Adaptive MIMO Decision Feedback Equalization for Receivers With Time-Varying Channels

Jihoon Choi, Heejung Yu, and Yong H. Lee, *Senior Member, IEEE*

***Abstract*—In an attempt to reduce the computational complexity of vertical Bell Labs layered space time (V-BLAST) processing with time-varying channels, an efficient adaptive receiver is developed based on the generalized decision feedback equalizer (GDFE) architecture. The proposed receiver updates the filter weight vectors and detection order using a recursive least squares (RLS)-based time- and order-update algorithm. The convergence of the algorithm is examined by analysis and simulation, and it is shown that the proposed adaptive technique is considerably simpler to implement than a V-BLAST processor with channel tracking, yet the performances are almost comparable.**

***Index Terms*—GDFE, MIMO-DFE, RLS algorithm, V-BLAST.**

I. Introduction

ALTHOUGH the V-BLAST architecture can provide high spectral efficiencies when multiple antennas are used at both the transmitter and the receiver [1], [2], its application to systems with time-varying channels is difficult because of an excessive computational load. Given M transmitter and N receiver antennas, the V-BLAST algorithm successively detects the M transmitted symbols from N received symbols at time n. It determines the optimal ordering for the successive detection and computes the nulling vectors at each iteration. For time-varying channels, the detection ordering and nulling vectors need to be updated for each time, plus the channel parameters should be tracked. These update and tracking operations in the time domain require excessive computations. To overcome this drawback, a simplified policy for updating and tracking is proposed in [3], where the V-BLAST detection is updated blockwise, and the channel tracking is interpolation-based, thereby creating a tradeoff between complexity and performance.

As an alternative approach to detecting multiple-input multiple-output (MIMO) systems in time-varying channels, the adaptive techniques in [4]–[7] may be employed.[1] By successively detecting the M transmitted symbols at each time, the adaptive decorrelating detector in [4] can suppress the co-channel interference caused by spatial multiplexing, but it requires channel estimation to determine the order of detection. The adaptive decision feedback equalizers (DFEs) in [5] and [6] are useful for reducing intersymbol interference (ISI) in MIMO systems over frequency-selective channels. However, they are not suitable for reducing the co-channel interference; in the DFEs, the M transmitted symbols at each time are simultaneously detected without considering the order of detection. The adaptive method in [7] is a blind technique, whereas the receivers in [4]–[6] are data aided.

In this paper, we develop another adaptive technique for detecting MIMO systems in time-varying channels. The proposed scheme is a data-aided adaptive MIMO-DFE based on the generalized DFE (GDFE) architecture [9]. For each time, the tap weight vectors are updated using an RLS-based time- and order-update algorithm and detection ordering determined according to a least squares error (LSE) criterion. Since the proposed algorithm does not require explicit channel tracking, its implementation is simpler than V-BLAST detection. Computer simulation results indicate that the proposed receiver can perform like V-BLAST with channel tracking and outperform the decorrelating detector in [4].

The remainder of this paper is organized as follows. The proposed adaptive MIMO-DFE is derived in Section II and its complexity compared with V-BLAST detection in Section III. In Section IV, the convergence of the proposed algorithm is demonstrated in a stationary environment. Finally, Section V compares the bit error rate (BER) performances of the proposed MIMO-DFE and a V-BLAST processor.

II. Proposed Adaptive MIMO-DFE

Fig. 1 shows the signal model and structure of the adaptive MIMO-DFE. The M symbols in $\mathbf{d}(n)$ are simultaneously transmitted through M antennas and received by N antennas to yield an N-dimensional vector $\mathbf{y}(n)$, which is given by

$$\mathbf{y}(n) = \mathbf{H}(n)\mathbf{d}(n) + \mathbf{u}(n) \tag{1}$$

where $\mathbf{H}(n)$ is an $N \times M$ channel matrix whose elements represent independent flat fading channels, and $\mathbf{u}(n)$ denotes

[1]The receivers in [4] and [5] were originally developed for multiuser detection in code division multiple access (CDMA) channels, but they can be applied to the MIMO systems because V-BLAST with the M transmitter and N receiver antennas can be viewed as an M-user CDMA system with spreading factor N [8].

Manuscript received December 27, 2003; revised January 4, 2005. This work was supported in part by the university IT research center program of the government of Korea. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Sergios Theordoridis.

J. Choi is with Media Lab. 1, Telecommunication Network Business, Samsung Electronics, Suwon-city, 442-600, Korea (e-mail: hevdark1@yahoo.com).

H. Yu is with the Next Generation Wireless LAN Research Team, Electronics and Telecommunications Research Institute, Daejeon, 305-350, Korea (e-mail: heejung@etri.re.kr).

Y. H. Lee is with the Department of Electrical Engineering and Computer Science, Korea Advanced Institute of Science and Technology, Daejeon, 305-701, Korea (e-mail: yohlee@ee.kaist.ac.kr).

Digital Object Identifier 10.1109/TSP.2005.857053

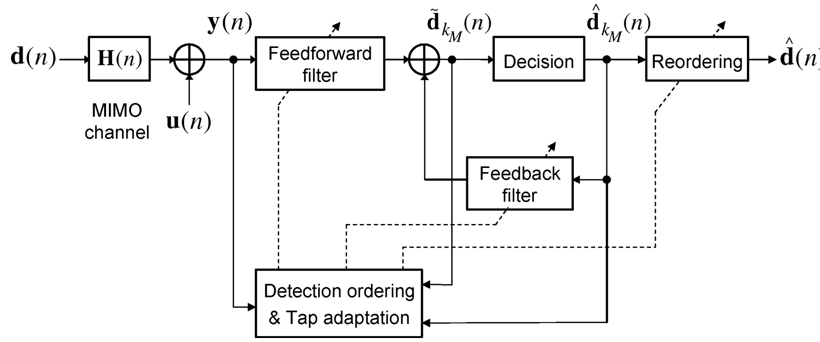


Fig. 1. MIMO-DFE structure.

