Low-Complexity Systolic V-BLAST Architecture
Toshiaki Koike-Akino

Abstract—In multiple-input multiple-output systems, an ordered successive interference canceller, termed vertical Bell laboratories layered space-time (V-BLAST) algorithm, offers good performance. This letter presents a low-complexity V-BLAST scheme suited for parallel implementation. The proposed scheme, using a greedy ordering, can achieve a performance comparable to that of V-BLAST with optimum ordering, while its computational complexity is lower than a linear detector.

Index Terms—MIMO, V-BLAST, greedy ordering

I. INTRODUCTION

Radio communications systems using multiple antennas at both transmitter and receiver, to form a multiple-input multiple-output (MIMO) system, have attracted attention, as a promising technique for achieving a significant increase in spectrum efficiency [1, 2]. A large number of MIMO detectors, such as linear detectors based on minimum mean-square error (MMSE) and based on zero-forcing (ZF), have been studied so far. Although linear detectors are generally low in complexity, their performance can be poor, especially in MIMO systems that use a small number of receiving antenna branches. To improve performance, a so-called vertical Bell laboratories layered space-time (V-BLAST) algorithm has been introduced; this performs successive interference cancellations in the appropriate order [3–8]. V-BLAST yields higher diversity gains and improves bit-error-rate (BER) performance.

The computational complexity of V-BLAST, when compared to a linear detector, is generally increased from $O[M^4]$ to $O[M^3]$. Here, $O[\cdot]$ denotes the order of complexity and $M$ denotes the number of antenna pairs. In time-varying MIMO channels, since the V-BLAST requires re-ordering and weight updating, the complexity reduction is significant for any realistic application. Hitherto, some advanced algorithms have been derived to decrease the complexity from $O[M^4]$ to $O[M^3]$ (see references in [5–10] for example). In [9], Wübben et al. have introduced norm-based sub-optimum ordering and a post-sorting algorithm to improve performance. In this letter, we propose an efficient recursion approach with greedy ordering that is suitable for parallel systolic-array implementation [11]. By using greedy ordering, the computational complexity of MMSE/V-BLAST is reduced and is lower than that of a simple MMSE linear detector with almost no performance degradation, even without optimum ordering. Through performance analyses, we confirm that the proposed scheme is significantly more advantageous than existing schemes that use other algorithms.

II. V-BLAST TECHNIQUE WITH AN MMSE CRITERION

A. MMSE Nulling

The subject of this letter is an uncoded $M \times N$ MIMO system, which uses $M$ transmitting and $N$ receiving antennas (assuming that $M \leq N$). The received signal is modelled as

$$y = Hx + w,$$

where $y \in \mathbb{C}^{N \times 1}$, $H \equiv [h_1, h_2, \ldots, h_M] \in \mathbb{C}^{N \times M}$, $x \equiv [x_1, x_2, \ldots, x_M]^T \in \mathbb{C}^{M \times 1}$ and $w \in \mathbb{C}^{N \times 1}$ denote the received signals vector, the channel matrix, the transmitted signals vector, and the additive Gaussian noise vector, respectively. A superscript $[\cdot]^T$ represents the transpose. Here, $h_m \in \mathbb{C}^{N \times 1}$, $x_m$ denote the channel vector and the transmitted signal, respectively, from the $m$-th transmitting antenna. Using the channel matrix $H$, the MMSE nulling weight matrix $G_M \in \mathbb{C}^{M \times N}$ is given in favour of minimizing the mean-square error between the actual transmitted signal $x$ and the filter output $G_My$ as follows:

$$G_M = \arg\min_{G \in \mathbb{C}^{M \times N}} E \| x - Gy \|^2$$

$$= H^\dagger (H H^\dagger + \sigma^2 I_N)^{-1} H \dagger R_m^{-1}$$

$$= (H^\dagger H + \sigma^2 I_M)^{-1} H^\dagger Q_M^{-1} H^\dagger,$$

where $\| \cdot \|$ denotes the Euclidean norm, $I_p \in \mathbb{R}^{p \times p}$ is the identity matrix, $E[\cdot]$ stands for the expectation function, and $\cdot^\dagger$ represents the Hermitian transpose. Here, it is assumed
that \( E[xx^\dagger] = I_M \) and \( E[ww^\dagger] = \sigma^2 I_N \). (If the system uses any linear precoding with a transmitting matrix \( P_T \) and a receiving matrix \( P_R \), we need only to substitute \( P_R H P_T \) for \( H \) to obtain the MMSE weight matrix for such precoding.)

