

Improved Approximate Maximum-Likelihood Receiver for Differential Space–Time Block Codes Over Rayleigh-Fading Channels

Poramate Tarasak, *Student Member, IEEE*, Hlaing Minn, *Member, IEEE*, and Vijay K. Bhargava, *Fellow, IEEE*

Abstract—In this paper, an approximate maximum-likelihood (ML) receiver for differential space–time block codes is investigated. The receiver is derived from the ML criterion and is shown to mitigate error floor occurring in a conventional differential receiver very well. Because the receiver employs knowledges of signal-to-noise ratio (SNR) and fading rate, we study mismatched cases when these parameters are not accurate. It is shown that the receiver is more sensitive to the mismatched parameters when the fading rate is high. Then, a union bound on the bit error probability is derived. The bounds show good agreement with the simulation results at high fading rate and at high SNR. Finally, a modified receiver, denoted as *multistage receiver*, is proposed to compensate the so-called *intra-block interference* caused by the time-varying characteristic of the channel within a transmission block. The multistage receiver offers further reduction of error floor of about half order of magnitude as compared with an approximate ML receiver.

Index Terms—Bit error probability, differential space–time block codes, maximum-likelihood receiver.

I. INTRODUCTION

ALAMOUTI'S space–time block codes [1] offer full spatial diversity while channel gains have to be known at the receiver. In some situations, such as when the mobile moves quickly, the channel varies rapidly and accurate channel estimates are difficult to obtain. Recently, differential space–time coding/modulation was proposed to achieve diversity gain without channel estimation [2]–[4]. Differential schemes in [3] and [4] are based on the group design of unitary matrices where their advantage is in the preservation of constellation. The differential scheme in [2] is based on orthogonal design with the advantage of its low differential detection complexity. However, conventional differential detection of these schemes has an approximately 3-dB performance degradation compared to coherent detection in a quasistatic fading channel. Further

degradation is observed in a time-varying channel and error floor appears at high signal-to-noise ratio (SNR) [5], [6].

It is known that multiple-symbol detection (MSD) [7], [8] and decision-feedback differential detection (DF-DD) [9]–[11] are effective detection schemes in time-varying fading channels for differential phase-shift keying (DPSK) in single-antenna systems. The basic idea is to extend the observation length to be larger than two consecutive signaling intervals and to gain additional information among received signals to improve detection. In time-varying channels, these receivers are closely related to a linear prediction receiver (e.g., [9], [12], and [13]). The linear prediction receiver finds the best transmitted sequence that minimizes Euclidean distance between the channel gains formed by the received signals and those from the prediction.

For group design differential space–time modulation, MSD in quasistatic fading channels has been considered in [14]. For a time-varying channel, MSD and DF-DD have been discussed in [5], where diagonal signal matrices have been chosen because they lead to a simple decoding metric. An approximate bit error rate (BER) for DF-DD was derived and shown to be in good agreement with the simulation results. With minimal increasing complexity, DF-DD reduces the error floor by approximately an order of magnitude. Suboptimum DF-DD using fixed linear predictor is discussed in [15], where it is shown that the suboptimum DF-DD performs close to optimum DF-DD at low fading rate (≈ 0.01). For differential space–time block codes (DSTBC) (orthogonal design), MSD is applied in [16] and [17] in a quasistatic channel. A performance improvement of 0.5–1.5 dB was shown with binary phase-shift keying (BPSK) and an increasing number of observation block intervals, from two to eight. For a time-varying channel, similar receiver to DF-DD with trellis structure has been treated in [18], where an approximate ML receiver was derived. The receiver in [18] can be considered as an extension mentioned in [5] and is able to reduce the error floor by approximately an order of magnitude as well. A drawback of the approximate ML receiver is in the assumption of fixed channel gains within a transmission block. This assumption leads to an irreducible error floor due to *intra-block interference*, caused by varying channel gains within a transmission block (see Section VI).

This paper focuses on the approximate maximum-likelihood (ML) receiver for DSTBC derived in [18] with BPSK constellation. The main contribution of this paper is twofold. First, we study the performance of the receiver under mismatched SNR and fading rate conditions and analyze the BER performance of the receiver by a union bound. Second, a modified receiver,

Manuscript received December 5, 2002; revised September 14, 2003 and November 19, 2003. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) under a Strategic Project Grant. This paper was presented in part at the IEEE International Conference on Communications (ICC), Anchorage, AK, May 2003.

P. Tarasak is with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC V8W 3P6, Canada (e-mail: ptarasak@ece.uvic.ca).

H. Minn is with the Department of Electrical Engineering, Eric Jonsson School of Engineering and Computer Science, University of Texas at Dallas, Richardson, TX 75083-0688 USA.

V. K. Bhargava is with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada.

Digital Object Identifier 10.1109/TVT.2004.823542

called *multistage receiver*, is proposed to mitigate intrablock interference by exploiting channel estimates obtained from the detection at the first stage. A modified receiver yields further reduction of the error floor left from an approximate ML receiver. The paper is organized as follows. Section II describes the system model and Section III briefly summarizes a conventional differential receiver and discusses an approximate ML receiver. Section IV studies the receiver with mismatched SNR and fading rate values by means of simulation. Section V evaluates the upper bound on the bit error probability of the approximate ML receiver and Section VI proposes a multistage receiver and presents its operation. Section VII presents simulation results and some discussion, while conclusions are given in Section VIII.

II. SYSTEM MODEL

The system model and notations follow from [18] with DSTBC mapping from [2]. The system consists of two transmit antennas and one receive antenna. The extension to more than two transmit antennas can be done by applying DSTBC from [19] and extension to more than one receive antenna is straightforward. The mapping from [2] can be rewritten in a matrix form similar to [17], as follows. A data vector $\mathbf{d}_k = [d_{2n-1} \ d_{2n}]^t$, which represents an n th pair of symbol, is selected from a unit energy M -ary (2^{nd} -ary) phase-shift keying (PSK) constellation according to the $2n_d$ data bits using Gray mapping. $(\cdot)^t$ represents a transpose operation. Then, a data matrix \mathbf{G}_n is determined by forming an Alamouti's matrix of a data vector and multiplying it by a unitary matrix as

$$\mathbf{G}_n = \begin{bmatrix} g_{2n-1} & g_{2n} \\ -g_{2n}^* & g_{2n-1}^* \end{bmatrix} = \frac{1}{2} \begin{bmatrix} d_{2n-1} & d_{2n} \\ -d_{2n}^* & d_{2n-1}^* \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}. \quad (1)$$

The factor $1/2$ in (1) ensures the average total transmit power from two transmit antennas to be one. A transmission matrix $\mathbf{D}_n = \begin{bmatrix} s_{2n-1} & s_{2n} \\ -s_{2n}^* & s_{2n-1}^* \end{bmatrix}$ is the Alamouti's format. $(\cdot)^*$ is a complex conjugate. Now, differential encoding can be written as

$$\mathbf{D}_n = \mathbf{G}_n \mathbf{D}_{n-1} \quad (2)$$

where an initial transmission matrix is $\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$. The symbols s_{2n-1} and s_{2n} are transmitted at signaling interval $2n-1$ from the first and second antennas, respectively. The symbol $-s_{2n}^*$ and s_{2n-1}^* are transmitted at signaling interval $2n$ from the first and second antennas, respectively.