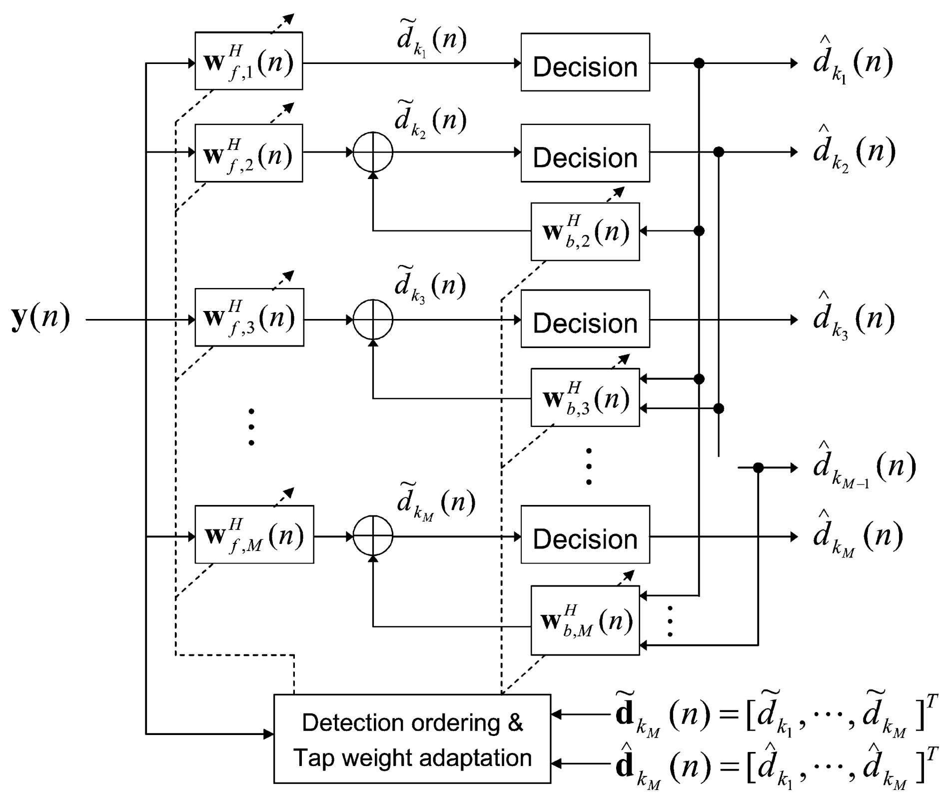


Fig. 2. Detailed architecture of MIMO-DFE.

noise. The received vector $\mathbf{y}(n)$ is passed through the feedforward filter that corresponds to the nulling vectors of V-BLAST detection [9]; then, the already-determined signals are cancelled via decision feedback filtering. The receiver successively detects the M transmitted symbols in $\mathbf{d}(n)$ (see Fig. 2). The detection is started with linear equalization, and the degree of decision feedback increases with the number of detected symbols. The filter tap weight vectors and order for detecting the M symbols are updated for each time. To be specific, the following notations are introduced: $\{\mathbf{w}_{f,i}(n)\}$ denote the N-dimensional feedforward weight vectors; $\{\mathbf{w}_{b,i}(n)\}$ denote the $(i-1)$-dimensional feedback weight vectors; $\mathbf{d}(n) = [d_1(n), d_2(n), \cdots, d_M(n)]^T$; $\hat{d}_{k_i}(n)$ and $\tilde{d}_{k_i}(n)$ denote the ith detected symbol and corresponding equalizer output, respectively, where $k_i \in \{1, 2, \cdots, M\}$. The output of the equalizers can be represented as

$$\tilde{d}_{k_i}(n) = \mathbf{w}_{f,i}^H(n)\mathbf{y}(n) + \mathbf{w}_{b,i}^H(n)\hat{\mathbf{d}}_{k_{i-1}}(n) \tag{2}$$

where $\hat{\mathbf{d}}_{k_{i-1}}(n) = [\hat{d}_{k_1}(n), \hat{d}_{k_2}(n), \cdots, \hat{d}_{k_{i-1}}(n)]^T$, and $\mathbf{w}_{b,i}^H(n) = 0$ when $i = 1$. For notational convenience, define

$$\mathbf{w}_{t,i}(n) = \begin{cases} \mathbf{w}_{f,i}(n), & i = 1 \\ \left[\mathbf{w}_{f,i}^T(n), \mathbf{w}_{b,i}^T(n)\right]^T, & i = 2, \cdots, M \end{cases} \tag{3}$$

and

$$\mathbf{y}_{t,i}(n) = \begin{cases} \mathbf{y}(n), & i = 1 \\ \left[\mathbf{y}^T(n), \hat{\mathbf{d}}_{k_{i-1}}^T(n)\right]^T, & i = 2, \cdots, M. \end{cases} \tag{4}$$

Then, (2) is rewritten as

$$\tilde{d}_{k_i}(n) = \mathbf{w}_{t,i}^H(n)\mathbf{y}_{t,i}(n). \tag{5}$$

When the channel is fixed, the optimal detection order under the minimum mean square error (MMSE) criterion can be written as

$$k_i = \arg\min_j E\left[\left|d_j(n) - \mathbf{p}_{i,j}^H\mathbf{R}_i^{-1}\mathbf{y}_{t,i}(n)\right|^2\right] \tag{6}$$

where $\mathbf{R}_i = E[\mathbf{y}_{t,i}(n)\mathbf{y}_{t,i}^H(n)]$, $\mathbf{p}_{i,j} = E[\mathbf{y}_{t,i}(n)d_j^*(n)]$, and $i = 1, 2, \cdots, M$. Given the detection order in (6), the optimal weight vector for the ith equalizer minimizing the mean square error (MSE) $E[|d_{k_i}(n) - \mathbf{p}_{i,k_i}^H \mathbf{R}_i^{-1}\mathbf{y}_{t,i}(n)|^2]$ can be represented as

$$\mathbf{w}_{t,i}(n) = \mathbf{R}_i^{-1}\mathbf{p}_{i,k_i}. \tag{7}$$

Note that $\mathbf{w}_{t,i}(n)$ is time invariant because the channel is fixed. MIMO-DFE with the optimal detection order in (6) and the weight in (7) is identical to V-BLAST associated with the MMSE criterion [9].

When the channel is time-varying, the weight and detection order need to be determined in an adaptive manner. The adaptation rules for the proposed receiver are derived in the following two subsections. The first part describes the rule for weight updating when the detection order $\{k_1, k_2, \cdots, k_M\}$ is given. This rule is then extended for updating both the weight and detection order in the second part. The adaptation rules are derived under the assumption that training vectors $\{\mathbf{d}(n)\}$ are available during the initialization period. After the initialization, the adaptive receiver is switched to the decision-directed mode. In what follows, the adaptation rules are only described for the decision-directed mode.

A. Weight Update for Fixed Detection Order

Suppose that the detection order $\{k_1, k_2, \cdots, k_M\}$ is known and fixed. In this case, the weight vectors $\{\mathbf{w}_{t,i}(n)\}$ in (3) can be obtained by solving the standard least squares (LS) problem. Specifically, the following LS cost function is minimized:

$$J_i(n) = \sum_{l=1}^{n} \lambda^{n-l} \left| \hat{d}_{k_i}(l) - \mathbf{w}_{t,i}^H(l)\mathbf{y}_{t,i}(l) \right|^2 \tag{8}$$

where λ is the forgetting factor that satisfies $0 < \lambda \leq 1$. The optimal tap weight minimizing $J_i(n)$ is given by

$$\mathbf{w}_{t,i}(n) = \mathbf{\Phi}_i^{-1}(n)\mathbf{z}_{i,k_i}(n) \tag{9}$$

where $\mathbf{\Phi}_i(n)$ is the time-averaged correlation matrix defined by

$$\mathbf{\Phi}_i(n) = \sum_{l=1}^{n} \lambda^{n-l}\mathbf{y}_{t,i}(l)\mathbf{y}_{t,i}^H(l) \tag{10}$$

and $\mathbf{z}_{i,k_i}(n)$ is the time-averaged cross correlation vector defined by

$$\mathbf{z}_{i,j}(n) = \sum_{l=1}^{n} \lambda^{n-l}\mathbf{y}_{t,i}(l)\hat{d}_j^*(l). \tag{11}$$

It is well known that the optimal weight in (9) can be calculated recursively using the RLS algorithm, which is summarized as follows.