The MMSE weight generation requires a matrix inversion: its computational complexity is generally a cubic order that accords with the matrix size [12]. Therefore, the formula expressed in (4) is often more useful than the one in (3) because the matrix size of \( Q_M \) is smaller than or equal to that of \( R_M \). Since the error covariance is expressed as

\[
E\left[(x - Q_M^{-1} H^\dagger y)(x - Q_M^{-1} H^\dagger y)\right] = \sigma^2 Q_M^{-1}, \quad (5)
\]

the most reliable signal with the smallest error rate can be predicted by observing the diagonal entries of \( Q_M^{-1} \). (Since we consider the same constellation for each transmitting antenna for the sake of simplicity, the smallest diagonal entry corresponds to the most reliable signal. For a per-antenna link adaptation [13], we should slightly modify its ordering metric.) On the other hand, when using the formulation of \( G_M = H^\dagger R_M^{-1} \), we should compute \( h_m^\dagger R_M^{-1} h_m \) for all \( 1 \leq m \leq M \) to estimate the most reliable signal, because the error covariance can be written as

\[
E\left[(x - H^\dagger R_M^{-1} y)(x - H^\dagger R_M^{-1} y)\right] = I_M - H^\dagger R_M^{-1} H. \quad (6)
\]

Hence, some researchers have used the representation in (4) to reduce complexity. Nevertheless, the expression in (3) is useful when deriving the proposed recursion algorithm because the proposed scheme does not need to compute \( R_M^{-1} \) explicitly, as is explained later in detail.

B. Basic V-BLAST Algorithms

As discussed above, we can choose the most reliable signal out of the MMSE nulling outputs. To simplify, it is supposed that the signals are optimally sorted in order of ascending reliability (the MMSE output for \( x_M \) is the most reliable and has the least error variance). At first, we detect the most reliable signal \( x_M \) using the MMSE nulling weights \( G_M \). Next, using the obtained hard-decision symbol \( \hat{x}_M = \text{dec}(G_M[M:y]) \), where \( \text{dec}() \) denotes the decision function and \( [A]_i \) is the \( i \)-th row vector of a matrix \( A \), the receiver subtracts the corresponding replica \( h_M \hat{x}_M \) from \( y \). If the decision is not erroneous, this replica subtraction leads to a better weight matrix \( G_{M-1} \) for the next reliable signal \( x_{M-1} \) using a deflated channel matrix \( H_{M-1} \) which is obtained by removing the last column of \( H \). Correspondingly, the V-BLAST algorithm successively detects multiplexed signals \( x_M, x_{M-1}, \ldots, x_1 \) in the most reliable order, and this improves the diversity gains significantly. Since V-BLAST requires \( M \) times weight generation and \((M - 1)\) times replica cancellation, the order of computational complexity is generally believed to be \( \mathcal{O}(M^2) \). It has been said that the complexity of the traditional V-BLAST scheme is inevitably greater than that of an MMSE linear detector. We propose an efficient technique that has a lower complexity than an MMSE linear detector and most existing low-complexity V-BLAST schemes [5–8].