With BPSK constellation, differential encoding in this manner preserves the constellation. The system achieves 1-b/s/Hz transmission rate and enjoys two orders of diversity. For higher modulation schemes, the approach in [18] and in this paper can still be applied. However, the complexity of the receiver will be higher due to constellation expansion.

The channel considered is a time-varying frequency-flat Rayleigh-fading channel with Doppler power spectrum according to Jakes' model. The channels corresponding to different transmit antennas are assumed to be independent and identically distributed. Let $a_i[n]$ denote the fading process

corresponding to the i th transmit antenna where $i = 1, 2$. Then, $a_1[n]$ and $a_2[n]$ are zero mean complex Gaussian random variables, each with unit variance and autocorrelation $R_{a_i}[m] = R_a[m] = J_0(2\pi f_d T m)$ where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind, $f_d T$ is the normalized fading rate or normalized maximum Doppler frequency.

The detector in [18] assumes constant channel gains within a transmission block, i.e., during the $(2n-1)$ th and $(2n)$ th signaling intervals. With this assumption, the received signal vector $\mathbf{r}_n = [r_{2n-1} \ r_{2n}]^t$ can be written as [18]

$$\mathbf{r}_n = \mathbf{D}_n \mathbf{a}_n + \mathbf{w}_n, \quad n = 0, 1, \dots, N-1 \quad (3)$$

where $\mathbf{a}_n = [a_1[2n-1] \ a_2[2n-1]]^t$ is the channel gain vector and $\mathbf{w}_n = [w_{2n-1} \ w_{2n}]^t$ is the noise vector. The elements w_{2n-1} and w_{2n} are zero-mean complex Gaussian random variables, each with variance $1/(2\text{SNR}) = \sigma_w^2/2$ per dimension.

III. CONVENTIONAL RECEIVER AND APPROXIMATE ML RECEIVER

Differential detection of DSTBC assumes fixed channel gains during two consecutive transmission blocks. It is known that a differential detector is optimal for a quasistatic channel [18]. The detector computes a vector $\mathbf{z}_{n+1} = [z_{n+1}^1 \ z_{n+1}^2]^t$ with $z_{n+1}^1 = r[2n+1]r^*[2n-1] + r^*[2n+2]r[2n]$ and $z_{n+1}^2 = r[2n+1]r^*[2n] - r^*[2n+2]r[2n-1]$. The decision rule of this receiver is [2]

$$\hat{\mathbf{d}}_{n+1} = \arg \min_{\mathbf{d}_{n+1}} \|\mathbf{z}_{n+1} - \mathbf{g}_{n+1}\|^2 \quad (4)$$

where $\mathbf{g}_{n+1} = [g_{2n+1} \ g_{2n+2}]^t$ and $\|\cdot\|^2$ is a square Euclidean norm of a vector. This receiver will be noted as a *conventional receiver* (CR) [18].

An approximate ML receiver is derived in [18]. It is *approximate* in the sense that it was derived based on the assumption of fixed channel gains during a transmission block, while an actual fading channel varies continuously. Detection applies Viterbi algorithm on a trellis representation of all possible transmitted sequences. The detected sequence, $\hat{\mathbf{D}}_0, \hat{\mathbf{D}}_1, \dots, \hat{\mathbf{D}}_{N-1}$, is the one that maximizes the log-likelihood function

$$l(\mathbf{r}|\mathbf{s}) = - \sum_{n=0}^{N-1} \left\| \mathbf{D}_n^H \mathbf{r}_n - \hat{\mathbf{a}}_n \right\|^2 \quad (5)$$

where $\hat{\mathbf{a}}_n = \sum_{k=1}^Q b_k^Q \mathbf{D}_{n-k}^H \mathbf{r}_{n-k}$ represents the Q th order prediction of \mathbf{a} . $(\cdot)^H$ is a Hermitian operation. $\{b_1^Q b_2^Q \dots b_Q^Q\}$ is a set of linear prediction coefficients of the process $\mathbf{y}_n = \mathbf{D}_n^H \mathbf{r}_n = \mathbf{a}_n + \mathbf{w}_n$ [18]. Later, we will omit the superscript Q in b_k^Q , as it is clear that a Q th order linear predictor is being used. The prediction coefficients can be determined from Cholesky decomposition of the matrix $\mathbf{M}^{-1} = \mathbf{B}^H \Xi \mathbf{B}$, where

$$\mathbf{M} = \begin{bmatrix} R_a[0] + \sigma_w^2 & R_a[-2] & \dots & R_a[-2Q] \\ R_a[2] & R_a[0] + \sigma_w^2 & \dots & R_a[-2(Q-1)] \\ \vdots & \vdots & \ddots & \vdots \\ R_a[2Q] & R_a[2(Q-1)] & \dots & R_a[0] + \sigma_w^2 \end{bmatrix} \quad (6)$$

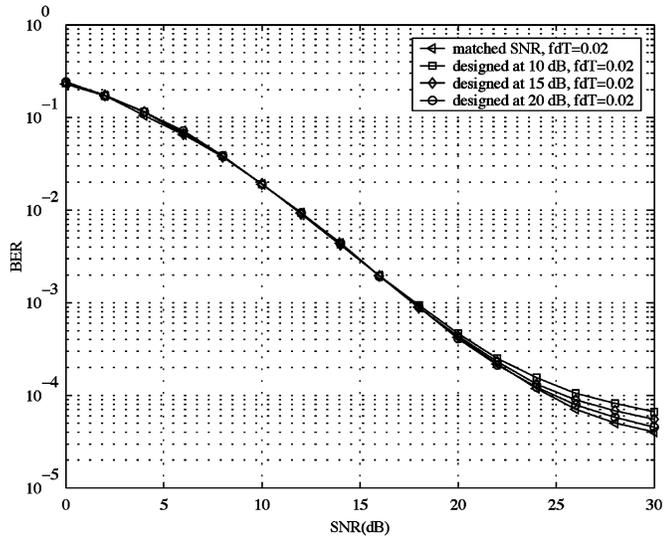


Fig. 1. Mismatched SNR effect to the BER when the design SNR are 10, 15, and 20 dB at $f_d T = 0.02$.

and

$$\mathbf{B} = \begin{bmatrix} -b_0^0 & 0 & 0 & \dots & 0 \\ -b_1^1 & -b_0^1 & 0 & \dots & 0 \\ -b_2^2 & -b_1^2 & -b_0^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -b_Q^Q & -b_{Q-1}^Q & -b_{Q-2}^Q & \dots & -b_0^Q \end{bmatrix}. \quad (7)$$

Ξ is a diagonal matrix containing inversed mean square values of prediction error as diagonal elements. After the receiver obtains the detected sequence $\{\hat{\mathbf{D}}_0, \hat{\mathbf{D}}_1, \dots, \hat{\mathbf{D}}_{N-1}\}$, it can determine the data sequence $\{\hat{\mathbf{G}}_1, \hat{\mathbf{G}}_2, \dots, \hat{\mathbf{G}}_{N-1}\}$ from (2). From (5), the detection problem can be stated to find the data sequence such that it minimizes the square error between the channel gains computed from the received signals and the predicted values.