Time-Update RLS Algorithm: Given $\{\mathbf{\Phi}_i^{-1}(n-1), i = 1, 2, \cdots, M\}$, evaluate for each i

$$\mathbf{q}_i(n) = \mathbf{\Phi}_i^{-1}(n-1)\mathbf{y}_{t,i}(n) \tag{12}$$

$$\mathbf{k}_i(n) = \frac{\lambda^{-1}\mathbf{q}_i(n)}{1 + \lambda^{-1}\mathbf{y}_{t,i}^H(n)\mathbf{q}_i(n)} \tag{13}$$

$$\mathbf{\Phi}_i^{-1}(n) = \lambda^{-1}\mathbf{\Phi}_i^{-1}(n-1) - \lambda^{-1}\mathbf{k}_i(n)\mathbf{q}_i^H(n) \tag{14}$$

$$\mathbf{w}_{t,i}(n) = \mathbf{w}_{t,i}(n-1) + \mathbf{k}_i(n)\xi_i^*(n) \tag{15}$$

where $\mathbf{k}_i(n)$ is the gain vector, and $\xi_i(n)$ is the *a priori* estimation error defined by

$$\xi_i(n) = \hat{d}_{k_i}(n) - \mathbf{w}_{t,i}^H(n-1)\mathbf{y}_{t,i}(n). \tag{16}$$

The time update RLS algorithm requires $O(MN^2)$ operations (multiplications and additions) for each n. The computational load for evaluating $\{\mathbf{w}_{t,i}(n)\}$ can be reduced by performing an *order*-update that utilizes the relation between the correlation matrix of the ith DFE $\mathbf{\Phi}_i(n)$ and that of the $(i+1)$th DFE $\mathbf{\Phi}_{i+1}(n)$. Next, the order-update algorithm is derived.

From (10), $\mathbf{\Phi}_{i+1}(n)$ is expressed as

$$\mathbf{\Phi}_{i+1}(n) = \begin{bmatrix} \mathbf{\Phi}_i(n) & \mathbf{z}_{i,k_i}(n) \\ \mathbf{z}_{i,k_i}^H(n) & \alpha_{k_i}(n) \end{bmatrix} \tag{17}$$

where

$$\alpha_j(n) = \sum_{l=1}^{n} \lambda^{n-l} \left| \hat{d}_j(l) \right|^2. \tag{18}$$

Using the matrix inversion lemma [10], $\mathbf{\Phi}_{i+1}^{-1}(n)$ is written as

$$\mathbf{\Phi}_{i+1}^{-1}(n) = \begin{bmatrix} \mathbf{\Phi}_i^{-1}(n) + c_i(n)\mathbf{w}_{t,i}(n)\mathbf{w}_{t,i}^H(n) & -c_i(n)\mathbf{w}_{t,i}(n) \\ -c_i(n)\mathbf{w}_{t,i}^H(n) & c_i(n) \end{bmatrix} \tag{19}$$

where

$$c_i(n) = \frac{1}{\alpha_{k_i}(n) - \mathbf{z}_{i,k_i}^H(n)\mathbf{w}_{t,i}(n)}. \tag{20}$$

The relation between $\mathbf{y}_{t,i}(n)$ in (4) and $\mathbf{y}_{t,i+1}(n)$ is expressed as

$$\mathbf{y}_{t,i+1}(n) = \left[\mathbf{y}_{t,i}^T(n), \hat{d}_{k_i}(n)\right]^T. \tag{21}$$

Using (19) and (21) in (12), the following recursive equation for $\mathbf{q}_i(n)$ is obtained:

$$\mathbf{q}_{i+1}(n) = \begin{bmatrix} \mathbf{q}_i(n) \\ 0 \end{bmatrix} + c_i(n-1)\xi_i(n) \begin{bmatrix} -\mathbf{w}_{t,i}(n-1) \\ 1 \end{bmatrix}. \tag{22}$$

The order-update algorithm for evaluating $\{\mathbf{w}_{t,i}(n), i = 1, \cdots, M\}$ is summarized below.

Order-Update Algorithm:

Step 1) Given $\mathbf{\Phi}_1^{-1}(n-1)$, calculate $\mathbf{q}_1(n), \mathbf{k}_1(n), \mathbf{\Phi}_1^{-1}(n)$, and $\mathbf{w}_{t,1}(n)$ using the time-update equations in (12)–(15).

Step 2) For each i, $1 \leq i \leq M-1$, first obtain $\mathbf{q}_{i+1}(n)$ using the order-update equation in (22). Then, $\mathbf{k}_{i+1}(n)$ and $\mathbf{w}_{t,i+1}(n)$ are evaluated from the time-update equations in (13) and (15).

The order-update algorithm is simpler to implement than the time-update algorithm for the following reasons.

- The recursion in (22) requires $O(N)$ multiplications, whereas the time-update in (12) needs $O(N^2)$ multiplications.
- The order-update only needs $\mathbf{\Phi}_1^{-1}(n-1)$ for start-up and does not require the update of $\{\mathbf{\Phi}_i^{-1}(n), i = 2, \cdots, M\}$ in (14).

The order-update algorithm needs $O(MN)$ operations for obtaining $\{\mathbf{w}_{t,i}(n), i = 1, \cdots, M\}$.

B. Updating Both Weight and Detection Order

To determine the weight $\mathbf{w}_{t,i}(n)$ and detection order k_i for the ith DFE, the set of candidate weight vectors and the LSEs in (8) for all candidate symbols are evaluated. Then, the detection order is determined by selecting the symbol associated with the minimum LSE, and the weight corresponding to the detection order is chosen from the set of candidate weights.

Suppose that $\{\hat{d}_j(l) | 1 \le j \le M, 1 \le l \le n\}$, $\mathbf{\Phi}_i^{-1}(n)$, and $\mathbf{z}_{i,j}(n)$ are given. Then, the set of candidate weight vectors is defined as $\{\mathbf{v}_{i,j}(n), j \in S_i\}$, where

$$\mathbf{v}_{i,j}(n) = \mathbf{\Phi}_i^{-1}(n)\mathbf{z}_{i,j}(n) \tag{23}$$

and $S_i = \{1, 2, \cdots, M\} - \{k_1, \cdots, k_{i-1}\}$ (S_i is a set encompassing all indices of the symbols to be detected from the ith iteration). Note that $\{\mathbf{v}_{i,j}(n)\}$ in (23) becomes the same as $\mathbf{w}_{t,i}(n)$ in (9) if j is replaced with the detection order k_i. The LSEs at the ith DFE are defined by

