EFFICIENT IMPLEMENTATION OF LINEAR MULTIUSER DETECTORS

X. F. Wang, W.-S. Lu, and A. Antoniou

Department of Electrical and Computer Engineering
University of Victoria
Victoria, B.C., CANADA, V8W 3P6
Email: {xfwang, wslu, andreas}@ece.uvic.ca

Abstract
A recursive algorithm for updating linear detectors in code-
division multiple-access systems is proposed. Based on this
algorithm, a window-based implementation with a signal-
based criterion for determining the window length is devel-
opled. Performance analysis and numerical experiments are
conducted that show the merits of the proposed implementa-
tion method.

1. Introduction
The prohibitive computational complexity of the optimum
multiuser detector [1] has motivated the search for subopti-
mum multiuser detectors with much reduced computational
complexity. An important family of suboptimum multiuser
detectors, namely, linear multiuser detectors, has been pro-
posed and extensively analyzed in [2]-[4]. In [5] linear detec-
tors have been shown to have a number of advantages.
However, linear detectors such as the decorrelating detec-
tor or the minimum mean-squared error (MMSE) detector
involve the tasks of computing and updating the inverse of
the correlation matrix or its modified version [4,5], which
require a considerable amount of computation. Furthermore,
the dimension of the correlation matrix in asynchronous sys-
tems is, in principle, the same as the message length, which
is usually very large. This implies not only a high computa-
tional burden but also an unacceptably large delay, if a
direct implementation is adopted.

In this paper, we describe a recursive algorithm for up-
dating the inverse of the correlation matrix. A convergence
analysis is conducted to show that in practice the recur-
sion can be done very quickly. Based on this analysis, a
window-based implementation is proposed. Unlike the con-
ventional window approach proposed in [7], where the win-
dow length is selected empirically, the proposed method uses
a signal-based criterion to select the window length in order
to achieve a given system performance. The proposed imple-
mentation has a computational complexity which grows
linearly with the number of users. A performance analysis
is presented, showing that the performance of the proposed
implementation can be arbitrarily close to the performance
of the decorrelating detector. Two numerical examples are in-
cluded to demonstrate the convergence rate of the Cholesky-
decomposition recursion and the system performance.

2. Linear Multiuser Detectors
We consider binary-phase-shift-keying (BPSK) transmis-
sion through an additive white Gaussian noise (AWGN) channel
shared by K asynchronous users in a direct-sequence code-
division multiple access (DS-CDMA) system. Let the mes-
sage length be N and denote the kth information bit of the
ith user and its amplitude by b_i(k) and \sqrt{\epsilon_k(k)}, respectively.
The concatenation of N successive output vectors of the K
matched filters, r, can be modeled as

\[ r = R_N Eb + n \]

where

\[
\begin{align*}
    r &= \begin{bmatrix} r^T(1) & \cdots & r^T(N) \end{bmatrix}^T \\
    b &= \begin{bmatrix} b^T(1) & \cdots & b^T(N) \end{bmatrix}^T \\
    E &= \text{diag}\{E(1) \cdots E(N)\} \quad E(k) = \text{diag}\{\sqrt{\epsilon_1(k)} \cdots \sqrt{\epsilon_K(k)}\}
\end{align*}
\]

n is an AWGN signal with zero-mean and variance \sigma^2 R_N, and

\[
R_N =
\begin{bmatrix}
    R_a(0) & R_a(1)^T & 0 & \cdots & 0 \\
    R_a(1) & R_a(0) & R_a(1)^T & \cdots & 0 \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & 0 & \cdots & R_a(0) \\
    0 & 0 & 0 & 0 & R_a(1) \\
\end{bmatrix}
\]

(2)

is an \((N \times N)\)-block 3-band matrix. Each \(R_a(m)\) for \(m = 0, 1, \ldots, K\) is a K \times K matrix whose \((k, l)\)th entry is given by

\[
R_{al}(m) = \int_{-\infty}^{\infty} g_a(t - \tau_k) g_l(t + mT - \tau_k) dt \quad m = 0, 1, \ldots
\]

where \(g_a(t)\) and \(\tau_k\) denote the normalized signature and the transmission delay for the \(k\)th user, respectively. For convenience of analysis, the transmission delays are arranged in ascending order, i.e., \(0 = \tau_1 \leq \cdots \leq \tau_K < T\), so as to obtain an \(R_a(1)\) which is an upper triangular matrix with zero diagonal entries.