The trellis structure of DSTBC can be defined as follows. Each trellis interval corresponds to a transmission matrix \mathbf{D}_n . For BPSK, the trellis consists of 2^{2Q} states with 2^2 emerging from each state and terminating at each state. Each state represents a Q -couple transmitted vector $\Gamma_n = [\mathbf{s}_n \mathbf{s}_{n-1} \dots \mathbf{s}_{n-Q+1}]$. The branch metric associated with each transition can be defined as [18]

$$\Lambda(\Gamma_n, \mathbf{d}_{n+1}) = \left\| \mathbf{D}_{n+1}^H \mathbf{r}_{n+1} - \sum_{k=1}^Q b_k^Q \mathbf{D}_{n+1-k}^H \mathbf{r}_{n+1-k} \right\|^2. \quad (8)$$

Exploiting the per-survivor processing technique, the number of states can be reduced to 2^{2P} with $P < Q$. Now, the branch metric $\Lambda(\Gamma_n, \mathbf{d}_{n+1})$ is determined from symbols associated with the transition and symbols along the survivor path terminating at the state Γ_n . The detection algorithm described above is referred to as the *Viterbi receiver* (VR). In this paper, we choose $P = 1$ for simplicity.

Note that the linear predictor embedded in DSTBC receiver can be considered as a *two-step* linear predictor, which forms the prediction of the value $\hat{a}_i[n]$ by a weighted linear combination

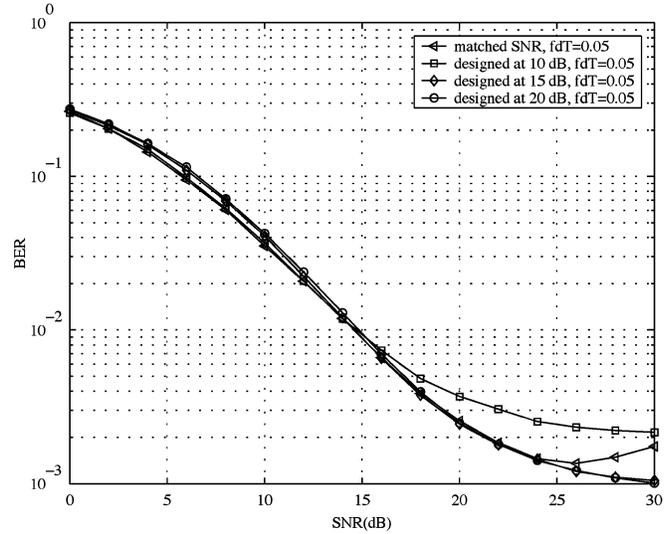


Fig. 2. Mismatched SNR effect to the BER when the design SNR are 10, 15, and 20 dB at $f_d T = 0.05$.

of the past input values $a_i[n-2], a_i[n-4], \dots, a_i[n-2Q]$. The prediction coefficients of this predictor are equivalent to that of a *one-step* linear predictor for single-antenna systems [20] with double maximum Doppler frequency. Therefore, the characteristics of a two-step linear predictor for DSTBC are similar to that of a one-step linear predictor for single-antenna systems. If we plot mean square prediction error versus SNR with different prediction order, we can obtain a rough guideline to select an appropriate order Q . From the plot (not shown), it is found that as the prediction order increases, the mean square prediction error reduces at SNR less than 20 dB. The most significant improvement occurs when the prediction order increases from two to three. At SNR greater than 20 dB, it seems that the linear predictor with order greater than five cannot virtually improve the mean square prediction error. Hence, the fifth order linear predictor is to be used in the later sections.

Note that the VR requires knowledges of Doppler frequency and SNR, which are assumed to be perfectly known at the receiver. In practice, some forms of Doppler frequency and SNR estimation have to be done and there will be errors between the actual and estimated values.

IV. VITERBI RECEIVER WITH MISMATCHED SNR AND FADING RATE

Although there might be some discussion on mismatched SNR or fading rate in other contexts, this section discusses the effect of mismatched SNR or fading rate on the performance of a VR with DSTBC. A particular difference lies on the effect of the so-called *intra-block interference* (defined in Section VI) to the performance, especially at high SNR. It should be noted that although the received signal has varying *instantaneous* SNR, the linear predictor uses fixed *average* SNR value as a design parameter. To gain better performance, it is possible to apply an adaptive linear predictor design (e.g., [9] and [11]) at the price of higher complexity.

Figs. 1 and 2 show the performances of VR using fixed-design SNR. At $f_d T = 0.02$, mismatched SNR value has little effect at

SNR below 20 dB. There exists only a small performance gap between mismatched cases and matched SNR at SNR higher than 20 dB. At $f_d T = 0.05$, the effect of mismatched SNR is more pronounced. At SNR higher than 15 dB, there is a large performance gap between the mismatched design of 10 dB and the matched SNR. However, with the design SNR equal to 15 and 20 dB, the receiver still performs quite well as compared with the matched SNR case.

It should be noted that, for the linear predictor designed with matched SNR, the BER increases with SNR for SNR higher than 25 dB. The reason for this behavior can be given as follows. At high SNR and high fading rate, the intrablock interference is relatively large as compared with the amount of noise. Consequently, the linear predictor designed with matched SNR is not truly optimal at high SNR. We have performed simulations in which the channel gains are kept constant during a transmission block and found that the above behavior did not occur. This verifies the given reason.

This behavior also depends on the order of the linear predictor. Different orders yield different SNR values at which the BER starts to increase with SNR. However, it is not clear at what SNR associated with what fading rate this will occur. From the results shown in Fig. 2, it can be seen that using a fixed lower SNR value for the design of linear predictor than the matched high-SNR value avoids this behavior. In later sections of this paper, when the receiver operates at SNR higher than 20 dB, we use fixed design SNR value of 20 dB.

Similar observation is found with mismatched fading rate effect. The fading rate mismatch has a more pronounced effect at $f_d T = 0.05$ in Fig. 4 than at $f_d T = 0.02$ in Fig. 3. At $f_d T = 0.02$, only a small discrepancy between the performance with mismatched fading rate and that with actual fading rate occurs at SNR higher than 20 dB. Even with the design at $f_d T = 0.01$ or 0.03 ($\pm 50\%$ mismatch), the receiver still works quite well as compared with actual fading rate. However, at $f_d T = 0.05$, the fading rate mismatch causes a large performance degradation. Interestingly, with mismatched fading rate higher than the matched value, the performance has more degradation than with the design fading rate lower than the matched value. For example, with design $f_d T = 0.0375$ (25% mismatch, lower than the actual value), the receiver still works quite well as compared with the design with matched fading rate. However, with design $f_d T = 0.0625$ (25% mismatch, higher than the matched value), the degradation is more distinct at SNR higher than 18 dB. This infers the effect of intrablock interference at high SNR seems to induce slower fading than the actual fading rate, so that the linear predictors using an improved approximate maximum-likelihood receiver for differential space-time block codes over Rayleigh-fading channels lower fading rate (designed at $f_d T = 0.025$, $f_d T = 0.0375$) value perform quite well.