$$\mathcal{E}_{i,j}(n) = \sum_{l=1}^{n} \lambda^{n-l} \left| \hat{d}_j(l) - \mathbf{z}_{i,j}^H(n)\mathbf{\Phi}_i^{-1}(n)\mathbf{y}_{t,i}(l) \right|^2 \tag{24}$$

where $j \in S_i$. Then, the detection order k_i is represented as

$$k_i = arg \min_{j \in S_i} \mathcal{E}_{i,j}(n). \tag{25}$$

When the channel is fixed and $\lambda = 1$, $(1/n)\mathcal{E}_{i,j}(n)$ converges to the MSE associated with $d_j(n)$ at the ith DFE as n increases, and thus, the detection order in (25) converges to the optimal order in (6). After some calculation, (24) is rewritten as

$$\begin{aligned}\mathcal{E}_{i,j}(n) &= \sum_{l=1}^{n} \lambda^{n-l} \left| \hat{d}_j(l) \right|^2 - \mathbf{z}_{i,j}^H(n)\mathbf{\Phi}_i^{-1}(n) \\ &\quad \times \sum_{l=1}^{n} \lambda^{n-l}\mathbf{y}_{t,i}(l)\hat{d}_j^*(l) \\ &= \alpha_j(n) - \mathbf{v}_{i,j}^H(n)\mathbf{z}_{i,j}(n)\end{aligned} \tag{26}$$

where $\alpha_j(n)$ and $\mathbf{z}_{i,j}(n)$ are defined in (18) and (11), respectively. This equation indicates that the LSE $\mathcal{E}_{i,j}(n)$ can be easily obtained once the candidate weight $\mathbf{v}_{i,j}(n)$ is given. In what follows, we develop an efficient order-update process for $\{\mathbf{v}_{i,j}(n)\}$.

The order update algorithm in the previous subsection cannot be directly applied to the evaluation of $\{\mathbf{v}_{i,j}(n)\}$ because the detection order is unavailable. Therefore, instead of using (22), $\mathbf{\Phi}_i^{-1}(n)$ and $\mathbf{z}_{i,j}(n)$ are separately updated in (23). Recall that $\mathbf{\Phi}_{i+1}^{-1}(n)$ can be obtained using the recursion in (19). To represent $\{\mathbf{z}_{i+1,j}(n)\}$ in terms of $\{\mathbf{z}_{i,j}(n)\}$, it is convenient to define the $(N+M)$-by-M cross-correlation matrix $\mathbf{G}(n)$ given by

$$\begin{aligned}\mathbf{G}(n) &= \sum_{l=1}^{n} \lambda^{n-l} \left[\mathbf{y}^T(l), \hat{\mathbf{d}}^T(l)\right]^T \hat{\mathbf{d}}^H(l) \\ &= \lambda \mathbf{G}(n-1) + \left[\mathbf{y}^T(n), \hat{\mathbf{d}}^T(n)\right]^T \hat{\mathbf{d}}^H(n)\end{aligned} \tag{27}$$

where $\hat{\mathbf{d}}(n) = [\hat{d}_1(n), \hat{d}_2(n), \cdots, \hat{d}_M(n)]^T$. When denoting the (i,j)th entry of $\mathbf{G}(n)$ by $g_{i,j}(n)$, then $\mathbf{z}_{i,j}(n)$ in (11) can be represented as

$$\mathbf{z}_{i,j}(n) = \left[g_{1,j}(n), \cdots, g_{N,j}(n), g_{N+k_1,j}(n), \cdots, g_{N+k_{i-1},j}(n)\right]^T \tag{28}$$

and $\mathbf{z}_{i+1,j}(n)$ becomes

$$\mathbf{z}_{i+1,j}(n) = \left[\mathbf{z}_{i,j}^T(n), g_{N+k_i,j}(n)\right]^T. \tag{29}$$

Furthermore, $\alpha_j(n)$ in (18) can be written as

$$\alpha_j(n) = g_{N+j,j}(n). \tag{30}$$

From (19) and (29), it can be shown that

$$\begin{aligned}\mathbf{v}_{i+1,j}(n) &= \mathbf{\Phi}_{i+1}^{-1}(n)\mathbf{z}_{i+1,j}(n) \\ &= \begin{bmatrix}\mathbf{v}_{i,j}(n) \\ 0\end{bmatrix} + \frac{g_{N+k_i,j}(n) - \mathbf{w}_{t,i}^H(n)\mathbf{z}_{i,j}(n)}{\mathcal{E}_{i,k_i}(n)} \\ &\quad \times \begin{bmatrix}-\mathbf{w}_{t,i}(n) \\ 1\end{bmatrix}\end{aligned} \tag{31}$$

where $j \in S_{i+1} = S_i - \{k_i\}$. Using (30), $\mathcal{E}_{i,j}(n)$ in (26) can be rewritten as

$$\mathcal{E}_{i,j}(n) = g_{N+j,j}(n) - \mathbf{v}_{i,j}^H(n)\mathbf{z}_{i,j}(n). \tag{32}$$

In the proposed adaptation rule, both the detection order $\{k_1, k_2, \cdots, k_M\}$ and the candidate weights $\{\mathbf{v}_{i,j}(n) | 1 \le i \le M, j \in S_i\}$ are obtained at each time n, under the assumption that the detected symbols at time n, $\{\hat{d}_j(n) | 1 \le j \le M\}$ are available. To satisfy this assumption, $\{\hat{d}_j(n)\}$ are evaluated using the weight and the detection order at time $n-1$. Specifically, we obtain

$$\hat{d}_{k_i}(n) = \text{decision}\left[\mathbf{w}_{t,i}^H(n-1)\mathbf{y}_{t,i}(n)\right] \tag{33}$$

where $\{k_i\}$ is the detection order at time $n-1$. The proposed rule, which will be referred to as Algorithm 1, is summarized as follows.

Algorithm 1

Step 1) Initialization:

$n = 0$

$k_i = i$, for all i

$\mathbf{\Phi}_1^{-1}(0) = \delta^{-1}\mathbf{I}$, $\mathbf{G}(0) = \mathbf{0}$

$\mathbf{w}_{f,i}(0) = \mathbf{1}$, $\mathbf{w}_{b,i}(0) = \mathbf{0}$, $\mathbf{v}_{1,i}(0) = \mathbf{1}$, for all i

where δ is a small positive constant.

Step 2) Iterative equalization and decision:

i) $n = n + 1$

ii) For $i = 1, 2, \cdots, M$, calculate $\hat{d}_{k_i}(n)$ in (33).

Step 3) Cross-correlation update: evaluate $\mathbf{G}(n)$ in (27).

