b^{(j)}_i)(w y_j - r_j) \]

- Steepest-descent or conjugate gradient algorithm \( \Rightarrow \) MBER solution
Block-data Based Adaptation

Estimate $p_s(y_s)$ based on training data $\{r(k), b_i(k)\}_{k=1}^{K}$ (kernel density estimation):

$$\hat{p}_y(y_s) = \frac{1}{K \sqrt{2\pi \rho_n} \sqrt{w^T w}} \sum_{k=1}^{K} \exp\left(-\frac{(y_s - \text{sgn}(b_i(k))y(k))^2}{2\rho_n^2 w^T w}\right)$$

where the radius parameter $\rho_n$ is related to the noise standard deviation $\sigma_n$

- $\hat{p}_y(y_s) \Rightarrow \hat{P}_E(w) \Rightarrow \nabla \hat{P}_E(w) \bullet$
- Gradient algorithm $\Rightarrow$ estimated MBER solution $\star$

Remark: This is analogous to estimated MMSE solution – sample estimates of autocorrelation matrix and cross-correlation vector replacing corresponding ensemble averages
Stochastic Gradient Adaptation

One-sample estimate of p.d.f. and instantaneous stochastic gradient $\Rightarrow$ LBER

- Re-scaling weight vector (to unit length)

$$w(k) := \frac{w(k)}{\sqrt{w^T(k)w(k)}}$$

- Detector output

$$y(k) = w^T(k)r(k)$$

- Weight update

$$w(k + 1) = w(k) + \frac{\mu}{\sqrt{2\pi} \rho_n} \exp \left(-\frac{y^2(k)}{2\rho_n^2}\right) \text{sgn}(b_i(k))(r(k) - w(k)y(k))$$

Step size $\mu$ and width $\rho_n$ are two algorithm parameters
Simulation

Example 1 Two equal-power users with \((+1,+1,-1,-1)\) and \((+1,-1,-1,+1)\), respectively, and the CIR transfer function \(H(z) = 1.0 + 0.25z^{-1} + 0.5z^{-3}\)

Data block: 100 samples, \(\text{SNR}_1 = \text{SNR}_2 = 16.5\ \text{dB}\), block adaptation for user 1:
\( \text{SNR}_1 = \text{SNR}_2 = 19 \text{ dB} \)

Stochastic gradient adaptation for \textbf{user 1}:

Average over 100 runs
Example 2 Four equal-power users with \((+1,+1,+1,+1,-1,-1,-1,-1)\), \((+1,-1,+1,-1,-1,+1,-1,+1)\), \((+1,+1,-1,-1,-1,-1,+1,+1)\) and \((+1,-1,-1,+1,-1,+1,+1,-1)\); the CIR transfer function \(H(z) = 0.4 + 0.7z^{-1} + 0.4z^{-2}\)

Data block: 1500 samples, \(\text{SNR}_i = 16\) dB for all \(i\), block adaptation for user 1:
SNR$_i = 15$ dB for all $i$

Stochastic gradient adaptation for **user 1**: Average over 50 runs
Conclusions

- MBER solution for linear multiuser detector can be superior over MMSE one
- LMS-style stochastic gradient adaptive MBER algorithms are available
- Our approach: Least Bit Error Rate, LBER
  - Kernel density estimate for p.d.f. of detector decision variable is natural and generic\(^1\)
  - Complexity is linear with detector length
  - Appear to have better performance in terms of convergence speed and steady-state BER misadjustment

\(^1\)We have extended the LBER to training nonlinear neural network multiuser detectors