V. VR ANALYSIS

Since the receiver is trellis based, we can apply a standard union-bound approach to derive the upper bound on the bit error probability. The analysis assumes ideally known fading rate and SNR and, therefore, prediction coefficients.

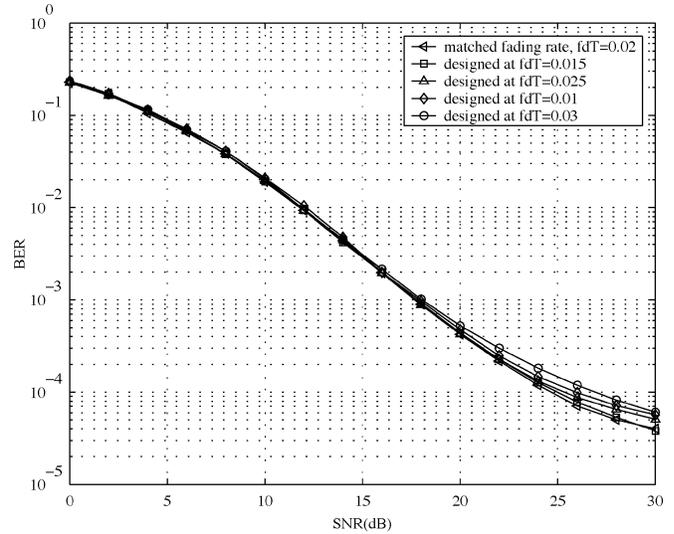


Fig. 3. Mismatched fading rate effect to the BER at $f_d T = 0.02$.

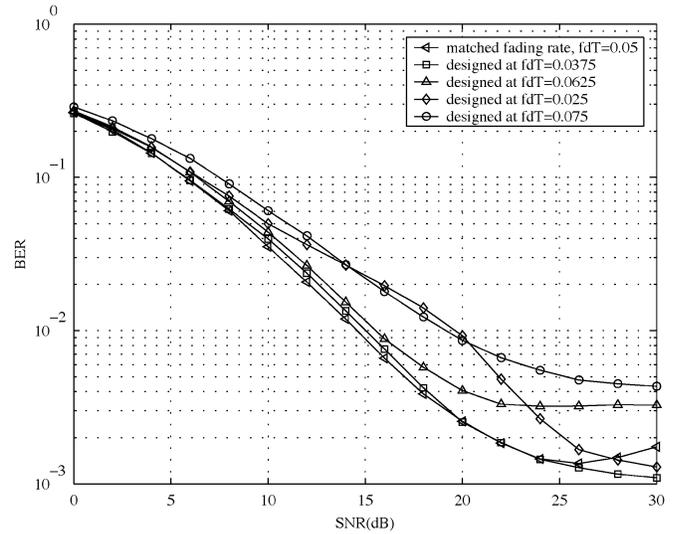


Fig. 4. Mismatched fading rate effect to the BER at $f_d T = 0.05$.

First, the branch metric (8) associated with the transmitted block \mathbf{D}_n is rewritten as

$$\begin{aligned} \Lambda(\mathbf{D}_n) &= \left\| \mathbf{D}_n^H \mathbf{r}_n - \sum_{k=1}^Q b_k \mathbf{D}_{n-k}^H \mathbf{r}_{n-k} \right\|^2 \\ &= \left\| \sum_{k=0}^Q \tilde{b}_k \mathbf{D}_{n-k}^H \mathbf{r}_{n-k} \right\|^2 \\ &= \sum_{k=0}^Q \sum_{k'=0}^Q \tilde{b}_k \tilde{b}_{k'}^* \mathbf{r}_{n-k'}^H \mathbf{D}_{n-k'} \mathbf{D}_{n-k}^H \mathbf{r}_{n-k} \quad (9) \end{aligned}$$

where $\tilde{b}_0 = 1$, $\tilde{b}_k = -b_k$. The error event of length L is defined as the event of the error sequence diverging from the transmitted sequence at epoch (block interval) n th and remerging to the transmit sequence at epoch $(n + L - 1)$ th on the trellis. Since the branch metric (9) is a function of the symbol block associated with the branch as well as the symbol blocks in

the prediction, the pairwise error probability (PEP) depends on the symbols on the error-event paths as well as $Q \times 2$ symbols prior to the error event. Without a loss of generality, suppose the error event starts from epoch zeroth, i.e., the error event of length L starts from \mathbf{D}_0 to \mathbf{D}_{L-1} . Therefore, we define $\mathcal{D}_L = \{\mathbf{D}_{-Q}, \mathbf{D}_{-Q+1}, \dots, \mathbf{D}_0, \mathbf{D}_1, \dots, \mathbf{D}_{L-1}\}$ and $\tilde{\mathcal{D}}_L = \{\tilde{\mathbf{D}}_{-Q}, \tilde{\mathbf{D}}_{-Q+1}, \dots, \tilde{\mathbf{D}}_0, \tilde{\mathbf{D}}_1, \dots, \tilde{\mathbf{D}}_{L-1}\}$.

Let $P(\mathcal{D}_L \rightarrow \tilde{\mathcal{D}}_L)$ denote the PEP of the error event of length L associated with the sequences \mathcal{D}_L and $\tilde{\mathcal{D}}_L$. Similar to [21], an upper bound on the BER can be obtained from the union bound on the number of error bits averaged over transmitted sequences and corresponding error sequences for all error events. The upper bound can be written as

$$P_b < \sum_{L, \mathcal{D}_L, \tilde{\mathcal{D}}_L} \times \frac{P(\mathcal{D}_L)P(\mathcal{D}_L \rightarrow \tilde{\mathcal{D}}_L)e(\mathcal{D}_L \rightarrow \tilde{\mathcal{D}}_L)}{2 \log_2 M} \quad (10)$$

where $P(\mathcal{D}_L)$ is the probability of transmitted sequence \mathcal{D}_L , $e(\mathcal{D}_L \rightarrow \tilde{\mathcal{D}}_L)$ is the number of error bits arising from this error event, and $2 \log_2 M$ is the number of information bits per trellis interval. The PEP of an error event of length L is expressed as

$$\begin{aligned} P(\mathcal{D}_L \rightarrow \tilde{\mathcal{D}}_L) &= P\left(\sum_{n=0}^{L-1} \Lambda(\mathbf{D}_n) > \sum_{n=0}^{L-1} \Lambda(\tilde{\mathbf{D}}_n)\right) \\ &= P\left(\sum_{n=0}^{L-1} \sum_{k=0}^Q \sum_{k'=0}^Q \mathbf{r}_{n-k}^H \right. \\ &\quad \times \left. (\tilde{b}_k \tilde{b}_{k'}^* (\mathbf{D}_{n-k} \mathbf{D}_{n-k}^H - \tilde{\mathbf{D}}_{n-k'} \tilde{\mathbf{D}}_{n-k'}^H)) \right. \\ &\quad \left. \times \mathbf{r}_{n-k} > 0\right). \end{aligned} \quad (11)$$