C. Revision of Fast V-BLAST Algorithms

Here, we reconsider the lowest-complexity algorithm proposed in [8]. This algorithm uses a recursive technique generating the inversion matrix \( Q_M^{-1} \) to reduce its complexity. We define \( Q_m = \left( H_m^\dagger H_m + \sigma^2 I_m \right) \) as an equation in which \( H_m = [h_{1}, h_{2}, \ldots, h_{m}] \). Using the deflated matrix \( Q_{m-1} \), the matrix \( Q_m \) is rewritten as

\[
Q_m = \left[ Q_{m-1} - \nu_m \right], \quad (7)
\]

where \( \nu_m \triangleq h_m^\dagger h_m \) and \( \nu_m \triangleq \|h_m\|^2 + \sigma^2 \). The inversion of the matrix above is obtained as

\[
Q_m^{-1} = \left[ Q_{m-1}^{-1} + \mu_m u_m u_m^\dagger - \mu_m u_m \right] \triangleq \left[ T_m \ t_m \right], \quad (8)
\]

where \( u_m \triangleq Q_{m-1}^{-1} h_m \) and \( \mu_m \triangleq 1/(\nu_m - u_m^\dagger u_m) \). Using these, we can recursively generate \( Q_M^{-1} \) with an initial matrix of \( Q_1^{-1} = 1/(\sigma^2 + \|h_1\|^2) \). The recursion above is more efficient for deriving \( Q_M^{-1} \) than the original algorithm presented in [8], in which the computation of \( Q_M \) is performed using the Sherman-Morrison formula[12].

In addition, matrix deflation from \( Q_m^{-1} \) to \( Q_{m-1}^{-1} \) that occurs in [8] can be further improved in the following manner. As shown in (8), we can compute \( Q_{m-1}^{-1} \) from \( Q_m^{-1} \) as

\[
Q_{m-1}^{-1} = T_m - \mu_m u_m t_m^\dagger, \quad (9)
\]

This is a simpler method than the one based on the Sherman-Morrison formula that is proposed in [8]. Although the derivation in (8) has also been introduced in [8] for the sole purpose of matrix deflation, the chief difference between it and our modification is that this expression is used for constructing \( Q_m^{-1} \) from \( Q_{m-1}^{-1} \) and for deflating \( Q_m^{-1} \) to \( Q_{m-1}^{-1} \).

The original scheme, which uses the Sherman-Morrison formula, in [8] requires \( \mathcal{O}(6M^2N) \) for constructing \( Q_M^{-1} \) and \( \mathcal{O}(2M^3) \) for deflating the matrix. The revised scheme described in this section can reduce the computational complexity to \( \mathcal{O}(2M^2N + 2M^3) \) when constructing \( Q_M^{-1} \) with the method described in (8) and for deflation to \( \mathcal{O}(2M^3/3) \) using the method described in (9). Therefore, our modification can reduce the complexity by \( \mathcal{O}(10M^3/3) \) for \( M = N \). This section is focused on the basic recursion idea of the fast algorithm in [8], which we modified to decrease complexity.

III. PROPOSED RECURSIVE ALGORITHM

A. Matrix Recursion

The matrix \( R_M \) is rewritten as

\[
R_M = \sigma^2 I_N + \sum_{m=1}^{M} h_m h_m^\dagger. \quad (10)
\]

With the term \( \sigma^2 I_N \) defined as \( R_0 \), we obtain the recurrence formula \( R_m = R_{m-1} + h_m h_m^\dagger \). The inversion matrix \( R_m^{-1} \) can be obtained in a recursive manner using the Sherman-Morrison formula with an initial matrix of \( R_0^{-1} = 1/\sigma^2 I_N \):

\[
R_m^{-1} = R_{m-1}^{-1} - \frac{R_{m-1}^{-1} h_m h_m^\dagger R_{m-1}^{-1}}{1 + h_m R_{m-1}^{-1} h_m}. \quad (11)
\]
We call this recursive technique a matrix recursion. Matrix recursion can reduce the complexity order of V-BLAST from quadratic to cubic [8]. However, the inversion matrix $R_m^{-1}$ is not explicitly needed. To achieve this, we need only the weight vector $h_m^\dagger R_p^{-1}$ for detecting $x_m$. More specifically, we do not require $h_m^\dagger R_p^{-1}$ when $m < p$. If we can predict the proper ordering in advance, we can omit these calculations.