A linear multiuser detector can be viewed as a linear
mapping, \(T\), as applied to the outputs of the matched-filter
bank. The linear mappings for the decorrelating detector
and the MMSE detector are

\[
T_L = R_N^{-1}
\]

and

\[
T_M = (R_N + \sigma^2 E^{-2})^{-1}
\]

3221 0-7803-4428-6/98 $10.00 © 1998 IEEE
respectively. As the linear mappings in (3) and (4) have the same structure, we shall focus our discussion on the decorrelating detector. With minor modifications, the results obtained are directly applicable to the MMSE detector.

3. A Window-Based Implementation Method

One approach to reducing the computational complexity is to minimize the need for recomputation. In [6], an algorithm for updating linear detectors was proposed for this purpose. However, only the synchronous case was considered. Another commonly used method is to detect the user information bits window by window, where in each processing window the corresponding linear mapping is a matrix of order $MK$ with $M \ll N$ [4,7]. In what follows, we show that much improved implementations can be achieved by adopting a signal-based criterion that determines the minimum number of transmission intervals in each processing window.

Since $R_N$ is positive definite with a probability 1 [3], there exists a decomposition

$$R_N = LL^T$$

(5)

where $L$ is a lower triangular matrix. From (2) and (5), it can be verified that $L$ has the form

$$L = \begin{bmatrix}
C_1 & D_1 & C_2 & 0 \\
D_1 & C_3 & 0 & 0 \\
0 & D_3 & \ddots & 0 \\
0 & 0 & \ddots & C_N
\end{bmatrix}$$

(6)

where $C_i$ is a $K \times K$ lower triangular matrix, and $D_i$ is a $K \times K$ upper triangular matrix with zero diagonal entries. It can be verified that $L^{-1}$ is a lower triangular matrix with the $(i,j)$th $(i \geq j)$ block $L_{ij} \in \mathbb{R}^{K \times K}$ given by

$$L_{ij} = (-1)^{i-j} C_i^{-1} \prod_{k=1}^{i-j} M_{i-k}$$

(7)

From (6), the following recursion can be deduced:

$$C_i C_i^T = R_{a}(0)$$

(8a)

$$C_i D_i^T = R_{a}(1)^T \quad 1 \leq i \leq N - 1$$

(8b)

$$C_{i+1} C_{i+1}^T = R_{a}(0) - D_i D_i^T \quad 1 \leq i \leq N - 1$$

(8c)

If we define $K_i = C_i C_i^T$, then (8) implies that

$$K_{i+1} = R_{a}(0) - D_i D_i^T$$

(9a)

$$= R_{a}(0) - R_{a}(1) K_i^{-1} R_{a}(1)^T \quad 1 \leq i \leq N - 1$$

(9b)

where $K_1 = R_{a}(0)$ and $K_i$ is positive definite. It was shown in [3] that the decorrelating detector approaches a $K$-input $K$-output linear time-invariant filter when $N \to \infty$. This implies that the limit $K = \lim_{N \to \infty} K_i$ exists. Furthermore, $K$ is positive definite if $\lim_{N \to \infty} R_N$ is positive definite, which is usually the case [3]. The convergence rate of $\{K_i\}$ can be used to understand the behavior of the decorrelating detector. From (9b), it follows that

$$K_{i+1} - K_i = R_{a}(1) K_i^{-1} (K_{i-1} - K_{i-1}) R_{a}(1)^T$$

(10)