The left-hand side of (11) can be written as a Gaussian quadratic form $\mathbf{z} = \tilde{\mathbf{r}}^H \mathbf{Y} \tilde{\mathbf{r}}$ where $\tilde{\mathbf{r}} = [\mathbf{r}_{-Q}^t, \mathbf{r}_{-Q+1}^t, \dots, \mathbf{r}_0^t, \dots, \mathbf{r}_{L-1}^t]^t$ is a $2(L+Q) \times 1$ complex vector and \mathbf{Y} is a $2(L+Q) \times 2(L+Q)$ Hermitian symmetric matrix. The matrix \mathbf{Y} is composed of an array of 2×2 block matrices \mathbf{y}_{ij} , $i, j = -Q, -Q+1, \dots, 0, \dots, L-1$ where

$$\mathbf{y}_{ij} = \begin{cases} \mathbf{0}_{2 \times 2}; & |i-j| > Q, \\ \sum_{m=u}^v \tilde{b}_{m-j} \tilde{b}_{m-i}^* (\mathbf{D}_i \mathbf{D}_j^H - \tilde{\mathbf{D}}_i \tilde{\mathbf{D}}_j^H); & \text{otherwise} \end{cases} \quad (12)$$

$\mathbf{0}_{2 \times 2}$ is a 2×2 zero matrix and the indices of the summation u, v are expressed as

$$\begin{aligned} u &= \begin{cases} 0; & i \leq 0 \text{ and } j \leq 0, \\ \max(i, j); & \text{otherwise} \end{cases} \\ v &= \begin{cases} L-1; & i \text{ and } j \geq L-Q-1, \\ \min(i+Q, j+Q); & \text{otherwise.} \end{cases} \end{aligned} \quad (13)$$

The characteristic function of a Gaussian quadratic form \mathbf{z} is given as [22]

$$\begin{aligned} \Phi_{\mathbf{z}}(\xi) &= \frac{1}{\det[\mathbf{I} - j2\xi \mathbf{R}_r \mathbf{Y}]} \\ &= \prod_{\forall \text{ eig}_i(\mathbf{R}_r \mathbf{Y}) \neq 0} (1 - j2\xi \text{eig}_i(\mathbf{R}_r \mathbf{Y}))^{-1} \end{aligned} \quad (14)$$

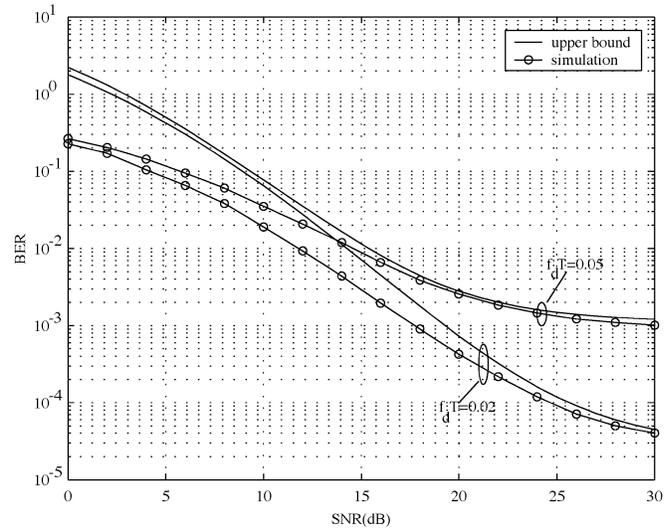


Fig. 5. Upper bound on the bit error probability versus simulation results of the VR.

where $\mathbf{R}_r = E[\tilde{\mathbf{r}}\tilde{\mathbf{r}}^H]$, $\det(\cdot)$ is a determinant of a matrix and $\text{eig}_i(\mathbf{R}_r \mathbf{Y})$ is the i th eigenvalue of the matrix $\mathbf{R}_r \mathbf{Y}$. The PEP in (11) can be evaluated by integration of the probability density function of \mathbf{z} from zero to infinity. Therefore, similar to the result in [21]

$$\begin{aligned} P(\mathcal{D}_L \rightarrow \tilde{\mathcal{D}}_L) &= \int_0^\infty \frac{1}{2\pi} \int_{-\infty}^\infty \Phi_{\mathbf{z}}(\xi) \exp(-j\xi \kappa) d\xi d\kappa \\ &= \frac{1}{2} + \frac{1}{2\pi} \int_{-\infty}^\infty \frac{\Phi_{\mathbf{z}}(\xi)}{j\xi} d\xi. \end{aligned} \quad (15)$$

To evaluate the integral in (15), we can apply a residue theorem that transforms an indefinite integral of a rational function to a summation of residues [23]. The residue theorem yields different forms of results according to the nature of eigenvalues of the matrix $\mathbf{R}_r \mathbf{Y}$.

Since \mathbf{R}_r is positive semidefinite and \mathbf{Y} is Hermitian, it can be shown that all eigenvalues of $\mathbf{R}_r \mathbf{Y}$ are real. Let \mathcal{P} denote a set of nonzero eigenvalues of $\mathbf{R}_r \mathbf{Y}$. For a length L error event, it is observed that there are $4 \times (L-1)$ nonzero eigenvalues. Among these nonzero eigenvalues, half is positive and the other half is negative. Furthermore, these nonzero eigenvalues fall into one of the following two cases.

Case 1) All nonzero eigenvalues are distinct. In this case, the poles of (14) are simple. By applying residue theorem, the PEP can be expressed as

$$P(\mathcal{D}_L \rightarrow \tilde{\mathcal{D}}_L) = 1 - \sum_{p_i \in \mathcal{P}, p_i < 0} \prod_{p_j \in \mathcal{P}, j \neq i} \frac{1}{1 - \frac{p_j}{p_i}}. \quad (16)$$

Case 2) There are $2 \times (L-1)$ distinct nonzero eigenvalues, each with multiplicity two. In this case, let $\tilde{\mathcal{P}}$ denote a set of *distinct* nonzero eigenvalues of $\mathbf{R}_r \mathbf{Y}$. Then, the PEP can be expressed as (17).