Step 4) Weight and detection order update:

i) Compute $\mathbf{q}_1(n)$, $\mathbf{k}_1(n)$ and $\mathbf{\Phi}_1^{-1}(n)$ using the time-update equations in (12)–(14). Then evaluate the following time-update equation corresponding to (15):

$$\mathbf{v}_{1,j}(n) = \mathbf{v}_{1,j}(n-1) + \mathbf{k}_1(n)\left\{\hat{d}_j(n) - \mathbf{v}_{1,j}^H(n-1)\mathbf{y}(n)\right\}^* \tag{34}$$

```
    for all j, 1 ≤ j ≤ M.
  ii) For i = 1,2,···,M−1:
    • Evaluate ℰ_{i,j}(n) in (32) for
      all j ∈ S_i, and obtain k_i in
      (25).
    • w_{t,i}(n) = v_{i,k_i}(n)
    • Evaluate v_{i+1,j}(n) for all
      j ∈ S_{i+1} using (31).
Step 5) Go to Step 2.
```

During a training period in which the true symbol vector $\mathbf{d}(n)$ is known, Step 2 ii) is unnecessary. When the channel is slowly varying, it may not be necessary to update the detection ordering for each time. In fact, simulation results for slowly varying channels showed that neighboring detection orderings from Algorithm 1 are often identical. The fact that detection orderings for $\mathbf{d}(n-1)$ and $\mathbf{d}(n)$ are the same indicates that the detection ordering is known when calculating $\mathbf{w}_{t,i}(n)$; thus, the weight $\mathbf{w}_{t,i}(n)$ can be efficiently updated using (22), (13), and (15), instead of Step 4 in Algorithm 1. These observations lead to the following adaptation procedure, called Algorithm 2, that sparsely updates the detection ordering.

```
Algorithm 2
 This algorithm is identical to Algo-
rithm 1 with the exception that Step 4
ii) in Algorithm 1 is replaced with the
following:
     ii) If n is a multiple of γ,
         which is a positive integer,
         then follow Step 4 ii) in
         Algorithm 1. Otherwise, for
         i = 1,2,···,M − 1, evaluate
         w_{t,i}(n) using (22), (13), and
         (15).
```

Algorithm 2 updates the detection ordering every γ symbol period and becomes Algorithm 1 when $\gamma = 1$.

III. Complexity Comparison

This section compares the complexities for implementing the proposed algorithms and the V-BLAST method with RLS tracking.

One of the most efficient algorithms for V-BLAST is in [8], which performs QR decomposition, inversion of an upper triangular matrix, reordering, and triangularization of the reordered matrices. When $M = N$, these operations need $(7/3)M^3$ complex multiplications and $(5/3)M^3$ complex additions, whereas RLS channel tracking requires $O(M^2)$ operations. Table I compares the complexities of Algorithm 1 and the corresponding V-BLAST-based method (the detection order is updated for each time). Both methods require $O(M^3)$ operations, yet the proposed algorithm provides some savings in computation. For example, as shown in Table I(b), when $M = N = 8(4)$, the number of complex multiplications is reduced by 25.3% (14.0%). In the case of Algorithm 2, where the detection order is updated every γ symbols, additional savings can also be made. Table II lists the number of operations for this case. Note that both methods become simpler to implement; however, in this case, the proposed Algorithm 2 only needs $O(M^2)$ operations when $\gamma \geq M$, whereas V-BLAST still requires $O(M^3)$ operations. Therefore, the computational savings of Algorithm 2 over V-BLAST can be more significant. For example, when $M = N = 8(4)$ and $\gamma = 8$, the number of complex multiplications is reduced by 36.3% (14.8%).

TABLE I
(a) Required Number of Arithmetic Operations per Symbol ($M = N$) in Algorithm 1 and Corresponding V-BLAST (Detection Orderings Are Updated for Each Time). (b) Number of Operations and Reduction Ratio When $M = 4, 8$, and 12

(a)

	Complex Multiplication	Complex Addition
V-BLAST Detector	$\frac{7}{3}M^3 + 5M^2$	$\frac{5}{3}M^3 + 4M^2$
Proposed Algorithm 1	$\frac{4}{3}M^3 + 7M^2$	$\frac{4}{3}M^3 + 5M^2$

(b)

	Operation	M=4	M=8	M=12
V-BLAST Detector	Complex Multi.	229	1515	4752
	Complex Add.	171	1109	3456
Proposed Algorithm 1	Complex Multi.	197	1131	3312
	Complex Add.	165	1003	3024
Reduction Ratio	Complex Multi.	14.0%	25.3%	30.3%
	Complex Add.	3.1%	9.6%	12.5%

TABLE II
(a) Required Number of Arithmetic Operations per Symbol ($M = N$) in Algorithm 2 and Corresponding V-BLAST (Detection Orderings Are Updated Every γ Symbols). (b) Number of Operations and Reduction Ratio When $M = 4, 8$, and 12, and $\gamma = 8$

(a)

	Complex Multiplication	Complex Addition
V-BLAST Detector	$\frac{1}{\gamma}(\frac{7}{3}M^3 + 5M^2)+$ $\frac{\gamma-1}{\gamma}(\frac{5}{6}M^3 + \frac{11}{2}M^2)$	$\frac{1}{\gamma}(\frac{5}{3}M^3 + 4M^2)+$ $\frac{\gamma-1}{\gamma}(\frac{5}{6}M^3 + 4M^2)$
Proposed Algorithm 2	$\frac{1}{\gamma}(\frac{4}{3}M^3 + 7M^2)$ $+\frac{\gamma-1}{\gamma}(\frac{15}{2}M^2)$	$\frac{1}{\gamma}(\frac{4}{3}M^3 + 5M^2)$ $+\frac{\gamma-1}{\gamma}(5M^2)$

(b)

	Operation	M=4	M=8	M=12
V-BLAST Detector	Complex Multi.	149	840	2442
	Complex Add.	122	719	2136
Proposed Algorithm 2	Complex Multi.	127	535	1266
	Complex Add.	88	377	912
Reduction Ratio	Complex Multi.	14.8%	36.3%	48.2%
	Complex Add.	27.9%	47.6%	57.3%

IV. Convergence Analysis

The aim is to demonstrate the convergence of Algorithm 1 when the channel is time-invariant. However, since directly proving the convergence is formidable, because of the detection order that is updated for each time, an indirect approach is taken based on the observations stated below.

- For time-invariant channels, the optimal detection order is fixed for all n.

- Assuming that the detection order of Algorithm 1 converges to the optimal order, Algorithm 1 approaches the RLS algorithm in (12)–(15).

In what follows, we first analyze the convergence of the RLS algorithm in (12)–(15), adopting the analysis in [11]. Then, the convergence of Algorithm 1 is indirectly demonstrated by comparing the analytical results with the corresponding simulation results for Algorithm 1.

A. Convergence of RLS Algorithm

For the analysis, the following assumptions were made.

- The noise vector $\mathbf{u}(n)$ is an independent and identically distributed (i.i.d.) complex Gaussian with a zero mean and variance of σ^2.
- The transmitted symbol vector $\mathbf{d}(n)$ satisfies $E[\mathbf{d}(n)\mathbf{d}^H(n)] = P_d\mathbf{I}_M$, where $\mathbf{I}_M$ is the M-by-M identity matrix.