By repeating (10) in conjunction with (8), we obtain

$$K_{i+1} - K_i = \prod_{j=0}^{p-1} M_{i-j} \cdot (K_{i+1-p} - K_{i-p}) \cdot \prod_{n=1-p}^{i-1} M_n$$

(11)

where $M_i = D_i C_i^{-1}$. It follows that

$$\frac{\|K_{i+1} - K_i\|}{\|K_{i+1-p} - K_{i-p}\|} \leq \prod_{j=0}^{p-1} M_{i-j} \|\prod_{n=1-p}^{i-1} M_n\|$$

(12)

Now if $i - p$ is sufficiently large such that $K_{i-p}$ is close to $K_i$, then from (8) $M_{i-p}$ is close to the limit $M = \lim_{i \to \infty} M_i$. In such a case, we have

$$\prod_{j=0}^{p-1} M_{i-j} \|\prod_{n=1-p}^{i-1} M_n\| \approx \|M^p\|^2$$

(13)

It can be shown that $\|M^p\| \leq c |\lambda|^p$, where $c \geq 1$ is a constant related to $M$, and $|\lambda|$ is the largest of the absolute values of the eigenvalues of $M$. Consequently, an approximation for the left-hand side of (12) can be obtained as $c^2 |\lambda|^{2p}$. This implies that $K_i$ converges linearly with $|\lambda|^2$ when $K_i$ is close to $K$. As will be demonstrated in our numerical experiments, in practice, for stable realizations of the decorrelating detector, $|\lambda| < 1$ is usually quite small and hence the convergence rate is high.

We now turn our attention to a window-based implementation of the decorrelating detector. Assuming that the window length is $M = 2p + 1$, equation (1) within a window can be written as

$$R_M E_i b_i = \tilde{r}_i - r_e + n$$

(14)

where

$$b_i = [b(i-p)^T \cdots b(i)^T \cdots b(i+p)^T]^T$$

$$\tilde{r}_i = r_i - [(R_a(1) E(i-p - 1) b(i-p - 1))^T 0 \cdots 0]^T$$

$$r_i = [r(i-p)^T \cdots r(i)^T \cdots r(i+p)^T]^T$$

$$E_i = \text{diag}(E(i-p) \cdots E(i) \cdots E(i+p))$$

$$r_e = [0 \cdots 0 b(i+p+1)^T E(i+p+1) R_a(1)]^T$$

and $R_M \in \mathbb{R}^{MK \times MK}$ has the same structure as $R_N$ except for the size. Hence, the decorrelating detector would estimate $b_i$ within the window as

$$\tilde{b}_i = R_M^{-1} \tilde{r}_i$$

(15)

Note that within the window, $r_e$ in (15) is not available and, therefore, a modified decorrelating detector is deduced from (15) by neglecting term $r_e$. This leads to

$$\tilde{b}_i = R_M^{-1} \tilde{r}_i$$

(16)

Denoting by $[y]_{i,k}$ the first $k$ elements of vector $y$, then $[b_i]_{k,k}$ contains the estimates of the user information bits
within the first \( k \) transmission intervals in \( B_k \). By choosing an appropriate window length \( M \), the difference between \([b_1, i + (p + 1)]^T K \) and \([b_1, i + (p + 1)]^T K \) can be made smaller than a given tolerance. To demonstrate this, we write the difference between \([b_1, i + (p + 1)]^T K \) and \([b_1, i + (p + 1)]^T K \) as

\[
[R_{M}^{-1} R_k]_{i + (p + 1)K} = [x^T L_{M1} \ x^T L_{M2} \ \cdots \ \cdots x^T L_{M,K + 1}]^T
\]

where \( x = C_{G}^{-1} R_k (i + (p + 1)) b_i (i + p + 1) \). If \([L_{M,x}^T x]^T \) for \( i = p + 1 \) is sufficiently small, the solutions \( b_i (i - p), b_i (i - p + 1), \cdots, b_i \) obtained from (16) will be sufficiently close to the corresponding solutions from (15). From (7), we have

\[
||L_{M,p+1}|| \leq ||C_{G}^{-1}|| \ \prod_{k=1}^{M-p-1} M_{M-k} \tag{17}
\]