The union bound of bit error probability (5) is a function of all lengths of error events, which is an infinite sum, and, therefore, it must be truncated. The union bound is still reliable if the

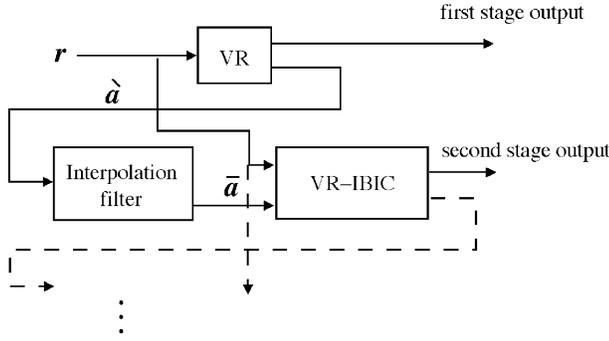


Fig. 6. Multistage receiver (MR) for DSTBC.

bound includes a finite number of dominant error events under certain conditions. When fading is moderately fast and SNR is high, short-length error events are dominant [21]. In this case, the union bound of bit error probability can be evaluated by including only short-length error events.

Fig. 5 shows the simulation performances of VR for DSTBC and the union bound on the bit error probability when $f_d T = 0.02, 0.05$. The bounds include length-two and length-three error events. Including error events with longer lengths does not affect the bound at SNR higher than 16 dB when $f_d T = 0.05$ and at SNR higher than 20 dB when $f_d T = 0.02$. The bound is rather loose when the fading rate is lower or SNR is lower, which is natural for a union upper bound

$$P(\mathcal{D}_L \rightarrow \tilde{\mathcal{D}}_L) = 1 - \sum_{p_i \in \tilde{\mathcal{P}}, p_i < 0} \left(\prod_{p_j \in \tilde{\mathcal{P}}, p_j \neq p_i} \frac{1}{\left(1 - \frac{p_j}{p_i}\right)^2} + \frac{2}{\prod_{p_j \neq p_i} \left(1 - \frac{p_j}{p_i}\right)^2} \cdot \sum_{p_j \neq p_i} \frac{1}{\left(1 - \frac{p_j}{p_i}\right)} \right). \quad (17)$$

VI. MULTISTAGE RECEIVER

Since VR was derived based on the assumption of fixed channel gains during a transmission block while the actual channel varies continuously, the main idea of the new receiver is to try to eliminate the effect of the actual channel gains varying within a transmission block.

With the actual channel, the elements of the received vector \mathbf{r}_n can be written as

$$\begin{aligned} r_{2n-1} &= s_{2n-1}a_1[2n-1] + s_{2n}a_2[2n-1] + w_{2n-1}, \\ r_{2n} &= -s_{2n}^*a_1[2n] + s_{2n-1}^*a_2[2n] + w_{2n}. \end{aligned} \quad (18)$$

Each received sample in (18) is affected from channel gains in its own symbol interval. Next, let us represent the channel gains at even symbol intervals relative to the channel gains at prior symbol intervals by defining $\Delta a_1 = a_1[2n] - a_1[2n-1]$ and $\Delta a_2 = a_2[2n] - a_2[2n-1]$. Now, r_{2n} in (18) can be written in terms of the channel gains at odd intervals as

$$\begin{aligned} r_{2n} &= -s_{2n}^* (a_1[2n-1] + \Delta a_1) \\ &\quad + s_{2n-1}^* (a_2[2n-1] + \Delta a_2) + w_{2n}. \end{aligned} \quad (19)$$

Suppose the Viterbi algorithm is working on the branch in which the transmitted block \mathbf{D}_n is associated, determining $\mathbf{D}_n^H \mathbf{r}_n$ with r_{2n-1} from (18) and r_{2n} from (19) yields the noise-corrupted channel gains

$$\begin{aligned} \hat{a}_1[2n-1] &= s_{2n-1}^* r_{2n-1} - s_{2n} r_{2n} \\ &= a_1[2n-1] + I1_n + s_{2n-1}^* w_{2n-1} - s_{2n} w_{2n} \\ \hat{a}_2[2n-1] &= s_{2n}^* r_{2n-1} + s_{2n-1} r_{2n} \\ &= a_2[2n-1] + I2_n + s_{2n}^* w_{2n-1} + s_{2n-1} w_{2n} \end{aligned} \quad (20)$$

where $I1_n = 0.5\Delta a_1 - \Delta a_2 s_{2n-1}^* s_{2n}$, $I2_n = 0.5\Delta a_2 - \Delta a_1 s_{2n-1} s_{2n}^*$. The values $I1_n$ and $I2_n$ represent *intra-block interference* (IBI). This means that even when the Viterbi algorithm is working on the branch in which the correct transmission block is associated, the computed channel gains are still affected by IBI. Hence, even if there is no noise, one can expect an irreducible error floor caused by IBI with the VR. The amount of IBI depends on how rapidly the channel varies. At a higher fading rate, the average power of IBI is higher. If the channel is fixed during one transmission block, $I1_n$ and $I2_n$ are zero and VR becomes an optimal ML receiver.

From the above discussion, if the channel gains are available at the receiver, the branch metric can be modified to mitigate the effect of IBI. Hence, the new receiver is proposed to operate in more than one stage. The first stage has no channel gains information available while the later stages obtain channel gains information from the prior stages.

Fig. 6 shows the components of the proposed *multistage receiver* (MR). The first stage is composed of only a VR. Each later stage is composed of an interpolation filter and a modified VR called *Viterbi receiver with intra-block interference cancellation* (VR-IBIC). The operation of the multistage receiver is explained as follows:

First stage: The VR performs as usual. In addition, channel estimates computed from the detected sequence are needed. Suppose the receiver obtains the detected sequence $\{\hat{\mathbf{D}}_0, \hat{\mathbf{D}}_1, \dots, \hat{\mathbf{D}}_{N-1}\}$. Then, it computes “rough” channel estimates $\hat{\mathbf{a}}_n = [\hat{a}_1[2n-1] \hat{a}_2[2n-1]]^t$ from

$$\hat{\mathbf{a}}_n = \hat{\mathbf{D}}_n^H \mathbf{r}_n, \quad n = 0, 1, \dots, N-1 \quad (21)$$

and sends them to the next receiver stage.

Second and later stages: The obtained rough channel estimates are present only at odd intervals and are contaminated by IBI, which appears like noise. To reduce the effect of IBI and to obtain channel estimates at even intervals, a low-pass interpolation filter is exploited before the channel estimates enter the VR-IBIC. This is done by padding a zero between each rough channel estimate value. Then, the inputs of the interpolation filter are $[\hat{a}_1[-1], 0, \hat{a}_1[1], \dots, 0, \hat{a}_1[2N-3]], 0$ and $[\hat{a}_2[-1], 0, \hat{a}_2[1], \dots, 0, \hat{a}_2[2N-3]], 0$. The output sampling rate of the interpolation filter is the same as the symbol rate. The filter outputs the “refined” channel estimates $\bar{\mathbf{a}}_n = [\bar{a}_1[n] \bar{a}_2[n]]^t$, $n = -1, 0, 1, \dots, 2N-2$ to be employed at the VR-IBIC.