The estimation error $e_{o,i}(n)$ of the ith ideal equalizer is defined by

$$e_{o,i}(n) = d_{k_i}(n) - \mathbf{w}_{o,i}^H \mathbf{y}_{t,i}(n) \tag{35}$$

where $\mathbf{w}_{o,i}$ is the optimal weight vector, which is constant over a fixed channel. It is assumed that $e_{o,i}(n)$ is white with a zero mean and variance of σ_i^2, which is written as

$$\begin{aligned}\sigma_i^2 &= E\left[\left|d_{k_i}(n) - \mathbf{w}_{o,i}^H \mathbf{y}_{t,i}(n)\right|^2\right] \\ &= P_d\left(1 - \left[\mathbf{h}_{k_i}^H, \mathbf{0}_{(i-1)}^H\right]\mathbf{w}_{o,i}\right) \\ &= P_d\left\{1 - \left[\mathbf{h}_{k_i}^H, \mathbf{0}_{(i-1)}^H\right]\mathbf{R}_i^{-1}\begin{bmatrix}\mathbf{h}_{k_i} \\ \mathbf{0}_{(i-1)}\end{bmatrix}\right\}\end{aligned} \tag{36}$$

where $\mathbf{h}_j$ is the jth column of a fixed channel matrix $\mathbf{H}$, $\mathbf{0}_j$ represents the j-dimensional null vector, and $\mathbf{R}_i$ is the ensemble-averaged correlation matrix given by

$$\mathbf{R}_i = P_d\begin{bmatrix}\mathbf{H}\mathbf{H}^H + \frac{\sigma^2}{P_d}\mathbf{I}_N & \mathbf{H}_i \\ \mathbf{H}_i^H & \mathbf{I}_{i-1}\end{bmatrix}. \tag{37}$$

Here, $\mathbf{H}_i = [\mathbf{h}_{k_1}, \mathbf{h}_{k_2}, \cdots, \mathbf{h}_{k_{i-1}}]$. Note that σ_i^2 in (36) is a function of the noise variance σ^2. Adopting the analysis in [11], we can show the following.

For $\lambda = 1$

$$\begin{aligned}E\left[\boldsymbol{\epsilon}_i^H(n)\boldsymbol{\epsilon}_i(n)\right] &= \sigma_i^2 \mathrm{tr}\left[E\left[\mathbf{\Phi}_i^{-1}(n)\right]\right] \\ &\cong \frac{\sigma_i^2}{n-(M+i)}\mathrm{tr}\left[\mathbf{R}_i^{-1}\right] \\ &\quad n \geq M+i+1\end{aligned} \tag{38}$$

where $\boldsymbol{\epsilon}_i(n)$ is the weight error vector given by $\boldsymbol{\epsilon}_i(n) = \mathbf{w}_{t,i}(n) - \mathbf{w}_{o,i}$.

$$E\left[|\xi_i(n)|^2\right] \cong \sigma_i^2 + \frac{M\sigma_i^2}{n-(M+i+1)}, \quad n \geq M+i+2 \tag{39}$$

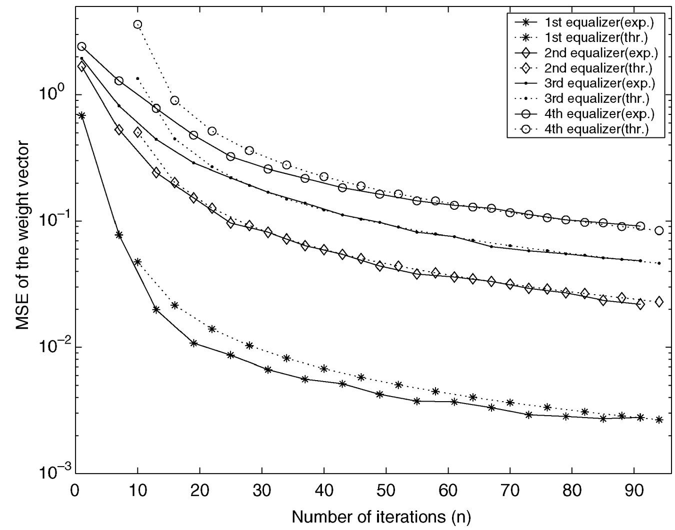


Fig. 3. MSE of tap weight vectors when $M = 4$, $N = 4$, and $E_b/N_0 = 15$ dB.

where $\xi_i(n)$ is *a priori* estimation error defined by $\xi_i(n) = d_{k_i}(n) - \mathbf{w}_{t,i}^H(n-1)\mathbf{y}_{t,i}(n)$.

In (38) and (39), the approximation can be replaced with the equality when $i = 1$, and the approximation error for $i \geq 2$ decreases as n increases. The mean-square value of $\xi_i(n)$ yields a learning curve for the RLS algorithm. Next, the theoretical results in (38) and (39) are compared with the corresponding simulation results for Algorithm 1.

B. Comparison Between RLS Theory and Behavior of Algorithm 1

The received signal vector $\mathbf{y}(n)$ was generated under the following assumptions: The elements of the channel matrix $\mathbf{H}$ were i.i.d. complex Gaussian with unit power and a zero mean; $\mathbf{H}$ was fixed during a burst; $\mathbf{d}(n)$ was composed of Hadamard sequences of length 128; and $\mathbf{u}(n)$ was i.i.d. complex Gaussian noise with a zero mean. The number of antennas was set at $M = N = 4$; the forgetting factor $\lambda = 1$; and $E_b/N_0 = 15$ dB. In the simulation, the empirical MSE values were obtained through 200 independent trials using Algorithm 1. The optimal detection ordering for $\mathbf{H}$ was obtained and used for evaluating the theoretical values in (38) and (39).

Fig. 3 compares the theoretical MSE values for the RLS weight vectors in (38) with the experimental MSEs for the weights from Algorithm 1. Fig. 3 shows a good agreement between the RLS theory and the experiment with Algorithm 1, when $n > 20$. The divergence between the theoretical and experimental curves in the first tens of iterations was caused by the approximation in (38). The learning curve corresponding to $E[|\xi_2(n)|^2]$ is shown in Fig. 4. The theoretical MSE in (39) reached a steady-state after about 14 iterations ($= 2 \times (M + N - 1)$) and was close to that of the experimental MSE from Algorithm 1. The results in Figs. 3 and 4 indicate that Algorithm 1 acted like a bank of RLS algorithms after about 20 iterations and converged to the optimal detection order.

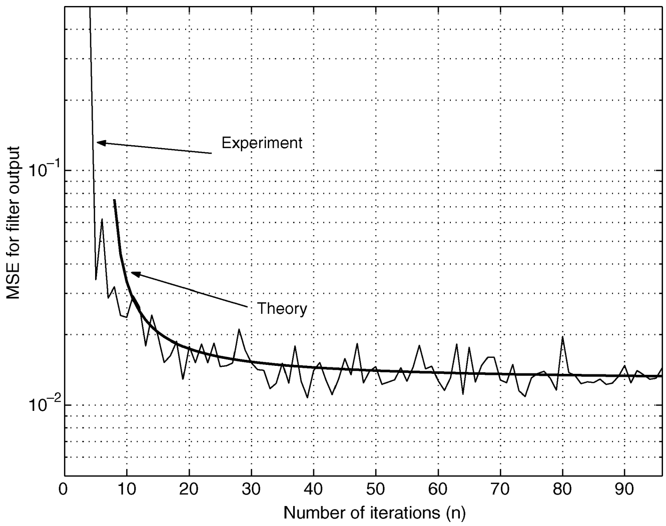


Fig. 4. Learning curves for second equalizer of Algorithm 1 when $M = 4$, $N = 4$, and $E_b/N_0 = 15$ dB.