B. Vector Recursion

Using (11), we can derive the weight vectors in a recursive manner without explicitly using $R_m^{-1}$.

$$g_{m,p}^\dagger \triangleq h_m^\dagger R_p^{-1} = h_m^\dagger R_p^{-1} - \frac{h_m^\dagger R_p^{-1} h_p h_p^\dagger R_p^{-1}}{1 + h_p^\dagger R_p^{-1} h_p}$$

$$= g_{m,p-1}^\dagger - \alpha_{m,p} g_{p,p-1}^\dagger,$$

(12)

where

$$\alpha_{m,p} \triangleq \frac{g_{m,p-1}^\dagger h_p}{1 + g_{p,p-1}^\dagger h_p}.$$  

(13)

We refer to this recursive technique for obtaining weight vectors as vector recursion. This scheme can systematically generate all weight vectors $\{g_{m,m}\}$. It is also suitable for parallel systolic array implementation [11]. Note that vector recursion can be derived by using the recursive formula for $Q_n^{-1}$ found in (8); since the same recursive formula is generated, its complexity is not different from the above recursion.

Note that replica cancellation can be performed by a scalar operation instead of a vector one:

$$\hat{x}_m = g_{m,m}(y - \sum_{p=m+1}^{M} h_p \hat{x}_p) = g_{m,m} y - \sum_{p=m+1}^{M} \alpha_{p,m}^* \hat{x}_p,$$

(15)

where $[\cdot]^*$ denotes the complex conjugate, and the factors $\alpha_{p,m}$ are already calculated in (13).

The recursion process for obtaining the weight vector $g_{M,M}^\dagger$ can provide all weight vectors and replica cancellation factors. Because we can omit weight generations for $g_{m,p}^\dagger$ when $m < p$, the computational complexity of V-BLAST becomes smaller than that of an MMSE linear detector (which requires $g_{m,M}^\dagger$ for all $1 \leq m \leq M$). This results from the assumption that ordering is performed in advance. For this reason, we propose a greedy ordering method, with which we determine the least reliable signal by observing $g_{m,p-1}^\dagger$ at each step.

C. Greedy Ordering

We can select the most reliable signal to be detected first, the one which has the maximum ordering metric given by $\gamma_{m,p} = \|g_{m,p}^\dagger h_m\|$ as in (6). Based on this metric, we propose the following greedy ordering. At the $p$-th recursion step, we decide on the least reliable signal to be detected:

$$\gamma_{m,p-1}' \triangleq \frac{(g_{m,p-1}^\dagger h_m)^2}{\sum_{i=p}^{M} \|g_{i,m-1}^\dagger h_i\|^2}.$$  

(16)

This represents a kind of signal-to-interference ratio (SIR).

D. Scalar Recursion

In addition, we can simplify the vector recursion scheme to a scalar recursion scheme, focusing on the fact that the receiver needs only the nulling outputs $\hat{g}_{m,m}^\dagger y$. Using (13), we can generate the filter output in a recursive manner:

$$\xi_{m,p} = \xi_{m,p-1} - \alpha_{m,p} \xi_{p,p-1},$$

(17)

where $\xi_{m,p} \triangleq \hat{g}_{m,m}^\dagger y$. Here, the factor $\alpha_{m,p}$ is rewritten as

$$\alpha_{m,p} = \frac{\eta_{m,p-1}}{1 + \eta_{p,p-1}},$$

(18)

where $\eta_{m,p} \triangleq \hat{g}_{m,m}^\dagger h_q$. For calculating $\alpha_{m,p}$, however, the following scalar recursion is needed:

$$\eta_{m,p,q} = \eta_{m,p-1,q} - \alpha_{m,p} \eta_{p,p-1,q},$$

(19)

Note that there is a relationship: $\eta_{m,p,q} = \eta_{q,p,m}^*$, which can omit superfluous computation. The ordering metric can be rewritten as

$$\gamma_{m,p-1}' = \frac{\eta_{m,p-1,m}}{\sum_{i=p}^{M} |\eta_{m,p-1,i}|^2}.$$  

(20)