Based on (17), it can be shown that the window length can be determined by inspecting the value of \([L_{M,p+1}^T] \). In summary, the window length can be selected as the smallest \( M = 2p + 1 \) such that \([L_{M,p+1}^T] \) is smaller than a given tolerance \( \epsilon \). Once the window length is determined, (16) is used to detect the user information bits window by window. In each window, the user information bits within the first \( p + 1 \) transmission intervals are detected, and the processing window is then shifted by \( p + 1 \) transmission intervals. In the light of the above analysis, an algorithm for updating a window-based decorrelating detector can be obtained as follows:

Algorithm 1

Step 1: Initialize \( K_0 = R_k (0), i = 1 \); set a tolerance \( \epsilon \) and the largest block length \( M = 2p + 1 \).

Step 2: Perform the Cholesky decomposition \( K_0 = C_{G} C_{G}^T \).

Step 3: Solve \( C_{G} D_{S} = R_k (1)^T \) for \( D_{S} \).

Step 4: If \( i \neq 3 \), go to Step 5; otherwise,

(a) compute \( M_3 = D_{S} C_{G}^{-1} \) and estimate \( \hat{\lambda} \), the largest of the absolute values of the eigenvalues of \( M_3 \);

(b) determine the block length as \( p = \min(p, p_1) \) with \( p_1 = \max(\lambda^2, \text{Int}[1 + \text{log}(\lambda^2)] / \epsilon) \), where \( \text{Int}[x] \) denotes the integer part of \( x \).

Step 5: If \( \text{Int}(i/2) > p \), stop; otherwise, set \( i = i + 1 \), compute \( K_i \) using (9a), and go to Step 2.

Two remarks on Algorithm 1 are in order. (a) The window length is determined in Step 4 by using the largest of the absolute values of the eigenvalues of \( M_3 \), \( \hat{\lambda} \), instead of \( \lambda \). This is because as indicated by our numerical experiments, \( M_3 \) is usually sufficiently close to \( M \). To estimate \( \hat{\lambda} \), various numerical methods are available, such as the power method [8]. Constant \( c \) can be estimated by the largest 1-norm of the rows or columns of \( M_3 \). (b) By fully exploiting the special structure of \( C_{G} \) and \( D_{S} \), the updating can be accomplished by using \( O(M(K - 1))/2 \) flops.

Once the linear detector has been updated, the process of user detection can be efficiently carried out by using the following algorithm:

Algorithm 2

Solve \( C_{G} y (i - p) = \hat{\tau} (i - p) \) for \( y (i - p) \)
for \( j = i + p + 1 : i + p \)

Solve \( C_{G} y (j) = \hat{\tau} (j) - D_{S,j-1} y (j - 1) \) for \( y (j) \)

Solve \( C_{G} b_i (i + p) = y (i + p) \) for \( b_i (i + p) \)
for \( j = i + p - 1 : i + p \)

Solve \( C_{G} b_i (j) = y (j) - D_{S,j+1} y (j + 1) \) for \( b_i (j) \)

It follows that the algorithm requires \( O((6p + 1)K^2) \) flops for processing one window. Since in each window only \( p + 1 \) bits for each user are detected, in total about \( 6K^2 \) flops per user bit are required, which is linear with respect to the number of users.