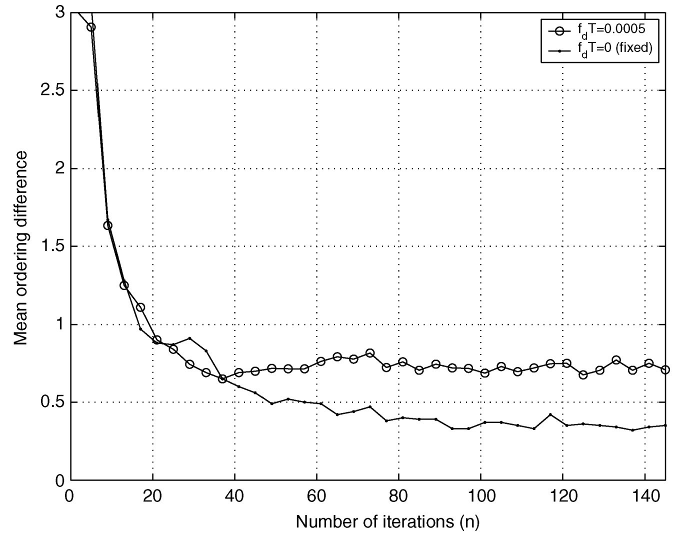


Fig. 5. Mean ordering error curves when $M = 4$, $N = 4$, and $E_b/N_0 = 15$ dB. Here, f_dT denotes the normalized Doppler frequency.

V. Simulation Results

This section compares the BER performance of the proposed adaptive MIMO-DFE with that of V-BLAST and the receivers in [3] and [4]. In the simulation, the following parameters were assumed: QPSK was used; the channels were independent Rayleigh flat fading, and each channel between a transmit and receive antenna pair was varied based on the Jakes model [12]; $\mathbf{u}(n)$ was i.i.d. complex Gaussian noise with a zero mean. The vectors $\{\mathbf{d}(n)\}$ were grouped into frames consisting of 160 vectors, where the first 32 vectors were training vectors. The number of antennas was set at $M = N = 4$. MMSE detection was used for all V-BLAST receivers, and the sub-block size was 4 for the receiver in [3]. For the V-BLAST receivers, the channel was estimated using the RLS algorithm, and the noise variance was obtained by

$$\hat{\sigma}^2(n) = \frac{1}{N}\sum_{l=1}^{n}\lambda^{n-l}(1-\lambda)\cdot\left\{\mathbf{y}(l)-\hat{\mathbf{H}}(l)\hat{\mathbf{d}}(l)\right\}^H \times\left\{\mathbf{y}(l)-\hat{\mathbf{H}}(l)\hat{\mathbf{d}}(l)\right\}. \quad (40)$$

In the adaptive decorrelating detector in [4], the order of detection was determined based on a zero-forcing criterion, using the estimated channel at the end of the training period. The detection order was fixed and used for processing all data vectors in a frame. The tap weights of the decorrelating detector were adjusted using the time-update RLS algorithm in (12)–(15).

To examine the performance of the proposed algorithm for determining the detection order, the following error measure was defined:

$$D(n) = \sum_{i=1}^{M} 1 - \delta(\hat{k}_i - k_i) \quad (41)$$

where $\hat{k}_i$ and k_i denote the detection orders determined by Algorithm 1 and the optimal V-BLAST receiver based on known channels, respectively, and $\delta(\cdot)$ is the unit sample sequence. $D(n)$ is called the ordering error. It becomes zero ($D(n) = 0$) when $\hat{k}_i = k_i$ for all i, and $0 \leq D(n) \leq M$. Fig. 5 shows the mean ordering error curves that were obtained by averaging $D(n)$ over 500 independent trials. The ordering errors converged to steady-state values after about 40 iterations. As expected, the steady-state ordering error for the time-varying channel was larger than that for the fixed channel. The steady-state ordering error is a cause of performance degradation in Algorithm 1.

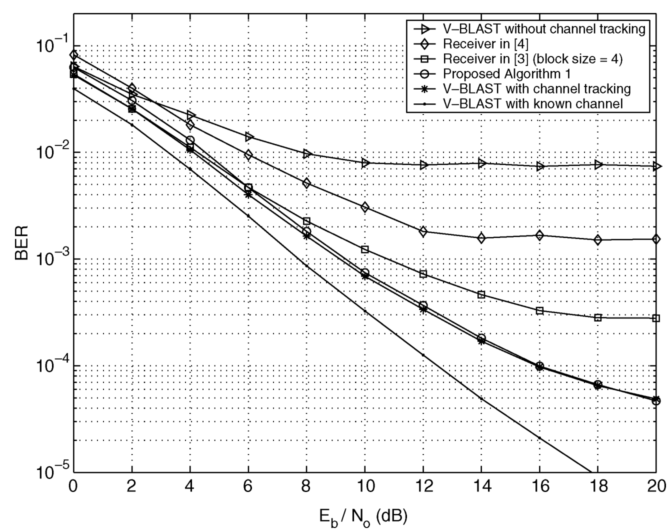


Fig. 6. Comparison of BER performances when $M = 4$, $N = 4$, and $f_dT = 0.0005$.

Fig. 6 shows the BER against E_b/N_0 when the normalized Doppler frequency $f_dT = 5 \cdot 10^{-4}$. The step size λ was 0.95, which was the optimal for the given f_dT. The proposed receiver exhibited an almost identical performance to the V-BLAST processor with channel tracking, yet outperformed the schemes in [3], [4], and the V-BLAST receiver without

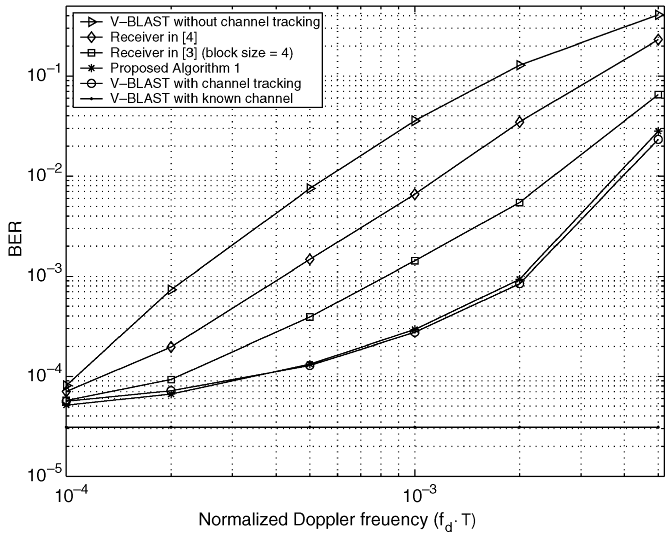


Fig. 7. Comparison of BER performances for various $f_d T$ values when $M = 4$, $N = 4$, and $E_b/N_0 = 15$ dB.