Using the scalar recursion, the systolic array architecture becomes a multi-layered structure as illustrated in Fig. 1. The upper layer can be computed using the lower layers. For preprocessing, we require $H^\dagger y$ and $H^\dagger H$. Its computational complexity is $O[2NM^3]$. As an essential element, the systolic computation requires an additional $O[2NM^3/3]$ complexity. For computing ordering metrics, we need an $O[M^3/3]$ complexity. The top-layered recurrence formula in (17) should be computed for every symbol, whereas the calculation of the recurrence formula in (19) is needed only once in a frame if the channel does not change. We summarize the proposed scalar recursion scheme in Algorithm 1. As shown in Fig. 1, the required memory load is $M$ for $\xi_{m,p}$, $M(M+1)/2$ for $\eta_{i,p,j}$, and $M(M+1)/2$ for ordering at the $p$-th stage. Note that $M(M+3)/2$ memories are used for real-valued numbers, and the others are for complex-valued numbers.

E. Adaptive Scalar Recursion for Fast Fading

For fast fading channels, the MMSE weight matrix should be updated frequently and optimum ordering may change during a transmission frame. For reducing complexity in such a channel, we present an adaptive scheme. The initialization process in the proposed algorithm dominates the overall complexity: $\eta_{i,0,j}$ ($H^\dagger H/e^2$) has an $O[2NM^2]$ complexity, which we reduce to a square-order complexity in this section.

It is assumed that the receiver will employ some kind of least-squares adaptive algorithms for tracking channel estimation. The least mean-square (LMS) algorithm or the recursive least-squares (RLS) algorithm is typically used for fast fading channels[11]. For both channel estimation schemes, a channel estimation is updated with the following:

$$H' = H + e b^\dagger,$$

(21)

where $e = y - H \hat{x} \in \mathbb{C}^N$ is an error vector and $\hat{x}$ is the hard decision for $x$. The vector $b \in \mathbb{C}^N$ is given as $b = \rho_w \hat{x}$ for LMS, and $b = \Phi \hat{x} / (\lambda_q + \hat{x}^\dagger \Phi \hat{x})$ for RLS, where $\Phi$ is updated as $\Phi' = (\Phi - b \hat{x}^\dagger \Phi) / \lambda_{ff}$ with an initial condition.

\[\text{IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, VOL. XX, NO. XX, MONTH 2009}\]
Algorithm 1 Proposed V-BLAST Scalar Recursion Algorithm with Greedy Ordering

— Initialization: $S_0 = \{1, 2, \ldots, M\}$
for all $i, j \in S_0$ do
  $\xi_{i,0} = h_i^j y/\sigma^2$
  $\eta_{i,0,j} = h_i^j h_j/\sigma^2$
end for
— V-BLAST Recursion with Greedy Ordering:
for $k = 1$ to $M$ do
  $p_k = \arg\min_{m \in S_k-1} (\eta_{m,k-1,m} / \sum_{\ell \in S_{k-1}} |\eta_{m,k-1,\ell}|^2)$
  $S_k = S_{k-1} \setminus \{p_k\}$
  for all $m \in S_k-1$ do
    $\alpha_{m,k} = \eta_{m,k-1,p_k} / (1 + \eta_{p_k,k-1,p_k})$
    $\xi_{m,k} = \xi_{m,k-1} - \alpha_{m,k} \xi_{p_k,k-1}$
  end for
  for all $q \in S_k$ do
    $\eta_{m,k,q} = \eta_{m,k,q-1} - \alpha_{m,k} \eta_{p_k,k-1,q}$
  end for
end for
— Successive Detection
for $k = M$ to $1$ do
  $\hat{x}_{pk} = \text{dec}(\xi_{p_k,k} - \sum_i \alpha_{i,k}^* \hat{x}_i)$
end for

de of $\Phi = [I_M | \cdots | I_M]$ ($\xi \gg 1$). Here, $\mu_{ag}$ is referred to as a step-size factor, and $\lambda_{ag}$ is a forgetting factor. Using (21), we can update $\eta_{i,0}$ as follows:

$$
A' = A + b (\xi - A \hat{x} + \|e\|^2 b/2\sigma^2) + (\xi - A \hat{x} + \|e\|^2 b/2\sigma^2) b^T,
$$

where $A = H^T H / \sigma^2$ (the $(i,j)$-th entry is $\eta_{i,0,j}$) and $\xi = [\xi_{1,0}, \xi_{2,0}, \ldots, \xi_{M,0}]^T$. The above computation may reduce the complexity of the initial process to a square order $O[8M^2]$.