4. Performance Analysis

Since the most inaccurate estimate of the user bits usually occurs at the central transmission interval of a processing window, the bit-error probability (BEP) of the detector is less than or equal to the BEP of the user bits at the central transmission interval. Denote the BEP of the \( i \)th bit of the \( k \)th user by \( P_k (i) \) and assume that \( i \) is the index for the central transmission interval of a processing window. If transmitting +1 or −1 is equally likely, then (16) implies that

\[
P_k (i) = P_k ([R_{M}^{-1} r_k]_{pK+k} = 0 | b_k (i) = 1) \tag{18}
\]

where \([y]_k \) represents the \( k \)th entry of vector \( y \). Assuming that \([L_{M,p+1}] < \epsilon \), it can then be verified that

\[
P_k (i) < Q\left( \frac{\sqrt{\epsilon_k (i)} - K\epsilon / \epsilon_{\text{max}}}{d_{pK+k}} \right) \tag{19}
\]

where \( Q(x) = \int_{x}^{\infty} (1/\sqrt{2\pi}) e^{-x^2/2} dx \), \( d_{pK+k} \) is the \((pK + k, pK + k)\)th entry of \( R_{M}^{-1} \), and \( \epsilon_{\text{max}} \) is the largest among \( \sqrt{\epsilon_1 (i + p + 1)}, \sqrt{\epsilon_2 (i + p + 1)}, \cdots, \sqrt{\epsilon_K (i + p + 1)} \).

Recalling the notation of the asymptotic efficiency defined in [1], a lower bound of the asymptotic efficiency of the \( i \)th bit of the \( k \)th user can be deduced as

\[
\eta_k (i) > \frac{(1 - K\epsilon / \epsilon_{\text{max}} / \epsilon_k (i))^2}{d_{pK+k}} \tag{20}
\]

From (20), we see that if \( \epsilon \) in Algorithm 1 is selected such that \( K\epsilon / \epsilon_{\text{max}} / \epsilon_k (i) < 1 \), the asymptotic efficiency of the proposed detector is close to \( 1/d_{pK+k} \), which coincides with the theoretical asymptotic efficiency of the decorrelating detector [2]. This observation serves as a guide to the selection of the tolerance in Algorithm 1. We can also see that the near-far resistance [2] of the proposed detector can always be made nearly optimal if \( \epsilon_{\text{max}} / \epsilon_k (i) \) is finite.

5. Numerical Examples

As a first example, we consider a two-user system. As will be demonstrated, the simplicity of the system allows us to carry out an analysis on issues such as convergence rate of \( K_i \), implementation of Algorithm 1, and system performance in an explicit manner. Assuming that \( R_1 (0) = \rho_1 \) and \( R_2 (1) = \rho_2 \), we have \( |\rho_1| + |\rho_2| \leq 1 \). By (8), a recursion for the largest of the absolute values of the eigenvalues of \( C_{G}^{-1} D_{S} \), denoted by \( |\lambda| \), can be found as

\[
|\lambda| = \frac{\rho_1 \rho_2}{\epsilon (1 - \rho_1^2 / \epsilon_2)} \tag{21}
\]
where \( x_{i+1} = 1 - \frac{\rho_{i}^2}{1 - \rho_{i}} \) and \( x_{1} = 1 \). This leads to

\[
x = \lim_{i \to \infty} x_{i} = \frac{1 + \rho_{1}^2 - \rho_{2}^2 \pm \sqrt{(1 + \rho_{1}^2 - \rho_{2}^2)^2 - 4\rho_{1}^2}}{2}
\]

(22)