Now, VR-IBIC computes the IBI from $\bar{I}1_n = 0.5\Delta \bar{a}_1 - \Delta \bar{a}_2 \hat{s}_{2n-1}^* \hat{s}_{2n}$ and $\bar{I}2_n = 0.5\Delta \bar{a}_2 - \Delta \bar{a}_1 \hat{s}_{2n-1} \hat{s}_{2n}^*$ where $\Delta \bar{a}_1 = \bar{a}_1[2n] - \bar{a}_1[2n-1]$, $\Delta \bar{a}_2 = \bar{a}_2[2n] - \bar{a}_2[2n-1]$

and \hat{s}_{2n-1} , \hat{s}_{2n} are the symbols associated with the branch in which the Viterbi algorithm is working. To remove the IBI, the branch metric (8) is modified as (22), where $\bar{\mathbf{I}}_n = [\bar{\mathbf{I}}_{1n} \bar{\mathbf{I}}_{2n}]^t$. In this manner, the IBI is removed from both the channel gains computed from the current received signals and from the prediction. After the removal, the sequence of channel gains becomes more suitable for a two-step linear predictor that has been derived without considering IBI.

If the later stage is to be continued, the rough channel estimates have to be computed again by (21) and are transferred to the later stage receiver.

$$\Lambda_{IC}(\Gamma_n, \mathbf{d}_{n+1}) = \left\| \mathbf{D}_{n+1}^H \mathbf{r}_{n+1} - \bar{\mathbf{I}}_{n+1} - \sum_{k=1}^Q b_k \left(\mathbf{D}_{n+1-k}^H \mathbf{r}_{n+1-k} - \bar{\mathbf{I}}_{n+1-k} \right) \right\|^2. \quad (22)$$

VII. RESULTS AND DISCUSSION

This section compares simulation results between CR, VR, and MR. For MR, the interpolation filter is a low-passed raised-cosine filter with a roll-off factor $\alpha = 0.2$. The cutoff frequency of the interpolation filter is chosen to be some small amount higher than $f_d T$ to avoid cutting high-amplitude spectrum at the edge of the fading spectral band. The cutoff frequency for $f_d T = 0.02$ is 0.0275 and for $f_d T = 0.05$ it is 0.075. The numbers of filter taps are 100 and 80 for $f_d T = 0.02$ and $f_d T = 0.05$, respectively. They are chosen to have enough taps such that the filter retains raised-cosine power spectrum. The number of stages is two.

In Fig. 7, the BER performances of DSTBC are compared when the receivers are MR, CR, and VR. We can see that although VR significantly reduces the error floor associated with CR, some amount of error floor due to IBI still exists, especially at $f_d T = 0.05$. At this fading rate, MR outperforms VR for SNR greater than 12 dB as the channel estimates from the first stage become more reliable. Nevertheless, at $f_d T = 0.02$, MR outperforms VR for SNR greater than 16 dB. The improvement of MR at $f_d T = 0.02$ is not as much as at $f_d T = 0.05$ because, on average, the amount of IBI is smaller at $f_d T = 0.02$. We conclude that MR can further reduce the error floor left from VR with two-stage receiver. As the number of stages is greater than two, however, MR provides virtually no improvement.

Fig. 8 compares the error floor of CR, VR, and MR evaluated at SNR = 30 dB at different $f_d T$. It is seen that MR significantly reduces the error floor from VR and CR. MR can achieve about half-order of magnitude reduction of error floor over VR.

VIII. CONCLUSION

In this paper, an approximate ML receiver for Tarokh's differential space-time block codes is investigated and the effect of SNR and fading rate mismatch is studied. It is shown that the mismatched cases are more sensitive when the fading rate is high. An upper bound on the bit error probability is derived and shown to be tight when fading rate and SNR are high. In order to improve the performance of an approximate ML receiver, the multistage receiver is proposed, which accounts for the intra-

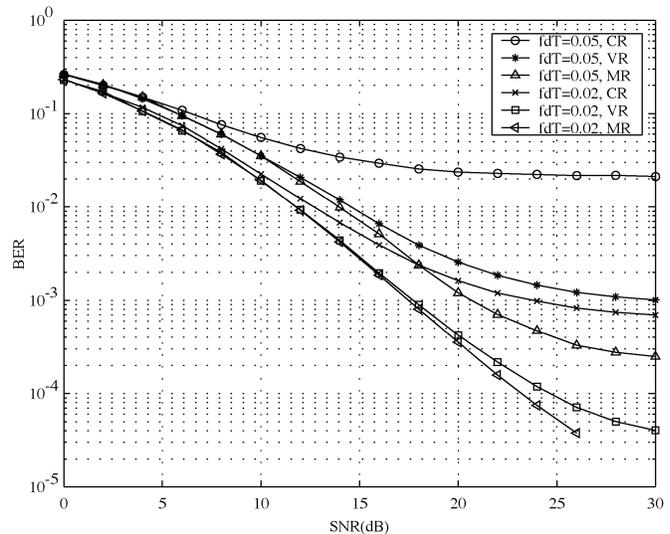


Fig. 7. Performance of DSTBC with BPSK at $f_d T = 0.02$ and $f_d T = 0.05$ with CR, VR, and MR with two stages.

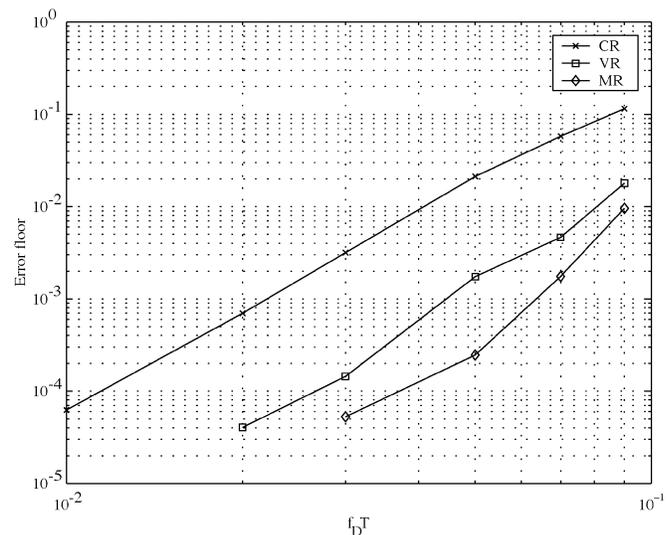


Fig. 8. Error floor of DSTBC with CR, VR, and MR with two stages evaluated at SNR = 30 dB.

block interference caused by the time-varying characteristic of the channel within a transmission block. With some additional complexity, multistage receiver provides further improvement over an approximate ML receiver. With a two-stage receiver, the error floor is reduced by about half order of magnitude left from an approximate ML receiver.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their constructive comments.

REFERENCES

- [1] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.
- [2] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1169–1174, July 2000.