**Fig. 1.** Scalar recursion for MMSE/V-BLAST in a systolic array with greedy ordering ($M = 3$).

**Fig. 2.** Asymptotic complexity in multiplications as a function of number of antenna pairs ($M = N$).

**IV. EVALUATION**

**A. Computational Complexity**

Table I lists the computational complexity for the original V-BLAST algorithm, the square-root algorithm [6], the fast algorithm [8], the modified fast algorithm in section II-C, the sorted QR decomposition [9], the proposed scalar recursion algorithm with greedy ordering, the adaptive scalar recursion algorithm with greedy ordering, and the MMSE linear detector using a scalar recursion for reference. This is the total number of real-valued multiplications and additions for the overall process including weight generations, ordering, interference cancellations and so on. Here, the Golub-Reinsch algorithm [12] is employed for a computationally stable way to compute an inversion matrix in the conventional V-BLAST algorithm [8]. The number of multiplications are plotted in Fig. 2 as a function of the number of antenna pairs for $M = N$. Table II shows their asymptotic complexities when $M = N$ is used. One can see that the proposed scheme offers a remarkable advantage for computational complexity. The adaptive scalar recursion algorithm for fast fading is also advantageous, especially for a large number of antenna pairs. Note that the proposed scalar recursion has a lower complexity than does a simple MMSE linear detector.

**B. Bit Error Rate Performance**

Figs. 3 and 4 show BER performance as a function of average $E_b/N_0$ per receiving antenna for uncoded QPSK 4×4 systems and uncoded 16QAM 4×4 systems, respectively, in frequency-flat Rayleigh fading channels. These figures plot the performance curves of V-BLAST scalar recursion with optimum ordering, V-BLAST scalar recursion with greedy ordering, V-BLAST based on the sorted QR decomposition, and MMSE linear detection. Note that the scalar recursion with optimum ordering performs in the same way as a
conventional MMSE/V-BLAST algorithm with optimum ordering. As shown in these figures, V-BLAST offers good BER performance when compared to an MMSE linear detector, especially for low level modulations. Note that the proposed algorithm with greedy ordering offers an almost comparable performance to a V-BLAST with optimum ordering, and it significantly outperforms an MMSE linear detector and the sorted QR decomposition BLAST algorithm although its computational complexity is lower than that of other detectors. The proposed scheme works well even for high level modulations, whereas the sorted QR decomposition suffers from a severe performance degradation.

V. Summary

First, we described a lower complexity version of the V-BLAST fast algorithm originally proposed in [8]. Then we proposed a more computationally efficient recursion scheme for MMSE/V-BLAST, which is suitable for hardware implementation, and greedy ordering. The proposed technique is less complex than most existing V-BLAST algorithms, because of its simplified recurrence formula. This proposal can readily deal with optimum ordering as well as greedy ordering. It should be noted that even though the proposed scheme is less complex than a simple MMSE linear detection, the proposed scheme outperforms the MMSE scheme.

ACKNOWLEDGMENT

This work is partly supported by JSPS Fellowships for Research Abroad, TAF and SCAT Japan.

REFERENCES

nication in a fading environment when using multi-element antennas,”

ordered successive receivers for space-time transmission,” IEEE VTC-


optimum sequential signal detection in a BLAST system” IEEE Trans.

extension of V-BLAST based on sorted QR decomposition,” IEEE VTC-

[10] R. Bohnke, D. Wubben, V. Kahn, and K.-D. Kammeyer, “Reduced com-
plexity MMSE detection for BLAST architectures,” IEEE GLOBECOM,


Johns Hopkins, 1996.

channel capacity using V-BLAST with rate and power adaptation,” IEEE