Letting \( i \to \infty \) in (21) and substituting (22) into (21) give \( |\lambda| = \lim_{i \to \infty} |\lambda_{i}| \). It can be readily verified that \( |\lambda| = 1 \) when \( |\rho_{1}| + |\rho_{2}| = 1 \), which corresponds to an unstable realization of the decorrelating detector. Fig. 1 depicts \( |\lambda| \) versus \( \rho_{1} \) and \( \rho_{2} \). It can be seen that \( |\lambda| \) is usually quite small. Specifically, \( |\lambda| < 0.4 \) if \( |\rho_{1}| + |\rho_{2}| < 0.9 \) whereas \( |\lambda| < 0.25 \) if \( |\rho_{1}| + |\rho_{2}| < 0.8 \). To illustrate the convergence rate of \( |\lambda_{i}| \), the smallest value \( i \) such that \( \frac{|\lambda_{i+1}|}{|\lambda_{i}|} < 0.01 \) is shown in Table 1. Also shown in the table is the smallest window length required to ensure that \( ||L_{M,\rho,\rho}|| < 0.01 \). As can be seen from Table 1, even in cases where the cross-correlation properties are poor, the convergence rate of \( |\lambda_{i}| \) is still fast and the window length required is moderate.

![Figure 1: \( |\lambda| \) in a two-user system.](image)

Table 1: Convergence rate of \( |\lambda_{i}|/(\text{smallest window length}) \)
in a two-user system.

<table>
<thead>
<tr>
<th>( \rho_{1} ) ( \rho_{2} )</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>2/5</td>
<td>2/5</td>
<td>2/5</td>
<td>2/5</td>
<td>2/5</td>
<td>2/5</td>
<td>2/5</td>
<td>2/5</td>
</tr>
<tr>
<td>0.2</td>
<td>2/5</td>
<td>2/5</td>
<td>2/5</td>
<td>2/7</td>
<td>2/7</td>
<td>3/7</td>
<td>3/7</td>
<td>3/7</td>
</tr>
<tr>
<td>0.3</td>
<td>2/5</td>
<td>2/5</td>
<td>2/7</td>
<td>2/7</td>
<td>3/7</td>
<td>3/7</td>
<td>3/7</td>
<td>3/7</td>
</tr>
<tr>
<td>0.4</td>
<td>2/5</td>
<td>2/7</td>
<td>2/7</td>
<td>3/9</td>
<td>3/9</td>
<td>3/9</td>
<td>3/9</td>
<td>3/9</td>
</tr>
<tr>
<td>0.5</td>
<td>2/5</td>
<td>2/7</td>
<td>2/9</td>
<td>3/11</td>
<td>3/11</td>
<td>3/11</td>
<td>3/11</td>
<td>3/11</td>
</tr>
<tr>
<td>0.6</td>
<td>2/5</td>
<td>2/7</td>
<td>3/11</td>
<td>3/11</td>
<td>3/11</td>
<td>3/11</td>
<td>3/11</td>
<td>3/11</td>
</tr>
</tbody>
</table>

As a second example, we consider a 10-user system. A rectangular chip waveform is assumed in all the simulations. The transmission delays are uniformly generated in one transmission interval. To examine the robustness of the detector in an environment with poor cross-correlation properties, a set of 31-chip sequences was randomly generated as spreading sequences. Extensive simulations have shown that even for a 10-dB received power imbalance \( \epsilon_{max}/\epsilon_{r} \), 7 or 9 transmission intervals for each window are sufficient to guarantee an asymptotic efficiency of above 80 percent of the theoretical asymptotic efficiency of the decorrelating detector. If a family of 31-chip Gold code is used as spreading sequences and a 10-dB received power imbalance is still assumed, the upper bound of the BEP of the proposed implementation with different window length is as illustrated in Fig. 2. As can be observed, a system with \( M = 9 \) exhibits a performance very close to the theoretical performance of the decorrelating detector.

![Figure 2: Bit-error probability of user 1.](image)

### 6. Conclusion

By exploiting the block 3-band structure of the correlation matrix, a recursive Cholesky-decomposition-based updating algorithm has been developed. Furthermore, a window-based implementation for the decorrelating detector, which includes a signal-based criterion for determining the window length, has been proposed. The performance analysis and the numerical experiments carried out show that the proposed implementation method achieves a performance which is very close to the theoretical one.

### 7. References