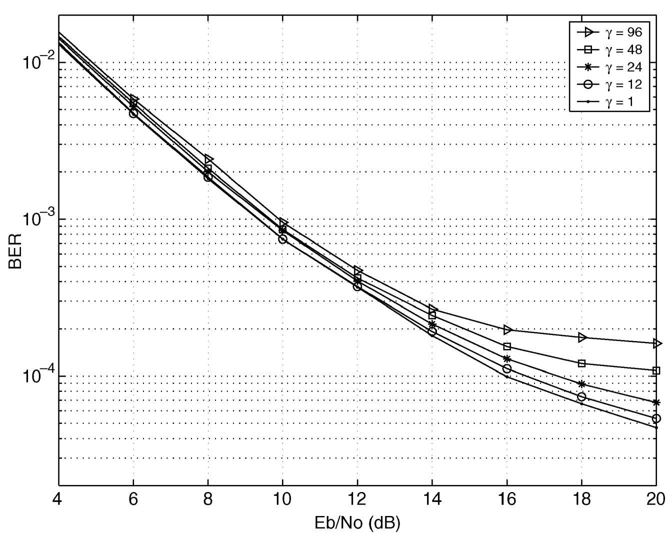


Fig. 8. BER performances of Algorithm 1 and 2 when $M = 4$, $N = 4$, and $f_d T = 0.0005$ (Algorithm 2 becomes Algorithm 1 when $\gamma = 1$).

channel tracking.[2] The BER values for the V-BLAST receiver with known channels, which is actually impractical, provided a performance bound. Comparing the receivers in [3], [4], and the V-BLAST without channel tracking, the receiver in [3] outperformed the others, and the V-BLAST performed the worst. This happened because the receiver in [3] updated both the tap weights and the detection order, whereas the decorrelating detector in [4] only updated the tap weights, and the V-BLAST employed fixed parameters. Fig. 7 presents the BER against $f_d T$ when $E_b/N_0 = 15$ dB. Again, the proposed and V-BLAST with channel tracking acted in a similar manner and outperformed the others. Fig. 8 shows the performance of the proposed algorithm when the detection ordering was periodically updated. When the period of detection order-update γ was less than a 24-symbol time, the performance loss caused by the sparse ordering was less than 1.2 dB at BER $= 10^{-4}$.

[2]The tap weights and detection order for the V-BLAST without channel tracking were determined during the training period and used for data processing without adaptation.

VI. Conclusion

Adaptive algorithms for MIMO-DFE were proposed for receivers with time-varying channels. The proposed receivers update the weight vectors through time- and order-update operations and adaptively determine the detection ordering according to an LSE criterion. With time-varying channels, the proposed techniques are simpler to implement than a V-BLAST processor with RLS channel tracking, yet the performances are comparable. Extending the proposed algorithms to frequency-selective fading channels remains to be investigated.

References

[1] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. IEEE ISSSE*, Pisa, Italy, Sep. 1998, pp. 295–299.

[2] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 11, pp. 1841–1852, Nov. 1999.

[3] Q. Sun and D. C. Cox, "Training-based channel estimation for continuous flat fading BLAST," in *Proc. IEEE ICC*, Helsinki, Finland, Jun. 2002, pp. 325–329.

[4] D. S. Chen and S. Roy, "An adaptive multiuser receiver for CDMA systems," *IEEE J. Sel. Areas Commun.*, vol. 12, no. 6, pp. 808–816, Jun. 1994.

[5] P. B. Rapajic and B. S. Vucetic, "Adaptive receiver structures for asynchronous CDMA systems," *IEEE J. Sel. Areas Commun.*, vol. 12, no. 5, pp. 685–697, May 1994.

[6] A. Maleki-Tehrani, B. Hassibi, and J. M. Cioffi, "Adaptive equalization of multiple-input multiple-output (MIMO) channels," in *Proc. IEEE ICC*, New Orleans, LA, Jun. 2000, pp. 1670–1674.

[7] G. Ginis, Y. Bar-Ness, and J. M. Cioffi, "Blind adaptive MIMO decision feedback equalization using Givens rotations," in *Proc. IEEE ICC*, Helsinki, Finland, Jun. 2002, pp. 59–63.

[8] W. Zha and S. D. Blostein, "Modified decorrelating decision-feedback detection of BLAST space-time system," in *Proc. IEEE ICC*, Helsinki, Finland, Jun. 2002, pp. 335–339.

[9] G. Ginis and J. M. Cioffi, "On the relation between V-BLAST and the GDFE," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 364–366, Sep. 2001.

[10] G. H. Golub and C. F. V. Loan, *Matrix Computations*, Third ed: Johns Hopkins Univ. Press, 1996.

[11] S. Haykin, *Adaptive Filter Theory*, Third ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

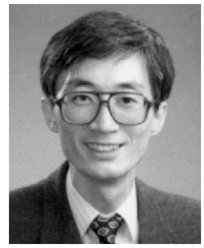
[12] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

Jihoon Choi received the B.S., M.S., and Ph.D. degrees from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 1997, 1999, and 2003, respectively.

From 2003 to 2004, he was with Department of Electrical and Computer Engineering, the University of Texas at Austin, where he performed research on MIMO-OFDM systems with limited feedback. Since 2004, he has been with the Samsung Electronics, Suwon, Korea. With Samsung, he has participated in Portable Internet (PI) and Enhanced Broadcast Multicast Service (BCMCS) projects. His research interests include MIMO, multicarrier modulation, equalization, and space-time coding for wireless communications.

Heejung Yu received the B.S. degree in radio science and engineering from the Korea University, Seoul, Korea, in 1999 and the M.S. degree in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2001.

Since 2001, he has been with the Electronics and Telecommunications Research Institute (ETRI), Daejeon. With ETRI, he has participated in IEEE 802.11 a/g and the next-generation wireless LAN (IEEE 802.11n) modem development projects. His research interests include OFDM, MIMO, synchronization, and equalization for wireless communications.



Yong H. Lee (S'81–M'84–SM'98) received the B.S. and M.S. degrees in electrical engineering from Seoul National University, Seoul, Korea, in 1978 and 1980, respectively, and the Ph.D. degree in electrical engineering from the University of Pennsylvania, Philadelphia, in 1984.

From 1984 to 1988, he was an Assistant Professor with the Department of Electrical and Computer Engineering, State University of New York at Buffalo. Since 1989, he has been with the Department of Electrical Engineering and Computer Science, Korea Advanced Institute of Science and Technology, Daejeon, Korea, where currently is a Professor. His research activities are in the area of communication signal processing, which includes synchronization, estimation, detection, interference cancellation, and efficient filter design for CDMA, TDMA, and orthogonal frequency division multiplexing systems. He is also interested in designing and implementation of transceivers.