- [3] B. L. Hughes, "Differential space-time modulation," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2567–2578, Nov. 2000.
- [4] B. M. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Trans. Commun.*, vol. 48, pp. 2041–2052, Dec. 2000.
- [5] R. Schober and L. H.-J. Lampe, "Noncoherent receivers for differential space-time modulation," *IEEE Trans. Commun.*, vol. 50, pp. 768–777, May 2002.
- [6] C. B. Peel and A. L. Swindlehurst, "Performance of unitary space-time modulation in a continuously changing channel," in *Proc. IEEE ICC '01*, vol. 9, June 2001, pp. 2805–2808.
- [7] P. Ho and D. Fung, "Error performance of multiple-symbol differential detection of PSK signals transmitted over correlated Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 40, pp. 1566–1569, Oct. 1992.
- [8] D. Divsalar and M. K. Simon, "Maximum-likelihood differential detection of uncoded and trellis coded amplitude phase modulation over AWGN and fading channels—metrics and performance," *IEEE Trans. Commun.*, vol. 42, pp. 76–89, Jan. 1994.
- [9] F. Adachi, "Adaptive differential detection using linear prediction for M -ary DPSK," *IEEE Trans. Veh. Technol.*, vol. 47, pp. 909–918, Aug. 1998.
- [10] R. Schober, W. H. Gerstacker, and J. B. Huber, "Decision-feedback differential detection of MDPSK for flat Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, pp. 1025–1035, July 1999.
- [11] R. Schober and W. H. Gerstacker, "Decision-feedback differential detection based on linear prediction for MDPSK signals transmitted over Ricean fading channels," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 391–402, Mar. 2000.
- [12] G. M. Vitetta and D. P. Taylor, "Maximum likelihood decoding of uncoded and coded PSK signal sequences transmitted over Rayleigh flat-fading channels," *IEEE Trans. Commun.*, vol. 43, pp. 2750–2758, Nov. 1995.
- [13] G. Colavolpe, P. Castoldi, and R. Raheli, "Linear predictive receivers for fading channels," *Electron. Lett.*, vol. 34, pp. 1289–1290, 25, 1998.
- [14] C. Cozzo, D. N. Mohile, and B. L. Hughes, "Multiple-symbol detection of differential space-time modulation," in *Proc. 34th Ann. Conf. Information Sciences and Systems*, Princeton, NJ, Mar. 2000.
- [15] C. Ling and X. Wu, "Linear prediction receiver for differential space-time modulation over time-correlated Rayleigh fading channels," in *Proc. IEEE ICC '02*, vol. 2, Apr. 2002, pp. 788–791.
- [16] P. Fan, "Multiple-symbol detection for transmit diversity with differential encoding scheme," *IEEE Trans. Consum. Elect.*, vol. 47, pp. 96–100, Feb. 2001.
- [17] C. Gao and A. M. Haimovich, "Multiple-symbol differential detection for space-time block codes," in *Proc. 36th Ann. Conf. Information Sciences and Systems*, Princeton, NJ, Mar. 2002.
- [18] E. Chiavaccini and G. M. Vitetta, "Further results on Tarokh's space-time differential technique," in *Proc. IEEE ICC '02*, vol. 3, Apr. 2002, pp. 1778–1782.
- [19] H. Jafarkhani and V. Tarokh, "Multiple transmit antenna differential detection from generalized orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2626–2631, Sept. 2001.
- [20] D. Makrakakis, P. T. Mathiopoulos, and D. P. Bouras, "Optimal decoding of coded PSK and QAM signals in correlated fast fading channels and AWGN: a combined envelope, multiple differential and coherent detection approach," *IEEE Trans. Commun.*, vol. 42, pp. 63–75, Jan. 1994.
- [21] B. D. Hart and D. P. Taylor, "Maximum-likelihood synchronization, equalization, and sequence estimation for unknown time-varying frequency-selective Rician channels," *IEEE Trans. Commun.*, vol. 46, pp. 211–221, Feb. 1998.
- [22] M. Schwartz, W. R. Bennett, and S. Stein, *Communication Systems and Techniques*. New York: McGraw-Hill, 1966.
- [23] A. D. Wunsch, *Complex Variables With Applications*. Reading, MA: Addison-Wesley, 1983.



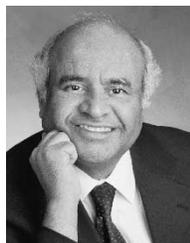
Poramate Tarasak (S'97) received the B.Eng. degree in electrical engineering from Chulalongkorn University, Bangkok, Thailand, in 1997 and the M.Eng. degree in telecommunications from Asian Institute of Technology (AIT), Pathum Thani, Thailand, in 1999. He is currently working toward the Ph.D. degree in electrical engineering at the University of Victoria, Victoria, BC, Canada.

Since September 2000, he has been a Research Assistant in the Department of Electrical and Computer Engineering, University of Victoria. His research interests include space-time coding and MIMO systems.



Hlaing Minn (S'99-M'01) received the B.Eng. degree in electronics from Yangon Institute of Technology, Yangon, Myanmar, in 1995, the M.Eng. degree in telecommunications from the Asian Institute of Technology (AIT), Pathum Thani, Thailand, in 1997, and the Ph.D. degree in electrical engineering from the University of Victoria, Victoria, BC, Canada, in 2001.

He was with the Telecommunications Program, AIT, as a Laboratory Supervisor in 1998. He was a Research Assistant from 1999 to 2001 and a Postdoctoral Research Fellow in the Department of Electrical and Computer Engineering, University of Victoria, in 2002. Since September 2002, he has been with the Erik Jonsson School of Engineering and Computer Science, the University of Texas at Dallas, as an Assistant Professor. His research interests include wireless communications, statistical signal processing, error control, detection, estimation, synchronization, and signal design.



Vijay K. Bhargava (S'70–M'76–SM'82–F'92) received the B.Sc., M.Sc., and Ph.D. degrees from Queen's University, Kingston, ON, Canada in 1970, 1972, and 1974, respectively.

He was with the University of Victoria, Victoria, BC, Canada, from 1984 to 2003 and with Concordia University, Montréal, PQ, Canada, from 1976 to 1984. Currently, he is a Professor and Head of the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, Canada. He is a coauthor of *Digital Communications by Satellite* (New York: Wiley, 1981) and coeditor of *Reed–Solomon Codes and Their Applications* (New York: IEEE Press, 1994) and *Communications, Information and Network Security* (Boston, MA: Kluwer, 2002). His research interests are in wireless communications.

Dr. Bhargava is a Fellow of the B.C. Advanced Systems Institute, Engineering Institute of Canada (EIC), and the Royal Society of Canada. He is a recipient of the IEEE Centennial Medal (1984), IEEE Canada's McNaughton Gold Medal (1995), the IEEE Haraden Pratt Award (1999), the IEEE Third Millennium Medal (2000), and the IEEE Graduate Teaching Award (2002). He is very active in the IEEE and was nominated by the IEEE Board of Directors for the Office of IEEE President-Elect and is a Past President of the IEEE Information Theory Society. He currently serves on the Board of the Communications Society and is an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS and IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.